

# ROLE OF RESEARCHERS IN PROGRAM DECISIONS

November 15, 1999

*Bill Barron*

Good Morning.

I am very honored to be here today, sharing the podium with John Bailar, at this the opening session of the first Research Conference sponsored by the Federal Committee on Statistical Methodology. I need to begin by making a public admission. Unlike Dr. Bailar, a professor at the University of Chicago who has extensive credentials as a researcher, I would not present myself to you as a researcher in the area of statistical methods. Practitioner, maybe, but researcher, no. So, when I was first approached about being here today to share the opening session with an eminent researcher, my initial reaction was, why me? Clearly, while I could obviously benefit from attending any number of the very interesting sessions that have been scheduled over the course of the next several days, I myself do not have a technical paper to present, and given how I spend my time during the course of a typical work day,-- euphemistically, what I do would be called problem solving not research -- I'm not going to have a technical paper to present any time soon.

Since most people are attending this conference because of their research backgrounds and are on the agenda because of their interests or recent work on an important technical topic, I suppose you could consider me an exception to the general rule. Although the conference organizers who asked me to speak are far too polite to say so, I'm really here, I think, because after more than thirty years at the Bureau of Labor Statistics, and now more than a year at the Department of Commerce and the Bureau of the Census, I have become in, a relative sense, something of a human dinosaur, and a transplanted human dinosaur at that. So in a research sense, I guess, I do have the credentials, or perhaps its more accurate to say, I have the characteristic, of being somewhat unusual, so I think there is a view that from this perspective alone perhaps I will have something interesting to say.

Somewhat more seriously, there is also the view that at this point in my career I have participated in the construction, design, implementation and management of some of the Federal Government's most important and most sensitive statistical programs. In some instances, major initiatives to improve some of these critical statistical measures have been designed, planned, and implemented very smoothly, at least from an external perspective. At the other end of the spectrum, efforts to improve or modernize a number of the statistical system's most critical products have become embroiled in controversy. By virtue of the positions I have held, I have had the opportunity to view these controversies from the inside out, so to speak, across agencies, across organizations, and across a very wide range of extremely important

statistical programs. Given this particular perspective, I have been asked to share my views on the role of researchers in program decisions. Well, from the point of view of this transplanted dinosaur, here goes.

I need to note the fact that there are people here from many government agencies in the United States, as well as from State government, academic institutions, from the private sector, including private sector research organizations, and from statistical agencies representing governments from around the world. While there are probably important differences in the specific roles researchers play in these many different organizations, my experience still tells me that at its essence, the critical role researchers play in these vastly different organizations very likely revolves around ensuring that an organization's methods, work, and outputs are state of the art in terms of current research methods and techniques. Obviously, I can speak first hand only about the work of statistical agencies of the U.S. Federal Government, where researchers can and should be expected to play a critical role in helping measurement programs keep pace with changes in measurement and analytic tools as well as the rapid change taking place in what is actually being measured.

As a context for presenting my views on the role of researchers in a Federal Statistical agency, I want to start by emphasizing the critical importance I attach to the overall work of Federal Statistical agencies themselves. I would like to do this by sharing with you some thoughts I provided to Secretary of Labor Lynn Martin back in the fall of 1991. At that particular point in time, the long-time, highly successful Commissioner of Labor Statistics, Janet Norwood had announced her departure as of January 1, 1992 and I had been asked to prepare a series of background papers on the Bureau of Labor Statistics, including the history of the agency and the position of Commissioner. A major point in the piece I prepared for Secretary Martin quoted the New York Times of July 1, 1884, which stated that the work of the BLS, which came into existence that year, ought to be in the hands of a person and, I quote, "...of a judicial turn of mind who has no interest in the results to be shown other than that of presenting the absolute truth and such conclusions as spring naturally from the facts and figures."

I used this quote to set forth what I believed to be the key selection criteria for the position of Commissioner. I went on to point out that this same commitment to having "...no interest other than that of presenting the absolute truth..." was, in my mind, the fundamental rationale for the existence of the BLS as a Federal agency. In my view, this same fact applies to the Census Bureau and the entire Federal Statistical System and its component agencies. Over the years, heads of various statistical agencies and other key observers as well as critics of the federal statistical system have cited the importance of a basic triumvirate of requirements for statistical agencies, namely timeliness, accuracy and objectivity. Clearly, all are critical to a successful statistical agency, and the advice I gave to Secretary Martin placed due emphasis on all three. However, I think it should be very clear that, while occasional lapses in timeliness can be understood and explained and while occasional, and, I hope, rare, errors can sometimes occur despite an agency's best efforts, lapses in objectivity are simply and without exception, completely unacceptable. In and out of the statistical system, we cannot over-emphasize that our commitment to the objective presentation of the facts is absolute. And it is this area of objectivity, where, I believe, researchers' greatest contributions to programs can be made.

On this point, and it should be no surprise to anyone, especially many of the people in this room, that the American Statistical Association has addressed this specific issue. In fact, the very first items on the list of the Official American Statistical Association Guidelines for Statistical Practice, as prepared by the ASA's Committee on Professional Ethics, state that statisticians should:

\*present their findings and interpretations honestly and objectively;

\*avoid untrue, deceptive, or undocumented statements; and

\*disclose any financial or other interests that may affect, or appear to affect, their professional statements.

This is good advice for researchers anywhere but this is particularly important for a Federal statistical agency and its staff. So, first and foremost, the role of a statistical agency, and the staff who work there, researchers and non researchers alike, indeed, everyone engaged in the work of the statistical agency, is to produce results which are timely, accurate, and, most important of all, objective.

In addition to surveys and statistical programs that are objective and statistically defensible, my experience is that statistical agencies funded with public funds also must have a basic commitment to programs, methods, and techniques that are efficient and cost effective, as well as explainable, understandable, and acceptable to the public. Please note that understandable and acceptable does not mean that every data user is necessarily happy with a particular result. For example, it has been my personal experience over the years, especially during times of economic stress, such as that which occurred during the course of the national elections in 1992 when I was Acting Commissioner of the BLS for 23 months, that with the issuance of every economic indicator, a large segment of folks were ecstatic, and another large segment of interested data users were not. I'll let you, if you want, guess, for any given statistic issued at that time, who was happy and who was less than pleased.

But what I would much rather have you understand, and I'm certain virtually all of you already understand this, is that it isn't the happiness of users we were striving for. Rather it was acceptance, acceptance of the fact that in a very intense environment we were doing the best possible job we could do, given the resources available, and the state of the art in survey methods, to measure such things as unemployment, inflation, employee compensation, and productivity. So while the results of some of the programs may not have met everyone's wishes, I believe we were meeting the expectations of the preponderance of our data users, by producing data that were as timely and accurate as we could make them, by managing and carrying out these programs in an objective, non-partisan fashion, and using the state of the art in measurement science given our knowledge and budget at the time.

While it is a responsibility that is shared with other critical occupations and functional groups within each agency, the research role in program decisions is to keep programs and decision makers armed with the latest methods and advice that research and modern research methods can provide. Obviously, I can only speak from my perspective and experience, and today in this setting, I would prefer to focus on successful, within-agency research programs. Thus it seems to me that the research programs that succeed, and the researchers who are successful in terms of seeing their work implemented, generally have the same basic priorities that I have strongly emphasized for agencies as a whole. To say this another way, from my perspective as a senior manager, my particular preference is for research, researchers, and research programs that are relevant to an agency's programs, problems and priorities, and for openers, this research work, like the work of the agency itself, must be timely, accurate, and objective. In this regard, researchers have the training to ensure that an agency's work and methods are current and state of the art.

Rather than to continue to discuss this in general terms, perhaps the best way for me to describe a successful model for in-house research efforts at a Federal statistical agency may be to pick and describe one particular, important success. Although I think it is generally true that research at the agency level

has, over the years, been severely underfunded, there have been, in spite of this reality, a number of important, very significant successes. For example, while it stands in contrast to the public commentary of some very prominent researchers, my own view is that the work to introduce probability sampling into the Consumer Price Index, while not problem free, was clearly an extremely significant effort. Along with the initiation of the Continuing Consumer Expenditure Survey and the now modified Point of Purchase Surveys, the introduction of probability sampling into the CPI program was, and still is, a very important success. This is an area where critics should focus on the vast array of measurement issues where work is still required rather than resort to unfortunate and inaccurate statements that imply that the CPI has never been improved since first introduced in 1917. Other significant research improvements I am directly familiar with would include introduction of new measures such as the Employment Cost Index, and the Census of Workplace Fatalities. However, for purposes of today's discussion, the particular program I would like to use as an outstanding example of successful research in a Federal statistical agency is the redesign of the Current Population Survey, a major undertaking of both the BLS and the Bureau of the Census during the period 1986 through 1994.

Let me take just a minute here to say a few words, for those of you not from the BLS or the Census Bureau, about what the Current Population Survey is and what its redesign entailed. The CPS was initiated almost 60 years ago, in March 1940, as the Monthly Report of Unemployment to provide up-to-date information on the U.S. labor force. The monthly unemployment and labor force data, usually issued on the first Friday of each month, is the most famous product of this 60,000-household survey. The CPS also produces important information such as data on characteristics of the population such as income and poverty, educational attainment, migration, and fertility. Originally conducted by the New Deal's Works Progress Administration, the CPS has been conducted by the Census Bureau since 1942 and its labor force component has been analyzed and released by the BLS since 1959.

The CPS redesign that was implemented in January 1994 was the first major modification of the survey since 1967--over a quarter century--and involved both major changes to questionnaire content and to data collection and processing. The advent of laptop computers made it possible to introduce a more sophisticated and complex questionnaire that could capture important changes taking place in the economy. Indeed, without advances in computer technology it would not have been possible to have a major CPS redesign.

What I would like to briefly illustrate is the role researchers played in establishing the relevance, feasibility, and credibility of the massive and important research effort required to implement the CPS redesign. By relevance, I mean building the case for the need for major modifications. Think of the dramatic social, cultural, and economic changes that occurred in the quarter century from 1967 to 1994. These changes included the massive movement of women into the labor force, a shift from a manufacturing-based to a service-sector job market, more multiple job holding and part time work, and other changes in the places where people worked and the type of work they did.

As early as 1979, 15 years before redesign was implemented, the National Commission on Employment and Unemployment Statistics had recognized the need for changes to the 1967-vintage CPS, even though it could not have envisioned the advances in computer technology that would make that possible. The members, advisors, and staff of that Presidential commission, which was chaired by Prof. Sar Levitan of George Washington University, represented some of the best research minds in the country. They began to make the case for change. Other researchers, working at the staff level throughout the early 1980's, conducted important investigations that pointed the way more specifically to the shortcomings of the

existing questionnaire. This research found that some concepts were not being measured precisely, or were not well explained in the existing questionnaire, or needed to be revised. It showed that there was a need for additional data on some important new topics. So years of extensive research went into building the case, that is, establishing the relevance and need for CPS redesign.

Researchers also contributed early on to establishing the feasibility of making major revisions to the CPS. I am thinking here of new methodologies introduced in the 1980's, namely cognitive theory and computer-assisted interviewing. Cognitive theory research applied the work of the behavioral sciences to dissecting the questionnaire to show us better ways to word and order questions. This type of research was essential to any strategy to reduce nonsampling error in the CPS. It also pointed the way to a more sophisticated and complex questionnaire that researchers believed could not be administered except via computer. That is why the research conducted on computer assisted telephone and personal interviewing was so essential to the overall success of CPS redesign. Indeed, one of the most important and interesting things about the CPS redesign project is how it brought together so many people from different disciplines, as well as from the two agencies, all making their unique contributions to the overall success of the project.

In addition to contributing to the relevance and feasibility of the CPS redesign, researchers also lent the effort extraordinary credibility. They did this not only because of their own credentials or the reputations of the agencies for which they worked, but because they designed and analyzed a major parallel or overlap sample that told us how planned changes to the CPS would affect the data. The overlap sample was conducted from July 1992 through December 1993, prior to the January 1994 start date for implementing the redesign. Its purpose was similar to that of the Dress Rehearsal Census - the Census Bureau conducted last year and completed analyzing this year for Census 2000. It was a way to avoid surprises but also to better understand the effect of the redesign. It was a relatively large sample, about 13,000 units, and so was costly. In fact, the costliness of the overlap sample was one of the reasons that it took so long after the recommendations of the Levitan Commission to begin work on the CPS redesign. During the early 1980's, with the widespread emphasis on tax cuts and reducing the size of government, it just wasn't possible to get money for such a large research project.

The analysis of the overlap sample indicated that the redesign could increase the unemployment rate by as much as one half of one percentage point. That doesn't sound large, but remember that this was 1993-1994, the beginning of an Administration, our current Administration in fact, that had been elected at least with occasional reference to the slogan "It's the economy, stupid." The Administration was also trying to obtain congressional approval of the North American Free Trade Agreement, and NAFTA's opponents warned it would lead to U.S. job losses. If the redesign led to changes that made it appear that the unemployment level was not going down, that could be very significant. Now, if there is one thing researchers should never, never, never do, it is to surprise the managers who must champion their research and fight for the funding for that research. And if there is one thing senior managers should never do, it is to surprise policy level. Our little research finding about the potential increase in the unemployment rate, could have had potentially explosive results. Therefore, we had a responsibility to inform key people about this and I recall some very interesting discussions with the leadership of the Department of Labor, the Council Of Economic Advisers, and other key administration leaders and White House staff on the history, timing and substance of the changes that were about to occur. Fortunately, these changes had all been announced and publicized well in advance of their introduction. I also think that it was very clear that the many years of professional contact and working relationships between the researchers and staff of the statistical agencies, and such noted technical experts such as the then new

BLS Commissioner Katharine Abraham and Drs. Larry Katz and Alan Krueger, Chief Economists, to Labor Department Secretary Robert Reich, were critical in establishing widespread acceptance and understanding of the significant changes that were about to occur.

We also informed the Congress in advance of releasing new unemployment numbers and I remember one Congressman (David Obey, who is now ranking Democrat on the House Appropriations Committee) emphasizing the need to undertake additional educational efforts so the public would not draw false implications about the economy as a result of changes in the survey. In fact, that was the whole purpose of doing the overlap sample--to prevent surprises--and I think that gave credibility to the whole research effort.

I will need to leave to others to explain why our carefully conceived overlap sample apparently significantly overstated the potential impact of the new process on the aggregate unemployment rate. For today, the important point I want to make is that the need to be prepared for the worst case scenario had been understood and at least we had prepared for that eventuality.

What was accomplished?

We introduced a rewritten and computerized questionnaire that took into account major shifts in the U.S. economy and society (especially the increasing importance of female workers and the service sector, and growth in alternative work schedules).

The computerized questionnaire permitted built-in skip patterns, tailoring of some questions, and dependent interviewing, that is, the use in the current interview of information obtained during a previous month's interview. Among the benefits of dependent interviewing, for example, was that it reduced the "spurious" changes in classification, such as at the detailed occupational level, that had occurred in the old instrument. Automation also eliminated errors stemming from incorrect skip patterns and from the printing and shipping of paper questionnaires.

Respondents find a more personalized interview, clearer questions, and a minimum of repetition from one month to the next. Retired and disabled respondents find the interview particularly streamlined.

As a result of cognitive testing, questions were made clearer by revising question wording and order, splitting complex questions into two or more simpler ones, and embedding recall strategies into the questions.

The measurement of certain labor force concepts was sharpened, including persons on layoff from a job, jobsearch methods, duration of unemployment, and industry and occupation. More accurate data were obtained for persons in the labor force, especially those with marginal work and job search activities.

Several definitions were modified to make them more consistent with existing concepts, including discouraged workers, part time for economic reasons, and new entrants and reentrants (categories of unemployment).

Questions were added so that information was available regularly for the first time on multiple jobholding and hours usually worked for all workers.

In sum, researchers contributed immensely to the successful implementation of CPS redesign. We have been using the new questionnaire and the new technologies now for five years and they have provided

new and improved data and greater efficiency. In a few years, researchers will show that the current questionnaire no longer adequately describes our Nation's labor force or that new technologies could be used to improve data collection further. And then we'll have another redesign. It's the nature of good research and good researchers to be constantly looking for improvements.

As an example of this, and as a final thought, I would like to mention that the Census Bureau's executive staff have been giving a lot of thought to the future of our organization and particularly to how we will position ourselves after Census 2000. We are calling this "Census Modernization" and researchers have had and will continue to have an important role in defining and implementing our ambitious plans.

There are two dynamics that are pressing the Census Bureau to pursue a modernization program. The first is what we call the "Boom and Bust" cycle. For eight years each decade, the Census Bureau is a small agency within the Commerce Department. During this period, our primary responsibility is to collect the economic and demographic data that impact economic policy and social policy at the Federal level data, that are essential to state, local and tribal governments, and that help shape or determine activities of businesses and the growth of communities. In the ninth year, the Census Bureau rapidly builds a nationwide infrastructure for the decennial census. In the tenth year, we briefly become the largest civilian agency in the government, count the population, undertake massive updating of core demographic data, and then quickly fade again into the background.

The second dynamic is that what we need to measure and how we measure is changing. We need to start measuring different new aspects of our lives such as e-business. We need to measure differently by modernizing our data collection and dissemination capabilities.

So we believe we have a tremendous opportunity to achieve efficiencies in non-decennial areas by taking advantage of decennial investments and to smooth the huge funding ramp-up associated with the decennial census. As part of the Decennial Census under the leadership of John Thompson and Jay Waite, we have designed systems and methodologies incorporating technologies that did not exist a decade ago, rebuilt partnerships and forged new ones, and spent considerable time and effort building confidence in a public increasingly suspicious of its government. Rather than tossing all that aside after 2001, we are looking at ways to integrate our decennial and survey activities and the systems that support them; build a stronger programmatic and technological infrastructure; stretch out and level out the decennial cycle; reduce the complexity, cost, and burden of the next decennial census; and actively explore and incorporate technological advances.

We have identified half a dozen major elements of modernization and researchers have already contributed significantly to success in some of them or will have future opportunities to do so. These are: 1) expand the American Community Survey, 2) draw new, updated, and coordinated samples for the major recurring household surveys, 3) update the census address list and associated geographic files throughout the decade, with community support, 4) enhance our data access and dissemination systems, 5) conduct early planning for the 2010 census so that we can provide adequate time for mid-decade testing and evaluation, and subsequent modification, and 6) explore evolving information technology, such as the Internet, Global Positioning Satellites, and Data Warehousing technology. There will be plenty of work over the next few years to keep our researchers very busy. I should mention that a seventh element of our modernization proposal is to improve, either through construction or renovation, the 60-year old Suitland facility so that our researchers and all of our other valuable staff will have a safe and comfortable environment in which to do their work.

In closing, I want to thank you for listening to a nonresearcher's views on what constitutes good research. The organizers of this conference have done an excellent job, and I want to commend them, for arranging a program that will display some of the finest examples of research now being conducted in Federal statistical and related agencies. I've touched on a number of qualities that I believe are essential to good research--objectivity, timeliness, accuracy, cost-effectiveness, statistical defensibility, explainability, relevance, feasibility, and credibility. While I place objectivity in a separate class by itself, I've made no attempt to order or classify these qualities into a coherent system. That's a challenge I leave with you, and the organization where you work. I hope some of my comments will assist you as you strive to meet that challenge.

Thank you.