FROM SURVEY DATA TO MULTIPLE TYPES OF DATA IN HISTORICAL AND REAL TIME[1]

Juanita Tamayo Lott, Census Bureau; Fritz Scheuren, NORC, University of Chicago;
Jay K. Keller, Census Bureau; David Banks, Duke University
Juanita.t.lott@census.gov

Key Words: federal requirements, data user expectations, federal statistical system, real-time data.

"Not everything that can be counted, counts. Not everything that counts can be counted." Albert Einstein

Background

By their nature, federal statistical agencies must be conservative and cautious. Their results, after all, are the bedrock upon which much of the rest of our information society is built**.** Yet, real-time, real-life decisions increasingly require the ability to quickly synthesize, integrate, and act on both historical and real-time data. A tension exists between being responsive and being reliable. This is not a new tension, although it is manifested today in new ways.

One of the ways this tension is manifested at federal statistical agencies is increasingly via work on the President's Management Agenda—notably federal information management decisions and actions with respect to crosscutting concerns like war, terrorism, health epidemics, and natural emergencies. Contrast how the private sector is dealing with these "new age" issues. By its nature, the private sector has been able to be more receptive and adaptive to the use and linking or integration of all sorts of data and information than has the public sector. We would argue that this must change, with the balance point being moved so that federal data systems become more flexible and more responsive than at present. How this will be accomplished is an open issue. Whatever happens, for federal data stewards, prudence and independence will be increasingly required in balancing governmental requirements for data with non-governmental expectations for data.

Given this backdrop, the present paper addresses one fundamental topic in addressing real-life, real-time, quickly made decisions -- the evolution of survey data requirements and methods for federal statistical systems in relation to the heightened expectations of various data users. Our focus is limited to delineating federal data requirements over time in relation to various types of data that are now routinely expected since survey data were heavily introduced in the 1940s. We specifically examine requirements for the Decennial Census of Population and Housing in relation to data user expectations. We then provide examples of statistical innovations that respond to the heightened expectations of data users amidst the new realities for governance. We close with some thoughts and characteristics that might be part of the effort needed to define a gold standard for social survey measurement in the 21st century. Specifically in the data rich world of today and tomorrow, we explore the role of a survey as a "Rosetta Stone" for other datasets, as well as its continuing role in obtaining data not otherwise available at all or in a timely manner.

The Purpose of Statistics

The word "statistics" is derived from the Greek word *statistiki* (**στατιστική)** that refers, following Aristotle (350 B.C.), to the study of political facts and figures. In *Social Statistics in Use*, Phillip Hauser (1975) noted that statistics were initially developed to serve the state and only widened later to serve other stakeholders in society. That is, statistics were used, initially, to help governments rule by understanding their population, economy, and environment. In our democracy we would say, paraphrasing Lincoln, so that 'the people can better rule themselves.'

Universal Principles

Over time the expansion of the basic uses of statistics grew so that by the 20th century the term "official statistics" came into use. In the 19th century, incidentally, such a coinage would have been seen as redundant. Keeping with this evolution in terminology, the United Nations (UN) established a set of 10 Fundamental Principles for Official Statistics (United Nations Statistics Division 2004). The first of these states:

> Official statistics provide an indispensable element in the information system of a

> democratic society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information.

The second principle deepens the meaning here, focusing attention on the need for scientifically based technologies and the sound interpretation of the results obtained. Specifically the second fundamental principle states:

> To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.

While all of the other UN principles have relevance too, it is enough for our current purposes to quote just one more, the fifth principle, which states:

> Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records.

The current paper focuses mainly on this last principle and how to integrate these diverse sources in an ever-changing world, despite the appropriate caution that official statisticians must have. One part of this ever-changing world is the extent to which other information providers are now looked to instead of---or in addition to---governmental agencies. In fact in our information age it has become obvious, for example, that the market share that the United States Census Bureau has as an information provider is shrinking. The key question may not be reversing this trend but preventing such a trend from becoming inefficient and counterproductive. What then is the role of a government information provider today? Where is our comparative advantage? One way to address this question is to look to our past. We will highlight the 20th century role of sampling as a source for new information and as a way to integrate what already existed.

Original Application to Sampling

Although well underway earlier (Duncan and Sheldon 1978), the development and use of sampling methods reached breakthrough speed and force for the federal statistical system in the 1940s (Scheuren 2003, Nathan 2001, Hansen 1987, Morgenstein and Marker 2000). This breakthrough was so successful that statisticians and other data analysts ever since have viewed survey research as the gold standard for 20th century measurement for public policy analysis purposes. Until the 1940's the dominant data sources that the federal government relied on were administrative records, like those from the Internal Revenue Service, the decennial censuses of population and housing, and other mainly 100% data sources.1

Initially these new survey datasets were not well integrated with the earlier sources, although this seems to always have been a goal. Because of their responsiveness, however, surveys, particularly since the 1970s became primary in many areas of public policy, with the result that earlier, slower sources became relegated to secondary roles. The need to integrate survey data---now primary---with what had become secondary, was on the minds of those early pioneers and led them to conduct matching studies to reconcile the sources and to find ways to use them jointly. Attempts were made, for example, to do this using both exact and statistical or synthetic matching to administrative records (*e.g.*, Jabine and Scheuren 1986, Herriott and Scheuren 1980, Steinberg 1972, Budd and Radner 1978) These matching methods were motivated by the desire to improve

[1] The Internal Revenue Service in 1916, almost from the inception of the income tax, began sampling its administrative records for information purposes as part of its Statistics of Income program. These were not strictly surveys, however. Regular surveys for administrative purposes did not start at the IRS until the 1950's with the inception of what was to become known as the Taxpayer Compliance Measurement Program. The Department of Agriculture before the 20th century was another agency that made many sampling applications in its efforts to predict crop yields. The Social Security Administration from its inception employed embedded samples inside its record systems. The most famous of these--and one that continues today--is the 1 percent Continuous Work History Sample (CWHS).

data quality of the surveys and to extend the data's analytical reach.

Such a broad movement cannot be encompassed in a single paper and so we have limited our focus to a discussion of data collected or obtained by the Census Bureau in its Decennial Censuses of Population and Housing, including the sample versions which, incidentally, also began in 1940.

Decennial Census of Population and Housing since 1940's

The 1940 Census is noted for its use of survey methods to obtain characteristics of the total population based on a sample. Prior to the 1940 Census, all census items were asked for each household. These items were used not just for apportionment and other political representation, but also for various requirements of governance. Due to concerns with cost, accuracy and timeliness of data collection and production, federal statisticians and analysts, particularly at the Census Bureau, demonstrated the utility of survey methods. At this point, United States demography could be viewed as homogeneous, mainly white, with married couple households comprising the largest proportion, and thus appropriate for survey methodology. The data requirements, overall, for the 1940 Census were consistent with prior censuses. The difference was that only a portion of United States households was surveyed for characteristics related to administering federal programs or enforcing federal laws and regulations. Data users were predominantly federal agencies, other governmental agencies, and university researchers.

The data requirements--- political representation, administration of federal programs, and enforcement of federal laws and regulations--- remained the same through the 1980 Census. In the 1980 Census a major methodological breakthrough---mailout/mailback enumeration---was fully implemented. Instead of a federal enumerator making house calls, census forms were mailed to households and one member completed the form for everyone in the household. The 1980 Census was the second mailout/mailback enumeration census, following the 1970 Census. However, the 1980 Census is more useful to examine for our purposes because it is the 1980 Census, not the 1970 Census, which more visibly captured the demographic and geographic consequences of heightened political representation and demographic heterogeneity of the 1960s and 1970s. Starting in 1962, a series of Supreme Court decisions upheld the principle of "one-person, one-vote."2 (*Gray v. Saunders* 1963) Various civil rights statutes, including the 1965 Voting Rights Act, were passed, requiring enforcement as well as re-examination of the administration of federal programs. In addition, the 1965 amendments to the Immigration and Naturalization Act facilitated the entry of new populations to the U.S. Finally, greater fertility rates of racial minority and Hispanic origin populations, both native born and immigrant, compared to the white population, also contributed to increasing demographic heterogeneity. To capture these various changes required outreach to various populations to participate in the 1980 Census. Moreover, there was a policy shift as noted by demographic statistician Calvin Beale, "The Great Society programs and legislation of the 1960s marked a distinct turning point in governmental use of social and economic data. These data became so much more applied, so much more in demand and of practical consequences as they were used to determine eligibility of people and communities for Federal money and assistance." (2004). Not surprisingly, engagement by new stakeholders changed and increased the user population for Federal data. Such engagement raised expectations of traditional Federal data users and added new ones.

By Census 2000, the new methodological phenomenon was the democratization of all data, not just primary census and survey data. The planning and implementation of Census 2000 occurred as the potential of electronic information collection, presentation (not just statistical) and dissemination challenged the prudence and independence of federal data stewards. Additionally, demographically, there was heterogeneity of data users and by sector---public, academe, nonprofit, and private. More so than with the 1980 Census, the interest of the mainstream and ethnic media and general public with Census 2000 expanded the pool of data users. Despite methodological and

[2] *Baker v. Carr* (1962) is usually cited as the most important case in redistricting law, but the phrase, "one man, one vote," which is so closely associated with the court's mandates on redistricting, actually came from the majority opinion in the lesser known case of *Gray v. Saunders* (1963). The language of "one man, one vote" set the stage for a future focus on both legislative and congressional apportionment.

demographic changes, however, the federal requirements for census and survey data remain relatively unchanged. The challenge of balance continues subsequent to the events related to September 11, 2001 and the planning for the 2010 Census that will be limited to short-term data and coverage questions providing apportionment and redistricting counts, as well as housing unit, sex, age, race, and Hispanic origin data. Detailed population and socio-economic characteristics, formerly captured in the sample portion of the decennial census, will be captured in the monthly American Community Survey, a major methodological change directed to continuous measurement of a heterogeneous population, particularly at sub-national levels.

Decennial Data Users and Their Expectations Over Time

Based on federal requirements going back to 1790, traditional data users for census data are the executive, legislative and judicial branches of the federal government. Since at least 1842 states have also been major data users, with the enactment of federal legislation related to districting within states (Anderson 1988) Decennial user expectations include a complete count of the U.S. population; political representation; description and monitoring of trends, gaps, and progress of the population, economy, and environment. Since the advent of surveys, censuses also serve as a benchmarking tool and sampling frame.

Quality expectations of data users are articulated statistically as accuracy and completeness of coverage. Quality expectations for governance include impartiality, reasonableness, and relevance. In addition, quality expectations of data users also relate to the quality of data producers. According to Ivan Fellegi, chief statistician of Statistics Canada, "Public confidence matters because the value of statistics to society directly depends on confidence in their producers. Since few users can actually replicate official statistics, their readiness to use them is ultimately a reflection of their confidence in the professional integrity of statisticians and their ability to carry out their function free of harmful political interference."(1999). In 2002, statistical agencies defined their own quality requirements as "accurate, timely, relevant, accessible, and reproducible" (Federal Register 2002). A related stewardship requirement is that such data should be reasonable in costs. The expectation for confidentiality of personal data has evolved over time, and now is at a very high level.

The 1940 Census was conducted as the United States was coming out of the Great Depression and entering World War II. The 1940 Census data were used to address the requirements of a nation at war. Primary data users included government statisticians and data analysts working in conjunction with non-government experts to deal with this specific situation. Just as James Madison foresaw conscription-related decisions as an application of 1790 Census data, the 1940 Census data were used to identify the pool of males for military duty. In this exercise, analysts discovered that the 1940 Census enumeration for black males was less than the number that registered for military duty for World War II. As commonly known, this finding led to specific research on the undercount of the black, and eventually other minority, populations. As importantly, it led to research on the broader issue of net undercoverage generally. From a quality point of view, the data user question/expectation was completeness and accuracy of the count and, later, correct racial classification. The performance expectation of data users was in the analytical and decision-making power of data.

The 1980 Census captured the new trends of increasing growth and diversity of the United States population by demography and geography in part due to greater attention to changes at subnational levels and to outreach to smaller but increasing populations, specifically racial and ethnic minorities. In preparation for the 1980 Census, the Census Bureau chartered three new federal advisory committees reflecting the federal requirements for data on Black, Hispanic, and Asian/Pacific Islander populations.3 Census historian, Margo Anderson, has called the 1980 Census a major demographic watershed that documented the reversal of the 20th century migration pattern of blacks from North to South, shifts of populations from old industrial cities to the suburbs and non metropolitan areas, and the increasing proportion of the Hispanic and Asian populations (Anderson 1988). The inclusion of new stakeholders from emerging demographic and geographic groups, including those represented by the

[3] In preparation for the 1990 Census, the Census Bureau chartered the American Indian/Alaskan Native Advisory Committee.

Census Bureau advisory committees, in the planning and conduct of the 1980 Census resulted in new data users with expectations for their own interests and uses. Access to data and ability to manipulate data, however, were still limited to sophisticated data users with mainframe computers.

By Census 2000, personal computers were available and affordable to individual data users. The Internet was free. With the explosion of electronic information, coupled with new generations who had only known a post-computer world, the expectations for statistics were at an all time high. Consumers, particularly vocal private sector ones, expected all the data, all the time, instantaneously, and for all levels of geography and all types of demographic groupings. The expectations have not diminished. A diversity of data users--from school children working on community reports, to interdisciplinary teams conducting research, to public officials making life and death decisions -- expect multiple types of data that are not only statistical but also multi-dimensional, visual, and dynamic for real-time and historical use. The Census Bureau has responded to these data user expectations with the on-line American FactFinder. In addition, data users expect to be able to link data sets and customize their analyses without necessarily observing the limits and caveats of the data.

Ongoing Statistical Innovations for Heightened Expectations and New Realities

Amidst these heightened data expectations and new realities, the Census Bureau continues to pioneer statistical and related innovations. Two are presented below. One addresses the expectation to relate geographic, housing, and demographic data. The other considers the complementary roles of survey data and administrative records. The way these new resources are being developed is described first, followed by some (still) general ideas on how these resources could be used in a new integrated survey program.

MAF/TIGER

In the early 1980s, the Census Bureau created the Topologically Integrated Geographic Encoding and Referencing (TIGER) system database to support the mapping needs for the decennial census and other programs. The design of the database adapted theories of topography, graphing, and mathematics to code all relevant geographic features in the United States and its territories (such as roads, railroads, and hydrography), as well as boundaries and other area identifiers needed to support the statistical programs of the Census Bureau. In addition to feature type and location information, attributes of the features were compiled, such as feature names and street address ranges, which provide the ability to assign individual addresses to geographic entities. The database contains codes for every geographic entity used for the Census Bureau's data tabulation process. The building of the TIGER database integrated a variety of encoding techniques such as automated map scanning, manual map digitizing, data keying, and computer file matching. The TIGER/Line files, extracts of selected information from the TIGER database, are released to the public and used extensively by state, local and tribal governments, other federal agencies, and private industry. Local governments have used the TIGER/Line data in applications requiring digital street maps. The private sector has used the data to create products that produce maps for government, business and the general public.

After its use in the 1990 census, the TIGER concept was upgraded. While a major improvement for 1990, TIGER did not, by design, contain individual addresses. The Census Bureau responded to this by developing the Master Address File (MAF), a permanent electronic list of residential and non-residential addresses for Census 2000. The MAF includes specific addresses that allow most census and household survey forms to be mailed, provides descriptions of living quarters locations enabling census enumerators to deliver forms and make follow-up visits to nonresponding addresses, and provides a mechanism to keep track of overall progress in completing each census and household survey. The MAF is linked to the TIGER database, but is restricted from public use by Title 13 of the U.S. Code.

The use of the MAF/TIGER system has greatly improved the accuracy of the data resulting from the Census Bureau's statistical programs. It also has improved the efficiency of the Census Bureau's operations. The system represents a far-reaching and successful use of automation and has resulted in a national resource used by other federal agencies, the Congress, numerous state, local, and tribal governments, and many private sector and academic organizations.

The Census Bureau continues to improve the system. Plans for the rest of this decade include improving the accuracy of the street locations in TIGER through the use of Global Positioning System (GPS) data, and perhaps ultimately associating GPS coordinates with every structure in the MAF. The geographic use of GPS has been under consideration by the Census Bureau for over a decade. In 1993, the U.S. Air Force launched the 24th Navstar satellite into orbit, completing the network of satellites now commonly known as the GPS, which allows the calculation of reference points accurate to a matter of meters (or, in advanced forms of GPS, even centimeters), giving every location on the planet a unique "address." Even before 1993, GPS was in use by the military, and was employed extensively during the Persian Gulf War helping ground troops find their way through the desert, helping naval vessels map mine fields and navigate through them, and helping Air Force and Navy aircraft deliver their weapons more accurately. In 1995, the National Academy of Sciences recommended increased civilian use of GPS, and the number of civilian GPS users began to exceed military users shortly thereafter, with many commercial markets emerging. Today GPS is used extensively by hikers and others participating in outdoor activities. GPS also is finding its way into cars, boats, planes, construction equipment, movie-making gear, farm machinery (for precision crop management), and laptop computers, and is likely on its way to becoming as basic to daily life as the cell phone, where new cell phones also incorporate GPS to support E-911 call location.

The Census Bureau seeks to complete its street locational task by 2008, with much of the work being accomplished through a major contract with the Harris Corporation. A spatially correct TIGER database will facilitate the acquisition and use of GPS coordinates for individual structures. Its implementation will improve navigation to work assignments; improve the adding or updating of road and other map features by field staff; improve the geocoding of addresses; improve the ability of field staff to successfully locate or return to a specific structure; improve address matching for unduplication; and improve quality control. Continued improvements will allow continued expansion of geographic partnership programs that update the MAF/TIGER databases, as the increased level of accuracy in TIGER will allow the use of high-accuracy state, local and tribal government Geographic Information System (GIS) files in ongoing updates of the database. This model, of government data and commercially value-added products, with each existing in separate environments, has been extraordinarily successful.

StARS (Statistical Administrative Records System)

At roughly the same time, the Census Bureau began development of its TIGER database in the 1980s, the agency began to invest in research related to the use of administrative records, which offered a number of tantalizing improvements and/or cost reduction in data collection and estimation. In the National Research Council's report, Modernizing the U.S. Census, the council panel concluded that administrative records data are "a major resource, both potential and realized, in the development and production of small area estimates" and further evaluated the "radical alternative" to traditional census-taking offered by an administrative records census. They recommended that research proceed on both fronts (Edmonston and Schultze, 1995).

One major product of the research is the Statistical Administrative Records System (StARS), a data warehouse consisting of seven major federal databases: the IRS 1040 Master file, the IRS Information Returns file, the Selective Service registration file, the Medicare Enrollment Database file, the Indian Health Service patient file, the Housing and Urban Development Tenant Rental Assistance System file, and the Social Security Administration Numident file.

There are a number of challenges to the use of administrative records in general, including dealing with errors specific to administrative records files, the mission-specific nature of these various administrative databases, discrepancies between administrative records data and data from other sources, unduplication and record linkage methodological issues, and the "point in time" limitations of the data.

Even so, an administrative records "experiment" in Census 2000 yielded promising results, and administrative records data are a key source of intercensal state and county estimates. A number of proposals exist for uses of administrative records in a decennial census, including direct substitution of administrative data for nonresponding households, imputed data, identifying

person duplicates, augmenting the Master Address File development process by reducing the amount of required address canvassing or reducing housing unit duplication, and simulating a complete administrative records census. Other data user expectations, such as using models to produce small area data and integrating statistical data with qualitative data, provide another impetus for statistical innovations.

Integrated Survey Program

One way to think about the integration opportunities that exist today between surveys and administrative records is to contrast what is now possible with the world that Neyman wrote about in his famous 1934 Royal Statistical Society paper (1934). After Neyman came to Washington in 1937, he found Hansen and many other Census Bureau employees already working on new sample survey methods (Hansen 1987, Scheuren 2003). Then ensued an unbelievable burst of creativity among a small group of pioneers (Nathan 2000), in the U.S. and elsewhere, who invented most of what practitioners still do today in government settings (Lohr 1997).

Sampling frames with few, if any, covariates, characterized the world Neyman and Hansen lived in a half-century ago. In many places, moreover, the frames had to be created in the course of survey development and seldom were they computerized to any degree. Even when the frames were available the covariates may not have had much known predictive power for the survey quantities of main interest. For computing reasons and for reasons related to the then general state of statistical theory, design and estimation problems were characterized in univariate terms (Cochran 1977). Enormous changes, of course, have occurred since then but we would argue that these have been piecemeal rather than wholesale and in some cases survey practice has lagged far behind other areas of statistics. It is time to look at how to capitalize on a setting that is becoming more common, at organizations such as the U.S. Census Bureau setting. This setting might be characterized in part as one where:

- Computation is cheap and readily available to an interdisciplinary team of both producers and their clients, allowing, for example, extensive design simulations to be possible;
- Frames are largely computerized and rich in variables that are predictive (Baier et al. 2004)
- Record linkage techniques, data warehousing and metadata tools are reliable enough in many cases to successfully use multiple sources at the design and analysis stages (Winkler and Scheuren 1997);
- Researchers steeped in modern statistics are available to work on both newly reformulated problems and on just plain new problems;
- Old sampling techniques like stratification, replication, balanced selection, and representativeness can be given new meaning and honed to new levels of efficiency;
- Graphical and other diagnostics, so common in mainline statistics, can be created to look at the properties of alternative design and estimation approaches;
- Conditional or client emphasis can be made central, rather than just the traditional unconditional or producer viewpoint; and
- There can be more emphasis on the direct use of multivariate models in small samples and for small domains in large samples.

While this may be a daunting list, each of the elements in the list is part of the practice of at least some practitioners today. To our knowledge, though, no one is doing all of them regularly in ensemble. Proposals to change this exist in commercial business surveys already (Baier et al. 2004). Progress on a more general front for household surveys was proposed some years ago (Scheuren 1993), but judged unfeasible at that time. Perhaps another general look is warranted today, given all the changes that have taken place in the decade since then. The Census Bureau takes this concept seriously enough that it includes in its Strategic Plan the objective to "produce new (integrated) information (products) using existing data sources by developing cutting edge techniques." (U.S. Census Bureau 2003)

Toward a New Gold Standard: Future Statistical Innovations for Heightened Expectations and New Realities

There are many ways to envision the development of a new gold standard for an integrated survey information program at the Census Bureau and/or elsewhere in the federal statistical system. One is to start with known problems and explore strategies to overcome or at least improve them. Another approach is to look at the new and

emerging opportunities for better statistical tools or more collaborative mechanisms that will speed up the rate of innovation. We provide a sample of each of these approaches below.

Survey Challenges

Response rates are falling, and data quality is eroding. What can be done? The declining response rate seems an irremediable part of our new cultural norms. Efforts to encourage participation through appeals to civic duty, the offering of minor compensation, or the use of Internet technology may not be able to keep pace with declining responses, though the Census Bureau and other statistical agencies continue to work together to examine trends and potential solutions. The problem with response rates is that they are often viewed as a marker for data quality. On the other hand, many responses that are not accurate yield low quality data. That leaves at least three possible alternatives:

- Improve compensation for survey participation. Commercial surveys often have stables of respondents, matched for demographics with the Census, who are paid a few hundred of dollars in exchange for conscientious completion of a monthly questionnaire. But these selected respondents are not a random sample, and this poses problems when expressing standard errors;
- Make better use of administrative records. This will probably require some changes in law and better use of record linkage. It may not fully address problems of data quality; and
- Make stronger use of Bayesian and other procedures. Here one uses data from historical surveys or administrative records or just common sense to develop inferences related to smaller sample sizes. Besides the advantage of being able to use smaller samples, Bayesian methods also avoid problems associated with multiple comparisons and simultaneous inference on many subgroups, since they do not expend an alpha level when making inferences. And Bayesian survey methods pave an easy path for sensitivity analysis, decision theory, and sequential acquisition of sample (these can all be done within the frequentist paradigm, but less naturally). The only real drawback to Bayesian techniques is that they invite disagreement from people who disbelieve the model or the assumptions used in the analysis, but this is not unusual in any statistical analysis, and perhaps it is a virtue to make the influence of one's assumptions so transparent.

Survey Opportunities

In practice, the federal system will probably continue making incremental changes in its surveys, trying to cope with increasing error rates and project expenses. It is completely plausible that in some large surveys, for example, perhaps only 30% of the sample give data that are actually relevant to the inquiry. The rest of the contacts may yield non-responses, wrong responses, mendacious responses, or improperly coded answers. To handle this problem we need to apply analytical tools that are robust to bad data, such as S-estimators (Ruppert 1992)]. Such procedures find substructure in the data that are well explained by a simple model, and which are resistant to gross distortions. There is good theory ready to be used for estimating means, variances, correlations, and regression coefficients. Regrettably, no such theory seems possible for estimates of proportions, which is one of the parameters most commonly wanted in the context of federal surveys.

Beyond the issue of response rate lies a fundamental problem in survey analysis. For example, despite the great strides made by the Quality Profile movement (Jabine), it is very hard to estimate Total Survey Error, which incorporates all the uncertainty due to nonresponse bias, response bias, household bias, model error, sampling error, and so forth. Instead, statisticians have been trained to carefully measure and report just the sampling error; this may be the least important error, but it is amenable to calculation. A major challenge for survey statisticians is to develop methods that enable some kind of approximate understanding of the overall inaccuracy of the estimates. Probably this work will require Bayesian analysis—the Bayesian hierarchical model offers a natural way to model components of variance and biases in subgroups. This kind of statistics has been comfortably adopted by statisticians in every branch of our field except federal programs, and it is past time to open the door in government to broad Bayesian thinking.

Survey opportunities will be best achieved through collaborative mechanisms. There are many examples of outstanding collaborative

mechanisms that already exist within the federal statistical system – like the Federal Committee on Statistical Methods4 We need more of these, especially ones that counter the continuing isolation of survey research from mainline statistics.

Conclusions and Next Steps

To recapitulate, in an information society, a fundamental challenge for statisticians generally is to provide information for real-time, real-life decisions that is both responsive and reliable. For federal data stewards specifically, prudence and independence are increasingly required in balancing governmental requirements for data with non-governmental expectations for data.

In the 20th century, survey methods were the gold standard for social measurement and given great prominence. At the same time, however, such work is more properly viewed in historical context as one complementary social measurement. That is, survey methods complement administrative records, which predate survey data, and geo-spatial data, which postdate survey data.

In the 21st century, computing power and electronically executed statistical methods allow integration opportunities across surveys, administrative records, statistical modeling, and graphical representations.

Incremental steps with bursts of intense discovery and innovation by statisticians and federal data stewards will continue toward a new gold standard of social measurement under the auspices of collaborative mechanisms within the federal statistical system.

References

Anderson, M. (1988), *The American Census, A Social History*, New Haven and London: Yale University Press.

Aristotle. (350 BC), *Politics* translated by Benjamin Jowett. http://classics.mit.edu/Aristotle/politics.html

Baier, P., Batcher, M., Hinkins, S., Liu, Y., Mushtaq, A., and Scheuren, F. (2004), *Model Ready Sample Designs and Client Ready Models*, Washington Statistical Society presentations, spring and fall, 2004.

Beale, C. (2004), Reflections on 50+ Years as a Demographic Statistician, Statistical Research Division Seminar, U.S. Census, Bureau, February 25, 2004. Reprinted in the April 2004 *AMSTAT News*.

Budd, E. and Radner, D. (1978), *Statistical Matching and US Income Distributions*. Bureau of Economic Analysis.

Cochran, W. (1977), *Sampling Techniques*, 3rd edition. Wiley: New York.

Duncan, J. and Sheldon, W. (1978), *Revolution in US Federal Statistics, 1926-1976*, U. S. Department of Commerce

Edmonston, B. and C. Schultze, editors (1995**),** *Modernizing the U.S. Census*, National Academy Press, Washington, D.C.

Federal Register (2002), June 2002, volume 67, No. 107, pp. 38467-38470.

Fellegi, I. (1999), Statistical Services—Preparing for the Future, *Survey Methodology*, December 1999, Vol. 25, No. 2, 113-128.

Gray v. Sanders (1963), March 18, 1963 (372 US 368).

Hansen, M. (1987), Reminiscences on the History of Survey Sampling and an Interview with Morris Hansen by Ingram Olkin. *Statistical Science.*

Hauser, Philip M. (1975), *Social Statistics in Use*, New York: Russell Sage Foundation.

Herriot, R. and Scheuren, F. (1980), The 1972 CPS-IRS-SSA Exact Match Project, *Report No. 11, Studies from Interagency Data Linkages*. Social Security Administration

Jabine, T. and Scheuren, F. (1986), Future of Administrative Records, *Journal of Business and Economic Statistics*. American Statistical Association.

[4] FCSM Publication series. Office of Management and Budget, www.FCSM.org

Jabine, T. Data Quality Profiles, *Federal Committee on Statistical Methodology Series.*

Lohr, Sharon (1997), *Sample Design and Analysis*, Duxbury Press. New York.

Morgenstein, D. and Marker, D. (2000), A Conversation with Joseph Waksberg, *Statistical Science*.

Nathan, G., ed. (2000), *Landmark Papers in Survey Sampling*. IASS Jubilee Commemorative Volume.

Neyman, J. (1934). On the different aspects of the representative method; the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, 558-606.

Ruppert, D. (1992), Computing S-Estimators for Regression and Multivariate Location/Dispersion, *Journal of Computational and Graphical Statistics*, 1, 253-270.

Scheuren, F. (2003), An excerpt from the 1949 history by Stephen that appeared in the May 2003 History Column of the *American Statistician*. The complete history by Stephen originally appeared in the *Journal of the American Statistical Association.*

Scheuren, F. (1993), Federal surveys integrated with administrative records. Presented at the 1992 Joint Statistical Association Meetings, American Statistical Association.

U.S. Census Bureau, *U.S. Census Bureau Strategic Plan FY 2004 – 2008,* September 2003, p. 5.

Steinberg, J. (1972), The 1963 Exact Match Project, *Report No. 1, Studies from Interagency Data Linkages*. Social Security Administration

United Nations (2004),"Fundamental Principles of Official Statistics." *United Nations Statistics Division – Good Practices,* June 16 2004. http://unstats.un.org/unsd/methods/statorg/FP-English.htm

Winkler, W. and Scheuren, F. (1997). Record linkage analysis issues, *Survey Methodology*.