



2001

Using U.S. Census Data to Study Population Composition

Barry Edmonston

Follow this and additional works at: <https://commons.und.edu/ndlr>



Part of the [Law Commons](#)

Recommended Citation

Edmonston, Barry (2001) "Using U.S. Census Data to Study Population Composition," *North Dakota Law Review*. Vol. 77 : No. 4 , Article 3.

Available at: <https://commons.und.edu/ndlr/vol77/iss4/3>

This Article is brought to you for free and open access by the School of Law at UND Scholarly Commons. It has been accepted for inclusion in North Dakota Law Review by an authorized editor of UND Scholarly Commons. For more information, please contact und.common@library.und.edu.

USING U.S. CENSUS DATA TO STUDY POPULATION COMPOSITION

BARRY EDMONSTON*

I. INTRODUCTION

Demography has been traditionally concerned with a particular range of population characteristics—referred to as population composition—that are core areas of interest.¹ Although population composition includes a variety of basic personal, social, and economic characteristics of the population, the fundamental topics of study by demographers usually include age, sex, racial or ethnic origin, and nativity.² Population composition data are routinely collected in surveys and administrative records.³ However, national censuses provide the benchmark data in the United States and other countries for measurement of trends in population composition.⁴

This article begins with an overview of the historical and constitutional basis for the United States' decennial census and the need for census results. It then describes the serious and complex problems and issues associated with taking a modern census—concentrating on what has been learned over the past decade about options regarding reforms to the traditional census, the collection of data on population composition in the census, and the continuing controversies surrounding efforts to improve the census. Next, the article describes the characteristics of a good census. The article then discusses how the demographic data provided by the census on sex, age, racial and ethnic origin, and nativity is used to study population composition. It concludes with discussion of the challenges for redesigning future United States decennial censuses.

* Director, Population Research Center and Professor, School of Urban Studies and Planning, Portland State University. B.A., 1968, University of Oregon; Ph.D., 1974, University of Michigan.

1. DAVID YAUKEY & DOUGLAS L. ANDERTON, *DEMOGRAPHY* 1 (2d ed. 2001) (defining demography as "the study of . . . the size and composition of populations according to diverse criteria").

2. *Id.* at 1-2.

3. *Id.* at 17.

4. *Id.*

II. NEEDS FOR CENSUS DATA

In 2000, the United States conducted its twenty-second population census—a record of taking a census every ten years since 1790.⁵ The taking of a U.S. census is an invitation to participate in an activity that produces results which have far-ranging implications for all residents.⁶ Participating in the census reminds all Americans of their obligations and stakes in an accurate count of the population.⁷

The U.S. Constitution specifically requires that a decennial census be taken for reapportionment of seats in the U.S. House of Representatives.⁸ In practice, the census also provides the information needed for congressional districts to be drawn for the 435 House members.⁹ The fundamental and original purpose of the census in the United States was to determine the population count of every state for apportioning seats in the House of Representatives.¹⁰ Census data are also used to determine the boundaries of congressional, state, and local political districts.¹¹ The U.S. census, from the very beginning, has added information about the population beyond the absolute minimum required for reapportionment and redistricting.¹² In debates about the early censuses, congressional leaders recognized that it would be useful to collect some essential information about the American population as part of the census process.¹³ The 1790 census asked each household questions about the age and sex of occupants and family relationships; the 1820 census added questions about nativity and industry; and subsequent censuses added questions about occupation, military service, income, literacy, providing the variety of important

5. See MARGO J. ANDERSON, *THE AMERICAN CENSUS: A SOCIAL HISTORY* 7-31 (1988) (giving a history of the U.S. census, including the uses of census data for reapportionment).

6. MODERNIZING THE U.S. CENSUS 1 (Barry Edmonston & Charles Schultze eds., 1995).

7. BARBARA EVERITT BRYANT & WILLIAM DUNN, *MOVING POWER AND MONEY* 179-80 (1995).

8. The Fourteenth Amendment to the U.S. Constitution provides that the “[r]epresentatives . . . shall be apportioned among the several States according to their respective numbers, counting the whole number of persons in each State, excluding Indians not taxed.” U.S. CONST. amend. XIV § 2. “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.” U.S. CONST. art. I, § 2, cl. 3. This requires a census or headcount of the U.S. population to be taken every ten years, following the first census of 1790. ANDERSON, *supra* note 5, at 9.

9. See MARGARET MIKYUNG LEE, LIBRARY OF CONGRESS, ORDER CODE RL30870, *CENSUS 2000: LEGAL ISSUES RE: DATA FOR REAPPORTIONMENT AND REDISTRICTING*, at CRS-1 (Mar. 12, 2001) (stating that 2000 census data were released for use in reapportionment) [hereinafter *DATA FOR REAPPORTIONMENT AND REDISTRICTING*].

10. ANDERSON, *supra* note 5, at 8-9.

11. Edmonston & Schultze, *supra* note 6, at 22.

12. See *id.* at 19-29 (discussing the evolution of uses of census data).

13. ANDERSON, *supra* note 5, at 11.

information that have come to be collected in recent censuses.¹⁴ Thus, the census results are fundamental for congressional apportionment and redistricting and also for providing information to thousands of people who make decisions about things such as about public and private sector issues involving health and education, transportation planning, community services, housing, consumer marketing, economic strategies, and social equity.¹⁵ Census results measure progress and give direction for future actions.¹⁶

With the previous thoughts in mind, it is possible to analyze the 2000 census. In many ways, the 2000 census was successful, especially given the difficulties of counting a large and increasingly diverse and mobile population.¹⁷ But two difficult issues emerged after that census.¹⁸ First, the costs of the 2000 census exceeded the costs of previous censuses, even allowing for population growth, inflation, and declines in mail response rates.¹⁹ Second, and most prominent for public concerns, the 2000 census—like prior censuses—counted some population groups less thoroughly than others.²⁰ This led to the growing concern that census costs might continue to climb with no likelihood of gains in the completeness of coverage.²¹ Similar issues followed the 1990 census, and a number of experts evaluated the need for census data and whether there were more effective ways to ameliorate the twin problems of coverage errors and high costs.²² Based on the 2000 census, discussions must continue regarding the requirements for a modernized census in 2010 and beyond.²³

A. REAPPORTIONMENT

The fundamental purpose of the census is to provide data for reapportionment of the U.S. House of Representatives.²⁴ Article I of the U.S. Constitution mandates that an “actual Enumeration” shall be made of the

14. See HENRY S. SHRYOCK ET AL., *THE METHODS AND MATERIALS OF DEMOGRAPHY* 16 (1976) (listing questions included in the population censuses of the United States from 1790 to 1970).

15. *Id.* at 1-2.

16. *Id.* at 3.

17. REPORT OF THE EXECUTIVE STEERING COMMITTEE FOR ACCURACY AND COVERAGE EVALUATION POLICY, RECOMMENDATIONS CONCERNING THE METHODOLOGY TO BE USED IN PRODUCING THE TABULATIONS OF POPULATION REPORTED TO STATES AND LOCALITIES PURSUANT TO 13 U.S.C. § 141(c), at i (Mar. 1, 2001).

18. U.S. CENSUS MONITORING BOARD PRESIDENTIAL MEMBERS FINAL REPORT TO CONGRESS 6-9 (Sept. 1, 2001) [hereinafter CENSUS MONITORING BOARD].

19. *Id.* at 6-7.

20. *Id.* at 9.

21. *Id.*

22. Edmonston & Schultze, *supra* note 6, at 44-56.

23. CENSUS MONITORING BOARD, *supra* note 18, at 9.

24. ANDERSON, *supra* note 5, at 9-10.

population every ten years.²⁵ There are two key phrases in Article I.²⁶ First, the interpretation of the word “Enumeration” is obviously a key to assessing alternative census designs for counting the population.²⁷ The second important consideration is that reapportionment must be effected for the entire country at one time—the census cannot count some areas in one year and other areas in another.²⁸

Since the 1930 census, the 435 seats of U.S. House of Representatives have been automatically reapportioned upon delivery of the population counts from the Census Bureau.²⁹ Once the Secretary of Commerce transmits the decennial census count for each of the states to the President, the reapportionment of the Congress is determined quickly.³⁰ The U.S. Constitution specifies that the number of congressional seats is to be apportioned to each state according to its population.³¹ The Constitution does not, however, prescribe a procedure for carrying out the apportionment.³²

The major issue in distributing congressional seats to states is one of equity. Regardless of the total population size of the United States, the number of congressional seats must equal 435.³³ Furthermore, each state

25. U.S. CONST. art. I, § 2, cl. 3.

26. See MARGARET MIKYUNG LEE, CONGRESSIONAL RESEARCH SERVICE, LIBRARY OF CONGRESS, ORDER CODE 93-177A, LEGAL ISSUES FOR CENSUS 2000, at CRS-1 to CRS-4 (Feb. 6, 1993) (discussing the meaning of the phrasing of the census clause of Article I of the Constitution) [hereinafter LEGAL ISSUES FOR CENSUS 2000].

27. *Id.* at CRS-1 to CRS-2.

28. ANDERSON, *supra* note 5, at 13-14.

29. See *id.* at 131-58 (providing a history of the use of census data for reapportionment).

30. See 13 U.S.C. § 141(b) (2000). “[A]pportionment of Representatives in Congress among the several States shall be completed within 9 months after the census date and reported by the Secretary to the President.” *Id.*

31. U.S. CONST. art. I, § 2, cl. 3.

32. *Id.* (stating only “[r]epresentatives . . . shall be apportioned among the several States . . . according to their respective Numbers”).

33. Until 1911, Congress increased the size of the House of Representatives at each reapportionment so that no state lost congressional seats. See ANDERSON, *supra* note 5, at 131-58 (providing a thorough review of reapportionment and the methods used for reapportioning congressional seats based on census data). Since 1911, the number of congressional seats has been fixed at 435; an addition of a congressional seat for one state means that another state loses a seat. *Id.* at 139. Congressional apportionment of seats in the House of Representatives is based on the method of equal proportions. *Id.* at 150-53. In this method, each of the fifty states receives one congressional seat, no matter what its population. *Id.* at 150-51 & n.22. Next, a priority value is calculated for the *n*th seat for each state (for example, *n* equals 2 for a state’s second seat after it has received the first seat), using the state’s apportionment population multiplied by the factor: $1/\sqrt{n(n-1)}$. *Id.* The factor decreases as the number of seats increases: the factor equals 0.71 for the second seat and 0.41 for the third seat. *Id.* The state with the largest population value receives the next congressional seat. MICHEL L. BALINSKI & H. PEYTON YOUNG, FAIR REPRESENTATION: MEETING THE IDEAL OF ONE MAN, ONE VOTE 50-51 (2d ed. 2001). When each state has exactly one congressional seat, the state with the largest population receives the fifty-first seat. *Id.* at 62. But, after it does, the population apportionment factor is recalculated and compared with the priority value for all other states to decide which state receives the fifty-second seat. *Id.* at 118. The repeated calculation of priority values and assignment of congressional seats continue until all 435 seats are distributed. *Id.* at 49. One state will receive the 435th congressional seat, and the state with the highest priority value for the 436th seat will not receive one. *Id.*

must have at least one seat, and partial representatives cannot be assigned to a state.³⁴ In practice, each state is given one representative, leaving 385 seats remaining.³⁵ Then, each seat is assigned in turn to the state with the highest ratio of population to assigned seats.³⁶ Once the 435th seat is assigned to a state with the highest remaining ratio, the apportionment process ends.³⁷

B. REDISTRICTING

Once congressional seats are assigned to each state, the second part of the redefinition of the House of Representatives involves defining the geographic boundaries for each district.³⁸ From the 1920s through the 1950s, the courts generally declined to intervene in the political thicket of redistricting, and congressional legislative districts became increasingly more unequal in population size.³⁹ Many states chose not to redistrict, unless they gained or lost seats, and those that did often paid little attention to achieving population equality across districts.⁴⁰ Very large deviations in population, generally favoring rural over urban and suburban areas, became quite common.⁴¹ After the 1960 round of reapportionment and redistricting, the largest congressional district in the United States had over five times the population of the smallest district; the twenty most populous districts had a combined population of 14 million, compared with a combined population of 4.6 million for the twenty smallest districts.⁴²

The landmark "one-person, one-vote" Supreme Court decisions, beginning in the 1960s, drastically changed the requirements for redistricting, with the Supreme Court continuing to view population equality among

34. U.S. CONST. art. I, § 3.

35. BALINSKI & YOUNG, *supra* note 33, at 2.

36. ANDERSON, *supra* note 5, at 150-51 & n.22.

37. Adding population to a state will not necessarily increase the likelihood of gaining an additional seat. BALINSKI & YOUNG, *supra* note 33, at 48. It is not correct, as some have thought, to note the average population per congressional seat and conclude that, if a state has an equivalent population addition, an extra congressional seat would have been awarded in the apportionment procedures. *Id.* at 68. The critical question for a state is whether a change in its priority value would affect the assignment of the last several congressional seats, if the population count for all states were corrected for estimated undercoverage. *Id.* at 69. Corrections to the census population count would typically change the congressional delegations for those few states with priority values close to the 435th cutoff. *Id.*

38. See David McMillen, *Apportionment and Districting*, in ENCYCLOPEDIA OF THE U.S. CENSUS 38 (Margo J. Anderson ed., 2000).

39. See Gordon E. Baker, *Whatever Happened to the Reapportionment Revolution in the United States?*, in ELECTORAL LAWS AND THEIR POLITICAL CONSEQUENCES 258 (Bernard Grofman & Arend Lijphart eds., 1986).

40. *Id.*

41. *Id.*

42. *Id.* at 286 & n.17.

congressional districts as a critical constitutional requirement.⁴³ Over the same period, the Supreme Court issued decisions that greatly affected state and local redistricting.⁴⁴

In 1975, Congress required the Census Bureau to provide, within one year after the census date, decennial census population tabulations to state officials for legislative reapportionment and redistricting.⁴⁵ In practice, the basic tabulations have come to mean tabulations for individual blocks, the smallest geographic entity in census data products, for decennial years.⁴⁶ Although other data may be used, states and local areas have come to rely almost exclusively on census data for redistricting.⁴⁷

The legislative redistricting requirements discussed above for census data determine several key features that a census must have.⁴⁸ The census must include an attempt at a complete enumeration of the entire country at one point in time.⁴⁹ The census must provide basic data—primarily age, race and Hispanic-origin—at the smallest possible geographic level; legislative districts can then be drawn by combining smaller blocks to meet court-mandated criteria for equal population across districts and for

43. In the first of these cases, *Baker v. Carr*, 369 U.S. 186, 232 (1962), which involved Tennessee state legislative districts, the Court held that reapportionment and redistricting matters were subject to judicial review under the Equal Protection Clause of the Fourteenth Amendment. In *Wesberry v. Sanders*, 376 U.S. 1, 8 (1964), the Court held that under Article I of the Constitution, congressional districts must be as nearly equal in population as practicable. In *White v. Weiser*, 412 U.S. 783, 792 (1973), the Court rejected a Texas congressional redistricting plan in which the smallest district was about 4% smaller than the largest district. Also, in a five-to-four decision in *Karcher v. Daggett*, 462 U.S. 725, 729 (1983), the Court rejected a New Jersey congressional redistricting plan in which the smallest district was only 0.7% smaller than the largest district. The Court held that the state could have avoided such a deviation, as it had rejected a plan with a population deviation of only 0.45%. *Karcher*, 462 U.S. at 730. Furthermore, the Court ruled that the state had failed to show that the deviation in its approved plan was needed to achieve a legitimate goal. *Id.* at 742; see also F.R. Parker, *Changing Standards in Voting Rights Law, in REDISTRICTING IN THE 1990S: A GUIDE FOR MINORITY GROUPS* 55-66 (William P. O'Hare ed., 1989) (reviewing redistricting court cases after the enactment of the Voting Rights Act of 1965).

44. See generally *Mahan v. Howell*, 410 U.S. 315 (1972) (finding Virginia's reapportionment statute constitutional because it advanced Virginia's policy of respecting political boundaries); *Kirkpatrick v. Preisler*, 394 U.S. 526 (1969) (holding that Missouri's reapportionment statute creating congressional districts varying in population from 420,000 to 445,000 did not meet the U.S. Constitution's one-man one-vote requirement); *Gaffney v. Cummings*, 412 U.S. 735 (1973) (finding Connecticut's reapportionment plan proper because population deviations were not significant).

45. Population Tabulation for State Legislative Apportionment, Pub. L. 94-171, 89 Stat. 1023 (1975) (codified as amended at 13 U.S.C. § 141(c) (2000)).

46. Blocks are usually defined in terms of housing units along the four sides of a city block. DOWELL MYERS, ANALYSIS WITH LOCAL CENSUS DATA: PORTRAITS OF CHANGE 66 (1992). They are identified using a three-digit number, unique within each census tract that usually includes about fifty blocks. *Id.* Blocks usually include about thirty to forty housing units and number about eighty to ninety persons. See *id.* at 18 (showing average population of one block in Altoona, Pennsylvania, numbers eighty-five).

47. GARRINE P. LANEY, CONGRESSIONAL RESEARCH SERVICE, LIBRARY OF CONGRESS, ORDER CODE 95-896 GOV, THE VOTING RIGHTS ACT OF 1965, AS AMENDED: ITS HISTORY AND CURRENT ISSUES, at CRS-11 (Aug. 20, 2001).

48. LEGAL ISSUES FOR CENSUS 2000, *supra* note 26, at CRS-1 to CRS-5.

49. *Id.* at CRS-1.

appropriate representation of minority groups under the Voting Rights Act.⁵⁰

C. OTHER NEEDS FOR CENSUS DATA

Given the legal requirements for a census to obtain such basic data as age, race and Hispanic-origin, what other content is needed?⁵¹ The census is unique in statistical activities of the federal government because it is reasonably accurate for small geographic areas and for small population subgroups.⁵² The census offers an important opportunity to cost effectively collect a range of important social, demographic, and economic data.⁵³ In order to refrain from burdening respondents too much, the census gathers most of the information not mandated by the constitution from only a sample of the population.⁵⁴ For example, in the 2000 census about one-in-six households received a “sample” census questionnaire.⁵⁵ The federal government sponsors many surveys that provide more in-depth information on a particular subject, but such surveys provide estimates only for the United States and sometimes for larger states or metropolitan areas.⁵⁶ Administrative records also provide useful information, but they typically pertain to a particular group of recipients or those receiving specific benefits.⁵⁷ The census alone provides a broad range of information encompassing the whole population, which can be cross tabulated for small geographic areas and small population subgroups.⁵⁸

A review of census data by the Office of Management and Budget in 1994 revealed an important conclusion: almost all items in the 1990 census were required by federal government agencies to meet specific federal mandates.⁵⁹ For some items, legislation requires the item and specifies that

50. *Id.* at CRS-4 (discussing the implication of legal requirements for census data).

51. See 13 U.S.C. § 141(a) (2000) (“The Secretary shall . . . take a decennial census . . . in such form and content as he may determine. . .”).

52. MYERS, *supra* note 46, at 13.

53. *Id.* at 63.

54. *Id.* at 83.

55. U.S. Census Bureau Frequently Asked Questions, available at <http://www.census.gov/dmd/www/genfaq.htm#sampling> (reporting that the Census Bureau used a one-in-six sample for the “long form” questionnaire for the 2000 census) (last modified Jan. 28, 2002).

56. The U.S. Census Bureau conducts a variety of surveys, often sponsored by other federal government agencies. See U.S. Census Bureau Surveys, available at <http://www.census.gov/main/www/surveys.html> (showing a list of current surveys by the Census Bureau) (last modified Jan. 10, 2002). These surveys, however, do not provide small area estimates for the United States. *Id.*

57. Edmonston & Schultze, *supra* note 6, at 167.

58. *Id.* at 22-27 and app. D through H, at 259-341 (discussing other needs for census data, including examples).

59. *Id.* app. M, at 438 (documenting information on the Office of Management and Budget review).

it be obtained from the census.⁶⁰ For example, the 1975 and later amendments to the Voting Rights Act require the Census Bureau to determine political jurisdictions that must implement procedures for bilingual voting to protect the rights of language minorities.⁶¹ These determinations are made using sample census data on citizenship, educational attainment, and English-language ability, together with information on age, race, and ethnicity.⁶² Data are reported for counties, municipalities, and minor civil divisions.⁶³ For other census items, legislation requires the item but does not specify the census as the source, although the census is the only practical source.⁶⁴

In addition to legislatively mandated items, the census serves important functions in the federal statistical system.⁶⁵ Some of these functions include providing estimates for program management and government reports, such as information on the foreign-born, relied on by the Immigration and Naturalization Service; denominators for vital rates, such as birth and death rates; and reweighting household sample survey data to reflect more accurately the distribution of the population.⁶⁶

D. SUMMARY OF THE NEEDS FOR CENSUS DATA

To summarize from the previous discussion, the federal government has a constitutional requirement to conduct a census or decennial enumeration for the purpose of apportioning the Congress.⁶⁷ The Constitution requires a complete count of the population at a single point in time.⁶⁸ This means that once every ten years there needs to be a canvassing of the entire country with the attempt to count every person.⁶⁹

Given that the Census Bureau is canvassing the entire country, the cost of asking additional questions, especially on a sample basis for some respondents, is greatly reduced.⁷⁰ If additional sample questions do not have unreasonable costs nor reduce overall response rates, then there is a

60. *Id.*

61. The Voting Rights Act Amendments of 1982, Pub. L. No. 97-205, 96 Stat. 205 (codified as amended at 42 U.S.C. § 1973b(f)(2) (1994)).

62. Edmonston & Schultze, *supra* note 6, app. M, at 438-39.

63. *Id.* at 440-44.

64. *Id.*

65. *Id.*

66. *Id.* at 24-25.

67. U.S. CONST. art. I, § 2, cl. 3.

68. *Id.*

69. *Id.*

70. Edmonston & Schultze, *supra* note 6, at 117.

strong basis for collecting additional important data at the time of the decennial census enumeration.⁷¹

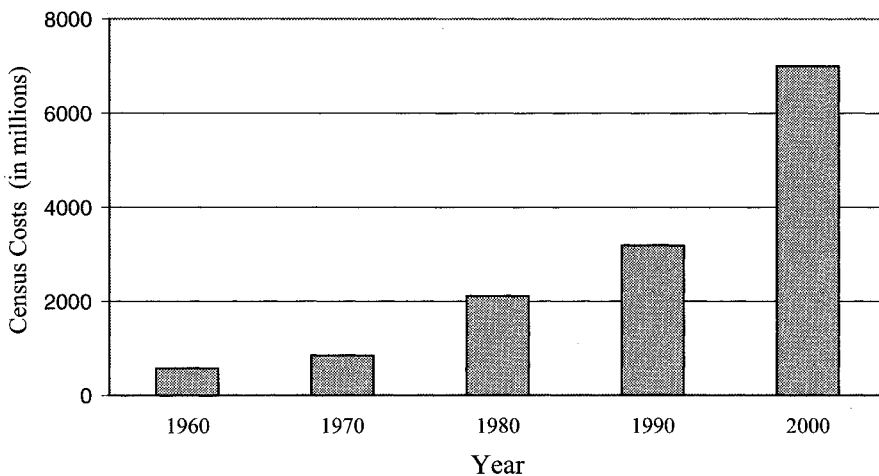
III. CONCERNS ABOUT THE CENSUS

There are several continuing concerns about various aspects of the decennial census: the rapid increases in census costs, the continuing worry that the census failed to count everyone, and a debate about whether the census questionnaire had become too long and complicated.⁷²

A. COSTS

The cost of census activities has increased sharply since 1970. In 2000 dollars, the 1960 census cost about \$574 million (see Figure 1).⁷³ The 2000 census cost \$7.0 billion, an increase of about \$6.4 billion or a 12-fold increase from 1960 after adjusting for inflation.⁷⁴

**Figure 1. Cost of the U.S. Decennial Census
in Constant 2000 Dollars, for 1960-2000**



71. *Id.* at 118.

72. *Id.* at 47-48.

73. Edmonston, 2001 computation, Figure 1; *see also* Edmonston & Schultze, *supra* note 6, at 48 tbl.3.1 (showing the cost of the 1960 census in 1990 dollars).

74. U.S. GENERAL ACCOUNTING OFFICE, REPORT TO CONGRESSIONAL REQUESTERS, DECENNIAL CENSUS: 1990 RESULTS SHOW NEED FOR FUNDAMENTAL REFORM 24-25 (June 1992) [hereinafter DECENNIAL CENSUS 1990]; *see also* Kenneth Prewitt, Director, Census Bureau, presentation at Bard College (Sept. 22, 2000) (transcript on file with author).

Taking inflation into account, census costs have increased for three reasons: decreases in the population mail questionnaire return rates, increases in population and housing, and an increasing use of highly labor intensive enumeration techniques.⁷⁵

Some of the cost growth can be attributed to the fall-off in mail response rates.⁷⁶ If a mail questionnaire was not returned to the census office, a field worker visited the address in an attempt to enumerate the household residents.⁷⁷ Under 2000 census procedures, as many as six visits could have been made.⁷⁸ Of the roughly \$6.4 billion inflation-adjusted cost increase between 1960 and 2000, somewhere between \$115 and \$275 million can be attributed to the fall in the response rate.⁷⁹

Some of the census cost increases can be attributed to growth in the population and the number of housing units to be counted.⁸⁰ Since census costs depend primarily on the expense of delivering a mail questionnaire to a household or having an enumerator visit a housing unit, it is more realistic to relate cost growth to the rise in the number of housing units rather than to population growth.⁸¹ Even when housing units are vacant and contain no household, there is a cost to ascertaining that fact; therefore, the number of housing units rather than the number of households is the relevant unit to consider for cost analysis.⁸² The number of housing units grew from 59 million in 1960, to 71 million in 1970, to 104 million in 1990, and to 116 million in 2000, accounting for about \$555 million of the 1960 to 2000 cost increases.⁸³

The remaining \$5.6 to \$5.7 billion in cost increases from 1960 to 2000 are primarily due to the Census Bureau's increased efforts to reduce census undercount through highly labor-intensive and expensive efforts to count every resident, including the procurement and installation of an entirely

75. DECENNIAL CENSUS 1990, *supra* note 74, at 24-25.

76. Edmonston & Schultze, *supra* note 6, at 44 tbl.3.1 (showing the 1970 to 1990 data); *see also* U.S. Census Bureau, Census 2000 Final Response Rates, *available at* <http://rates.census.gov> (showing 2000 data) (last visited Feb. 22, 2002) [hereinafter Census 2000 Final Response Rates]. The 1960 census used interviewers who visited each housing unit to collect census information. Edmonston & Schultze, *supra* note 6, at 48. The 1970 census was the first decennial census to use a primarily mail-out, mail-back questionnaire; the mail response rate was 78%. *Id.* The mail response rate declined to 75% in 1980, to 65% in 1990, and then increased slightly to 67% in 2000. *Id.*; Census 2000 Final Response Rates, *supra* note 76.

77. Edmonston & Schultze, *supra* note 6, at 44.

78. MARGO J. ANDERSON & STEPHEN E. FIENBERG, WHO COUNTS? THE POLITICS OF CENSUS-TAKING IN CONTEMPORARY AMERICA 294 (paperback ed. 2001).

79. Edmonston & Schultze, *supra* note 6, at 48.

80. *Id.* at 54.

81. *Id.*

82. *Id.*

83. *See id.* at 44-56 (detailing a comprehensive analysis of the factors responsible for recent census cost increases).

new technology to assist in the data collection process.⁸⁴ From 1960 onward, the Census Bureau has responded to important outside pressures to produce a “better” census: there was an increased demand for accurate population counts at very detailed geographic subdivisions and in hard-to-enumerate areas for purposes of congressional and legislative redistricting and otherwise carrying out the Voting Rights Act and its amendments.⁸⁵ At the same time, public cooperation with the census process, as measured by the mail response rate, declined and was lowest precisely in the areas where the pressures for an accurate count were greatest.⁸⁶ The Census Bureau responded by pouring on resources in highly labor-intensive enumeration efforts to count every person.⁸⁷

Although it may be convincingly argued that the value of the census data far exceeds projected costs, Congress and the Census Bureau are responsible for insuring that every effort be made to contain costs while providing data of the highest possible quality.⁸⁸ An increase in census costs per housing unit in 2000 dollars, from about \$10 in 1960 to more than \$60 in 2000, invites public discussion about the possibility of designing a more cost-effective census.⁸⁹

Two general observations form the background for debate about a future census.⁹⁰ First, there are no conceivable changes in the collection of census data that will simultaneously meet all of the following objectives: (1) continue a highly intensive census effort, relying principally on physical enumeration and labor-intensive follow-up techniques, to overcome the consequences of declining mail response rates; (2) provide detailed and reliable small-area data for redistricting and the Voting Rights Act; (3) provide the other housing and demographic data widely demanded for cross-tabulation at the level of small geographic areas; (4) reduce the differential undercount; and (5) keep costs from growing rapidly.⁹¹ There is, in short, no magic bullet.⁹²

Second, from an inspection of recent trends in census costs, efforts to decrease differential coverage and to deal with decreasing mail response rates, especially through highly labor-intensive enumeration techniques,

84. *Id.* at 55-56.

85. *Id.* at 56.

86. *Id.* at 44.

87. *Id.* at 46.

88. *Id.* at 46-47.

89. *See id.* at 48 tbl.3.1 (showing a household unit cost of \$9 in 1960) and CENSUS MONITORING BOARD, *supra* note 18, at 25 (showing a household unit cost of \$62 in 2000).

90. Edmonston & Schultze, *supra* note 6, at 56.

91. *Id.*

92. *Id.*

have been a key factor in driving up census costs.⁹³ Moreover, efforts to improve differential coverage have had an increasingly diminishing return—more money was spent in 1990 than in 1980, but there was no gain in coverage.⁹⁴ Differential coverage improvement efforts may have been carried to the point at which additional effort and expense may yield little or no improvement in coverage or in decreasing subgroup differences in net undercount.⁹⁵ Expensive efforts to improve census coverage are understandable given such forces as the impetus of the Voting Rights Act to provide detailed data on race and ethnicity at the block level.⁹⁶ Nevertheless, it is appropriate to ask if this continued effort to improve differential coverage, which has been unsuccessful so far and has increased census costs, is necessary for future censuses.⁹⁷

B. CONTENT

Historically, the census has collected additional content beyond the minimum set of items needed from all households for the constitutional purposes of reapportionment and redistricting.⁹⁸ Since 1960, most of the additional data have been collected on a separate sample form sent to a fraction of households.⁹⁹ These additional data are widely used and serve many important public purposes.¹⁰⁰ The nation needs the breadth of information that these census data provide, especially for small areas and small population groups.¹⁰¹

The process for determining the content of the census, both the short form and the sample form, has involved bringing together federal agencies and balancing their data needs against considerations of questionnaire length and feasibility.¹⁰² Views of other data users have also been sought.¹⁰³ This process has worked well.¹⁰⁴

93. *Id.*

94. *Id.*

95. *Id.*

96. *Id.*

97. *Id.*

98. *Id.* at 188. In general, the number of items for which data are obtained in the census and the major content areas have remained about the same since 1960. *Id.* at 187. For basic information on U.S. census questionnaires, including a count of the number of questions on census questionnaires in 1960, 1970, 1980, and 1990, see *id.* app. A, at 187-227. The census questionnaire comes in two versions: a short-form that every household receives and a longer sample questionnaire that is distributed to one-in-six households. *Id.* The number of housing questions has been reduced over time; hence, the burden of the short form on the population has declined somewhat over time. *Id.*

99. *Id.* at 188.

100. *Id.* at 117.

101. *Id.* at 132.

102. Constance F. Citro, *Content Determination*, in ENCYCLOPEDIA OF THE U.S. CENSUS 92 (Margo J. Anderson ed., 2000) [hereinafter *Content Determination*].

103. *Id.*

Given the importance of the broad range of data for small areas and small population groups that the census currently collects, questions also arise as to whether those data should be collected as part of the census or by some other means. Is the census the right vehicle to collect sample information? Do the added questions increase the costs of the census or impair the quality of the data?

The argument for the view that the sample form is a problem for the census can be formulated as follows: respondents find the long form unduly burdensome because of its length and complexity; this burden lowers the overall mail response return rate; this effect increases census costs; and the lower mail return rate may also contribute to population undercoverage.¹⁰⁵

Questions about the sample form are of intense interest to almost every user of census data and elicit strong opinions on all sides.¹⁰⁶ Many data users are impassioned in their defense of the sample form; they implicitly reject the idea that it hurts the basic census enumeration and argue in support of the rich range of data on various topics that it provides for small areas and small population groups.¹⁰⁷ Others see the sample form as a

104. *Id.*

105. See Edmonston & Schultze, *supra* note 6, app. A, at 188-90 (addressing questionnaire receipt and response).

106. The 2000 short form included seven questions, six related to population characteristics—name and telephone number, relationship to head of household, sex, age, Latino or Hispanic-origin, and race—and one related to housing—whether the housing unit is owner or renter-occupied. Constance F. Citro, *Long Form*, in *ENCYCLOPEDIA OF THE U.S. CENSUS* 92 (Margo J. Anderson ed., 2000). This was a reduction from the 1990 short form which contained six questions. *Id.* The sample form incorporates the questions on the short form and asks a total of fifty-two questions, compared to fifty-seven questions on the 1990 form. *Id.* One new question has been added since 1990: the sample form includes a question on grandparents who are primary caregivers for their grandchildren in order to collect information required by the welfare reform legislation. See *id.* at 274 tbl.1 (asking “[w]hether and how long responsible for grandchildren in home”). Copies of the 2000 census questionnaires are available at <http://www.census.gov/dmd/www/infoquest.html> (last visited Jan. 24, 2002).

The number of items on the sample questionnaire, population and housing combined, grew from forty-nine items in 1960, to sixty-four items in 1970, to sixty-five items in 1980, and declined to fifty-eight items in 1990. Edmonston & Schultze, *supra* note 6, app. A, at 192-94 tbl.A.2. Furthermore, the burden of the sample questionnaire has progressively declined in terms of the sampling rate. *Id.* app. A, at 189. Twenty-five percent of households were designated to receive the sample form in 1960, 20% in 1970, 19% in 1980, and 16% in 1990. *Id.* For a list of the number of census questions in the 1960 to 1990 censuses, see *id.* app. A, at 192-94 tbl.A.2.

Over time, the type of information on the sample questionnaire changes. *Id.* The 1990 census no longer asked about such items as washing machines and other appliances that most households owned. *Id.* But, over time, questions have been added on disability and ancestry. *Id.* The greatest number of questions are related to employment, occupation, and income. *Id.*

The content of the decennial census has been determined in recent decades by a council of federal agency representatives coordinated by the Office of Management and Budget (OMB). *Id.* at 114. The Census Bureau has determined the broad dimensions of the census questionnaire, the overall length of the short and sample form, and has evaluated questions in terms of their fitness and feasibility in a census context. *Id.* Within this framework, federal agencies have argued for items to serve congressionally-mandated data needs and to serve their program and policy needs. *Id.* Congress also plays a role because the Secretary of Commerce is required to provide Congress with the list of topics proposed for inclusion no later than three years before Census Day and to provide the proposed list of specific items no later than two years before Census Day. *Id.* at 115.

107. SHRYOCK ET AL., *supra* note 14, at 83-102.

threat to the cost and quality of the census data that are needed to serve the basic constitutional requirements.¹⁰⁸ In light of this controversy, we next review the evidence of the effects of the sample form on census costs, mail return rates, and coverage.

The sample form adds costs to the census in a number of ways, including: extra printing costs, extra postage, additional follow-up for every percentage point that the mail return rate for the sample form is less than that for the short form, additional editing and follow-up for item non-response, coding of such items as industry and occupation, and additional data processing and publication costs.¹⁰⁹ However, the evidence shows that the sample form, which is essentially a large sample survey on top of the massive effort undertaken for the complete (short-form) census, represents a marginal addition to total census costs.¹¹⁰ Moreover, the costs associated with the sample form do not explain the escalation in census costs that have occurred in recent decades.¹¹¹

The total costs of the sample form, including follow-up and all other costs, were about 9% to 10% (\$230 to \$250 million) of the \$2.6 billion costs of the 1990 census.¹¹² Cost estimates provided by the Census Bureau to the National Research Council's Panel on Census Requirements for the Year 2000 and Beyond suggest that the cost of the long form in the context of the 1990 census methodology may range from 11% to 12% of the total costs.¹¹³ Although cost estimates for the sample form for the 2000 census are not yet available, they are likely to be in the range of 10% to 20% because there were relatively few changes in the production and administration for the long form in the 2000 census compared to 1990 procedures.¹¹⁴ The overall evidence is that the marginal cost of incremental data on the census sample questionnaire is quite low relative to their usefulness.¹¹⁵

The sample form increases the burden of the census on the population, which may in turn reduce response rates relative to the short form.¹¹⁶ The effect of the sample form on mail return rates in 1980 was minimal: the mail return rate was 81.6% for the short form and 80.1% for the sample

108. Edmonston & Schultze, *supra* note 6, at 46-47.

109. *Id.*

110. *Id.* at 100.

111. *Id.* at 50-51.

112. *Id.* at 117 (citing S. M. Miskura, *Estimating the Full Cycle Costs For the Simplified Questionnaire Test (SQT)*, in 1 DESIGN 2000, ch. 30, No. 6 (1992)).

113. *Id.* at 122.

114. *Id.*

115. *Id.* at 117.

116. *Id.* at 113-39 (detailing an extensive discussion on the relationship of census content and mail return rates).

form, a difference of 1.5 percentage points.¹¹⁷ The sample form had a somewhat greater effect on mail return rates in 1990.¹¹⁸ The sample-form mail return rate was 4.5 percentage points below the short-form mail return rate (70.4% versus 74.9%).¹¹⁹ There is some evidence that that sample-form/short-form differential in return rates was greater in hard-to-enumerate areas.¹²⁰ Since only one-sixth of all households received the sample form, however, the difference in return rates reduced the overall mail return rates for the nation by less than one percentage point.¹²¹

Overall, the evidence is clear that the problems with mail return rates experienced in the 1990 census characterized the sample form almost to the same degree as the short form.¹²² In recent years, the Census Bureau has conducted extensive experiments with different form types and lengths and with other aspects of the mailout process.¹²³ Their research has identified some promising ways to increase mail returns by households for both the short and sample forms.¹²⁴ It may be that implementing the various improvements to the mailout process, including making the short form shorter and more user-friendly, will widen the differential between the short-form and the sample-form mail return rates.¹²⁵ One must be cautious, however, in assessing the available evidence because none of the tests of improvements to the short and sample form and the comparative effects on mail return rates has yet been conducted in anything approaching a census environment.¹²⁶

C. COVERAGE

Although the census count of the United State's population has never been complete, public concerns about the incompleteness have increased in recent decades.¹²⁷ The census is the sole basis for apportionment of congressional seats and is relied on heavily for the distribution of federal funds.¹²⁸ Improved statistical and demographic techniques permit the Census Bureau to estimate the incompleteness of the census with a greater

117. *Id.* at 118.

118. Because data are not yet available on the 2000 census, the most recent detailed evidence is from the 1990 census.

119. Edmonston & Schultze, *supra* note 6, at 119.

120. *Id.*

121. *Id.* at 119.

122. *Id.*

123. *Id.*

124. *Id.* at 119-21.

125. *Id.* at 121.

126. *Id.* at 121-22.

127. *Id.* at 140-41.

128. *Id.* at 140.

accuracy than in the past.¹²⁹ Thus, concern about census incompleteness springs, ironically, from the improved professional work of Census Bureau staff and the extraordinary expectations of the public for a “complete” census count.¹³⁰

Some undercount of the population occurs in the censuses of all countries.¹³¹ This section documents what is known about U.S. census undercoverage and discusses two major implications of census incompleteness: political representation at the congressional level and the distribution of federal resources that are allocated on the basis of population.

Coverage estimates, which measure the extent to which the census counts all the people, are made by two methods.¹³² One method is to conduct a larger sample survey in conjunction with the decennial census, match all individuals in the survey to those reported in the census, and then estimate the number of unenumerated people in the census by age, sex, and race.¹³³ The second method, demographic analysis, is to develop an estimate of the population independent of the census, using birth and death records for previous years, immigration and emigration data, and previous censuses.¹³⁴

Demographic estimates are the primary means for comparing censuses over time for the nation as a whole.¹³⁵ Demographic analysis of coverage shows that the net national undercount, the number of people omitted minus the number overcounted, was estimated at 7.0 million in 1940, 6.3 million in 1950, 5.6 million in 1960, 5.5 million in 1970, 2.8 million in 1980, and 4.7 million in 1990.¹³⁶ The undercount rate dropped steadily from 1940 (5.4%) to 1980 (1.2%), before rising in 1990 (1.8%) for the first time in fifty years.¹³⁷

Based on the criterion of net undercount, the 1990 census was somewhat worse than the 1980 census.¹³⁸ According to the 1990 estimates by demographic analysis, almost three-fourths of the net national undercount was nonblacks (primarily whites).¹³⁹ The rate of undercount, however, was

129. *Id.* at 140-41.

130. *Id.* at 55.

131. See SHRYOCK ET AL., *supra* note 14, at 29-47 (discussing census data, including errors that occur in population censuses).

132. See Edmonston & Schultze, *supra* note 6, at 30-43 (discussing methods for measuring census coverage and its estimates). This volume provides the source for the census coverage figures cited in this section.

133. *Id.* at 30.

134. *Id.* at 30-31.

135. *Id.* at 31.

136. *Id.*

137. *Id.*

138. *Id.*

139. *Id.* at 33.

over four times higher for blacks than for nonblacks, 5.7% and 1.3%, respectively.¹⁴⁰ In the 1990 census, the undercount rate for both men and women was also about four to five times higher for blacks than for nonblacks; it varied from 8.5% for black men to 0.6% for white women.¹⁴¹

Demographic analysis of the difference between black and nonblack net undercount rates shows modest increases from 1970 to 1990.¹⁴² In 1940, the black undercount rate was 3.4 percentage points higher than the nonblack rate.¹⁴³ The difference has increased since 1940; it reached 4.4 percentage points in 1990, which is higher than at the beginning of massive efforts by the Census Bureau in 1970 to narrow the difference.¹⁴⁴

Special decennial census surveys in 1990 and 2000 reveal that net undercount rates are also higher for Asian and Pacific Islanders, Hispanics, American Indians and Alaskan Natives than for whites.¹⁴⁵ Comparing net undercount rates to whites in the 2000 census, rates were more than 40% higher for Asians, more than four times higher for Latinos, and seven times higher for on-reservation American Indians.¹⁴⁶ The undercount of Asian and Hispanic groups is likely to have been influenced by the relatively large numbers of people who are foreign-born and who may not have understood census questionnaires and procedures.¹⁴⁷

There are several implications for undercount for minority groups.¹⁴⁸ In political representation and funding based on population, undercounted groups get less credit for their population than they are due.¹⁴⁹ Political districts for undercounted areas, drawn relative to population, are “overpopulated,” meaning they have more people than the official data report, than comparable accurately counted districts.¹⁵⁰ “Overpopulated” districts result in underrepresentation of minority areas (i.e., fewer districts)

140. J. Gregory Robinson et al., *Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis*, 88 J. AM. STAT. ASS'N 1061, 1061-79 (1993).

141. Edmonston & Schultze, *supra* note 6, at 33.

142. See *id.* at 32 tbl.2.1 (showing the difference in the net population undercount rate between blacks and all other population groups, 1940 to 1990).

143. *Id.*

144. *Id.*

145. Press Release, U.S. Census Bureau, Statement by William G. Barron, Jr. on the Current Status of Results of Census 2000 Accuracy and Coverage Evaluation Survey (July 13, 2001).

146. *Id.*

147. Jorge Chapa, *Hispanic/Latino Ethnicity and Identifiers*, in *ENCYCLOPEDIA OF THE U.S. CENSUS* 245 (Margo J. Anderson ed., 2000).

148. Jeffrey S. Passel, *What Census Adjustment Would Mean*, *POPULATION TODAY*, June 1991, at 6.

149. *Id.* at 6-7.

150. Michael Cohen, *Coverage Evaluation*, in *ENCYCLOPEDIA OF THE U.S. CENSUS* 101 (Margo J. Anderson ed., 2000).

at all levels of government—federal, state, and local—that base political representation on population size.¹⁵¹

1. *Effect of Undercoverage on Congressional Apportionment*

Underenumeration in the census has serious political, economic, and social implications.¹⁵² The decennial population count, reported in the census, affects the state apportionment of seats in the U.S. House of Representatives and the geographic boundaries for congressional districts, state legislative districts, and city council districts.¹⁵³ Under the “equal proportions” methods for federal apportionment, a shift of relatively few people could result in a change in a state’s representation.¹⁵⁴

For all state and local districts, the possibility that undercoverage will have an effect on a district’s boundaries depends on the size of the district; the coverage rates by age, sex, and race; the distribution of the population by age, sex, and race; and the undercoverage rates of contiguous districts.¹⁵⁵ If undercoverage were eliminated, population counts would increase in areas with a large number of high-undercount groups, such as minorities and inner-city residents.¹⁵⁶ An estimate of what the 1990 census results would have been, if statistical methods had been used to adjust the count, can be obtained by using Census Bureau survey information to make a “correction” for each state’s population.¹⁵⁷ If used in 1990, a correction, using a set of adjustment factors taking age, sex, race, region, and urban-rural characteristics into account, would have been made for each of 7 million census blocks in the nation.¹⁵⁸

What would have happened to the 1990 reapportionment if the state populations had been corrected for undercoverage? It is difficult to determine precisely how the application of corrections for the estimated undercoverage in the 1990 census would have affected actual congressional reapportionment, because the adjustment would have been done for small geographic areas.¹⁵⁹ If correction factors had been applied to each state’s population data, three states would have gained a congressional seat—Georgia a twelfth seat, Montana a second seat, and California a fifty-third

151. *Id.* at 95.

152. *Id.* at 100.

153. Passel, *supra* note 148, at 7.

154. *Id.*

155. *Id.* at 7-8.

156. *Id.* at 8.

157. *Id.*

158. *Id.*

159. *Id.*

seat—and three states would have lost a seat—Oklahoma, Pennsylvania, and Wisconsin.¹⁶⁰

North Carolina received the 435th seat in the reapportionment based on the 2000 census.¹⁶¹ If North Carolina's population had been 3,807 persons fewer, or 0.04% of the state's total, it would not have qualified for the 435th seat.¹⁶² At the same time, Utah failed to qualify to gain another seat by a population of 856, or 0.04% of its total population.¹⁶³

If correction factors had been used to correct each state's population for undercount in the 2000 census, there would have been a shift of one out of the 435 congressional seats.¹⁶⁴ In 2000, Texas would have gained one seat with Ohio losing one seat.¹⁶⁵ It should be noted that very small population differences affect the assignment of congressional seats close to the cutoff—regardless of whether the decisions are based on population figures including or excluding corrections for undercoverage.¹⁶⁶ Because of California's large population size, relatively small percentage changes in its population can add or subtract a congressional seat; California would have either gained one or lost two congressional seats in 2000, for example, if its total population were two percentage points different.¹⁶⁷

Congressional redistricting would be affected to a greater extent than apportionment because virtually all congressional districts, except for those in single-state districts, would have their boundaries changed by adjusted census block data.¹⁶⁸ Moreover, a census that is corrected for undercoverage in the physical enumeration would affect the redistricting for state legislatures and city councils that rely on decennial census data.¹⁶⁹

2. *Effect of Undercoverage on Federal Distribution of Funds to States*

The undercount also affects the distribution of federal and state funds, which are allocated based on population.¹⁷⁰ Funds for education, health, transportation, housing, community services, and job training are allocated to geographic areas on the basis of population size and social and economic

160. *Id.* at 7. Similar analysis using 2000 census data has not been done.

161. Barry Edmonston, *The Case for Modernizing the U.S. Census*, TRANSACTION SOCIAL SCIENCE AND MODERN SOCIETY, Nov./Dec. 2001, at 47, tbl.2 [hereinafter SOCIETY].

162. *Id.*

163. *Id.*

164. *Id.*

165. *Id.*

166. *Id.*

167. *Id.*

168. *Id.* at 48.

169. *Id.*

170. *Id.*

factors.¹⁷¹ In 1990, the federal government disbursed about \$125 billion to state and local governments, and nearly half of this amount was distributed using formulas involving census population data.¹⁷² Several studies have examined the effect of adjusting for census undercount on the distribution of funds to state and local governments.¹⁷³ All studies of the 1970, 1980, and 1990 censuses concluded that the impact of census population adjustment on grant allocations would have been small.¹⁷⁴

The total federal allocation of grants involving census counts in 1989 was \$58.7 billion, or about \$236 per capita for eligible population jurisdictions.¹⁷⁵ For several reasons, however, adjusting the allocation for undercount would not simply result in an additional \$236 per unit of net undercount.¹⁷⁶ First, population is only one of several factors in many federal formula grant allocations; in such programs, an increase of population results in only a partial increase in funding.¹⁷⁷ Second, although many grant allocations increase with population gains, there are some programs in which funding is *reduced* with population growth.¹⁷⁸ Finally, and most importantly, federal grant formulas are largely fixed in their total amount.¹⁷⁹ In reality, as the total U.S. population is increased by correcting for estimated undercount, smaller amounts of funds per capita would be available for allocation.¹⁸⁰ If, for example, a fixed sum was apportioned among geographic areas on the basis of population size alone and the population of every geographic area doubled, there would be no change in funds allocated to any area, only a reduction in the per capita amount.

Overall, there were 108 federal programs that had state and local obligations of \$58.7 billion in 1989, based on population counts.¹⁸¹ Five major programs accounted for \$51.8 billion of the total.¹⁸² Of the major programs, Medicaid is the largest, and the Highway Planning and Construction Program is the second largest.¹⁸³ The overall obligation per capita

171. See Edmonston & Schultze, *supra* note 6, at 40-43 (describing the use of census data for federal fund allocation).

172. Michael P. Murray, *Census Adjustment and the Distribution of Federal Spending*, 29 DEMOGRAPHY 319 (1992).

173. See *id.* at 319-21 (reviewing earlier studies).

174. *Id.* at 320.

175. *Id.* at 319-32.

176. *Id.* at 320.

177. *Id.* The Community Development Block Grant Program in the Department of Housing and Urban Development is an example of such a program.

178. *Id.*

179. *Id.*

180. *Id.*

181. *Id.* at 319.

182. *Id.*

183. *Id.* at 322, 327.

for all federal programs was \$236 in 1989.¹⁸⁴ When each state and local jurisdiction count is corrected for estimated undercount, some states and local areas lose and some gain.¹⁸⁵ The overall amount per net undercounted person among gaining areas is about \$56, considerably less than the average per capita obligation.¹⁸⁶ Only 34% to 41% of areas would gain because of corrected population counts.¹⁸⁷ Many governments with only modest population undercounts would not, in fact, actually gain additional federal grant money.¹⁸⁸ The effect of adjustment on redistribution of federal funds would be modest: only 0.32%, or about \$190 million, of the total federal obligations would be altered by correcting the population count.¹⁸⁹

By definition, the distribution of money under these programs would change if there were a differential change in the population count.¹⁹⁰ The effect of the undercount on each state's share of a fixed total of funds distribution depends on state characteristics.¹⁹¹ Moreover, the amount of money gained and lost is obviously related not only to the estimated undercount rate, but also to population size.¹⁹²

In federal funding allocation programs, social and economic factors as well as population counts are used.¹⁹³ Using these other factors demonstrates the importance of enumeration and accurate data concerning people counted, in order to advance optimal program planning and equitable distribution of funds.¹⁹⁴ The reduction of population undercount and the improvement of accuracy of collected data are both important to the Census Bureau, which needs to provide accurate data to ensure the fair funding of federal programs.¹⁹⁵

IV. WHAT IS A GOOD CENSUS?

There are major national policy choices involved in considering requirements and techniques for the decennial census.¹⁹⁶ Some options for

184. *Id.* at 319.

185. *Id.* at 329 (stating that adjustment leads to reallocating 0.46% of the total grants from these programs with no state losing more than 1.5% of its original grant and no state gaining more than 3%).

186. *Id.* at 330.

187. *Id.*

188. *Id.*

189. *Id.*

190. *Id.* at 320 (stating all formula-based grants have a fixed sum of money to distribute; thus as population count rises, the amount available to each person shrinks).

191. *Id.* at 321.

192. *Id.* at 320.

193. *Id.* at 324-26.

194. *Id.* at 326.

195. *Id.* at 330.

196. DATA FOR REAPPORTIONMENT AND REDISTRICTING, *supra* note 9, at CRS-31.

conducting future censuses illustrate the dilemma in thinking about how future censuses should look.¹⁹⁷ A first choice involves how to implement the basic requirements of court decisions mandating equality in the population size of voting districts.¹⁹⁸ Both the court decisions and the Voting Rights Act—grounded firmly in the Constitution, including the Fourteenth and Fifteenth Amendments—mandate census data requirements.¹⁹⁹ They specify that the decennial census must be carried out at a single point in time, once every ten years, and must attempt to count every resident.²⁰⁰

A second choice involves the level of information beyond basic demographic detail that must be provided for small areas or small population groups.²⁰¹ A large number of programs and funding allocations are based on census data.²⁰² In theory, many of these programs and funding allocations could rely on either a large ongoing household survey or administrative records.²⁰³ The required detail for the decennial census is limited to a few key demographic characteristics—including age, sex, and race—that must be tabulated for small geographic areas.²⁰⁴

Third, there are tradeoffs involving costs, accuracy, and the acceptability of statistical adjustment.²⁰⁵ A census in which physical enumeration efforts are reduced through the use of a sample-based follow-up would lower overall costs.²⁰⁶ A census relying more on statistical techniques for adjustment would improve census coverage at the national level and for most state and metropolitan areas, while reducing differential undercoverage for race and Hispanic-origin groups.²⁰⁷ As described in the next section, the conventional 1990 census could be redesigned in order to improve population coverage and reduce costs.²⁰⁸ Such a redesigned census would not be perfect—no census in any country counts every person—but

197. *Id.* at CRS-26 to CRS-27.

198. *Id.*

199. *Id.* at CRS-4 to CRS-6. A comprehensive survey of legal issues bearing on the requirements for the U.S. census, including citation of relevant court decisions, can be found in this report. *Id.*

200. *Id.* at CRS-1 to CRS-2.

201. *Id.* at CRS-4.

202. Edmonston & Schultze, *supra* note 6, at 1.

203. *Id.* These household surveys would need to be large enough, for example, to provide estimates for states and larger urban areas.

204. *Id.* at 46.

205. *Id.* at 56.

206. *Id.*

207. *Id.* at 137.

208. *Id.* at 107-10.

would offer substantial improvements.²⁰⁹ A census needs to be supported by a broad array of Americans in its design and outcome.²¹⁰

A. A MODERNIZED CENSUS

There are two broad approaches to counting the population: the traditional census and a redesigned census.²¹¹ The traditional approach, used in the 2000 census, relies completely on intensive physical efforts to count the entire population.²¹² In recent decades, direct enumeration has been based primarily on mail-out, mail-back questionnaires.²¹³ In some areas and for all households that do not return a questionnaire, the Census Bureau sends out an enumerator to collect information for the household or for people not living in households.²¹⁴

The traditional census approach begins with construction of an address register, including elaborate procedures to improve its comprehensiveness.²¹⁵ Census forms are then mailed to a comprehensive list of residential addresses, with instructions to mail back the completed questionnaire.²¹⁶ Not all households return their completed questionnaire within a reasonable period of time.²¹⁷ For households that do not respond to the main questions about the number of family members and their key demographic characteristics,²¹⁸ census enumerators undertake an intensive follow-up effort to determine whether the unit is occupied and, if so, to contact the household and obtain responses.²¹⁹ Repeat visits are made, administrative records are

209. *Id.*

210. *Id.* at 110-11.

211. *Id.* at 76.

212. *Id.*

213. *Id.*

214. The process of conducting the decennial census is a complex operation. *Id.* app. B, 228-38 (presenting an overview of the U.S. census process). The major components of the 2000 census involved the following: develop a master address file for every household in the nation; mail out a questionnaire to almost every household; monitor the mail-back for each questionnaire; enumerate selected rural areas and American Indian lands; develop sampling for the long-form census questionnaire; follow-up all nonresponding households; plan a national campaign for coverage improvement; evaluate the coverage of the census for major groups and regions; record, edit, and process census data; hire and train an efficient staff and organize national field operations; develop geographic tools and materials; design research and experimentation; and develop data products in a variety of formats to meet the needs of a broad range of users. *Id.* at 228. These activities are carried out over a period of time, concurrently and sequentially. *Id.*

215. *Id.* at 238 (citing U.S. DEP'T OF COMMERCE, CENSUS BUREAU, PLANNING AND CONDUCTING THE 1990 DECENNIAL CENSUS (June 1991)). A similar discussion of the 2000 census is not yet available.

216. *Id.* at 229.

217. *Id.* at 230-31.

218. The mail response rate for the 2000 census was 67%, meaning that about one-third of households did not return their mail census questionnaire and needed to be contacted by a census enumerator. *Releases*, U.S. Census Bureau, available at <http://www.census.gov> (last visited Mar. 4, 2002).

219. Edmonston & Schultze, *supra* note 6, at 231.

sometimes examined, and special programs to contact particular groups are carried out; these groups include such as homeless people, people in institutions, dormitories, barracks, and other people who do not live in regular household settings.²²⁰ The process is continued for an extended period of time to physically enumerate every household and all the people in every household.²²¹

Extensive special programs have been directed toward coverage improvement in recent censuses.²²² These programs are expensive, both in absolute terms and often in terms of the cost per person or housing unit.²²³ These special programs have included, for example, a follow-up of people reporting a change of address to the U.S. Postal Service during the census enumeration period, a campaign to find persons missed from the census by contacting community organizations or visiting places frequented by transients, and matching administrative records to census lists for selected areas.²²⁴ The results from the returned mail questionnaires, enumerator follow-up, and intensive special coverage improvement efforts are combined to produce the actual census count of the U.S. population—reported by the Census Bureau for the nation and for subdivisions down to the block level.²²⁵

The traditional approach of direct enumeration has been subject to the two basic criticisms discussed earlier: high and rapidly rising costs and high differential undercount.²²⁶ The traditional approach has been pushed well beyond the point at which it adds to the overall accuracy of the census count.²²⁷ As noted earlier, the costs of taking the census have escalated sharply, even after allowing for inflation and population growth.²²⁸

Furthermore, the 2000 census produced a net undercount of 1.2% for the nation as a whole.²²⁹ This net undercount included overcounting in some areas and among some groups, which was more than offset by undercounting among other areas and groups.²³⁰ There are no new procedures available that could substantially reduce differential undercount, working within the traditional census approach.²³¹ Historically undercounted popu-

220. *Id.* at 232-34.

221. *Id.* at 234.

222. *Id.* at 235-36.

223. *Id.*

224. *Id.*

225. *Id.* at 237.

226. *Id.* at 76.

227. *Id.* at 55.

228. *Id.*

229. CENSUS MONITORING BOARD, *supra* note 18, at 29 tbl.2.

230. *Id.*

231. Edmonston & Schultze, *supra* note 6, at 78.

lation groups will likely not witness improved coverage through a more intensive, expensive traditional census.²³²

The alternative approach, called a redesigned census, would combine an initial stage of direct counting with various statistical estimation techniques.²³³ Correctly designed, this approach would rely on a large independent survey to produce statistical estimates—by area, racial group, and other relevant demographic characteristics—of the net undercount or overcount contained in the census data.²³⁴ By designing the survey as an integral component of the census, the census could eliminate all operations that add relatively little to accuracy but have high unit costs.²³⁵ The traditional labor-intensive physical enumerative design could be redesigned to simultaneously reduce costs and improve accuracy and coverage.²³⁶

In theory, a redesigned census would contain costs, reduce error in the population count, and improve data quality.²³⁷ Data quality improvement could be accomplished through four components: increased use of statistical techniques, improvement of mail response rates, reengineering to achieve cost reductions, and increased development of public support.²³⁸

B. STATISTICAL ESTIMATION

A good census would use statistical estimation to simplify and truncate individual follow-up after reasonable efforts to count everyone had been expended, using evaluation surveys to complete the count and reduce differential undercount.²³⁹ The evidence from recent censuses is that traditional reliance on physical enumeration has been pushed well beyond the point at which it adds to the overall accuracy of the census.²⁴⁰ Moreover, such traditional methods still result in a substantial undercount of minority populations.²⁴¹ It is likely that statistical estimation can be used, in combination with the mail questionnaire and follow-up of nonrespondents on a reduced scale, to produce a better census at reduced costs.²⁴²

The use of statistical methods for sampling nonrespondents and surveys to complete the count recognizes that modern statistical procedures

232. *Id.*

233. *Id.* at 78-79.

234. *Id.* at 79.

235. *Id.* at 79-80.

236. *Id.*

237. *Id.* at 107-11.

238. *Id.*

239. *Id.* at 93.

240. *Id.* at 86.

241. *Id.* at 78.

242. *Id.* at 95.

could produce better data for the country as a whole and for large areas and population groups by reducing the differential undercount.²⁴³ The use of statistical techniques, however, would somewhat increase the variability for small areas.²⁴⁴ If the major concern with improving the population coverage in the nation's decennial census lies in the national and major population groups, then it is important to recognize that these data are improved at some increase in the variability of small-area data.²⁴⁵

Efforts to follow up individually those who fail to return the mail questionnaire would be simplified and truncated after a reasonable effort, and statistical sampling would be used to estimate the number and characteristics of the nonrespondent households that remain.²⁴⁶ In addition, data quality surveys would be undertaken to improve the overall count and reduce differential undercount.²⁴⁷

Statistical estimation techniques have long been used in the census for a number of purposes, including closeout procedures, determining vacant units, and imputation of missing census responses.²⁴⁸ Using statistical methods for sampling nonrespondents and surveys to complete the count is formal recognition that modern statistical procedures can improve the census process, reduce costs, and produce better data by reducing the differential undercount for the country as a whole and for large areas and population groups.²⁴⁹

243. *Id.* at 99.

244. *Id.* at 100.

245. *Id.*

246. *Id.* at 86-88.

247. The Census Bureau has long used sampling in its regular national household surveys. For a historical review, with citations, of the Census Bureau's work on statistical methods and sample surveys, see TOMMY WRIGHT & JOYCE FARMER, U.S. CENSUS BUREAU, A BIBLIOGRAPHY OF SELECTED STATISTICAL METHODS AND DEVELOPMENT RELATED TO CENSUS 2000, STATISTICAL RESEARCH REPORT SERIES NO. RR 2000/02 (Aug. 10, 2000). Sampling has had an important role in the decennial census through its use in collecting additional information from a sample of the population. *Id.* at iii-iv. Sampling has two potential uses in a redesigned census. Edmonston & Schultze, *supra* note 6, at 86-87.

The first is to estimate the number and characteristics of people who were not found, even after attempts at nonresponse follow-up, and people who were counted more than once or should not have been counted. *Id.* Sampling for nonresponse follow-up takes place within the actual process of physical enumeration. *Id.* The number of nonresponding households and their specific address is known in the census, so the purpose of sampling from nonresponding households is to estimate the number of people at each address and their characteristics. *Id.*

The second use of sampling in a redesigned census is to evaluate the completeness of the physical enumeration. *Id.* A large household survey would be conducted, independent of the decennial census enumeration, in order to provide modeled estimates of net census undercoverage for small areas. *Id.* This data quality survey would not check on undercoverage for specific addresses, but it would provide undercoverage estimates at the level of census blocks. Tommy Wright, *A One-Number Census: Some Related History*, 283 SCIENCE 491 (Jan. 1999) [hereinafter *A One-Number Census*]. The Census Bureau refers to this use of sampling to complete the count as an integrated coverage measurement survey. Constance F. Citro, *Sampling in the Census*, in ENCYCLOPEDIA OF THE U.S. CENSUS 328 (Margo J. Anderson ed., 2000).

248. *A One-Number Census*, *supra* note 247, at 492.

249. Edmonston & Schultze, *supra* note 6, at 84-85.

C. REDUCING COSTS

Greater reliance on statistical methods could reduce census costs by 10% to 20% of overall census costs when applied to a traditional census design.²⁵⁰ But if reliance is placed on statistical methods to improve the population coverage, all procedures for enumerating the population can then be reconsidered.²⁵¹ Such a fundamental rethinking of the census has the greatest potential for cost reductions.²⁵²

A thorough rethinking of census procedures could consider issues of data collection, data processing, and data dissemination.²⁵³ For example, how many district offices are needed, how long do they need to be open, what staffing and supervisory personnel are needed, and how much of the product line can shift from print to electronic media? Overall, the National Academy of Sciences' Panel on Census Requirements for the Year 2000 and Beyond concluded that the use of statistical estimation and census reengineering could reduce overall census costs by about one-half;²⁵⁴ applied to 2000 census costs of \$7.0 billion, this suggests possible cost saving of about \$3.5 billion through the implementation of a modernized census design.

V. THE MEANING OF POPULATION COMPOSITION

A meaningful discussion about what kinds of census data are needed requires a more detailed discussion about the types of data used for demographic analysis. This section examines the uses of census data for the study of population composition, an important part of demographic analysis. Population composition refers to the distribution within a population of one or more individually defined traits that influence population comparisons or demographic phenomena.²⁵⁵ The individual traits or variables to which population composition refers are usually those believed to have major significance for the uses of population data.²⁵⁶ Thus, while there are a large number of variables of potential interest for study in demography, there are a limited number that are relevant for use in compositional studies.²⁵⁷

^{250.} *Id.* at 96. The 10% to 20% figure is based on an examination of 1990 census costs, where it was estimated that 1990 census costs of \$2.6 billion could be reduced by \$300 to \$400 million through the greater use of statistical methods. *Id.*

^{251.} *Id.* at 100.

^{252.} *Id.* at 101.

^{253.} *Id.* at 107-11.

^{254.} *Id.* at 110.

^{255.} YAUKEY & ANDERTON, *supra* note 1, at 5.

^{256.} *Id.* at 5-6.

^{257.} *Id.* at 6.

To a large degree, compositional elements that are relevant for study are those characteristics of individuals that are enumerable by conventional census-taking techniques.²⁵⁸ In earlier U.S. censuses, enumerable meant that non-professional census personnel could collect the information.²⁵⁹ In recent U.S. censuses, enumerable means that adults in a household are able to provide information on a mail questionnaire.²⁶⁰

The individual characteristics to which composition generally refers include sex, age, marital status, place of birth, education, occupation, labor force status, industry, relation to head of household, and other such features.²⁶¹ This article limits its focus to a general discussion of age, sex, nativity, and race and ethnicity.

A. AGE AND SEX

Age and sex are pivotal characteristics in analyses of composition.²⁶² The distributions of other characteristics are usually contingent in one way or another on age and sex distributions.²⁶³ This results not only from the fundamental social and economic importance of sex and age but also from their independent variability.²⁶⁴ As inherent biological properties of individuals, age and sex do not determine marital status, education, or other demographic characteristics.²⁶⁵ But other demographic processes are often linked to the life cycle and, hence, develop a pattern that is often revealed by the age and sex profiles.²⁶⁶

1. Age

A basic procedure in analyzing age structure is to examine what proportion of the population is at each stage of the life cycle, how these proportions have been changing, how they differ from place to place, and what factors are responsible for the age composition and its changes.²⁶⁷ The median age of the population gives a quick and approximate measure of the age of a population.²⁶⁸ The median age indicates the age that divides the population into two equal parts: 50% of the population is above the

258. *Id.*

259. *Id.* at 28.

260. *Id.*

261. *Id.* at 19-21.

262. *Id.* at 19-20.

263. *Id.*

264. *Id.* at 5-6.

265. *Id.*

266. *Id.*

267. *Id.* at 77-79.

268. *Id.* at 78-79.

median age and 50% is below it.²⁶⁹ From 1820 to 1990, the median age of the U.S. population rose from 16.7 years to 33.0 years.²⁷⁰ Therefore,

during a period of seventeen decades the median age increased by 16.3 years, or about one year per decade. The rise in the age was most rapid between 1900 and 1950, [rising from 22.9 in 1900 to 30.2 in 1950] caused principally by fertility declines. Between 1950 and 1970, the trend of the previous 150 years was interrupted; the median age in 1960 and 1970 was lower than it was in 1950. [This unique reversal is attributed to] the reversal in the long-term downward trend in fertility that took place starting in 1944 to 1946.²⁷¹

In recent decades, the median age has increased from 28.0 in 1970 to 35.3 in 2000.²⁷²

If present low birth rates continue, the median age will increase and the absolute and relative population who are sixty-five years of age and older will expand. Current population projections indicate that the median age of the U.S. population will increase [from 35.3 years in 2000] to thirty-eight years in 2050.²⁷³

Despite the fact that the

median age is a useful measure for indicating generally the age level of a population, it cannot give detailed information about age structure or the distribution of the population among various stages of the life cycle. A more informative picture is to study the proportions of the population for a general set of age groups, such as childhood (age 0 to 8), youth (age 9 to 17), adulthood (age 18 to 64), and the elderly (age 65 and older).²⁷⁴

Using these age categories, Table 1 displays percentage distribution of the U.S. population changes by age group.²⁷⁵ More than a century ago, in 1880, the U.S. population was much younger with a larger proportion of the population in the childhood and youth ages, fewer in the adult years, and dramatically fewer in the elderly years.²⁷⁶ At present, “about six out of ten

269. *Id.* at 79.

270. Barry Edmonston, *Composition of the Population*, in *ENCYCLOPEDIA OF THE U.S. CENSUS 76 & fig.1* (Margo J. Anderson ed., 2000) [hereinafter *Composition of the Population*].

271. *Id.*

272. *Id.*

273. *Id.*

274. *Id.*

275. *Id.* at 76 tbl.1.

276. *Id.*

residents are adults, about one out of four are children or youth, and about one out of eight are elderly.”²⁷⁷

Table 1. Percentage Distribution of the U.S. Population by Major Life Cycle Age Groups, 1880-2050.²⁷⁸

Major Stages of the Life Cycle	Childhood	Youth	Adult	Elderly	Total
Ages (in years)	0 to 8	9 to 17	18 to 64	65+	
Year					
1880	24	19	53	3	100
1900	19	21	56	4	100
1950	16	16	60	8	100
2000	11	14	63	12	100
2050	12	12	56	20	100

In 2050, the U.S. population will likely have slightly fewer relative persons in the childhood and youth years and fewer persons in the adult years.²⁷⁹ The decreases will be counterbalanced by unprecedented growth of persons in the elderly years, increasing to about one in five persons in the population.²⁸⁰

2. Sex

For the first time in history, the United States had a preponderance of women in 1950.²⁸¹ Until 1950 males had outnumbered women.²⁸² The extent of the imbalance is only slight.²⁸³ Instead of being divided fifty-fifty, males were 48.7% of the total population in 1990.²⁸⁴

It is conventional to describe a population's sex composition in terms of the sex ratio, or the number of males per 100 females.²⁸⁵ From 1820 to 1910, the sex ratio of the population of the United States fluctuated between

277. *Id.*

278. *Id.* This table is based on a historical reconstruction of the U.S. population by Passel and Edmonston. *Id.*

279. *Id.*

280. *Id.*

281. *Id.* at 76-77 fig.2.

282. *Id.*

283. *Id.*

284. *Id.*

285. *Id.*

102 and 106.²⁸⁶ From 1910 to 1950, it steadily declined and dropped to 99 in 1950 and to 96 in 2000.²⁸⁷

The native and the foreign-born have had a distinct sex ratio history: The native-born population had a sex ratio near 100 between 1870 and 1950, and was always below the general sex ratio for the nation. In 1950 its sex ratio was [exactly] the same as the national average.

The sex ratio for the foreign-born population shows a preponderance of males until 1960. During the years of heavy immigration to the United States at the turn of the twentieth century, immigrants were predominantly male. Consequently, the sex ratio for the foreign-born population during that time was extraordinarily high. In 1850 for example, the foreign-born population had a sex ratio of 124. In 1910, it was 129. In the period from 1920 to 1960, as the volume of immigration was curtailed, the number of women immigrants came to equal and surpass that of men. By 1960 the sex ratio for the foreign-born had declined to 94.²⁸⁸

Women will continue to outnumber men in future years, but the relative surplus of women will remain relatively unchanged.²⁸⁹ The principal cause of the tendency of women to outnumber men is the greater longevity of women.²⁹⁰ If present demographic trends continue and if the gap in life expectancy narrows between women and men, the sex ratio of the overall population will remain at about 96 between 2000 and 2050.²⁹¹

The sex ratio for the overall population represents a balancing between the sex ratios for the foreign and native-born components of the population.²⁹² Because U.S. immigration includes a surplus of women, the foreign-born population will become more predominantly female as immigration continues and as the foreign-born population ages, since a greater number of older women than men will survive.²⁹³ The sex ratio of the native-born population is expected to shift slightly upward, from 96 in

286. *Id.*

287. *Id.*

288. *Id.*

289. *Id.*

290. *Id.*

291. *Id.*

292. *Id.*

293. *Id.*

2000 to 98 in 2050, as the gap in life expectancy between males and females narrows.²⁹⁴

C. NATIVITY

Nativity has a social significance that reaches into many spheres of study.²⁹⁵ The foreign-born have a distinctive age structure and tend to be selected from several major countries of origin.²⁹⁶ Furthermore, the foreign-born often evidence differences in English-language abilities, fertility, settlement patterns, and other social and economic characteristics.²⁹⁷

Immigrants have arrived in the United States throughout its entire national history, with a great deal of variation from the peaks of the 1900s and the 1980s and 1990s, to the valleys of the 1930s.²⁹⁸ Immigration increased steadily in the decades after World War II because the United States enjoyed a high degree of political freedom and economic prosperity, compared with Europe and many other countries.²⁹⁹ The 1965 changes in U.S. immigration law prompted even further increases, as the United States began to receive large numbers of new immigrants from Asia and Latin America.³⁰⁰

Immigrants affect the composition of the population in several ways.³⁰¹ They change the racial/ethnic makeup of the population if they differ from the resident population.³⁰² Immigrants *always* affect the generational composition, since they increase the size of first generation (i.e., the foreign-born population).³⁰³ The foreign-born population derives principally from past levels of immigration but is also influenced by the effects of emigration and mortality.³⁰⁴

The size of the foreign-born population in the United States reflects the changing course of immigration over time.³⁰⁵ In 1850, approximately 2.2 million foreign-born persons resided in the United States, constituting 9.7% of the total population.³⁰⁶ With a continuing heavy volume of immigration

294. *Id.*

295. *Id.* at 77.

296. *Id.*

297. *Id.*

298. *Id.*

299. *Id.*

300. *Id.*

301. *Id.* at 77-78 fig.3 (showing the number of foreign-born population in the United States from 1850 to 2050).

302. *Id.*

303. *Id.*

304. *Id.* at 78 fig.3.

305. *Id.*

306. *Id.*

in the late nineteenth century, the foreign-born population grew steadily, reaching a peak of 14.4 million in 1930.³⁰⁷ At the same time, the foreign-born population increased as a proportion of the U.S. population, reaching the peak proportion of 14.8% in 1890.³⁰⁸

With the diminution of immigration from about 1918 to 1946, the foreign-born population decreased in both numbers and proportions, as mortality reduced the aging wave of immigrants.³⁰⁹ By 1970, the foreign-born population had decreased to 10.5 million, accounting for only 5.1% of the total population.³¹⁰

The large increase in immigration that began in the 1960s produced a rapid turnaround in the forty-year decrease of the foreign-born population.³¹¹ By 2000, the number of foreign-born persons residing in the United States reached the highest levels in the history of the country—more than 28 million.³¹² Relative to the rest of the population, however, the foreign-born population is less than two-thirds of the highest levels attained from 1860 to 1920: just over 10% of the population was foreign-born in 2000 versus 13% to 15% at the end of the nineteenth century.³¹³

Population projections suggest that the foreign-born population of the United States will increase from 21.3 million in 1990, to 28.4 million in 2000, and 50.4 million in 2050, assuming current trends in fertility, mortality, and international migration.³¹⁴ Under similar assumptions for the growth of the native-born population, these projections forecast an increase of the foreign-born population, as a percentage of the total population, from 8.6% in 1990, to 10.4% in 2000, and 14.2% in 2050.³¹⁵

D. RACE AND ETHNICITY

Race and ethnic characteristics of the U.S. population have long had special importance.³¹⁶ Although “race” is not now regarded as a meaningful biological concept, this fact should not overlook the social and economic correlates of self-reported racial and ethnic identity in U.S. society.³¹⁷ Population statistics provide much of the basis for factual study

307. *Id.*

308. *Id.*

309. *Id.*

310. *Id.*

311. *Id.*

312. LISA LOLLOCK, U.S. CENSUS BUREAU, THE FOREIGN-BORN POPULATION IN THE UNITED STATES, CURRENT POPULATION REPORT NO. P20-534 (Mar. 2000).

313. *Composition of the Population*, *supra* note 270, at 78 fig.3.

314. *Id.*

315. *Id.*

316. *Id.* at 78.

317. *Id.* at 78-79.

of the conditions under which racial and ethnic groups live and the ways in which their relative positions are changing.³¹⁸ For much demographic analysis of the U.S. population, ethnic and racial origin are basic variables that must be included as statistical controls before studying the effect of other variables.³¹⁹

Throughout most of its history, the United States has been essentially a biracial society, comprised of the white majority and a black minority, along with a small American Indian population.³²⁰ In 1790, the United States' population was recorded as 81% white and 19% black.³²¹ This situation is now changing.³²² The United States is becoming a more diverse society.³²³

The first year for which adequate data on the five "racial" groups have been developed is 1900.³²⁴ At the turn of the century the U.S. population was 87% white, 12% black, and about 1% other racial groups.³²⁵ The 1900 census counted about 240,000 American Indians, 240,000 Asians (primarily Chinese, Japanese, and Hawaiians), and 660,000 Hispanics (predominantly Mexican-origin population).³²⁶

All racial groups increased numerically from 1900 to 1960, although the relative shifts were small.³²⁷ By 1960, the white population's share had

318. *Id.* at 79.

319. Historically, the decennial census has included questions on race and ethnicity, although the specific questions asked and the categories for tabulating the answers have changed every ten years. Sharon M. Lee, *Racial Classifications in the U.S. Census: 1890-1990*, 16 *ETHNIC RACIAL STUD.* 75, 83-84 (1993). These changes have occurred because of shifts in the racial and ethnic makeup of the population, changes in social attitudes and political concerns, and the evolving needs of the federal government for data. *Id.* at 91-92. The growing racial and ethnic diversity of the American population, changing attitudes about race and ethnicity, and the increasing use of census data have converged to make census questions on race and ethnicity the focus of attention and controversy. *Id.* Those questions now play a special role in debates over census content, census methods, and public cooperation. Edmonston & Schultze, *supra* note 6, at 85-111.

320. *Composition of the Population, supra* note 270, at 79 (noting that the U.S. census did not distinguish other racial groups until 1850).

321. *Id.*

322. *Id.*

323. *Id.* at 79 fig.4 (presenting information on the percentage composition of the U.S. population, by race and Hispanic origin from 1850 to 2050).

324. Racial categories for classification of federal data changed in the 2000 census. Claudette Bennett, *Race: Questions and Classifications*, in *ENCYCLOPEDIA OF THE U.S. CENSUS* 314-15 (Margo J. Anderson ed., 2000). Prior to the 2000 census, respondents were asked to choose one racial self-identification from a list that included white, Asian and Pacific Islander, black, and American Indian and Alaskan Native. *Id.* at 315 tbl.1 (showing the categories). In the 2000 census and future federal data collection, respondents were asked to identify more than one racial category, if appropriate, and the category for Asian and Pacific Islanders was separated into two categories. *Id.* at 317. For copies of the 1990 and 2000 census questionnaires, see *Questionnaires*, U.S. Census Bureau, available at <http://www.census.gov/dmd/www/2000quest.html> (last modified Jan. 28, 2002).

325. *Composition of the Population, supra* note 270, at 79 fig.4.

326. *Id.*

327. *Id.*

dropped only slightly, from 87% to 85%.³²⁸ The proportion of blacks in the population had decreased slightly, from 12% to 11%.³²⁹ The other three racial/ethnic groups had increased their combined proportion from about 1% to 4%.³³⁰

Since 1960, the shifts in origins of immigrants have led to substantial absolute and relative increases in the Asian and Hispanic populations.³³¹ Over the last forty years, the combined proportion of Asians, American Indians, and Hispanics increased from 4% in 1960 to 18% in 2000.³³² These three groups, which accounted for just over one million people in 1900, had grown to a total of more than fifty-one million by 2000.³³³

As a result of rapid growth of the Asian and Hispanic groups, the composition within the minority population is also changing.³³⁴ The minority population of the United States was almost exclusively black during the nineteenth century, with only a small number of persons of other races.³³⁵ While the black population has increased at a steady rate, the Asian and Hispanic groups have grown at much more rapid rates during the past four decades.³³⁶ Consequently, the proportion of minority population that is black has been dropping steadily, decreasing from 70% in 1960 to 41% in 2000.³³⁷

With a continuation of current levels of immigration, the racial/ethnic composition of the United States will continue to change.³³⁸ Two groups are likely to experience substantial growth during the next fifty years: Asians and Hispanics.³³⁹ The Asian population will grow at rates exceeding 1% for the next fifty years, increasing from 12 million in 2000 to 41 million in 2050.³⁴⁰ The Asian population would become about the same population size as the black population around 2050.³⁴¹

The Hispanic population—assumed to have the largest share of immigration, in part because of its predominant share of illegal immigration—will grow substantially over the next fifty years.³⁴² The Hispanic popu-

328. *Id.*

329. *Id.*

330. *Id.*

331. *Id.*

332. *Id.*

333. *Id.*

334. *Id.*

335. *Id.*

336. *Id.* at 79-80.

337. *Id.* at 80.

338. *Id.* at 79-80 fig.4.

339. *Id.*

340. *Id.*

341. *Id.*

342. *Id.*

lation will increase from 35 million in 2000, or 12% of the total population, to 72 million, or 20% of the population, in 2050.³⁴³ Under these assumptions, the Hispanic population will pass the black population to become the largest minority group in the nation shortly before 2010.³⁴⁴

VI. FUTURE CENSUSES

Future censuses will face continuing challenges from several issues that affected the 2000 census, including pressures to improve population coverage, reduce costs, and collect more timely data.³⁴⁵

A. DRIVING FACTORS FOR CENSUS REDESIGN

Redesign of the U.S. decennial census is inevitable.³⁴⁶ The pressures for change, to reduce costs and improve accuracy, have been persistent and widespread throughout the 1990s.³⁴⁷ Congressional leaders of both political parties agreed on one fact: the 2000 census needed to be different from previous censuses.³⁴⁸ But while there was consensus on the need for change, there was disagreement on the specific changes that should be implemented in the 2000 census.³⁴⁹

Change in the U.S. census is not new.³⁵⁰ The census has evolved considerably over the past 200 years, changing as American society became larger and more diverse and employing new methods as technology improved.³⁵¹ Earlier U.S. censuses were conducted by federal enumerators who canvassed the country on foot or horseback, counting people on tally sheets.³⁵² In the 1880s and 1890s, Herman Hollerith pioneered the use of punch cards and electric tabulating machines—the predecessor of early IBM equipment—for the more efficient processing of the 1890 U.S. census.³⁵³ The 1960 census witnessed the first widespread use of mail-out, mail-back census questionnaires in order to reduce the costs of having an enumerator visit every U.S. household.³⁵⁴

343. *Id.*

344. *Id.*

345. See CENSUS MONITORING BOARD, *supra* note 18, at 9-14 (providing recommendations for future censuses).

346. *Id.* at 13-14.

347. Edmonston & Schultze, *supra* note 6, at 75.

348. SOCIETY, *supra* note 161, at 53.

349. *Id.*

350. *Id.*

351. *Id.*

352. *Id.*

353. GEOFFREY D. AUSTRIAN, HERMAN HOLLERITH: FORGOTTEN GIANT OF INFORMATION PROCESSING 58-73 (1982).

354. SHRYOCK ET AL., *supra* note 14, at 33-34.

Based on the evidence discussed earlier, the 2000 census required modernization for two central reasons.³⁵⁵ First, cost savings are necessary.³⁵⁶ The census has become exceedingly expensive on a per capita basis.³⁵⁷ Using the conventional approach, census costs will continue to grow, probably without any improvement in the accuracy or quality of the census.³⁵⁸ Moreover, the Census Bureau will encounter increasing challenges to recruit a large number of satisfactory enumerators, given the temporary and part-time nature of enumerator employment.³⁵⁹ Even if a census were redesigned to rely on statistical methods and to reduce the dependence upon enumerators, there are serious questions about the Census Bureau's ability to recruit an adequate workforce of enumerators.³⁶⁰

Second, coverage improvements are needed.³⁶¹ Although progress was made in reducing the overall net undercount rate for censuses through 1980, the net undercount rate rose in 1990.³⁶² More importantly, the differential undercount rate between minorities and whites failed to narrow from 1940 to 1990, despite the best efforts of the Census Bureau to reduce the undercount of minorities.³⁶³ By 1990, there was an all-time high undercount of 1.8 million black residents, substantially disenfranchising them of political representation.³⁶⁴ For the national population, a total of 4.7 million persons were not counted—affecting political representation as well as denying federal funds to many poorer rural areas, cities, and towns.³⁶⁵

B. THE 2000 CENSUS SITUATION

In response to widespread criticism of the 1990 census and relying on advice from the Congress and scientific advisory panels, the Census Bureau proposed fundamental reforms for conducting the 2000 census.³⁶⁶ The overreaching goals for the 2000 census, set by the Census Bureau, were to reduce costs, improve accuracy, and to reduce the persistent differential undercounts of the minorities and the poor.³⁶⁷

355. SOCIETY, *supra* note 161, at 53.

356. *Id.*

357. *Id.*

358. *Id.*

359. Edmonston & Schultze, *supra* note 6, at 50.

360. *Id.*

361. SOCIETY, *supra* note 161, at 53.

362. *Id.*

363. *Id.*

364. *Id.*

365. *Id.*

366. ANDERSON, *supra* note 5, at 191-92.

367. *Id.* at 193.

As discussed in previous sections, the Census Bureau has been unable to reduce the continuing gap in differential undercounts with a traditional census approach, even though the 2000 census was the most intensive and costly decennial census ever conducted.³⁶⁸ A redesigned census, however, could simultaneously reduce census costs and improve coverage, while substantially eliminating differential undercoverage.³⁶⁹

There is widespread agreement about several of the new methods that the Census Bureau has used in the 2000 census.³⁷⁰ First, census questionnaires were simpler and clearer, with new ways used to increase the mail response rate.³⁷¹ Even modest increases in response rates saved the Census Bureau millions of dollars in follow-up costs.³⁷² Second, new partnerships with local officials improved the accuracy and completeness of the count.³⁷³ Local officials checked the accuracy of the census mailing address lists and thereby reduced an important source of census undercoverage.³⁷⁴ Third, new technology was used to weed out duplicate census questionnaires.³⁷⁵ This technology allowed the Census Bureau to supplement mailed questionnaires by placing them in public places—such as post offices—for mail return.³⁷⁶ This helped to reduce the undercount for traditionally difficult to enumerate groups.³⁷⁷ Finally, the Census Bureau spent \$100 million on paid advertising to help convince the public to participate in the 2000 census.³⁷⁸ The Census Bureau relied totally on donated advertising in previous censuses.³⁷⁹

But there are strong political concerns about the use of statistical methods to complete the count.³⁸⁰ Although public criticism of statistical methods that could be used in the census often treated all methods as the same, proposed statistical sampling would be used in three different stages in a modernized census.³⁸¹ First, a longer census questionnaire would be mailed to a sample of all housing units—in the same way that was done in

368. SOCIETY, *supra* note 161, at 53.

369. *Id.*

370. *Id.* at 42.

371. *Content Determination*, *supra* note 102, at 93-94.

372. CENSUS MONITORING BOARD, *supra* note 18, at 5.

373. *Id.*

374. *Id.* at 11-12.

375. John H. Thompson, *Organization and Administration of the Census*, in *ENCYCLOPEDIA OF THE U.S. CENSUS 299* (Margo J. Anderson ed., 2000).

376. *Id.*

377. *Id.*

378. Margo Anderson, *Advertising and the Census*, in *ENCYCLOPEDIA OF THE U.S. CENSUS 13* (Margo J. Anderson ed., 2000).

379. *Id.*

380. SOCIETY, *supra* note 161, at 53.

381. *Id.* at 50-52.

each previous census since 1960.³⁸² There does not seem to be debate about the statistical aspects of a sample census questionnaire in the census.³⁸³

Second, among those households who do not return a census questionnaire, after several attempts to contact them by mail, an enumerator would then contact a sample of non-responding households.³⁸⁴ The Census Bureau knows about the non-responding households; the purpose of the sample is to estimate the characteristics of people living in these households.³⁸⁵ The use of a sample for non-response follow-up will save a great deal of money.³⁸⁶

Third, in order to complete the count, a large separate sample of the U.S. population would be selected to provide an independent estimate of census undercount.³⁸⁷ The only conceivable method for estimating census undercoverage at the sub-national level, where demographic analysis is not reliable, is this quality control sample.³⁸⁸

These last two stages for statistical methods in a modernized census are the center of public debate.³⁸⁹ The debate about statistical methods for a modernized census does not seem to relate to the methods themselves, about which there is strong agreement among statisticians and demographers who have presented their findings in reports of the National Academy of Sciences.³⁹⁰ Rather, the debate has focused on the potential for manipulation of statistical methods so that the census counts might be favorable to one political party or the other.³⁹¹ While the 2000 and previous censuses counted some people without physical enumeration, the current debate has not centered on the required evidence for physical enumeration.³⁹² The present debate centers on persistent qualms about the expanded role that statistical methods would play in future, possibly modernized censuses.³⁹³ Although the Census Bureau has longstanding credibility for

382. *Id.* at 50.

383. *Id.* at 49.

384. *Id.* at 50.

385. *Id.*

386. *Id.*

387. *Id.*

388. *Id.*

389. *Id.* at 50-51.

390. CENSUS MONITORING BOARD, *supra* note 18, at 1.

391. *See id.* at 9 (recommending the removal of partisan politics from the census process).

392. Passel, *supra* note 148, at 6. For example, people who did not respond to mail questionnaires were sometimes counted if neighbors or letter carriers could provide sufficient information about a household and its members.

393. *Id.* at 8.

the confidentiality and quality of its data, there are demands for additional guarantees for data integrity for the new uses of statistical methods.³⁹⁴

One process implemented in the 2000 census was designed to guarantee census data integrity: a monitoring group of eight outside experts was convened to advise Congress.³⁹⁵ This group monitored census operations to insure that there was no manipulation of statistical estimates, while maintaining the required independent role of the Census Bureau.³⁹⁶

C. IMPROVING CENSUS DATA ON POPULATION COMPOSITION

As discussed in the previous section, census data on key variables of population composition—age, sex, nativity, and race and ethnicity—have been subject to undercoverage in previous censuses.³⁹⁷ In some cases, the estimates of undercoverage are large.³⁹⁸ An intensive examination of 2000 census undercoverage in seven rural California communities, with a large number of Mexican immigrants, reports overall net undercount levels of 23%.³⁹⁹

If all groups were undercounted in the census, there would be discussion of overall net undercoverage, but there would be little debate about issues of equity.⁴⁰⁰ Because some groups, however, are undercounted more than others, there are serious concerns about the effects of differential coverage error.⁴⁰¹

The ultimate goal of the decennial census should be to produce accurate final census numbers.⁴⁰² Because there is no evidence that inaccuracies in the final numbers can be reduced substantially through changes in the direct count procedures, recent debate has focused on the extent to which statistical methods can be used to supplement the direct census enumeration.⁴⁰³ Adjustment, if used, would improve the accuracy of the characteristics of the population as well as the final total population numbers.⁴⁰⁴ The quality of the data for population composition would be

394. CENSUS MONITORING BOARD, *supra* note 18, at 14.

395. *Id.* at 2.

396. *Id.*

397. *See* Cohen, *supra* note 150, at 100-101 (discussing census undercoverage).

398. *Id.* at 101.

399. ILENE J. JACOBS & EDWARD KISSAM, THE CALIFORNIA ENDOWMENT COMMUNITIES FIRST PROGRAM, CENSUS 2000 UNDERCOUNT OF IMMIGRANTS AND FARMWORKERS IN RURAL CALIFORNIA COMMUNITIES 7 tbl.2 (Aug. 1, 2001).

400. SOCIETY, *supra* note 161, at 45.

401. *Id.* at 50.

402. *A One-Number Census*, *supra* note 247, at 491.

403. *Id.*

404. SOCIETY, *supra* note 161, at 50.

improved, along with population numbers for the nation and other large geographic areas.⁴⁰⁵

VII. CONCLUSION

After four censuses, 1970 to 2000, in which the Census Bureau has spent large amounts of money on traditional approaches, we know that only fundamental reform can improve census coverage.⁴⁰⁶ While a redesigned census, as described earlier, is not a magic wand for census ills, a redesigned census that incorporates modern statistical methods would improve the nation's census by reducing costs and improving census coverage for the nation and major ethnic groups.⁴⁰⁷ As demonstrated by the analysis of several key variables on population composition, there is a clear need for accurate census data.⁴⁰⁸ A modernized census that used new statistical methods to deal with undercoverage would improve census information on age, sex, racial or ethnic origin, and nativity—fundamental variables for the study of population composition.⁴⁰⁹

405. *Id.* at 51.

406. SOCIETY, *supra* note 161, at 53.

407. *Id.*

408. *See generally* *Composition of the Population*, *supra* note 270, at 75-80.

409. *Id.*
