Working Paper No. 8
25 April 2003

ENGLISH ONLY

UN STATISTICAL COMMISSION and
UN ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES (EUROSTAT)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE-EUROSTAT Work Session on Population and Housing Censuses
(Ohrid, The former Yugoslav Republic of Macedonia, 21-23 May 2003)

Session II – Invited paper

THE ROLE OF THE AMERICAN COMMUNITY SURVEY IN RE-ENGINEERING THE UNITED STATES CENSUS OF POPULATION AND HOUSING

Submitted by the U.S. Census Bureau[1]

I. This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. The paper reflects the views of the author, not necessarily the Census Bureau.

I. Introduction

1. The U.S. Census Bureau's mission is to provide timely, relevant, and high-quality data about the people and economy of the United States. The decennial census, which has included the enumeration of the nation's population and housing as well as the collection of detailed demographic and socioeconomic information, is the foundation of the federal statistical system. Approximately 5-in-6 of the nation's housing units report only basic demographic and housing data (about 7 questions) on a "short form". The remaining 1-in-6 of the nation report the same basic demographic and housing data in addition to more detailed housing, social, and economic information (about 70 additional questions) on a "long form". The country is apportioned, states are redistricted, policies are developed, plans are made, voting and civil rights laws are supported, and billions of dollars of federal funds are allocated based on the results of the decennial census. Further, sample surveys are calibrated based on the decennial census. The quality of the decennial census directly affects the quality of subsequent surveys and censuses throughout the federal statistical system.

2. Three major areas benefit when the decennial census long form is replaced with an ongoing survey - the decennial census itself, the Intercensal Population Estimates Program which produces count estimates for a variety of purposes, and for a wide variety of data users. By removing the highly resource intensive and detailed long form data collection from the 2010 census design, the

[1] Paper prepared by Deborah H. Griffin.

design can be simplified, thereby reducing risk and potentially improving coverage of the country's population. An ongoing survey that can provide key information for every county and many sub-county levels can help improve estimates of change since the last census for the Intercensal Population Estimates Program. Perhaps of greatest consequence, in response to the rapid growth of a large and diverse population and stakeholders' needs, an ongoing survey can result in more timely data for all users.

I.1 Decennial census planners continue to confront the challenge of change

3. Over the decades, the Census Bureau has encountered and subsequently overcome numerous policy, operational, technical, and fiscal challenges, thus ensuring the continuing integrity of the federal statistical system. Census 2000 challenges were many. While evaluations of data quality are still underway, Census 2000 is widely considered the most accurate census to date in terms of overall coverage. Evaluations indicate that the coverage error in Census 2000 was very low (U.S. Census Bureau, 2003). However, this success cannot be repeated without incurring unacceptable risk and cost (U.S. Census Bureau, 2001).

4. The Census Bureau remains committed to learning from prior censuses and striving for ongoing operational improvement of survey methods. Just as in other decades, demographic and technological change has prompted the Census Bureau to make bold decisions and take innovative approaches. While planning for the 1940 census, the Census Bureau determined that it was neither cost-effective nor necessary to ask all questions of every person. Planners developed the census sample as a major enhancement to the 1940 census and that design has been used through 2000. In planning for the 1970 census, the Census Bureau determined that it could no longer effectively enumerate the population through personal visits alone. The Census Bureau reengineered the design by moving to a self-response mailing strategy as the primary data collection mode in 1970–a strategy still in place and designed to improve data accuracy and reduce the number of costly personal visits.

5. Successful enumeration in 2010 will require just such innovations. Challenges facing planners of the 2010 census are formidable. Demographic and technological changes are occurring too rapidly not to consider bold innovation. For example, enumeration complexity is increasing due to a growing, diverse population and increasingly diverse living arrangements; data collection is further complicated due to escalating privacy and confidentiality concerns; and the Census Bureau must plan for a dynamic technological and political environment with limited resources (Killion, 2001).

6. Consequently, managers are reengineering the 2010 decennial census via a strategy that harnesses the synergy among three Census Bureau programs–modernization of the Master Address File or MAF/TIGER , the American Community Survey implementation, and the streamlined 2010 census design. While a full discussion of the MAF/TIGER modernization is beyond the scope of this report, it is important to recognize that a complete and accurate address system is crucial to the success of either of the other two strategic programs. The MAF/TIGER provides the so–called frame or address pool on which the decennial census and all demographic surveys–including the ACS, are based. The accuracy of the MAF/TIGER is widely considered to be the single best forecaster of a successful, complete decennial census and ACS implementation. This is especially true in terms of coverage of housing and population.

1.2 The American Community Survey–the response to stakeholder needs

7. Over 10 years ago, in response to congressional and other stakeholder demands for more timely and relevant data, the Census Bureau began examining a new approach for gathering long form data. Over time, rapid demographic change has outpaced the usefulness of the decennial long form. Consequently, in lieu of the static, once-in-a-decade snapshot of the nation's population, Census Bureau experts began researching the feasibility of an ongoing survey to collect and

disseminate timely demographic and socioeconomic data. This research culminated in the Continuous Measurement program, now referred to as the American Community Survey (ACS).

8. The Census Bureau began developing the ACS in the mid 1990s and has been collecting ACS data since 1996. Data collection began in four test sites and has since expanded to additional test sites and to a national sample. A national sample of about 800,000 addresses called the Census 2000 Supplementary Survey (C2SS) was conducted as part of Census 2000 using ACS methods. The purpose of the C2SS was to demonstrate the operational feasibility of collecting long form data at the same time as, but separate from, the decennial census. Analysis of the C2SS has led to the conclusion that the ACS is a viable replacement for the census long form (U.S. Census Bureau, 2001 and U.S. Census Bureau, 2002).

9. When fully implemented, the ACS will be in every county in the United States, as well as in Puerto Rico, and will survey an annual sample of three million housing units. The ACS will produce information on content items similar to the decennial census long form for communities across the country, including small areas such as census tracts, small towns, American Indian Reservations, Native Alaskan villages, and rural areas. The ACS will provide yearly estimates of demographic, housing, social, and economic characteristics for all states, as well as for all cities, counties, metropolitan areas, and population groups of 65,000 people or more. For smaller areas, such as census tracts, three to five years of data will be necessary to accumulate sufficient sample to produce estimates. Areas with between 20,000 and 65,000 people can use data averaged over three years and areas of less than 20,000 people (such as census tracts, rural areas, small towns, and some American Indian reservations) will require five years to accumulate a sample size that is similar in size to that of the decennial census long form. These multi-year averages will be updated every year to give data users important measures of change over time.

II. BENEFITS OF THE AMERICAN COMMUNITY SURVEY

II.1 For the Decennial Census

10. The American Community Survey is crucial to a successful 2010 Census design. Collecting long form data throughout the decade via the ACS will have a profound effect on both 2010 Census design plan and outcome. The collection of sample data adds substantial burden and complexity to the decennial census. Implementing the ACS means that staff responsible for planning the decennial census can more sharply focus on its constitutional mandate–to accurately count the population and housing to apportion the House of Representatives. The ACS–supported by a complete and accurate address system–will simplify the 2010 census design, resulting in improvements in both coverage and data quality.

II.1.1 The American Community Survey will improve planning and simplify the 2010 decennial census design

11. Planning for a short-form-only census in 2010 allows innovation and streamlining in ways and on a scale not possible when long form data collection is included. Planners can focus on the basics–determining the best methods to count the nation's large, growing, and diverse population and housing. An implemented ACS directly and indirectly supports 2010 planning and improves the design. Currently the Census Bureau maintains a Planning Database that provides important information to aid in identifying areas warranting outreach or special enumeration methods. For example, data can be provided indicating where languages other than English are spoken throughout the country. Specific areas can be flagged as potentially "hard-to-enumerate" in advance of the census based on demographic and socioeconomic characteristics and previous census response rates. Data from the ACS can refresh the data in the Planning Database making it even more valuable. This knowledge enables planners to better focus decennial census data collection and outreach research and design efforts.

12. Because ACS staff will maintain a continuous presence in the local areas, rapport and relationships will have already been established, thus facilitating the 2010 census enumeration activities. For example, ACS staff and community partners will have adequate time to identify and resolve issues. They will be able to refine data collection or outreach processes and procedures for use in the decennial census. In addition, an ongoing community presence will continue to foster relationships and partnerships among government and community officials.

13. Collecting only short form information allows the Census Bureau to simplify and improve upon most decennial census operations, including questionnaire design and printing, data collection, and data capture. The absence of the long form will substantially reduce the amount of paper required and will facilitate expanded electronic reporting. For example, the long form accounted for about half of all the paper in Census 2000 and its size (about 45 pages) increased production difficulties. Not having a long form greatly simplifies printing, assembly, and postal service distribution activities. Additionally, focusing on enumeration enables the Census Bureau to take full advantage of state-of-the-art technology to streamline data collection and processing activities. For example, enumerators may collect short form data using a handheld computer that enables fast and accurate data collection with less paper. Less paper translates into simpler, faster data processing, substantially reducing workloads. Also of critical importance is the impact that the long form has on mail response. In Census 2000 the long form was returned at a much lower rate in all parts of the country (Stackhouse, 2002). This implies a greater follow-up workload and increased costs. A census with only short forms could be more cost efficient if more interviews could be collected by mail.

14. Finally, full implementation of the ACS will allow the Census Bureau to try out new technologies and adapt them to make data collection and processing more efficient and cost effective. For example, the census is exploring Internet data collection and collection of data through voice recognition on the telephone. The results will provide early insight into issues associated with the expanded use of the Internet and other electronic data collection modes, identifying areas for additional research and testing.

II.1.2 Implementing the American Community Survey–supported by the MAF/TIGER modernization–will potentially improve decennial census coverage

15. In addition to aiding decennial census planners, the ongoing collection of long form data via the implementation of the ACS, supported by an accurate MAF/TIGER system, potentially will lead to improved coverage in the 2010 census enumeration. The ACS presence in counties across the U.S. prior to 2010 will provide current and accurate information about where to target traditionally hard-to-enumerate groups and areas. For example, should a new community of non-English speaking people be identified, regional staff can better respond by ensuring appropriate outreach and other special data collection methods. This is in stark contrast to Census 2000 planning. Planners were dependent on data from the 1990 long form to support the Planning Database. While the Planning Database demonstrated its potential, the data were often too old to reliably target neighborhoods. Regional managers and staff had to depend on community specialists to identify such areas–a much less systematic approach (U.S. Census Bureau, 1998). With ongoing ACS data collection, current and detailed demographic and socioeconomic profiles of geographic concentrations will be available to support census enumeration, as well as to continuously update the Planning Database.

16. The MAF/TIGER modernization provides the foundation upon which the Census Bureau's censuses and surveys will depend. MAF/TIGER is a mission-critical program. The MAF/TIGER modernization enhances and improves the current system by incorporating global positioning system technology to ensure positional accuracy and by moving from a Census Bureau-developed database to a commercial-off-the-shelf system.

17. To take advantage of a fully implemented ACS, the Census Bureau requires a complete and accurate MAF/TIGER. On the other hand, a complete and accurate MAF/TIGER depends on the implementation of the ACS. The ongoing presence of ACS staff enables the timely update of the address list, including the identification of new construction and newly converted group quarters. An accurate MAF/TIGER helps ensure that the data collected by the ACS correctly reflect the population and housing characteristics for small geographic areas and groups. In 1990 and in 2000, enumeration errors were attributed to errors in the census address file that affected both coverage and data quality. These errors included housing units not being included, included more than once, or being placed in the wrong geographical location (known as geocoding error). For example, the Census Bureau estimates that about one-third of the persons missed in the1990 census were due to missing entire housing units (Hogan, 1993). In addition to missing housing units, overcounts occur when housing units are listed more than once. In Census 2000, decennial managers had to design a complex, unplanned operation to try to identify and delete housing units that had been improperly duplicated in the MAF. By providing continuous updates to the address list throughout the decade, MAF accuracy will be improved for 2010.

18. Together, the ACS implementation and the MAF/TIGER modernization reduce the complexity of the 2010 census design. These two strategic programs allow flexibility for incorporating new and innovative data collection and processing methods. Such methods include a targeted second mailing or mailing forms in two languages to selected small areas. Expending resources throughout the decade in an informed, focused, and proactive manner will optimize the 2010 census planning environment and lay the foundation for future censuses and surveys. The ACS implementation, supported by a modern and accurate MAF/TIGER, will provide better ground truth–not only a crucial factor in a successful 2010 census but also an invaluable contribution to governments and communities.

II.2 For the Intercensal Population Estimates Program

19. The ACS estimates are weighted to population and housing benchmarks - either the most recent decennial census results or the most recent intercensal estimates. The Intercensal Population Estimates Program develops and disseminates annual estimates of the total population and the distribution by age, sex, race, and Hispanic origin for the Nation, states, counties, and functioning governmental units. The accuracy of the intercensal estimates is key to overall ACS and survey accuracy. Improvements in the population estimates will also lead to improvements in demographic analysis. Demographic analysis is a critical tool used to evaluate coverage in the census.

20. The Census Bureau has developed the Program for Integrated Estimates (PIE) to research and introduce enhancements to the intercensal estimates. The PIE program will integrate information from Census 2000, the ACS, and administrative records to produce improved population and housing estimates for all areas, including small areas. A fully implemented ACS will provide annual distributions for every county and many sub-county levels that will facilitate improved estimates of change. Complete information of this type is not currently available for the entire population at sub-county levels. ACS distributions can be used along with other data to improve estimates of the components of annual change.

21. A comprehensive research and production program has begun to use ACS in the production of more accurate and reliable population estimates for the Nation, states, counties, and all governmental units.

II.3 For Federal Agencies and Other Data Users

22. The American Community Survey can better reflect a changing America. The value of decennial census long form data to inform federal decision-making and to the federal statistical system cannot be overstated. It forms the basis for the annual allocation of billions of dollars in federal funds and can guide planners and policymakers at all levels of government including the

smallest of communities. Decennial census long form data, however, has been subjected to the constraints inherent in a single, massive data collection effort. The ACS, on the other hand, has been specifically designed to ameliorate those constraints, by collecting long form data throughout the decade. The ongoing presence of ACS data enables informed planning throughout America. It is the most effective and efficient tool available to measure the complexities of the shifting American landscape in a timely manner. In this information-based economy, federal, state and local decision makers, as well as private business and nonprofit organizations, need current, reliable, and comparable data to chart the future.

II.3.1 The American Community Survey has been designed to collect detailed demographic and housing information

23. The ACS design and implementation stands in stark contrast to the traditional decennial long form survey. The ACS design is premised on the ongoing operational improvement of survey methods to ensure data quality while the decennial long form, once designed and implemented, cannot be readily adjusted. For example, because of schedule and budget concerns, the long form in Census 2000 did not include a content follow-up, thereby reducing data completeness. In contrast, following up with respondents to obtain answers to questions is an integral part of the ACS methodology.

24. Collecting data from nonresponse households is a formidable task. Because the long form is part of the decennial census enumeration, Census Bureau managers are dependent on the hiring and training of hundreds of thousands of temporary employees. Due to schedule constraints, these employees undergo limited training. In contrast, ACS staff are permanent professional interviewers. These professional interviewers are typically experienced and have undergone extensive training that includes methods to collect information from reluctant respondents. This better level of preparedness is expected to lead to both higher response rates and higher interview quality. When mail response rates are low, nonresponse follow-up workloads are high and it is difficult to complete census enumeration activities on schedule. In especially hard-to-count areas, collection of sample data can detract from the goals of an accurate count.

25. Another major contrast between the two approaches is the data collection methodology. The Census 2000 long form data are collected as part of the one-time census enumeration activity. The 20 million long forms distributed in Census 2000 were mailed along with the 100 million short forms in mid-March, 2000. Follow-up for housing units not returning forms by mail began in late April, 2000 and lasted about nine weeks. While the snapshot of America was taken as of April 1, 2000, readying the long form data for release to data users occurred from about October 2000 until the summer of 2002. Until the actual release of the Census 2000 long form data in 2002-2003, users had to rely on 12-13 year old data from the 1990 Census. The ACS will drastically shorten the "wait" for data.

26. In contrast, the ACS data are collected each month. Three-month cycles of data collection are employed for each monthly sample–a combination of mail out/mail back questionnaires, and telephone and personal visit nonresponse follow-ups. Each month a sample of housing unit addresses is drawn from the MAF, and questionnaires are mailed at the beginning of month one. A second (targeted) questionnaire is mailed about 3 weeks later to the addresses that did not respond by mail. During month two, housing units that did not respond by mail and have telephone numbers are called. Finally, in month three, a subsample of remaining nonrespondents is drawn, and a personal visit by professional interviewers is made to those housing units. The quality of data is further improved by using computers to collect the data. The use of the computerized questionnaire, rather than a paper questionnaire, allows the Census Bureau to incorporate consistency checks of the data into the collection process.

27. The ACS divides a huge nationwide workload into manageable pieces over a longer time frame. ACS staff are able to gather the information and conduct the required analyses and quality

checks in a controlled manner. Consequently, while data release and distribution for the decennial long form take two years, ACS data products can be released in about six months. Such efficiency can be achieved because the program is not burdened by the production of the products required for apportionment.

II.3.2 The American Community Survey development program provides current, timely information essential for governing

28. Decennial census long form data have played an indispensable role in governing the nation for decades. Every question on the form has a specific federal legislative or judicial mandate or requirement. When the Census Bureau reported to Congress on March 30, 1998, on the specific Census 2000 questions, the Census Bureau provided selected citations from about 130 laws that use specific long form items. As one example, there are 25 citations alone for the long form disability questions. Without the decennial long form, the Congress could lack the information needed to effectively manage its programs. However, as stated earlier, increasingly rapid demographic change has rendered the social, economic, and housing profiles from the decennial long form less useful for meeting congressional and other stakeholder needs. This is why federal agencies and communities have urged the Census Bureau to develop the ACS.

29. Already, the ACS development program is beginning to assist some federal managers administer their programs. For example, the 107th Congress has mandated the use of ACS data for determining the number of children and youth with limited English proficiency for use in computing formula grants to states. Similarly, the ACS will provide community data on children with disabilities and specific functional limitations for use in the distribution of more than 4 billion dollars in grants to states for programs for children with disabilities.

30. When fully implemented, the ACS will enable informed decision making across the full spectrum of federal programs. The ACS will collect information that is comparable within and across states for program evaluations and use in funding formulas. Because ACS information will be provided every year, trends over time and results of some federal programs may be measured. Helping to define the needs and the resources of the people living in the U.S., a fully implemented ACS will provide the tools required for the business of governing in an increasingly complex and dynamic society at the federal, state, and local community level.

II.3.3 Communities continue to benefit from the American Community Survey development program

31. To date, empirical evidence of the usefulness of ACS data is primarily found at the community level and the positive response of participating communities has been strong. Indeed, the usefulness of the information at the local level is indicative of the value that the fully implemented ACS will bring to the federal and state governments. For example, the fully implemented ACS will continue to help communities establish goals using objective information, increase program accountability by measuring results, and measure program performance for future planning. While the Census Bureau and community researchers are continuing to evaluate and understand the program's results, the preliminary assessment is that the information has surpassed expectations as to its quality and usefulness.

III. ISSUES

32. While the benefits are extensive, the transition to the ACS from the long form includes recognition of some concerns. Data users, comfortable with receiving data for all geographic areas at one point in time, must adapt to the concept of moving averages. These are the 3-year and 5-year accumulations of data required for areas with fewer than 65,000 people. ACS data for some items are expected to differ slightly from data that would be collected in a census. This is due to the different rules and reference periods needed in a survey versus used in a decennial census. The

Census Bureau will assist data users, most critically Federal agencies, to understand how ACS will differ from the data they have traditionally used. Nonetheless, transition "pains" are anticipated.

33. Another area of concern is cost. Due to its size and design, the ACS will be an expensive undertaking. Research and evaluation efforts will continue to search for cost efficiencies but a long term fiscal commitment to this program is critical to its success.

IV. CONCLUSIONS

34. Implementing the ACS is integral to the Census Bureau's mission to produce timely, relevant, and high-quality demographic, housing, and socioeconomic data. The Census Bureau has a constitutional and statutory mandate to enumerate the population and housing of the U.S. as well as to collect data needed by policymakers at multiple levels of government. Given rapid demographic change and the ever-increasing demand for timely and relevant information, the Census Bureau must move away from the massive and nearly overwhelming effort to collect the detailed data from one in six households every ten years. The rational for designing, developing, and implementing the ACS is to provide this critical data throughout the decade, facilitate an accurate decennial census in 2010, and improve the Intercensal Population Estimates Program.

BIBLIOGRAPHY

Hogan, Howard, "The 1990 Post-Enumeration Survey: Operations and Results," Journal of the American Statistical Association, September 1993, p. 1056.

Killion, Ruth Ann, "2010 Census Planning: Demographic and Technological Change Escalating Complexity," July 2001.

Stackhouse, Herbert F. and Treat, James B., "Census 2000 Response and Return Rates–National and State by Form Type," Internal Census Bureau document. DSSD Census 2000 Procedures and Operations Memorandum Series L-10, February 12, 2002.

U.S. Census Bureau, "The Census 2000 Language Program: Lowering Barriers to Census Participation–A Business Case Analysis," July 31, 1998.

U.S. Census Bureau, "Potential Life Cycle Savings for the 2010 Census," April 30, 2001.

U.S. Census Bureau, "Meeting 21st Century Demographic Needs - Implementing the American Community Survey: Demonstrating Operational Feasibility," July 2001.

U.S. Census Bureau, "Meeting 21st Century Demographic Needs - Implementing the American Community Survey: Demonstrating Survey Quality," May 2002.

U.S. Census Bureau, "Technical Assessment of A.C.E. Revision II," March 12, 2003.