

## Population Shifts and Demographic Methods

*C. Matthew Snipp<sup>1</sup> and Juanita Tamayo Lott<sup>2</sup>*

Demographers and other analysts typically assume the stability of benchmark population estimates for making subsequent comparisons and estimates of change over time. However, fundamental changes in statistical classifications thwart this assumption and make such comparisons difficult, if not impossible. We review several important developments in the United States federal statistical system and discuss their effect on the data for racial and ethnic minority populations. We also present data from studies showing that these changes have potentially large impacts for certain smaller racial and ethnic minority populations. The conclusion notes several difficult challenges to data users posed by these findings.

*Key words:* United States Census; race; ethnicity; tabulations.

This article illustrates how shifts in the size, composition, and distribution of racial and Hispanic origin populations between the 1970 and 2000 Censuses have been affected not only by changes in the meaning of basic concepts over time and space but also by changes in statistical methods. That is, while tracking distributions and trends over time and space, it is equally important to track changes in methods for allocating and tabulating basic demographic variables. This article argues that the size of small populations can be greatly affected by variation in benchmark distributions, specifically racial classifications and tabulation procedures. These two sources of variability are an underappreciated problem for population studies using demographic analysis.

### 1. Demographic Methods

Demographic methods are one of the oldest tools available to enumerate various populations. This involves estimating the size of the population with a simple balancing formula where:  $P_1 = P_0 + B_{0,1} - D_{0,1} + NM_{0,1}$ , where  $P_1$  is the total population at the present time;  $P_0$  is the total population at some earlier time,  $B_{0,1}$  represents the number of

<sup>1</sup> Stanford University, Department of Sociology, Stanford, CA 94305-2047, U.S.A. Email: snipp@stanford.edu

<sup>2</sup> Senior Statistician (retired), U.S. Census Bureau. Emails: juanita.tamayo.lott@gmail.com and jtlott@alumni.uchicago.edu

**Acknowledgments:** An earlier version of this article was presented at the 2003 annual meeting of the American Statistical Association. It reports the results of research and analysis undertaken by U.S. Census Bureau staff. It has undergone a U.S. Census Bureau review more limited in scope than that given to official U.S. Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. We appreciate the thoughtful reviews of Claudette Bennett, Mark Holdrege, Nicholas Jones, Jay K. Keller, Alphonso Mason, Nampeo McKenney, Jeffrey Passel, J. Gregory Robinson, Fritz Scheuren, Laurel Schwede, Philip Steel, and Kathleen Styles.

births in the interval between the present and some earlier time;  $D_{0,1}$  represents the number of deaths in the same interval between 0, 1, and  $NM_{0,1}$  is the net migration number for the relative gain (loss) of persons entering or leaving the population (nation-state) in the same period. This simple formula can be made more complex by adding ascribed group characteristics that allow estimates to be obtained for subpopulations such as racial and ethnic minorities.

In the United States, information about race, or color, along with age and sex, has been collected for over 200 years. With expansion of electronic computing ability, demographic methods are complemented by other statistical methods, notably survey research and modeling. Censuses and surveys have facilitated the collection of demographically related characteristics such as marital status, life expectancy, and nativity, as well as achieved characteristics such as educational attainment, language fluency, and occupation. Survey methods also have facilitated the observation of characteristics which can be ascribed or achieved, such as citizenship. Models have facilitated inferences about and simulations of selected populations.

Demographers have written volumes about how to use demographic analysis to manipulate a set of marginal distributions for the purpose of anticipating what those distributions might look like at some point in the future. In contrast, they have not paid as much attention to the sources of variability in the benchmarks they use. But as the old expression “garbage in, garbage out” suggests, the stability of benchmark populations can no longer be taken as a given, especially where data for racial and ethnic minority populations are involved. These sources of variability in benchmark distributions – racial classifications and tabulation procedures – are an under-appreciated problem for population studies using demographic analysis.

The American Indian population represents a truly unique case in some respects. In other ways, it exemplifies the methodological and conceptual challenges that must be confronted in the future. In particular, the American Indian<sup>3</sup> population holds a unique legal status in the United States of America that no other racial or ethnic minority group can claim. The experience of American Indians, however, is not so different with regard to the matter of how one goes about the business of creating a relatively unambiguous racial classification; one that settles the question of who does and does not belong to a particular racial or ethnic group. While it is easy to see how this has become a contentious matter for American Indians, it could be just as contentious for other groups. In particular, American Indians may be an important bellwether for other groups insofar as they have much higher rates and a longer history of interracial marriages than other groups. This is a pattern that other groups are in the process of emulating. Since the 1960's, interracial marriages have been increasing at an accelerating rate (Harrison and Bennett 1995). In the not-too-distant future, rates of inter-racial and inter-ethnic marriages for Asians and Hispanics could reach the same level as those for American Indians.

Even today, about one-third of all Hispanic children have a non-Hispanic parent. Between high rates of inter-marriage of the U.S.-born population and high rates of

<sup>3</sup> For the sake of editorial brevity, we include Alaska Natives whenever we use the term “American Indians” instead of repeatedly using the lengthier phrase “American Indians and Alaska Natives.”

immigration and inter-marriage of Asians and Latin Americans, whose reports of race and ethnicity are more fluid than those of the traditional U.S. White and Black populations, there most certainly will be a swell to the numbers of interracial persons in coming U.S. generations. Moreover, while large variations are observable in data for Native Hawaiians and other Pacific Islanders, and for American Indians and Alaska Natives, variation is also possible for larger groups, such as Blacks in smaller geographic areas where intermarriage rates are high.

In the face of these growing numbers, the old paradigm for tracking race and ethnicity in American society is going to become ever more obsolete and difficult to handle. In the case of American Indians, the tabulation procedures to bridge past trends and data to the future amount to pounding a square peg into a round hole, with about the same precision and finesse. Based on some of the data in this article, especially the data for persons identifying themselves as belonging to "Other" races, the experiences of American Indians presage the future of other racial and ethnic minority groups in America – Asian-Americans/Pacific Islanders, Latina/o Americans, and, to a lesser degree, African-Americans.

## **2. Population Shifts from 1970 to 2000: Continuity and Change**

From the first census of the U.S. populations in 1790 through the 2000 Census, patterns of continuity and change in the profile, understanding, and meaning of a growing U.S. population have persisted (Lott and Keller 1999).

### *2.1. 1970 Census*

The 1970 Census of Population and Housing was a historical first in census history. It was the first full implementation of a mailout/mailback census. That is, the census forms were completed by a household respondent rather than by an enumerator as had been the case in all prior censuses. This change in methodology is one potential reason for variation in the counts of racial groups over time, particularly American Indians. The White population and married couple households were dominant, reflecting the demographic composition of U.S. households. Ninety percent of the U.S. population was White, 9% Black, and 1% all other races, including American Indians/Alaskan Natives and Asian/Pacific Islanders. The proportion of the foreign-born population was at a historical low of 4.7%. Clearly, the 1970 Census was not a good baseline for projecting the 1980, 1990, and 2000 Censuses.

To their credit, the federal agencies in concert with the federal statistical system, particularly the U.S. Census Bureau, understood the coming demographic shift to a growing non-White population beyond the Black population (U.S. Commission on Civil Rights 1973; Lott 1998). Since the 1940 Census, particular statistical attention had been given to the Black undercount within the context of Ansley-Coale's early undercount work and the 1940 undercount of Blacks based on matching 1940 Census data to military enlistment administrative records. For the 1970 Census, reporting of race and ethnicity was limited. Hispanics were identified by Spanish origin and Spanish surname. American Indians were segregated in a separate report. On the other hand, with statehood conferred on Hawaii and Alaska, growing attention was paid to Asian Americans/Pacific Islanders who comprised sizable portions of the populations of those states.

## 2.2. 1980 Census

In the 1980 Census, attention focused on documenting the growth of historically discriminated minority populations and new U.S. settlers. This attention was spurred by numerous pieces of legislation including the passage of the 1965 amendments of the Immigration and Nationality Act, Civil Rights legislation from the 1950s through the 1970s, including the Voting Rights Act of 1965 and its amendments, and the Indian Civil Rights Act of 1968. Concern about the undercount of Blacks and other racial/ethnic minorities was articulated by advocates for these groups as well as demographers, statisticians, and urban officials. "Hispanic" origin was added as a special identifier of ethnicity, separate from the race question. Hispanic ethnicity was deemed to be distinct from race because representatives of the Hispanic population made a persuasive case before the President's Cabinet on Mexican Americans and the Federal Interagency Committee on Education that Hispanics could be of any race.

In preparation for a complete and accurate count of racial and Hispanic origin populations in the 1980 Census, the U.S. Census Bureau created four federal advisory committees on the Black, Hispanic, American Indian/Alaskan Native and Asian/Pacific Islander populations, respectively. These committees were also in place for the 1990 Census and the 2000 Census. They are currently involved in preparations for the 2010 Census. However, for the 2010 Census the Pacific Islander populations have an advisory committee separate from the Asian population advisory group.

The 1980 Census was also taken within the context of the original 1977 Office of Management and Budget statistical policy on standards for the collection of data on race and ethnicity, known as Directive 15. This standard was timely for program, policy, and research considerations, given the increasing numbers, proportion, and within group diversity of Blacks and other non-White groups. These were partly due to higher fertility rates, growing immigration rates, and growing suburbanization of non-Whites, which followed the similar migration of Whites out of urban areas since the 1950s.

## 2.3. 1990 Census

The 1990 Census was a watershed event in the measurement of race and ethnicity in American society. It produced the greatest amount of detailed data on the largest number of races and Hispanic origin populations and subpopulations. On the other hand, it raised expectations for current and emerging groups to be identified and counted as distinct populations. There emerged a variety of groups opposed to the procedures employed by the U.S. Census Bureau to count and tabulate the United States population by race and national origin. In particular, two broad sets of interests coalesced to instigate a concerted attack on the way race and ethnicity were measured in the census. One set were groups who wished to be identified on the census but, for one reason or another, were omitted from the race question. For example, Taiwanese and Arab American advocacy groups argued that their communities should be included among the selections appearing on the race question, viz. that Taiwanese and Arab should be included as a choice along with Chinese, Korean, and Aleuts, among others.

An equally vigorous opposition to the census arose within a loose network of organizations concerned with the well-being of multiracial families (Farley 2001). These

organizations consisted mostly of families in which one spouse was White and the other was African-American. One of the larger and better-known of these organizations was based in Atlanta, GA, and was known as RACE, an acronym for Re-classify All Children Equally.

They had two strenuous objections to specific census procedures and to Directive No. 15 more generally. The instructions on the 1990 (and earlier) Census required respondents to mark one race only for each member of the household. The child of a Black mother and a White father, for example, could be identified as “White” or as “Black” but not as both. In cases where a respondent neglected to read these directions, or purposely wrote two or more responses, the U.S. Census Bureau implemented a complicated series of “editing” procedures that reassigned multiracial responses to one of the five single race responses.<sup>4</sup> These procedures deeply offended the members of organizations such as RACE. In particular, they found that these rules forced them to choose a single identity for their offspring, thereby “privileging” the race of the mother or father over the heritage of the other parent. Not only did they find these procedures offensive, they argued that such measures were a source of deep distress to them and their children (Farley 2001). The empirical evidence for this resistance appeared when about 500,000 respondents did not follow the instruction to identify with one race only, and marked two or more responses instead (Wallman et al. 2000).

Amidst these challenges to traditional racial and ethnic classification and tabulation, it is easy to lose two significant results of the 1990 Census. In addition to capturing very detailed demographic and geographic data on the U.S. population, the 1990 Census facilitated analytical ability not only to focus on historically under-represented populations, but also to shift the focus to hard-to-enumerate populations by use of demographic characteristics within geographic/spatial units. Demographic and geographic/spatial analysis of the 1990 Census by U.S. Census Bureau staff for coverage evaluation and improvement measurement targeted the following characteristics:

- vacant housing units
- multi-housing units
- renter occupied units
- housing units with more than one person per room
- households that are not husband/wife households
- housing units without telephone
- no highschool graduates
- persons below poverty level, receive public assistance income
- unemployed
- linguistically isolated household, and
- householder moved in unit 1989 or 1990.

These variables formed the basis for the Census Bureau Planning Database (PDB), which assembled a range of housing, demographic and socioeconomic variables from the 1990

<sup>4</sup>The five single races approved by U.S. Office for Management and Budget (OMB) for the 1990 Census included a special “Some Other Race” category, which was not stipulated by Directive 15.

Census that correlated with nonresponse and undercount. The database provides a systematic way to preidentify potentially difficult to enumerate census tracts for both censuses and surveys. While the Planning Database did not include race or ethnicity as variables, it shed innovative light on how to address the differential undercount. Using 2000 census data, the Planning Database confirmed the important distinction that race and Hispanic origin may be associated with some characteristics more than others but are not synonymous with them. For example, in Census 2000, selected Hard-to-Count variables – renter units, crowded housing units, no phone, persons at poverty level, and linguistically isolated households – were highly associated with tracts where race/origin group, except Non-Hispanic Whites, represented 50% or more of the total tract population. Some variation existed across groups. For example, for “crowded units” this ranged from 3.6% for African-Americans to 16.5% for Hispanic origin; for “no phone” this ranged from 1.4% for Asian/Pacific islander to 28.0% (Bruce and Robinson 2003, Table 2).

#### *2.4. Classification and Tabulation Changes Between the 1990 Census and the 2000 Census*

These complaints about the measurement of race and ethnicity in the 1990 Census eventually resonated in Congressional offices and prompted action in the form of hearings in the House of Representatives. In 1993, the House Subcommittee on Census, Statistics, and Postal Personnel, chaired by then Representative Thomas Sawyer (D-Ohio), held four informational hearings on the OMB Statistical Standards on Race and Ethnicity, Directive No. 15 (Office of Management and Budget 1997). No doubt aware of these objections, officials from the Office of Management and Budget declared their intentions to conduct a careful review of Directive 15 and of the methods the federal government used to collect data about race and ethnicity (Edmonston and Schultze 1995).

The OMB review of Directive 15 was launched with a conference organized under the auspices of the National Academy of Sciences in February 1994 (Edmonston et al. 1996). At this meeting, papers and commentary on racial measurement were presented by experts with various perspectives. In attendance were social scientists, private sector officials, local, state, and federal civil servants, and representatives of various interest groups such as multiracial families. The findings from this meeting were assembled and eventually published in *Spotlight on Heterogeneity* (Edmonston et al. 1996). In addition to this meeting, OMB assembled a large inter-agency task force to undertake the official review of Directive 15. This task force included representatives from more than thirty agencies within the federal government with a stake in the collection of racial data. OMB also organized hearings around the country to give local groups and individuals an opportunity to express their concerns about Directive 15 (Lott 1998).

In addition to these efforts, OMB officials worked closely with the U.S. Census Bureau to determine the likely effect on the federal statistical system of a revision of Directive 15. As part of the planning process for the 2000 Census, the U.S. Census Bureau fielded several large tests to determine the effects of modifying Directive 15, especially its effect on the decennial census. The first test was conducted in the May 1995 Current Population Survey, in which a multiracial response item was included among the choices for racial self-identification. In 1996, the U.S. Census Bureau conducted two additional large

surveys – the National Content Test and a survey specially designed to test alternative specifications of questions for race and Hispanic origins. The latter survey was designated the Race and Ethnicity Targeted Test (RAETT) and included approximately 120,000 households (OMB 1997; Hirschman et al. 2000).

The results of these tests suggested that in the aggregate, changing the race question to allow multiracial responses would not seriously skew the racial profile of the United States. Indeed, these surveys indicated that only about two or three percent of the total population would select a multiracial response if given this option. However, there was also evidence that some small groups of Asians and American Indians would be affected to a much greater degree than the population as a whole (Tucker et al. 2002; Hirschman et al. 2000). That is, these small populations included a sizable number of mixed race persons who would identify themselves as such if given the opportunity. As a result, the statistical profile of these groups might be significantly altered by the inclusion of a multiracial response option (Tucker et al. 2002).

After consideration of extensive testimony, research, and deliberation, OMB officials decided to issue a revision of Directive 15 that would address many of the concerns and complaints lodged against it. In October of 1997, a Federal Register notice announced new standards for the classification of race and ethnicity. These standards were mandated for adoption by all federal agencies, and their contractors and grantees, effective January 1, 2003 (OMB 1997).

The 1997-revised standard established a new and slightly different set of racial categories along with new guidance for the collection of racial and ethnic data. The categories it mandated were the following:

- American Indian or Alaska Native (including Central and South America)
- Asian
- Black or African American
- Native Hawaiian or other Pacific Islander
- White

Besides these categories of race, the revision also mandated that information be collected for the Hispanic (or Latino) population as a separate and distinct ethnic group. OMB also mandated that respondents should be instructed to “mark one or more” or “choose one or more” responses to queries about their race. That is, respondents were to be allowed the opportunity to express multiple responses about their race if they so desired (OMB 2000).

Besides this significant format change, the revised categories were noteworthy in several respects. The revision provided Native Hawaiians with a significant victory in their quest for recognition. Native Hawaiian activists seeking recognition for their status as indigenous people with inherent rights of sovereignty had initially sought to be included within a broad category designated as “Native American.” Some American Indian groups opposed a broad category of “Native American,” presumably because they were concerned this might dilute their own claims as indigenous people (Edmonston et al. 1996). OMB concurred but granted the Native Hawaiians a category that would set them apart from the polyglot collection of ethnicities formerly subsumed within the initial category of “Asian and Pacific Islander.” Thus, the difference between the 1978 initial standard on race and ethnic classification and the revised standard formalized twenty years

later in 1997 was the addition of “Native Hawaiian or other Pacific Islander,” which was not a new category *per se*, but a disaggregation of the initial Asian or Pacific Islander category.

Some American Indians also opposed the phrase adding Central and South American Indians. Groups representing American Indian interests argued that Central and South American Indians did not have reservations or a special legal or political status similar to of those tribes in the 48 states, and thus should not be enumerated with them. This was consistent with the 1975 report by the Federal Inter-Agency Committee on Education that led to Directive 15. OMB noted this opposition but ruled to add the phrase “(including Central and South America),” noting that the number of Central and South American Indians in the United States is small and this change would have little effect (OMB 1997).

Finally, OMB reaffirmed the original Directive 15 that explicitly set Hispanics and Latinos apart as an ethnic group, as opposed to a racial group. The guidelines suggest that whenever possible, two questions should be used to collect information about race and ethnicity, respectively. The race question should use at least the minimum five racial categories, and a second “ethnicity” question should be used to determine Hispanic or Latino origins. The guidelines indicated that a single race question might be acceptable in some instances – including Hispanic or Latino as a response – but it was also clear that it did not accept the proposition that Hispanic or Latino was acceptable as racial designation.

In the October 1997 Federal Register notice, OMB offered a modest rationale for making its revision of Directive No. 15: that improving the federal statistical standard on classification and collection of racial and ethnic data is desirable, and the system in place since 1978 did not fully capture the diversity of American society. However, OMB offered relatively little guidance with respect to how the changes should be managed. A committee of data specialists was assigned the task of working out the details of implementation in the wake of the revised standard. Predictably, a large part of this burden fell to the U.S. Census Bureau because the first large-scale implementation of the 1997 revision would be the 2000 Census.

The 1997 OMB standard and their own experience pretesting multiracial questions provided U.S. Census Bureau officials with a straightforward method for asking respondents about their race. The 2000 Census race question includes minimum categories mandated by OMB along with an additional set of categories for Asians and for Pacific Islanders, similar to the additional categories that appeared on previous censuses. On April 1, 2000, U.S. residents were asked to publicly declare their race along with the usual information about residence and family relationships.

## 2.5. 2000 Census

The shift from undercounted populations to hard-to-enumerate populations continued in the 2000 Census. The population changes which had begun to be captured gradually since the 1970 Census became more dramatized and publicized with the 2000 Census. This was due in part to: (1) the very open and public dialogue on the measurement of race and ethnicity preceding the 2000 Census; (2) the coincidence of this discussion with a general interest in and reassessment of multiple identities by the United States population (such as work/home identity, peer/family identity, gender/sexual identity, and



group/individual identity); and (3) public awareness of and heightened expectations and ownership of the 2000 Census.

For population changes related to race and Hispanic origin, a fundamental question arose, which was: “*How homogeneous are populations classified by a common ancestry but possessing different life experience, a variety of socio-economic statuses, different nativity status distributions, and different patterns of concentration and dispersion by residence in different geographical areas?*” The 2000 Census data with respect to age distribution/life expectancy, fertility rates/marital status, citizenship/nativity ratios, racial composition, national origin composition, immigration composition and distribution, and household composition and distribution indicated that there are spatial, temporal, and cohort dimensions of basic demographic data, particularly for race and Hispanic origin.

The production of the 2000 Census data presented officials within the federal government with the first real challenge arising from the 1997 revision of Directive No. 15. Specifically, Census officials were responsible for delivering to Congress and state legislatures across the country by April 2001, detailed population counts for redrawing legislative districts. These data, known as the PL 94-171 files, contain information about the age, race and sex of persons for geographic areas about the size of Census tabulation blocks in all fifty states.

Managing the sheer volume of these files was a problematic task for the U.S. Census Bureau staff. The reporting of sex was simple, and the PL 94-171 files could include a truncated age distribution (under and over 18 years). However, the format for reporting race was a daunting task. As pointed out by U.S. Census Bureau advisory committees and others, there were 63 unique combinations that could be formed from the response items on the race question alone. Overlaying these 63 items with Hispanic/Non-Hispanic yields a race/ethnic reporting format consisting of 126 unique categories. Past censuses strained to report personal and family characteristics by race, using the five categories stipulated by the original version of Directive No. 15. Under the revised standard, the number of possible iterations was multiplied twenty-five fold.<sup>5</sup>

For the PL 94-171 files, the resolution of this dilemma would seem to require some reduced number of combinations for which data could be reported. In March 2000, the OMB issued a policy memorandum about how multiple racial data should be allocated for civil rights enforcement. Specifically, the bulletin provided two essentially separate pieces of guidance – one for aggregating data and one for allocation when utilizing data for civil rights enforcement and monitoring. The provisional aggregations proposed in this document listed the following groups, as one way that racial data should be reported:

1. American Indian or Alaska Native
2. Asian
3. Black or African American
4. Native Hawaiian or Other Pacific Islander

<sup>5</sup> The U.S. Census Bureau resolved some of this complexity by reporting data for persons who identify with one race only along with data for persons who identify with two or more races. In some cases, “two or more races” is treated as a residual category (e.g., “persons reporting two or more races”) and in others, race-specific combinations are reported, e.g., “White alone” and “White in combination with some other race.” Needless to say, these numbers cannot be aggregated and summed to 100%.

5. White
6. American Indian or Alaska Native and White
7. Asian and White
8. Black or African American and White
9. American Indian or Alaska Native and Black or African American
10. > 1 percent: Fill in if applicable \_\_\_\_\_
11. Balance of individuals reporting more than one race
12. Total

Items 10 and 11 were meant for areas, principally in California and Hawaii, with unique racial combinations that were rare in other parts of the country. However, because this list was not exhaustive of all the possible combinations of multiracial responses, OMB offered this additional guidance for the aggregation of racial data:

- Responses in the five single race categories are not allocated [to other racial groups].
- Responses that combine one minority race and white are allocated to the minority race.
- Responses that include two or more minority races are allocated as follows:

“If the enforcement action is in response to a complaint, allocate to the race that the complainant alleges the discrimination was based on. If the enforcement action requires assessing disparate impact or discriminatory patterns, analyze the patterns based on alternative allocations to each of the minority groups.” (OMB 2000).

Knowing or not, the OMB adopted a policy based on a limited form of hypodescent<sup>6</sup> to resolve the question of whether mixed race persons were entitled to protection under the various civil rights laws. Heeding these recommendations in the final design for the PL 94-171 file, the U.S. Census Bureau staff proposed a reduced number of categories. However, after consulting with other agencies of the federal government that would depend heavily on these data, especially the Office of Civil Rights in the Justice Department, the original plan was found to be unsatisfactory. A subset of racial combinations useful for all agencies in all places was impossible to specify in advance. Put another way, data needed to be available to complete the blanks left in Items 10 and 11, and the guidance left open the possibility that data might be needed for other multiracial combinations as well. Consequently, a decision was made to release these files with the detailed combinations of race and ethnicity.

The PL 94-171 files were released in the spring of 2001 followed by additional data based on the 100 percent short form questionnaire. In subsequent releases, the U.S. Census Bureau opted for more abbreviated formats for reporting multiracial responses. In most cases, these reports have been limited to the five racial categories stipulated by OMB, separate reports for Hispanic and Non-Hispanic persons, and a residual category reported for persons reporting “Two or more races” in some instances, and in other instances, a set

<sup>6</sup>The term “hypodescent” in the United States refers to the belief that any amount of African American ancestry is sufficient for designating an individual as African American. In popular vernacular, this is sometimes described as the “one-drop rule.” It also is important to understand that these rules did not modify or otherwise alter the information provided by individual respondents.

of categories for persons reporting a race “alone” or a race “in combination with one or more races.” Needless to say, this format has the virtue of being concise and economical. However, it also obfuscates an enormous amount of detail about the racial heterogeneity of the multiracial population. For the foreseeable future, a simple resolution of these issues remains out of sight.

### 3. Demographic Analysis

Demographic analysis covers the total U.S. population, not just household and/or civilian populations. Demographic analysis represents a macro-level approach. Population benchmarks are developed for the census date by analysis of various types of demographic data essentially independent of the census, such as administrative statistics on births, deaths, immigration, and Medicare enrollments, as well as estimates of emigration and undocumented immigration. Internal consistency is an important aspect of the independence of demographic analysis. The foundation of the demographic method is the logical and longitudinal consistency of the underlying demographic data. Demographic analysis follows the process of population change as it occurs, starting with births, then incrementing or decrementing cohort size with subsequent information on mortality and net migration. The administrative data for demographic analysis are virtually complete (no samples involved) and available annually for the core components of births, deaths, immigration and Medicare enrollments.

Demographic analysis has played an important role in evaluating the overall quality of the decennial census as well as providing inter-censal estimates of the United States population subdivided by age, race, sex, and geography. Prior to the 2000 Census, inter-censal estimates were unable to respond to the OMB revised classification because the aforementioned administrative records were not available in a consistent format. However, beginning in January 2003, federal agencies and particularly the National Center for Health Statistics, which tracks births and deaths, adopted this standard for their record systems.

The American Indian population offers some important insights into the complications that are likely to affect demographic analysis for racial and ethnic groups, especially small populations with relatively large numbers of mixed race persons. Prior to 1997, alternative estimates for the American Indian population could be mostly attributed to definitional differences used for enumeration – tribal enrollment versus self-identification for example. However, the OMB standard substantially complicates the problem of inconsistency across estimates, and by extension, inconsistent population projections derived from demographic analysis. For example, there are 2.5 million persons who report their race as American Indian only and another 1.6 million who report American Indian in combination with another race, for a total of 4.1 million persons who identify as American Indian. Race-specific population estimates must carefully heed these differences or risk making substantial errors in their estimates. A similar problem would occur in small area estimates containing substantial numbers of multiracial persons.

As is becoming well known, the 1997 modification of OMB Directive No. 15 introduced the option of allowing respondents to indicate one or more choices for their race. The obvious virtue of this methodology, and the principal reason for introducing it, is that it

corresponds more closely to empirical reality than a forced choice question. That is, there are large numbers of persons who may claim more than one race, and this question acknowledges this reality. However, while the question itself is fairly simple, the problems associated with tabulating and presenting these problems are much more complex than they appear at first glance. One is that percentages distributions showing the racial composition of an area are no longer constrained to sum to 100 percent as they did in the past. Another is that temporal comparisons of these data with earlier series are substantially more complex.

These complexities are especially problematic for making population projections for a number of reasons, especially for the short-term foreseeable future. Birth and death data are still collected using a single race, forced choice question, and it will be a number of years before reliable series and rates will be available for multiracial populations. An exacerbating consideration is that there is no reason to assume homogeneity in the multiracial population. In particular, there is no reason to assume that birth rates for Black and White persons are the same as birth rates for Asian and White persons. Until reliable vital events data become available for specific types of multiracial persons, it will be necessary to make *a priori* assumptions for such groups that they should be matched with the data for one or another monoracial population. For instance, mixed-race Black-White persons could be assigned the birth and death rates for either the Black or the White population.

As previously mentioned, the fact that under the new standard, population totals can exceed 100 percent due to double (or triple or quadruple, etc.) counting of mixed-race persons is problematic in its own right. In addition, population projections and estimates must carefully partition the population in a way that does not permit such double counting, or aggregate population totals may significantly exceed the actual number of persons that might be expected from rates due to births, deaths, and migration. For example, allocating multiracial persons to single race groups for the purpose of making population projections could lead to an overestimate for one group and a deficit for another. To illustrate, suppose all Black-White persons were counted as Black. This would inflate the Black population and correspondingly reduce the White population, and vice versa if the allocation rule were reversed.

One approach to these problems involves the construction of a “bridge” that simultaneously ensures that population totals do not exceed 100 percent and allows a more parsimonious match with existing race vital event data. A bridge is a strategy for tabulating multiracial responses in a way that is more or less comparable with older data series based on monoracial forced choice responses. There are several ways this can be done. Unfortunately, each of these bridge tabulation procedures has particular disadvantages associated with it, and there are no clear criteria for choosing one procedure over another. As the data for American Indians show, most problematic is that choosing one procedure over another can result in significantly different marginal distributions, which ultimately can result in significantly different population projections.

Tucker et al. (2002) analyzed these procedures in detail, especially with regard to their influence on population counts for racial and ethnic minority groups. The results of this study have extremely important methodological implications for users of racial and ethnic data, and for this reason it is worthwhile reviewing its findings. Tucker and his group identified a variety of ways that data containing multiracial responses can be retabulated to

approximate a distribution of monoracial categories. However, these ways can be grouped under two general approaches. They designated one set of procedures as “Deterministic Whole Assignment” and the other as “Deterministic Fractional Assignment.” The former involves tabulations in which whole persons are assigned to one or another monoracial category. That is, it is a set of decision rules for assigning multiracial individuals to monoracial groups. This approach has the virtue of yielding population counts that are whole numbers. However, these decision rules amount to a set of procedures for selectively ignoring the information provided by multiracial persons about their race.

The latter approach is less intuitive but it entails decision rules that literally divide a person across different monoracial groups. So, for example, under this scheme one-half of an individual might be allocated to the total count of one group, and one-half to the total count of another group. The virtues and vices of this approach are a mirror image of whole assignment. In particular, fractional assignment will yield population counts that are not whole numbers – what is the meaning of a number such as 62.75 Black people in a neighborhood? Where is the 0.75 Black person to be found? In this regard, such an approach disembodies race as a characteristic of individuals and treats it as a group or area trait. This perspective runs counter to the way that race is usually understood in American society, i.e., as a trait of individuals. However, the virtue of this approach is that it does incorporate more clearly the information provided by multiracial persons.

Within each of these broad approaches, there are a number of different variations that potentially have an effect on the counts of racial and ethnic minority populations. The balance of this article will describe each of these variations, their advantages and disadvantages, and their effect on American Indian population counts. The application of these procedures to actual surveys is described in detail below.

### *3.1. Deterministic Whole Assignment: Smallest Group Method*

This tabulation procedure assigns persons who identify themselves as White and a minority group to the particular minority population. For example, persons who identify as White and as Black are counted as Black for enumeration purposes. When two or more minority populations are identified, the respondent is assigned to the smallest one. This method favors minority populations insofar as it enlarges groups with small numbers. However, for a small group with a large number of mixed race persons – e.g., American Indians – it is possible, if not likely, that this method could lead to highly inflated numbers that are far out of proportion with past population estimates.

### *3.2. Deterministic Whole Assignment: Largest Group Other than White*

This tabulation rule assigns responses that include White to the minority group indicated by two-race, mixed-race persons. For example, a respondent who reports his or her race as Black and White for tabulation purpose would be counted as Black. The process is more complicated for persons who report more than two races. In this instance, the decision rule stipulates that tri-racial responses be assigned to the statistically most common non-White race. For example, a Black-White-American Indian respondent would be counted as Black. Procedurally, this approach suppresses the population of very small groups, and in extreme cases may even cause their number to become too small for analysis. However,

this method is also less likely to artificially inflate small group populations or have a significant effect on larger minority populations. Conceptually this tabulation procedure represents a curious throwback to old ideas about hypodescent – anyone who reports minority and White is excluded ipso facto from enumeration in the White population. For this reason, it is easy to understand why many would object to the implementation of this rule on political grounds.

### *3.3. Deterministic Whole Assignment: Largest Group*

This rule assigns responses from two or more racial groups to the group with the largest population with respect to monoracial responses. For example, Black-American Indian respondents would be counted as Blacks. White-Black respondents would be counted as Whites. This procedure would obviously enlarge the White population but in proportional terms, the increase would most likely be small. In contrast, this procedure might cause a sizable decrement in small populations heavily intermarried with Whites, e.g., American Indians and some Asian-American groups.

### *3.4. Deterministic Whole Assignment: Plurality*

This coding rule can be applied only when follow-up information is obtained about which race a respondent would like to be counted. When multiracial respondents report more than one race, they must be asked which the race they most strongly identify with is. The proportion choosing each of the two possibilities was calculated and all responses in a particular multiple race category were assigned to the group with the highest proportion of responses on the follow-up question. For example, for persons who identify as White and American Indian, if the race that was chosen as their main race was American Indian, then all persons who reported White as American Indian were assigned to the American Indian population. This is perhaps the most complicated and least intuitive tabulation method. It is also the least feasible for most applications. Most public data sources do not have a follow-up question about “main” race and for this reason alone, this method is unlikely to be widely used.

### *3.5. Deterministic Fractional Assignment: Equal Fractions*

This tabulation procedure assigns fractions of persons to groups according to the numbers of multiracial responses given by respondents. For instance, biracial responses are tabulated by assigning 0.5 persons to each group. A Black and White person would result in 0.5 persons added to the White population and 0.5 persons added to the Black population. In a like manner, tri-racial and higher order responses are divided into equal fractions and assigned to their respective groups. As noted above, this method has the virtue of making use of information provided by multiracial respondents. Yet, this approach also bears some resemblance to the kinds of allocations associated with blood quantum classifications. Instead of making a determination of ancestry, this procedure simply assumes that anyone who identifies as bi-racial is equally divisible with respect to his or her race. Some respondents may object to being counted as half-White, for instance, when most of their ancestors were of another race.

### 3.6. *Deterministic Fractional Assignment: Unequal Fractions*

Responses are tabulated with this procedure by applying some type of a priori partitioning scheme. It can be a variant of the pluralistic scheme described above whereby fractions of respondents are assigned to whatever groups are most commonly identified as main races. So, for instance, if two-thirds of the White-American Indian population responded that American Indian was their “main” race – the race they most closely identify with – then one-third of a White-American Indian multiracial respondent would be counted toward the aggregate total of the White population and two-thirds toward the total for the American Indian population. Needless to say, this is an extremely cumbersome and highly complex method for tabulating multiracial responses. For nontechnical users, it is likely to be extremely difficult to comprehend. Nonetheless, it does have some virtue insofar as it can rest on an empirical distribution of “main” responses for determining fractional assignments.

## 4. **Methods and Findings**

To explore the actual effects of these tabulation procedures, the federal working group applied them to several different surveys (Tucker et al. 2002). These surveys did not implement a question about race in the same format as specified by the revised OMB standards. Indeed, these surveys were fielded prior to the issuance of the OMB revision. Nonetheless, the questions they asked incorporated a multiracial response similar enough to the OMB standard that it was possible to test these tabulation alternatives. Specifically, these data were obtained from a supplement to the May 1995 CPS, the National Health Interview Survey (1993–1995), and a survey obtained from a 1998 Washington State population survey. It is beyond the scope of this discussion to report these results in detail but they are very instructive about the impacts of these procedures.

The summary findings from these surveys substantiate the points made at the outset of this discussion. That is, for large groups such as the White population, the results are fairly consistent across surveys and across methodologies. Regardless of which tabulation procedure is used, the results are relatively invariant. Similarly for groups with relatively small numbers of inter-racial marriages, such as the Black population, the choice of tabulation procedure also does not have a large consequence for the results. However, for small groups with significant numbers of inter-racial unions – specifically American Indians, and Asian Americans to a lesser degree – the choice of tabulation methods is a more consequential decision.

The federal working group generated a large number of tables with these data that clearly demonstrate these findings. And while it is not possible to examine all of the data, it is worth looking at how tabulation rules affect the most extreme case: American Indians. The data in Tables 4 to 6 are adapted from the analyses of the federal working group and show the percentages of American Indians within each sample using alternative tabulation rules.

Table 1 simply displays the percentage distributions by race for American Indians, Whites, and others using the several procedures described. The conclusion that can be drawn from these data is quite unsurprising. Regardless of which method is used, the White population is relatively unaffected. In the Current Population Survey (CPS) data, Whites constitute 82.4 percent of the sample using the old monoracial forced choice

survey format. In the survey supplement containing a question allowing multiracial responses, the number of White respondents diminishes slightly, ostensibly because these individuals are now allocated to other groups under the various tabulation rules. Apart from this small decrease of one or two percent the various tabulation rules yield almost invariant results, ranging from 80.4 to 81.0 percent.

For American Indians, the results are predictably more variable. The multiracial format yields a somewhat larger number of American Indians and varies more, from 0.8 to 1.2 percent, across methodologies. While these percentages may seem very small, it is important to keep in mind that the American Indian population makes up a very small share of the total U.S. population from which the CPS is drawn. Furthermore, an increase in population from 0.7 percent in the old monoracial format to the 1.2 percent yielded by the smallest group method is more substantial than it appears. In fact, it is about a 71 percent increase. An increase of this magnitude applied to the 1990 American Indian population estimated at 1.98 million would have given a figure of about 3.4 million instead.

All of the bridging methods have in common that they reassign the race of multiracial respondents, albeit in different ways. The data in Table 2 show how persons classified as American Indian under the old rule would be reallocated under the new procedures. For the sake of comparison, readers should note that 96 to 97 percent of persons classified as White by the monoracial format were also classified as White by the multiracial methods. This obviously did not happen for American Indians. For example, 22 to 25 percent of American Indians identified by the old-style CPS race question were identified as White in the multiracial format, and about 10 percent were identified as Black. It should be kept in mind that the sample numbers for American Indians in the CPS are very small, about 300. Nonetheless, these findings suggest very clearly that the data for American Indians are much less consistent across tabulation methods than the data for other populations, especially Whites. This is further confirmed by the rates of misclassification shown in Table 3.

The data in Table 3 represent the degree of inconsistency between the tabulations obtained from the procedures outlined above and the reference distribution obtained from the monoracial forced choice survey question. Respondents are deemed misclassified when there is a mismatch between these instruments. Again, the results are not surprising. For Whites, the misclassification rates hover in the 3 to 4 percent range, regardless of the tabulation method employed. For American Indians and for "Others," the results are disappointing. The misclassification rates for both of these groups are in the 35 to 40 percent range. However, there is somewhat more variability for American Indians than for the "Others." The "Other" group is noteworthy because it likely includes a sizable number of multiracial persons who in lieu of checking a single box, prefer to identify themselves as something other than what appears in the standard list of categories.

One likely reason for the shortcomings in these methodologies is that they are thwarted by the presence of relatively large numbers of multiracial persons, especially among American Indians and "Others." In contrast, these procedures work very well for the White population that has relatively few such individuals. One might plausibly wonder about the effect of rising numbers of multiracial respondents. It is an empirical question whether, in



Table 1. Percent distribution of Whites and American Indians for bridge tabulation methods, May 1995 Current Population Survey

Race groups	Reference distribution	Deterministic whole assignment				Deterministic fractional assignment	
		Smallest group	Largest group other than White	Largest group	Plurality	Equal fractions	NHIS fractions
White	82.4	80.4	80.4	81.0	80.7	80.7	80.7
American Indian	0.7	1.2	1.0	0.8	0.8	1.0	0.9
Other <sup>a</sup>	16.9	18.4	18.6	18.2	18.5	18.3	18.4
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Source: U.S. Bureau of the Census. 1999. *The Bridge Report: Tabulation Options for Trend Analysis*, Appendix D, Table 5-B, unpublished report.

<sup>a</sup>Includes Blacks, Asian and Pacific Islanders, and Others.

Table 2. Percent distribution of race classification by bridging methods for persons who reported their race as American Indian or Alaska Native in the May 1995 Current Population Survey

Race assigned by bridging methods	Deterministic whole assignment				Deterministic fractional assignment	
	Smallest group	Largest group other than White	Largest group	Plurality	Equal fractions	NHIS fractions
White	24.5	22.2	24.5	24.5	23.3	24.1
Black	10.3	10.2	10.3	10.3	10.2	10.3
American Indian	62.9	62.9	60.4	60.4	61.7	60.7
Asian & P.I.	1.0	2.0	2.0	2.0	2.0	2.0
Other	2.8	2.8	2.8	2.8	2.8	3.0
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: U.S. Bureau of the Census. 1999. *The Bridge Report: Tabulation Options for Trend Analysis*, Appendix D, Table 6-A, unpublished report.

Table 3. Percent distribution of all respondents misclassified by bridge tabulation methods, May 1995 Current Population Survey

Race reported	Deterministic whole assignment				Deterministic fractional assignment	
	Smallest group	Largest group other than White	Largest group	Plurality	Equal fractions	NHIS fractions
White	3.6	3.6	3.3	3.3	3.4	3.4
Black	4.4	3.9	4.7	3.9	4.4	4.3
American Indian	37.1	37.2	39.6	39.6	38.3	39.3
Asian or Pacific Islander	5.9	5.8	6.4	6.6	6.1	6.2
Other	40.6	40.6	40.6	40.6	40.6	40.1
Total (all groups)	5.0	4.9	4.7	4.7	4.8	4.8

Source: U.S. Bureau of the Census. 1999. *The Bridge Report: Tabulation Options for Trend Analysis*, Appendix D, Table 10-A, unpublished report.

the future, high rates of inter-marriage will produce large numbers of persons who report a multiracial identity. However, one can still pose the question how these methodologies might work if there are larger numbers of multiracial respondents. To address this question, a report produced by OMB shows the outcome of increasing the numbers of multiracial respondents by factors of 2, 4, 6, and 8 (OMB 2000 Appendix C, Table 13A). This is a heuristic simulation of what might happen to the racial composition of the sample if there are larger numbers of multiracial respondents. The results of this exercise for American Indians are shown in Table 4.

Increasing the numbers of multiracial respondents leads to a small diminution of the White population. However, the American Indian population “grows” from 0.7 percent of the population to about 3.4 percent of the total under the smallest group procedure. Under all of these scenarios, the American Indian population will increase. However, it is only under the methodologies that do not privilege an ethnic minority identity (“largest group” and “plurality”) that the increase is negligible. In short, there is a relatively large population that can, if they chose to do so, claim to have some American Indian ancestry, and can thereby claim this as one of their racial identities. If these individuals decide in large numbers to claim this identity, the results in Table 4 suggest that we can expect to see the skyrocketing growth in the American Indian population witnessed since 1970 continue well into the future.

The recently released Census Quality Survey (CQS) yields additional insights into the challenges of bridging from earlier racial classifications to the one used in the 2000 census. This survey consisted of two panels. One panel consisted of respondents who were asked to respond with one race only and in a subsequent interview given the option of reporting one or more races. The second panel reversed the ordering of these questions. The respondents’ reports were also matched with those they provided in the 2000 Census. The first results from this survey provide two important insights about racial data from the 2000 Census.

One is that the data show a high degree of stability in the reports of persons who claim only one race, and a relatively high degree of instability among persons who report two or more races. For example, among persons who reported two or more races in the 2000 Census, about 40 percent changed their race or reported only one race in the CQS. For persons reporting their race only as Black or as White, about 98 percent consistently reported their race in subsequent interviews as shown in Table 5.<sup>7</sup>

The CQS is perhaps most fascinating with respect to the choices made by persons who reported more than one race in the census or in the initial interview of the CQS, and then were asked to respond to a question in which they were allowed to select only one race for their response. The data from the CQS indicate a complex, almost idiosyncratic set of responses that appear to reflect social perceptions of racial hierarchies in American society such as the “one drop rule” (hypodescent). For example, African-Americans who also reported their race in combination with White or American Indian most often selected “Black” as their race when permitted to report only one race.

<sup>7</sup> Mode effects may partially explain the instability of these results. However, they are not addressed in the CQS report.

Table 4. Percent of American Indian and Alaska Native respondents tabulated by bridge methods if multiple race responses increase by factors of 2, 4, 6, and 8 in the May 1995 Current Population Survey

Factors of increase	Reference distribution	Deterministic whole assignment				Deterministic fractional assignment	
		Smallest group	Largest group other than White	Largest group	Plurality	Equal fractions	NHIS fractions
1	0.7	1.2	1.0	0.8	0.8	1.0	0.9
2	0.7	1.5	1.3	0.8	0.8	1.1	0.9
4	0.7	2.2	1.7	0.8	0.8	1.4	1.0
6	0.7	2.8	2.1	0.8	0.8	1.7	1.1
8	0.8	3.4	2.5	0.8	0.8	2.0	1.2

Source: U.S. Bureau of the Census. 1999. *The Bridge Report: Tabulation Options for Trend Analysis*, Appendix D, Table 13, unpublished report.

Table 5. Consistency in the percent of non-Hispanic persons reporting their race in the 2000 Census and Panel A of the Census Quality Survey

	CQS initial contact (“Mark One or More Races”)		
	Single race	Two or more races	Total
Single race	98.7	1.3	100.0
Two or more races	60.1	39.9	100.0

Source: U.S. Census Bureau 2003. *Census Quality Survey to Evaluate Responses to the Census 2000 Question On Race: An Introduction to the Data, Census 2000 Evaluation B.3*. Washington, D.C.: U.S. Bureau of the Census.

## 5. Conclusion: The Continuing Challenges of Population Shifts and Demographic Methods

The experience of American Indians is not so different with regard to the matter of how one goes about the business of creating a relatively clear-cut racial classification; one that settles the question of who does and does not belong to a particular racial group. We have shown for the American Indian population how variability of racial classification and tabulation has become a complicated matter for American Indians. It could be just as complicated for other populations. Moreover, as we have already discussed earlier, independently of the classification and tabulation procedures for statistical purposes, federal civil rights enforcement rules insist on tabulations that favor the racial minority group, i.e., a limited form of hypodescent.

Notably, the guidance provided by the Office of Management and Budget explicitly states that allocation for enforcement purposes should not be confused with various allocation methods under consideration for statistical purposes such as “bridging” to past data collections to conduct trend or time series analysis (OMB 2000a). As shown above in the case of PL 94-171, the ability to produce useful racial and ethnic data that address both statistical and other equally important considerations is a difficult undertaking under the revised standard.

A related challenge is that the racial definitions for statistical purposes may vary from those used in federal program planning, implementation, and distribution of funds, especially for American Indians. In addition, given the foreseeable diversity within and across existing racial and ethnic populations (traditional civil rights groups, emerging immigrant groups, and multiracial persons), standard racial categories readily understood by the U.S. population are problematic.

A fourth challenge is that shifts in size, composition and distribution by race have been accompanied by greater geographic mobility of the U.S. population across states and physical residences. This geographic and residential mobility and diversity is recognized in examining complex households and relationships in the planning for the 2010 Census. That is, the concept of “residence rules” may undergo the same sort of reassessment (Schwede 2003) as the concept of “race” did in the planning of the 2000 Census. Given the above, basic demographic variables such as race, may no longer be sufficient for describing and understanding the diverse U.S. population. This is a relevant finding for federal policies and programs for the U.S. population.

Table 6. Percent distribution of single race responses reported by non-Hispanic persons who reported two or more races in Panel A of the Census Quality Survey

CQS initial contact ("mark one or more races")	CQS Recontact ("Choose one race") with follow-up probe									
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more (same)	Two or more (different)	Missing	Total
White – Black	11.9	33.8	0.0	0.0	0.0	17.4	20.5	1.8	14.5	100.0
White – AIAN	50.1	0.7	26.7	0.2	0.0	1.0	8.1	1.0	12.2	100.0
White – Asian	36.9	0.3	0.0	24.3	0.3	4.0	18.5	2.1	13.7	100.0
White – NHOPI	31.5	0.4	0.0	0.2	44.0	2.0	8.9	3.6	9.3	100.0
White – SOR	69.7	1.0	0.0	0.0	0.1	10.1	3.6	2.1	13.4	100.0
Black – AIAN	0.2	53.7	14.0	0.0	0.0	2.4	12.1	4.8	12.9	100.0
Black – Asian	0.1	56.2	0.0	10.1	0.0	3.7	11.5	5.2	13.2	100.0
Black – NHOPI	0.9	60.6	0.0	7.5	19.7	4.3	0.0	7.1	0.0	100.0
Black – SOR	4.5	71.3	0.0	4.0	0.0	10.4	1.3	5.1	3.4	100.0
AIAN – Asian	27.4	2.2	24.5	30.3	2.0	1.5	5.4	2.3	4.4	100.0
AIAN – NHOPI	0.0	0.0	7.8	0.0	69.0	0.0	12.0	0.0	11.2	100.0
AIAN – SOR	49.2	2.9	16.5	0.0	0.0	5.3	0.0	4.3	21.9	100.0
Asian – NHOPI	1.9	0.9	0.0	25.4	47.0	0.5	10.0	2.2	12.2	100.0
Asian – SOR	10.2	3.2	4.6	51.3	0.0	11.3	5.7	11.3	2.3	100.0
NHOPI – SOR	0.0	33.6	0.0	12.4	20.0	3.5	0.0	28.8	1.7	100.0
Three or more	14.8	14.4	3.4	6.5	17.5	10.8	6.1	11.1	15.4	100.0
White – Black	29.3	13.0	6.6	10.1	5.2	6.8	12.6	3.5	12.9	100.0
Total	50.1	0.7	26.7	0.2	0.0	1.0	8.1	1.0	12.2	100.0

Source: U.S. Census Bureau 2003. *Census Quality Survey to Evaluate Responses to the Census 2000 Question On Race: An Introduction to the Data, Census 2000 Evaluation B.3.* Washington, D.C.: U.S. Bureau of the Census.

Based on recent population shifts and geographical mobility, the U.S. Census Bureau has undertaken the next level of necessary statistical analysis to improve the identification of the U.S. population with the introduction of the Census Bureau Planning Database (PDB). As mentioned earlier, the PDB is an outgrowth of demographic analysis to measure and compare coverage of the U.S. population by the Decennial Census of Population and Housing data against historical and current demographic administrative records, namely vital statistics, immigration data, and Medicare enrollment records. In coverage evaluation of the 1990 census data, the Planning Database identified other characteristics of undercoverage, including housing tenure with renters less likely to be counted than homeowners, households that are not husband/wife households, persons below poverty level and receiving public assistance, unemployment, and linguistically isolated households. While these socio-economic characteristics may have intersected with demographic characteristics of race and Hispanic origin they were not synonymous. This finding serves as another measure of the growing racial and national origin diversity of the U.S. population, not just across groups but also within them. Furthermore, adding socio-economic and housing variables to basic demographic ones by census tracts improves the ability to identify and track difficult-to-enumerate populations by demography and geography over time. A logical research challenge will be to examine the Planning Database for its potential ability to complement or offset the variation provided by most recent racial classifications and tabulation procedure.

These challenges also underscore the importance of research that will produce “bridges” connecting contemporary and future racial classifications with older ones. Absent such bridges, assessing change over time is at best an imprecise exercise and impossible at worst. Fortunately, some progress has been made toward this end. Analysts from the U.S. National Center for Health Statistics have developed a probabilistic model for aggregating a multiracial classification into a smaller set of single race categories (Ingram et al. 2003). Although this classification appears to work reasonably well, it is of limited value because researchers wishing to use this model for other data sources must have access to detailed geographic data not readily accessible to most users of public data.

Another important bridging study has been recently published by Liebler and Halpern-Manners (2008). They begin with the basic approach used by the NCHS researchers and adapt it for an application to a public data source, the United States 2000 decennial census public use microdata sample. They provide evidence to support their argument that their approach is superior to other approaches, including those discussed in this article. However, a crucial shortcoming in both the NCHS and the Liebler and Halpern-Manners bridging models is that they can be used only to bridge backwards in time. Specifically, they require the convenient fiction that multiracial persons do not exist either in the past or the present. In effect, these bridging models are tools for erasing or otherwise reallocating the two or more races reported by multiracial respondents.

Acknowledging the empirical reality of multiracial people, presently and in the past, a more ideal solution would allow analysts to “bridge forward.” That is, a bridging model that would untangle older single race classifications into a set of categories consistent with modern classifications that allows respondents to choose more than one race. Needless to say, bridging forward is a much more complex task than bridging back by collapsing categories into single race responses.



In times of major population changes, it behooves data users to revisit their roles as data stewards and researchers (Harrison 2002). This article has shown that shifts in the characterization of a basic demographic characteristic, “race,” from monoracial to multiracial can result in significant changes in fundamental methods of classification and tabulation. Such changes, in turn, can greatly affect findings of the count of specific policy-relevant subpopulations and across the existing racial and ethnic populations of traditional civil rights minority groups, emerging immigrant groups, and multiracial persons. The federal statistical system must continue to explore how best to gauge and address these demographic realities.

## 6. References

- Bruce, A. and Robinson, J.G. (2003). *The Planning Database: Its Development and Use as an Effective Targeting Tool in Census 2000*. October 7, draft. Population Division, U.S. Census Bureau.
- Edmonston, B. and Schultze, C. (eds) (1995). *Modernizing the U.S. Census*. Washington, DC: National Academy Press.
- Edmonston, B., Goldstein, J., and Lott, J.T. (eds) (1996). *Spotlight on Heterogeneity*. Washington, DC: National Research Council.
- Farley, R. (2001). *Identifying with Multiple Races: A Social Movement that Succeeded but Failed?* Population Studies Center Research Report No. 01-491. Ann Arbor, MI: Population Studies Center, University of Michigan.
- Harrison, R. (2002). *Inadequacies of Multiple-Response Race Data*. *The New Race Question: How the Census Counts Multiracial Individuals*, Joel Perlman and Mary C. Waters (eds). New York: Russell Sage Foundation, 137–160.
- Harrison, R. and Bennett, C. (1995). *Racial and Ethnic Diversity*. *State of the Union: America in the 1990s, Volume Two: Social Trends*, Reynolds Farley (ed.). New York: Russell Sage Foundation.
- Hirschman, C., Alba, R., and Farley, R. (2000). *The Meaning and Measurement of Race in the U.S. Census: Glimpses into the Future*. *Demography*, 37, 381–394.
- Ingram, D.D., Parker, J.D., Schenker, N., Weed, J.A., Hamilton, B., Arias, E., and Madans, J.H. (2003). *United States Census 2000 Population with Bridged Race Categories*. *Vital Health Statistics, Series 2*, no. 135.
- Liebler, C.A. and Halpern-Manners, A. (2008). *A Practical Approach to Using Multiple-race Response data: A Bridging Method for Public-Use Microdata*. *Demography*, 45, 143–155.
- Lott, J.T. (1998). *Asian Americans: From Racial Category to Multiple Identities*. Walnut Creek, CA: Alta Mira Press.
- Lott, J.T. and Keller, J.K. (1999). *Continuity and Change: The Decennial Census in the 21st Century*. Paper presented at the American Statistical Association/Joint Statistical Meetings, August 7–12, Baltimore, MD.
- Office of Management and Budget (1997). *Recommendations from the Interagency Committee for the Review of the Racial and Ethnic Standards to the Office of Management and Budget Concerning Changes to the Standards for the Classification of Federal Data on Race and Ethnicity*. *Federal Register*, 62(131, July 9), 36874–36946.

- Office of Management and Budget (2000). Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity, Appendix C. Washington, DC: Executive Office of the President, Office of Management and Budget.
- Office of Management and Budget (2000a). OMB Bulletin No. 00-02: Guidance on Aggregation and Allocation of Data on Race for Use in Civil Rights Monitoring and Enforcement, March 29.
- Schwede, L. (2003). Complex Households and Relationships in the Decennial Census and in Ethnographic Studies of Six Race/Ethnic Groups. The Census 2000 Testing and Experimentation Program, Ethnographic Studies. Statistical Research Division, U.S. Census Bureau.
- Tucker, C., Miller, S., and Parker, J. (2002). Comparing Census Race Data under the Old and New Standards. The New Race Question: How the Census Counts Multiracial Individuals, Joel Perlman and Mary C. Waters (eds). New York: Russell Sage Foundation, 365–390.
- U.S. Census Bureau (2003). Census Quality Survey to Evaluate Responses to the Census 2000 Question on Race: An Introduction to the Data. Census 2000 Evaluation B.3. Washington, DC: U.S. Census Bureau.
- U.S. Commission on Civil Rights (1973). To Know or Not to Know: Collection and Use of Racial and Ethnic Data in Federal Assistance Programs. Washington, DC: U.S. Government Printing Office.
- Wallman, K.K., Evinger, S., and Schechter, S. (2000). Measuring Our Nation's Diversity: Developing a Common Language for Data on Race/Ethnicity. *American Journal of Public Health*, 90, 1704–1708.

Received November 2004

Revised June 2008