The US Decennial Census: Political Questions, Scientific Answers

KENNETH PREWITT

WHY HAVE PREPARATIONS for the US population and housing census in 2000 been the target of sharp partisan battles, battles that will also affect how the census counts will be used? The answer takes us, first, back to the political origins of the decennial census.

The census in the US Constitution

The US Constitution, written in 1787, includes a provision (in Article I, Section 2) for a decennial census:

> Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers. . . . The actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they shall by Law direct.

America's founding political elites crafted this paragraph with two nation-building purposes in mind. The bold new Constitution they were writing would replace a weak confederacy with a centralized and stronger federal government. But as with any engineering of a new government, the first and difficult task is to build consensus on how power will be allocated. Geography, not estates or social classes, was the building block. The confederacy that was being pushed aside had given each of the original 13 states equal voting power. This favored smaller states; their leaders at the Constitutional Convention naturally insisted that the principle of state equality be upheld in the new government. The more populous states, such as New York and Virginia, balked at such power sharing, arguing the fairness of

allocating political power proportionate to population size. A bicameral legislature offered the compromise. In one branch, the Senate, each state would be equal in its voting power; in the other, the House of Representatives, voting would be allocated proportionate to population. Of course "proportionate to population" meant that there must be a count of the population; consequently, the Constitutional innovation of providing for a census.

But why a census every ten years? Because of the second nation-building task. Geographic expansion was much on the mind of the new nation. Its restless people were already crossing the Appalachian mountain range, spreading westward into the Ohio Valley and down the Mississippi River to the Gulf of Mexico. Some, among them drafters of the new Constitution, even imagined a nation reaching across the continent to the Pacific Ocean. A vast territory would thus be added to the new nation. But what was to be the status of these soon-to-be-acquired territories? Would they be annexed as colonies or join the nation on equal footing with the original 13 states? The prospect of an empire with colonies did not sit well with the principles of equality for which the war of independence had been waged. New and equal states it was to be, and the nation added states steadily throughout the nineteenth century. The decennial census measured population growth and its geographic dispersion and thereby served to regulate the pace at which the western and southern territories were added as new states to the Union.[1]

In serving these two Constitutional purposes—reallocation of power as the population grows and regulating the expansion of the union—the decennial census was intended to serve state-building, that is, "political" purposes; no sense can be made of current census controversies without appreciating this basic fact.

A troubling shift in the politics of the census

The decennial mechanism worked as anticipated by the founders. Every ten years, with one exception, seats in the House of Representatives have been reallocated as the population increased and moved westward. The exception is 1920. Southern, rural members of Congress did not welcome the 1920 census results that documented the massive wartime population movement from the rural south to northern cities. Reapportionment was delayed until after the 1930 census.[2]

The 1920s debate about census results was intensely partisan; and this was not the first political battle over the census. George Washington rejected a formula designed by New York's Alexander Hamilton for allocating seats after the 1790 census, a step that led to a substitute formula more favorable to agricultural interests proposed by Thomas Jefferson of Virginia. Partisan concerns dominated during the pre–Civil War period as the acceptance of new states to the Union was carefully paced to maintain the balance between pro- and antislavery forces. For the most part, however, across

nearly two centuries the partisan battles over the decennial census focused on how the census counts were to be used, not on how the census was to be conducted. Although there were grumbles about the completion and accuracy of counts as early as 1790, the counts were accepted as the best approximation available, and politics intruded at the point of interpretation and application.

This circumstance has changed in the last 15 years, with far-reaching and troubling consequences for the US decennial census. Political leaders now argue that the census can be designed to predetermine the apportionment outcome in a way that favors one or the other of the political parties. This invites partisan disagreement not just about how the census count should be interpreted or applied, but how the census should be designed.

This shift in political thinking began with legal action focused on the 1980 census and gradually gained momentum throughout the decade. It was brought into sharp focus when, in 1987, political appointees of the Republican administration in the Department of Commerce (the governmental home of the Census Bureau) overrode the professional judgment of senior Census Bureau staff, who are nonpolitical civil servants, about a key design element for the 1990 census. There was concern among a number of congressional members and mayors of large cities, mostly Democrats, that the Commerce Department had based its ruling less on technical or budgetary grounds, as it claimed, than on partisan calculations. The Department's decision was litigated, eventually leading to a court-enforced agreement that the Census Bureau could implement a modification of its recommended design for the 1990 census, but leaving unresolved how the count generated by that design might be used. In this agreement the court was protecting two centuries of census practice: the *collection* of the census information should be the responsibility of a professional Census Bureau, but the *application* of the numbers would be decided in the legal/political arena.[3]

Following the completion of the 1990 census, it fell to the Secretary of Commerce to decide whether to use the census results that had been produced by the disputed procedure. The Secretary, rejecting the recommendation of the Census Bureau Director, decided not to use those results, reasoning that:

> [T]he choice of the adjustment method selected by the Bureau officials can make a difference in apportionment, and the political outcome of that choice can be known in advance. I am confident that political considerations played no role in the Census Bureau's choice of an adjustment model for the 1990 census. I am deeply concerned, however, that adjustment would open the door to political tampering with the census in the future.[4]

This passage, as best I can determine, is the first instance in American political history in which a high government official gives voice to the speculation that the nonpartisan, professionally managed Census Bureau might

choose a data-collection methodology so as to favor one political party over another.

Within a few years the partisan concern over what the Census Bureau might do had become the facile accusation that the proposed methodology for Census 2000 was in fact designed to achieve a given partisan outcome. A widely circulated memo to this effect was sent by Jim Nicholson, writing as Chairman of the Republican National Committee, to the state Republican Party leaders on 20 May 1997:

> I am contacting you to recruit your assistance in addressing an issue of unusual importance to the future of the Republican Party. At the heart of the matter is one of the federal government's most fundamental Constitutional functions: the United States census. At stake is our GOP majority in the House of Representatives, as well as partisan control of state legislatures nationwide.
>
> The Clinton Administration is implementing a radical new way of taking the next census that effectively will add nearly *four and one-half million Democrats* to the nation's population. This is the political outcome of a controversial Executive decision to use a complex mathematical formula to estimate and "adjust" the 2000 census. . . .
>
> The GOP would suffer a negative effect in the partisan makeup of 24 Congressional seats, 113 State Senate seats and 297 State House seats nationwide. . . . An adjusted census could provide Democrats the crucial edge needed to prevail in close contests to control several state legislative chambers.[5]

Readers will seriously underestimate the importance of this memo if they conclude that it is only Republican Party officials who have concluded that the Census Bureau could and would propose a data-collection methodology with a given partisan outcome in mind. Census methodology has become fair game across the political spectrum—a deeply disturbing development.

What, then, brought us to this point?

The stakes are raised: Social justice and federal funds

To understand the current controversy over the decennial census, we first take note of developments in the years following World War II that substantially raised the political stakes.

The 1960s introduced notions of "social justice" to the politics of reapportioning seats in the House of Representatives following each decennial count. A 1962 Supreme Court ruling advanced the "one person, one vote" principle, which prohibited legislative districts of unequal population size. Prior to this ruling the districts from which members of Congress were selected could have hugely different numbers of voters resident, in effect di-

luting the voting power of the mostly urban districts of high population and enhancing the voting power of the mostly rural districts of low population. The politics behind this court ruling continue the long tradition of partisan battles over census results that primarily affect geographic interests—rural versus urban, southern versus northern.

Partisan differences in this period also took into account differing demographic divisions in American society that partially overlapped the traditional geographic battle lines. Here the key legislative initiative was the Voting Rights Act of 1965, designed to allow equal access by racial minorities to the voting booth. Among other provisions, federal law now required that legislative district boundaries be drawn in a nondiscriminatory manner. To enforce one person, one vote and nondiscriminatory voting districts, the Department of Justice requires data in rich geographic detail. These data are generated by the decennial census. A breakdown of the population by race and ethnicity is now provided at the block level, and these data are used to determine whether various voting districts have been designed to dilute the voting power of racial minorities. Additional social justice responsibilities adhered to the decennial census in the wake of ongoing civil rights and equal employment legislation. The census provides the denominator against which to determine whether racial, gender, and ethnic patterns in employment, housing, education, and other social services reflect discriminatory intent. Federal departments making use of census data in this connection include, especially, the Departments of Labor, Housing and Urban Development, Education, and Health and Human Services.

The postwar decades added yet a further task to the decennial census: allocation of federal funds. The relationship of census counts to resources was an on-and-off matter for most of the nineteenth and early twentieth centuries. There was a brief use of the census for taxation in the early nineteenth century, but this quickly fell away. And the early twentieth century saw scattered instances of using census-based formulas to distribute federal funds. The task of resource allocation increased in importance in the 1950s and 1960s. As summarized by Margo Anderson:

> Congress expanded federal aid to state and local governments through the grant-in-aid system. Congress enacted programs that used statistically based funding formulas to cover federal support to states and localities for school lunch programs, airport construction, and hospital construction in 1946, and for water pollution control in 1948. In the 1950s and 1960s, the grant-in-aid [system] was used to supply federal funds for the interstate highway system, housing assistance, anti-poverty programs, employment and training programs, urban redevelopment, and water and sewer projects.[6]

As we begin the twenty-first century, the annual dollar amount of federal funding that uses census data as one of the criteria for program design

TABLE 1 Summary of annual funds allocated by federal agencies through funding formulas and eligibility requirements based on census data as of 1999

Federal agency	Total amount (dollars)
Department of Housing and Urban Development	7.2 billion
Department of Transportation	25.4 billion
Department of Education	12.6 billion
Department of Health and Human Services	133.5 billion
Department of Justice	1.6 billion
Environmental Protection Agency	2.3 billion
Department of Agriculture	10.7 billion
Department of Commerce	351 million
Department of Labor	4.5 billion
Total	198.2 billion

SOURCE: General Accounting Office, US Government. >http://www.gao.gov<

and allocation approaches $200 billion (see Table 1). Barring radical change in how the federal government works, the allocation of more than $2 trillion will be affected by Census 2000 counts before the next decennial census in 2010. This already huge number does not include state and local spending and private-sector investments, which together would add many millions more.

Given these stakes—a socially just election system and huge flows of federal funds—it is no surprise that there is a partisan edge to the focus on census numbers. But this in itself does not explain the particular battle that erupted over census methodology. For that explanation we retrace our steps, returning to the 1940s and the differential undercount.

The differential undercount

The significant expansion of the purposes to which census data are put occurs in the wake of the "discovery" of what has come to be known as the *differential undercount.* A census undercount is a measure of how many persons short of 100 percent are not recorded in a census. An undercount of some magnitude has occurred in every US decennial census since 1790, when George Washington complained that a combination of citizen resistance and flaws in the enumeration procedure produced an official count that fell short of the true number of residents in the new country. In no decennial census since has the enumeration process been able to account for every resident in the United States, nor does it do so in any other nation. Demographers everywhere assume that a census is an approximation of the true count, perhaps an overestimation but more likely an underestimation.

TABLE 2 Percent net undercount for black and white males in decennial US censuses, 1940–90

	1940	1950	1960	1970	1980	1990
Black	8.4	7.5	6.6	6.5	4.5	4.4
Nonblack	5.0	3.8	2.7	2.2	0.8	1.2
Overall	5.4	4.1	3.1	2.7	1.2	1.6

NOTE: The net undercount is the result of the total overcount minus the total undercount; the point estimates presented here are based on demographic analysis.
SOURCE: U.S. Census Bureau, summarized in "The Census: A history," Population Resource Center, 1999. Additional documentation available at >www.census.gov<

Insofar as the benefits of a census are allocated on a share basis—as is true if the benefit is a fixed number of congressional seats or a fixed amount of federal funds—an undercount distributed equally across geographic units and population groups does not result in inequitable outcomes. Inequity or injustice emerges only if some areas or groups are counted at lower rates than others: that is, if the undercount is differential. That we know to be the case for the US census.

Although census professionals had long accepted that the undercount rate differed from one demographic group to another, circumstances related to World War II provided the first systematic measure. In the early 1940s the government initiated mandatory, universal selective service registration. Though obviously not its intention, this universal registration provided statisticians with two independent counts of males between ages 21 and 35: the count recorded in the 1940 census and the count of those registered for potential enlistment to military service. Comparison of these counts provided the first reliable measure of how many persons, at least in this demographic group, had been missed in the census.

Save for one factor, this finding would have attracted little interest beyond the demographers and statisticians concerned with improving census practice. What attracted wider interest was that African-American males of draft age had been missed in the census at much higher rates than white males. Here was the first systematic evidence of a differential undercount.[7]

Using a variety of techniques, the Census Bureau began to study the differential undercount—which, we now know, has persisted across the last half-century, as illustrated in Table 2.

The first important national examination of the implications of a differential undercount was a 1967 Conference on Social Statistics and the City. The summary report drew the by-then-obvious conclusion:

> Where a group defined by racial or ethnic terms, and concentrated in special political jurisdictions is significantly undercounted in relation to other groups,

> then individual members of that group are thereby deprived of the constitutional right to equal representation in the House of Representatives and, by inference, in other legislative bodies. They are also deprived of their entitlement to partake in federal and other programs designed for areas and populations with their characteristics. In other words, miscounting the population could unconstitutionally deny minorities political representation or protection under the Voting Rights Act. It could also deny local jurisdictions grant funds from federal programs.[8]

These themes are now commonplace in public discussion about census design issues—repeated on editorial pages and in congressional testimony, scholarly studies, civil rights newsletters, and Census Bureau documents. As a result, no one with even a passing interest in the decennial census can have failed to grasp that fair and equitable outcomes are jeopardized if the counts are more complete for some groups than others. And when the undercounted are racial and ethnic minorities, it is self-evident that the politics of representation will become framed as a civil rights issue.

With significant issues of fairness at stake, what was the Census Bureau to do about the persistent differential undercount?

Census methodology: Dual system estimation

There are two ways to measure the differential undercount. One method compares census counts with population estimations based on demographic analysis. These estimations take into account births and deaths from vital statistics and rates of immigration and emigration and calculate the number of persons resident in the United States. These data, at the national level, can be used to estimate the population of selected racial groups, which can then be compared to census counts of the same groups to calculate whether underenumeration in the census differs by race.

The second method derives from sample survey data collected independently of the census, but for the same time frame, that are then matched with census records. This matching process is the basis for calculating the rate of under- and overenumeration of population groups in the census counts. This statistical method, known in wildlife studies as capture/recapture, has a major advantage over demographic analysis. It permits the calculation of undercount rates for many more population groups and in greater geographic detail than national or state estimates.[9]

Moreover, dual system estimation allows the initial census count to be adjusted to correct for any detected under- or overcounting in the original census. Nicholson, quoted earlier, had dual system estimation in mind when he wrote of "a controversial Executive decision to use a complex mathematical formula to estimate and 'adjust' the 2000 census." In short, dual

system estimation and the correction of the initial census count are at the center of the partisan dispute over methodology for the 2000 census, as they were also in the 1990 census and to a lesser extent in 1980.

My task here is not to assess the strengths or weaknesses of dual system estimation when applied to a national census. But note should be taken that a number of distinguished statisticians have expressed reservations about its application for reasons that have nothing to do with prospective partisan outcomes. Citing operational and statistical considerations, these experts conclude that trying to correct for the under- and overenumeration can introduce new sources of error in the census.[10] Other statisticians agree with the judgment of the Census Bureau that an adjusted census based on dual system estimation yields a closer approximation of the true population count than an unadjusted census.[11]

Note should also be taken that adjusted counts move legislative seats and federal funds only at the margin. That is, although all of the seats in the House of Representatives (435) and significant federal funds are distributed on the basis of census counts, the changes in shares produced by adjusted counts are relatively small. The political debate, however, often focuses on the large totals rather than the comparatively small changes in shares, and thus tends to attribute exaggerated consequences to undercount rates.[12]

Census 2000 and the differential undercount

Early in its planning for Census 2000, the Census Bureau concluded that the differential undercount, which has been measured over a half-century but probably has characterized every decennial census for 200 years, could not be eliminated or even substantially reduced through traditional enumeration methods. No US population census can reach and solicit a response from every resident in the country. There will be evasion, refusals, and fraudulent responses; there will be missed housing units and incomplete counts within houses; there will be undetected errors in census operations. In short, there will be an undercount. And some population groups—the linguistically isolated, the less socially integrated, recent immigrants, those without legal status, the highly mobile, the poor and uneducated, those indifferent to civic responsibility—are more likely to be missed. The undercount, in short, will be differential.

The Census Bureau accordingly proposed to use a post-census sample and dual system estimation as part of its methodology. It further proposed to apply dual system estimation techniques to adjust the under- and overcounts. For reasons by now obvious, this proposal was strongly opposed by Republican Party leaders in Congress. The Speaker of the House sued the Department of Commerce to preclude the use of sampling for apportionment.[13]

In January 1999, a closely divided Supreme Court ruled that the federal statute governing the Census Bureau prohibited sampling for reapportionment purposes. The ruling appears to require the use of sampling (if feasible) for purposes other than reapportionment. These other purposes include the use of census data for redistricting and for the formulas that distribute federal funds.

The Census Bureau then designed census operations to conduct as full a basic enumeration as possible in order to deliver state-by-state apportionment counts by the statutory deadline of 31 December 2000. In addition, the Bureau plans to collect data from a large independent sample of approximately 300,000 households. Results from this sample will be matched to the census counts in those households, and estimates of under- and overcounting will be calculated. These estimates can then be used to correct the original enumeration-based census, and the corrected counts will then be provided for purposes of redistricting and federal funding formulas. It is likely that legal action will be taken by states and cities that believe they will be either harmed or helped by the use of the corrected census counts. In particular, there is likely to be a legal effort to prohibit the use of adjusted census data in the state-by-state redistricting process. It is also likely that legal action will be initiated by advocates for the use of corrected numbers. As of this writing, therefore, the only certain outcome is that apportioning congressional seats across the 50 states will not involve any correction for the differential undercount.

Additional Census 2000 design issues

Although the differential undercount has attracted the most political attention, other design issues have not gone unnoticed, and here I make brief mention of three topics: the cost of the decennial census, the questions asked in the census, and the response rate or level of public cooperation with the census.

Following the 1990 census, widely described as "less accurate" than the one that preceded it (see Table 2), the Census Bureau was instructed to prepare for a census in 2000 that would be more accurate and less expensive. The only design available to meet this exacting charge is one that employs sampling methods. Basic enumeration is labor intensive, therefore expensive; it cannot reach everyone, therefore inaccurate. The initial design for Census 2000 recommended sampling at two points. One use, as noted above, was for dual system estimation to correct for the differential undercount. Sampling was also recommended as part of the initial enumeration procedures. In this application, the census would enumerate 90 percent or more of every census tract and then estimate the characteristics of the remaining nonresponders on the basis of a sample survey.

The Supreme Court ruling that prohibited sampling for apportionment required the Census Bureau to abandon sampling during nonresponse followup as well as for dual system estimation. In other words, the earlier instruction that Census 2000 be less expensive and more accurate was impossible to meet.

In the tradeoff between cost and accuracy, the Census Bureau opted for the latter and found the US Congress to be in agreement. The Bureau's strategy was to increase enumeration accuracy by increasing public cooperation, both in the mailback phase of the census and in the nonresponse followup phase. In this effort Census 2000 includes a number of new operations: three first-class mailings to every housing unit (an alert letter, the census form, a reminder card); a costly paid-advertising campaign; extensive promotion activities such as a census-in-schools effort and paid partnership specialists; questionnaire assistance centers staffed by 15,000 paid employees; assistance guides in 49 languages; and so forth. These operations have been designed in response to an assessment that the population is harder to count than it was in 1990—the result of a higher proportion of new immigrants and of linguistically isolated households, an increase in nonstandard living arrangements in which no one feels responsible for others in the household, demanding work schedules and greater mobility, and declining levels of civic engagement. The several innovations in publicity/promotion/assistance operations substantially increased the cost of the census, well above the standard increase that can be attributed to inflation and to growth in population and housing units since 1990. Full cycle costs of Census 2000 will exceed $6.5 billion.

A related effort to increase mailback response rates was to simplify the census form, making it more user-friendly and reducing respondent burden. The former was achieved by shifting from a Fosdic data scanning technology (which is more demanding on the respondent) to intelligent character recognition that permits a questionnaire design easier to use in a self-enumeration context.

To reduce respondent burden, the census in 2000 eliminated a number of questions and shifted others from the short to the long form.[14] No questions are asked unless they are required by law or are necessary for an existing federal program. Short-form data can be used to estimate individual and household characteristics at the block level, and thus for the most part are restricted to those items necessary for enforcement of the Voting Rights Act of 1965. Long-form information is used to estimate characteristics at higher levels of geography. Of course, what is or is not asked in a census or, in some instances, what is on the short or long form, attracts keen political attention.

One such example in 2000 had to do with the shift of the marital status question from the short form to the long form, which raised concern in

Congress that the Census Bureau was taking a stand on (and slighting) family values. The Census Bureau has no policy position on marriage or on any other of the dozens of items on the long form: education, veteran status, language, occupation, income. It has simply been determined that information on these items is not needed at the block level. The political cry over the placement of the marital item—which led to a "Sense of the Senate" resolution endorsing family values and urging that the census respect these values—is interesting only in what it reveals about the political symbolism of a census.

Of far greater relevance to social policy is a provision in Census 2000 to allow respondents to self-identify as belonging to more than one racial group.[15] It is doubtful whether more than a few percent of the population will describe themselves in Census 2000 as multiracial, but the proportion will increase as interracial marriage becomes more common. This expected change in self-identification has long-term and unpredictable consequences for race-based social policy. Laws prohibiting racial or ethnic discrimination in education, housing, employment, or how voting districts are drawn assume a small number of fixed racial or ethnic groups, primarily black, Hispanic, and Asian, though sometimes also including American Indians, Alaskan Natives, and Native Hawaiian and Pacific Islanders.[16] Enforcement often requires a numerator against which to measure observed practice: for example, if 12 percent of the population is black, but only 6 percent of, say, medical professionals are black, might this indicate racial discrimination? Census data are frequently the source of the numerators used to enforce nondiscriminatory social policies. With the proliferation of different multiracial groups in society and the general blurring of racial boundaries, the future of such enforcement is unclear.[17]

An underlying problem facing Census 2000 is the decline in mailback response rate. Since 1970 the decennial census has depended on a mailout/mailback design. The census starts with a comprehensive address file and then delivers the census form to every residence in the country.[18] In 1970 approximately 78 percent of households returned the form by mail. This dropped slightly to 75 percent in 1980, and then more sharply to 65 percent in 1990. Early planning for Census 2000 suggested that the response rate would be approximately 61 percent, and on that estimate the Census Bureau has budgeted and designed its operational requirements.

Reversing the decline in public cooperation has become a major challenge. Low response rates negatively affect the quality of census data and increase costs. They are also indicative of a lack of civic obligation among large segments of the population. In fact, some thoughtful observers question whether the use of sampling in the nonresponse phase might further weaken civic obligation as persons concluded that the experts, deploying advanced statistical methods, could count the population irrespective of their cooperation. Although this concern is now moot, because sampling for

nonresponse will not be used in 2000, it helps raise the broader question of whether the decennial census can be presented as a "civic ceremony" based on the idea of civic responsibility.

It remains to be seen whether the decline in response rate can be reversed, but Census 2000 will make the effort to do so. It will heavily publicize the unsatisfactory 1990 mailback response rate, and, in a campaign labeled "90 Plus Five," challenge communities across the nation to improve their 1990 response by five percentage points. The overall promotion/publicity effort in the 2000 census might be viewed as a very large applied social science project. To what extent can this effort "talk the nation" into a higher level of public cooperation than our models indicate to be the natural level?

This section has briefly covered three issues that are central to Census 2000, albeit not of a character as to have invited the same kind of intense partisan battle associated with using dual system estimation to correct for the differential undercount. I return to that issue and offer concluding comments on the politics of Census 2000.

The political battle revisited

To summarize the argument advanced in this article, the decennial census was established in 1787 to serve the political purpose of allocating seats in the House of Representatives. In the course of two centuries the census has had added to it two other major political tasks: federal fund allocation and enforcement of nondiscriminatory elections. When a census carries this much political weight, it is not surprising that partisan differences have emerged from time to time. Historically, these differences generally have concerned the application of census counts. In recent years a significant shift has occurred, with a heated battle emerging over the methodology of census-taking.

This battle was initiated by the charge that the decennial census could be designed to predetermine partisan outcomes in the reapportionment and redistricting processes. As Census 2000 was planned, the charge shifted from "could be" to "is." The accusation that the Census Bureau has acted and would act on a narrow partisan agenda fails to take into account two counter forces.

First, neither the traditions nor the competencies of the Census Bureau are consistent with advancing a partisan agenda. By tradition the Bureau is resolutely professional and apolitical.[19] Its peer community is the professional statistical community worldwide. It earns respect in that community by providing accurate data. Only persons uninformed about the importance of peer approval to professionals could believe that the Census Bureau would exchange its high standing among statisticians and demographers for a short-term political purpose. Moreover, the Census Bureau

enjoys the confidence of the American public—a confidence that is indispensable for public cooperation with its surveys and studies (the decennial census being only one of many Census Bureau activities). Even if its traditions were to allow it, the Census Bureau does not have the competence to predetermine partisan outcomes. It has no statistical expertise in reapportionment or redistricting and no expertise on trends in voting behavior. To predetermine partisan outcomes, the Census Bureau would need to bring to bear such expertise when it selected data-collection methodologies several years in advance of when census counts are used for reapportionment or redistricting.

A second consideration is that the decennial census is designed and executed under a spotlight. For Census 2000, for example, four congressional committees have oversight responsibility; continuous examinations are undertaken by the General Accounting Office, an investigatory arm of Congress, and also by the Inspector General, an investigatory arm of the Executive Branch; a specially appointed Census Monitoring Board (well-funded and well-staffed) is given the explicit charge to report to the Congress and the President as to whether there is any evidence of political bias in census operations; a half-dozen major standing advisory committees, including one appointed by the National Academy of Sciences, pay close attention to details of census design and operations; and several dozen journalists track census issues while many other private-sector groups have taken a lively interest in the preparations for and implementation of Census 2000. None of these oversight processes, advisory groups, or public watchdog efforts has identified any instance of partisan intention in the Census 2000 design—not for lack of looking, but because none is to be found.

Although Census 2000 was designed without partisan intent or consideration, the question remains whether statistical adjustment could be implemented in a manner that favors a particular outcome. Statistical adjustment involves a number of technical choices. Concern has been voiced that these choices might reflect partisan considerations, that is, might take into account how the adjustment could benefit particular regions, states, or cities. Recognizing this concern, the Census Bureau is taking extraordinary steps to protect the statistical integrity of decisions at every step in the adjustment process. Internal to the Bureau is a committee comprised of senior statisticians and survey managers that reviews and documents every technical decision made by those directly responsible for implementing dual system estimation. A careful record is kept of these deliberations. Census design decisions are widely shared—through documents, congressional hearings, and in open forums with the oversight panel of the National Academy of Sciences. This process of internal deliberation and external transparency will continue through every step of implementation and will be available for public inspection. In short, the Census Bureau will continue to take every step possible to demonstrate that, in its design and implementation, sta-

tistical adjustment is based on the best technical judgment available and never involves partisan considerations.

Notes

This article is based on the W. Edwards Deming Lecture of the American Statistical Association, 14 August 1999, Baltimore, MD.

1 The Constitutional clause, cited above, mandating the decennial census also anticipated that states would be taxed on the basis of population size. This practice fell into disuse early in the nineteenth century and the reference to taxes was deleted in the Fourteenth Amendment to the Constitution, ratified in 1868. James Madison, ever alert to the need to balance differing political interests, noted in Federalist Paper No. 54 the benefit to an accurate census of attaching both representation and taxation to its count: "It is of great importance that the States should feel as little bias as possible to swell or reduce the amount of their numbers. . . . By extending the rule to both [representation and taxation], . . . the States will have opposite interests which will control and balance each other and produce the requisite impartiality." In this and in so many other areas, Madison was prescient. Now that a greater population size registered in the decennial census conveys only benefits and no penalties, there is a strong urge across thousands of jurisdictions to "swell" but never to "reduce the amount of their numbers."

2 For an excellent treatment of this event, as well as for a broad and instructive overview of the census in historical perspective, see Margo J. Anderson, *The American Census: A Social History* (New Haven: Yale University Press, 1988).

3 A complete account of this event in census history, on which I have relied, appears in Margo J. Anderson and Stephen E. Fienberg, *Who Counts? The Politics of Census-Taking in Contemporary America* (New York: Russell Sage Foundation, 1999).

4 Robert A. Mosbacher, "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* (Vol. 56, no. 140, 22 July 1991), p. 33583. The passage cited presents only one of the reasons offered by the Secretary; other considerations were technical and operational.

5 Emphasis in original. This estimation of how many Republican seats would be "lost" has not been documented in any independent analysis. Most students of reapportionment and redistricting have concluded that it is nearly impossible a priori to calculate partisan shifts in legislatures that could follow a decennial census. As indicated later in this article, the Census Bureau does not collect information on political party affiliation, does not conduct analysis of possible partisan outcomes, and has no position on the soundness of Nicholson's forecast.

6 Anderson, *The American Census*, cited in note 2, p. 203.

7 Historians and demographers have estimated undercounts, including differential undercounts, for earlier decennial census results. See, for example, John W. Adams and Alice Bee Kasakoff, "Estimates of census underenumeration based on genealogies," *Social Science History* 15, no. 4 (winter 1991): 527–544.

8 David Heer (ed.), *Social Statistics and the City* (Cambridge, MA: Joint Center for Urban Studies of the Massachusetts Institute of Technology and Harvard University, 1968), p. 11.

9 An excellent nontechnical overview of how the Census Bureau uses dual system estimation can be found in Tommy Wright and Howard Hogan, "Census 2000: Evolution of the revised plan," *CHANCE: A Magazine of the American Statistical Association* 12, no. 4 (fall 1999): 11–19. For more technical treatments see the references cited in Wright and Hogan or consult the web site of the Census Bureau.

10 See Lawrence D. Brown et al., "Statistical controversies in Census 2000," *Juri-*

metrics 39 (summer 1999): 347–375. Critics of dual-system estimation focus on "distributional accuracy" or how well the census data distribute population shares across various geographical units. The Census Bureau believes it has a Constitutional responsibility to improve "numerical accuracy," that is, to count everyone possible, even though its operations cannot guarantee distributional accuracy. It also believes that a census cannot be planned with the explicit goal of distributional accuracy without serious damage to numerical accuracy.

11 See, for example, the report of the National Research Council's Panel on Alternative Census Methodologies, *Measuring a Changing Nation: Modern Methods for the 2000 Census* (Washington, DC: National Academy Press, 1999). For a more general treatment see Anderson and Fienberg, *Who Counts?*, cited in note 3.

12 Adjustment of a state's population for underenumeration by x percent would increase that state's claim on federal funds by roughly x percent *if* the population counts of all other states remained unchanged. If the population of all other states, too, were adjusted by the same x percentage, the allocation of funds would remain unaffected. As the undercount would tend to differ from state to state, some states (those with above-average adjustment for undercount) would gain, and others (with below-average adjustment for undercount) would lose. The percentage deviations from the average in practice are likely to be modest for most states. For additional analysis see *Formula Grants: Effects of Adjusted Population Counts on Federal Funding to States: Report to Congressional Requesters* (General Accounting Office, US Government, February 1999).

13 The term "apportionment" refers to the allocation of the 435 seats in the House of Representatives to the 50 states on the basis of census counts. It is to be distinguished from "redistricting," which refers to the drawing of legislative boundaries within each state, the geographic areas so constructed being electoral districts from which Representatives are chosen. Census data are used in redistricting because states must conform with the "one person, one vote" ruling and the nondiscriminatory provisions of the Voting Rights Act mentioned above. Redistricting sometimes also refers to the drawing of legislative boundaries for election to state legislatures, county governments, city councils, and school boards.

14 For Census 2000 only seven questions are asked of 100 percent of the households. Most households receive these questions on the short form. A sample of one out of six households receives the long form, which includes the basic seven items as well as approximately 35 more questions asking for demographic and housing information.

15 This decision is pursuant to a 1997 statistical policy directive of the Office of Management and Budget (OMB).

16 The OMB directive and Census Bureau practice treat Hispanic as an ethnicity, recognizing that persons of Hispanic origin can be white, black, Asian, or Native American Indian.

17 Census 2000 identifies five discrete racial groups: white; African-American, black, or Negro; Asian; Native Hawaiian or other Pacific Islander; and American Indian or Alaskan Native. It also allows respondents to select an Other category, making a total of six. There are 63 possible combinations to how the race question can be answered.

18 This sentence summarizes and oversimplifies a number of operations. In 2000, for example, about 1.5 percent of the addresses will be visited and directly enumerated by a census-taker; in rural areas, which account for about 20 percent of the addresses, Census Bureau employees will deliver the questionnaire. For the remaining addresses, primarily in cities and towns, the form will be mailed through the post office system.

19 Of a professional staff that numbers in the thousands, three positions are political appointees: the Director and one person each in public information and in legislative relations. Only the Director is presidentially appointed and congressionally confirmed. For instructive discussion of Census Bureau professionalism see Harvey M. Choldin, *Looking for the Last Percent: The Controversy Over Census Undercounts* (New Brunswick, NJ: Rutgers University Press, 1994).