Statistical Policy
Working Paper 26

Seminar on Statistical Methodology in the Public Service

Part 1 of 2

Federal Committee on Statistical Methodology

Statistical Policy Office
Office of Information and Regulatory Affairs
Office of Management and Budget

August 1997

THE FEDERAL COMMITTEE ON STATISTICAL METHODOLOGY
(November 1996)

MEMBERS

Nancy J. Kirkendall, Chair
Office of Management and Budget

Susan W. Ahmed
National Center for Education Statistics

Yvonne M. Bishop
Energy Information Administration

Cynthia Z. F. Clark
Bureau of the Census

Steven Cohen
Administration for Health Policy Research

Lawrence H. Cox
Environmental Protection Agency

Zahava D. Doering
Smithsonian Institution

Gerald Gates
Bureau of the Census

Robert Fay
Bureau of the Census

Daniel Kasprzyk
National Center for Education Statistics

Daniel Melnick
National Science Foundation

Robert P. Parker
Bureau of Economic Analysis

Charles P. Pautler, Jr.
Bureau of the Census

David A. Pierce
Federal Reserve Board

Thomas J. Plewes
Bureau of Labor Statistics

Wesley L. Schaible
Bureau of Labor Statistics

Rolf R. Schmitt
Bureau of Transportation Statistics

Monroe G. Sirken
National Center for Health Statistics

Alan R. Tupek
National Science Foundation

Denton R. Vaughn
Social Security Administration

Robert Warren
Immegration and Naturalization Service

G. David Williamson
Centers for Disease Control and Prevention

EXPERT CONSULTANTS

Robert Groves
Joint Program in Survey Methodology

Graham Kalton
Joint Program in Survey Methodology

DEDICATION
to
Maria Elena Gonzalez

The Seminar on Statistical Methodology in the Public Service was dedicated to the memory of Maria Elena Gonzalez, the founder and leader of the Federal Committee on Statistical Methodology (FCSM) from its inception in 1975 until her death in 1996. Maria worked -- often quietly -- but always tenaciously to improve the quality of the statistics our government provides.

No single professional accomplishment is more closely associated with Maria Gonzalez than her stewardship of the FCSM. This novel approach to leveraging the talents of our statistical community was fully Maria's creation. It is a tribute to her vision that the current environment renders interagency cooperation ever more important and relevant. We are pleased to honor Maria by vowing that the FCSM will continue to be a focal point for collaboration among Federal Agencies on matters of statistical methodology and statistical practice.

Each of us is better for having known Maria Gonzalez. Her boundless energy and intellect, exemplified in her leadership of the FCSM, are at once a legacy and a challenge to us.

Katherine K. Wallman
Chief Statistician
Office of Management and Budget

PREFACE

The Federal Committee on Statistical Methodology was organized by the Office of Management and Budget (OMB) in 1975 to investigate issues of data quality affecting Federal statistics. Members of the committee, selected by OMB on the basis of their individual expertise and interest in statistical methods, serve in a personal capacity rather than as agency representatives. The committee conducts its work through subcommittees that are organized to study particular issues. The subcommittees are open by invitation to Federal employees who wish to participate. Since 1978, 26 Statistical Policy Working Papers have been published under the auspices of the Committee.

On November 12-13, 1996, the Council of Professional Associations on Federal Statistics (COPAFS) hosted a "Seminar on Statistical Methodology in the Public Service." Developed to capitalize on work undertaken during the past fifteen years by the Federal Committee on Statistical Methodology and its subcommittees, the seminar focused on a variety of topics that have been explored thus far in the Statistical Policy Working Paper series and on work on statistical standards done by the Statistical Policy Office. The subjects covered at the seminar included:

Electronic Dissemination of Federal Statistics
What the Public Needs to Know About Federal Statistics
Training Federal Statisticians
Measuring Customer Satisfaction
One-Stop Shopping for Federal Statistics
Reviewing and Reporting Quality in Survey Data
Performance Measurement in Statistical Agencies
Survey Integration: Initiatives in Heath Data
Sharing Data for Statistical Purposes
Reinventing Economic Classifications
Using Administrative Records for Statistical Purposes
Validation of Cognitive Questionnaire Pretesting Methods

Each of these topics was presented in a two-hour session that featured formal papers and discussion, followed by informal dialogue among all speakers and attendees.

Statistical Policy Working Paper 26, published in 2 parts, presents the proceedings of the "Seminar on Statistical Methodology in the Public Service." In addition to providing the papers and formal discussions from each of the twelve sessions, this working paper includes Victor Cohn's keynote address, "What We the Public Need to Know," and comments by Graham Kalton and Nancy J. Kirkendall at the closing session, "Toward an Agenda for the Future."

We are indebted to all of our colleagues who assisted in organizing the seminar, and to the many individuals who not only presented papers but who also prepared these materials for publication.

Table of Contents

Tuesday, November 12, 1996

Part 1

KEYNOTE ADDRESS

Session 1 - ELECTRONIC DISSEMINATION OF FEDERAL STATISTICS

Session 2 - WHAT THE PUBLIC NEEDS TO KNOW ABOUT FEDERAL STATISTICS

Session 3 - TRAINING FEDERAL STATISTICIANS

Session 4 - MEASURING CUSTOMER SATISFACTION

Session 5 - ONE-STOP SHOPPING FOR FEDERAL STATISTICS

Session 6 - REVIEWING AND REPORTING QUALITY IN SURVEY DATA

Wednesday, November 13, 1996

Session 7 - PERFORMANCE MEASUREMENT IN STATISTICAL AGENCIES

Part 2

Session 8 - SURVEY INTEGRATION: INITIATIVES IN HEALTH DATA

Session 9 - SHARING DATA FOR STATISTICAL PURPOSES

Session 10 - REINVENTING ECONOMIC CLASSIFICATION

Session 11 - USING ADMINISTRATIVE RECORDS FOR STATISTICAL PURPOSES

Session 12 - VALIDATION OF COGNITIVE QUESTIONNAIRES PRETESTING METHODS

TOWARD AN AGENDA OR THE FUTURE

Part 1

KEYNOTE ADDRESS

WHAT WE, THE PUBLIC, NEED TO KNOW

THE NUMBERS YOU COLLECT -- WHAT WE NEED TODAY AND TOMORROW

An address to the Seminar on Statistical Methodology in the Public Service, sponsored by the Council of Professional Associations on Federal Statistics.

By Victor Cohn, former science editor of the Washington Post, 1994-1966 scientific and public affairs fellow of the American Statistical Association.

You are important. You are important to us, the public, to us, the media, to business, to public policy makers, everyone.

We hear that numbers are dull. A turnoff. But we are ruled by numbers, by our incomes, our expenditures, our taxes, by the prices we pay and on and on. The numbers you collect, the numbers you convey to the public and how you convey them affect my life and the lives of millions. They affect public decisions, business decisions, my decisions.

Statistics uses numbers. But statistics, I have learned, is really about accurate and useful information. It is about straight thinking. The better and clearer we all can be, the clearer and easier become many lives.

What do we, the public, and we of the media want of you? And I don't say it's easy.

First of all and most of all,we want honest numbers. Accurate, of course -- and I know that's not always easy -- or at least your best possible estimates. But also politically honest numbers, data uninfluenced by the political or the personal considerations of the people you work for, or those who influence the people you work for. My hero is Janet Norwood, who spent 25 years at the Bureau of Labor Statistics, much of it as commissioner, and steadfastly resisted all attempts to paint rosy or unrosy pictures for the benefit of whatever Democratic or Republican administration.

Next, how can you explain what you do? How can you help us -- us, the public, and us, the media -- discern what to possibly believe and what to question when assertions that include or should include statistics, are thrown at us? How can we separate the truth from the trash? Or the probable truth from the probable trash?

I am here because I happened to write a little book called*NEWS & NUMBERS: A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields.** It has received a

pretty good reception in the journalistic comunity, and, to my amazement, has gone into a sixth printing and even been recommended -- in the*Journal of the American Statistical Association, Chance* and medical, scientific and other journals -- to statisticians, physicians, scientists and even public policy makers.

Well, I have a confession to make. I was asked to write an article about the book for the quarterly publication of the National Association of Science Writers, and I not only wrote the article but I also wrote the headline, which was: MATH ILLITERATE WRITES GUIDE TO STATISTICS. Yes, the only thing I remember about high school math is that the instructor, a kind Mr. Peterson back in Minneapolis, said he'd give me a "D" for the last semester if I promised not to take any more math. I promised and avoided math, also physics and chemistry, all through high school and college. With this wonderful preparation, I naturally went into science writing.

I tell you this because my problem, like the problem of all reporters and of the public, became just this. When it comes to any controversial issue -- and what issue isn't, whether government, welfare, education, crime, health, risk, the environment or whatever? -- one side says, The sky if falling, the other, It isn't! One side says, We have a wonderful fix, the other, You'll ruin things!

Does this sound familar about Washington? Does it sound familiar about our recent presidential race? Who to believe? How can we assess the credibility of what we hear?

I suspect strongly that this is not a problem of the media and the public alone. I know that I am addressing many trained statisticians and some who may not have statistics degrees, but are experienced in dealing with statistics. Surely many of you sometimes have to try to decide whether to believe or use some statistics and statements that are thrown at you. And you may sometimes wish for some simple, in part non-statistical questions and concepts that you can use -- either verbally or as you read -- to help separate the probable truth, something you can probably believe, from the probable trash.

When I was writng*NEWS & NUMBERS,* many fine statisticians and epidemiologists helped me find some simple, yet revealing questions.

Such as:

How do you know? Are you just telling us something you "know" or have "observed" or "found to be true"? Or have you done or found any studies or experiments?

What are your data? Your numbers? How or where did you get them?

How sure can you be about them? Are there any possible flaws or problems in your conclusions? Salespeople are sure. Honest folks admit uncertainties.

How valid -- and we have to explain to the media and the public that in science, valid means accurate -- are your numbers and conclusions?

How reliable -- and we have to explain that this means how reproducible -- are they? Have results been fairly consistent from study to study?

What is your degree of certainty or uncertainty by accepted tests?

Who disagrees with you? And why?

The interesting thing is that the very way someone answers, or dodges, such questions can reveal much.

I believe that we can then go a long way toward helping non-statisticans discern the probable facts from the probable trash, a long way toward judging claims and statistics that are thrown at us, by explaining that there are six basic concepts that apply to all science, all studies and virtually all knowledge of society and the universe. And that remembering these can help anyone ask, "How do you know?" with a considerable degree of sophistication.

I can't go into detail here on these, but they are:

1. UNCERTAINTY -- the certainty of uncertainty, the fact that all science is almost always uncertain, or uncertain to a degree. Nature is complex, people are more so, research is difficult. Almost all anyone can say about atoms or cells or human beings or the biosphere is: there is a strong possibility that such-and-such is true, and we may know more tomorrow.

This can help us tell people why things so often seem settled one way today and another tomorrow, and why so much is debated, whether the effects of global warming, a pesticide, a low-fat diet or a medical treatment. Yet uncertainty need not impede crucial action if society understands and uses the rest of these principles.

2. PROBABILITY -- the use of probability and how scientists live with uncertainty by measuring probability. And how the laws of probability and chance tell us to expect some unusual, even impossible sounding events. And that there is something else to remember when someone says, "How do they know this stuff isn't causing harm?" Science cannot prove a negative. No one can prove that Little Green Men from Mars have not visited earth. The burden of proof should be on those who say something is true.

3. POWER -- the power of large numbers in finding something if it's there, say an increase in cancer in workers exposed to some substance, and how we must be wary of studies with only a small number of cases. How we should ask too about any association's statistical strength, its odds.

4. BIAS -- the danger of bias. How bias in science means introducing spurious associations and reaching unreliable conclusions by failing to consider other influential factors -- confounding variables or co-factors, in plain language, other explanations. We can learn to watch for bias by asking, "Are there any other possible explanations?"

5. VARIABILITY -- its ubiquity. How a common pitfall of science is that everything measured or studied varies from measurement to measurement. Every human experiment, repeated, has at least slightly and sometimes markedly different results.

6. Finally, the fact that there is a **HIERARCHY of STUDIES,** from the least to the generally most believable, starting with simple anecdotes, going on to more systematic observation or "eyeballing," then proceeding to true experiments, comparing one population or sample with another, under controlled or known conditions.

All this tells us that one study rarely proves anything, that when someone tells us, "I've done a study," we should ask, "What kind? How confident can you be in the results? Were there any possible flaws?" Et cetera. It tells us that we must seek out the most credible evidence, the most likely probabilities. It tells us that we must look for a consensus among the best studies and the best, most neutral observers, those who remain scientists rather than salespeople.

Now, how can you get our message out? How can you get it to the public and the media? How can you make your information understandable and useful?

Reporters, the most conscientious, have a lot of constraints and demands. What we need are numbers and conclusions that are put to us clearly and in language we can understand. We're not the only ones. John Tucker at the Board of Mathematical Sciences at the National Research Council recently told me that part of the reason statisticians or people of any description responsible for collecting and disseminating numbers get let off the table, get left out of decision making, is that the very people they work with or work for don't understand their numbers.

Authors or investigators sometimes fail to include some numbers that any of us would want for intelligent decision-making. Researchers wrote in the*New England Journal of Medicine* that "one to three drinks" a day may help protect against heart attacks. They defined a drink as 13.2 grams of alcohol. But with whiskey, wine and beer all at various proofs or alcohol content, neither the journal's article or editorial nor any news reports I read or saw told in common language how much booze, wine or beer one should consume to drink no more than the prescribed grams. A free-pouring drinker could down far more.

The media and the public have a lot of trouble understanding risks. Part of the problem is a failure to understand the difference between a plain number and a rate, and the fact that a rate has to mean so many per so many per unit of time. A*Washington Post* headline once read, "Airline Accident Rate Is Highest in 13 Years." The story, like many others misusing the word "rate," reported no rate at all, merely death and crash totals. A correction had to be printed pointing out that the number of accidents per 100,000 departures -- the actual rate, the "so many per so many" -- had been declining year after year.

Another problem: the word "safe." I believe we should pretty much quit using it. When the FDA or EPA or some other agency describes something as safe, it is almost never completely so.

We commonly have to accept some risk to achieve some gain. Current example: air bags in our cars. Instead of calling a prescription drug, for example, safe, maybe we should start to say "relatively safe" or "as safe as possible" or some such phrase and try to indicate in the best numbers and rates we can the degree of safety or risk, the rate of adverse events. When we faril to do this,when we announce something as "safe" and that turns out to be wrong -- the frequent case -- we create the public confusion and lack of confidence that exists today.

Some practical rules for dealing with the media?

Let me quote and add to some stated by Dr. John Bailar, former statistical adviser to the*New England Journal of Medicine*, now head of biostatistics and epidemiology at the University of Chicago.

* "Give it all the time it takes." Indeed do this, I would add, when you are phoned or approached by a reporter who seems interested in learning something and doing a good job, and obviously devoted some time and effort to learning at least a little something before speaking to you. But if you are suddenly quizzed by someone who seems to know nothing and is interested only in a quick quote or sound bite -- and that's all the exposure you're going to get -- think carefully about what you want to say, say just it, and don't say much else if you think that's the honest and important message that should be conveyed.

This may sound cynical but the superficial reporter deserves no better. The public may deserve a great deal better, so it's important to say the right thing.

* "Explain in short, simple sentences." And use plain language. Statistics and mathematics have their own useful languages, but these aren't the languages that the rest of us use. Words like "normal," "mean," significance," "validity," "reliability," "power" have different meanings in statistical and common language.

* Help your audience understand the issues by familiar examples or phrases. Instead of saying "the U.S. homicide rate is 10.4 per 100,000 per year," say "one person in every 10,000 will be a homicide victim.

* If you say "homicides were up by X percent," add the actual number, the "10.4 per 100,000" that tells people their own chances.

* "If you don't know or aren't sure, say so. Your stock will rise."

* Be prepared to answer the questions: What would you recommend that the public do about this? What would you do, or what are you doing? What would you want your family to do?

But be clear whether you are responding from an official position or stating your personal opinion. "Distinguish between what is the science," the facts, and "what do they mean?" -- how they should be interpreted in your opinion.

* "Any constraints on use -- that is, 'off the record' -- must be on the table before the words are out. Afterwards is too late. . . . Be clear about what you mean by 'off the record,' and make sure there is agreement."

In Washington, at least, not necessarily elsewhere, such phrases have a pretty precise meaning.

"Off the record" means "you can't use this," though the reporter may be free to get the same information from another source.

"On background" means "you can use it, but don't use my name" -- and get an understanding on how it may be attributed, whether to a Department of Something official" or "Congressional aide" or "government official" or whatever.

"Deep background" means you can use it but only on your own authority, not with attribution to anyone.

I don't mean to encourage you to have such recourse. Have the courage to say what you mean and to stand behind what you say. Statements from anonymous sources lack the credence of "Max Smith said . . . "

* If you would increase your chances of being reported accurately, provide a copy of your paper and charts and perhaps other background material. Tell your information people to add a copy of the full paper to their press release.

* Offer to review what is written for accuracy of facts -- and facts alone, not writing style or interpretation. But do not demand control. Put it generously -- "I'd be glad to have you call me if you want to recheck any of the facts or the way you've put it." If you are called, don't try to wiggle out of what you have previously said, unless you think there's something plainly wrong. Again, have the courage of your convictions, even when they sound blunt when shown or read back to you.

In doing all this, use your public information peple, but find the good ones, those who not only have shown they know how to get ink and tube time, but also have taken the trouble to learn your subject, and come sniffing around from time to time to find out what your are dong. Value and try to promote these people.

But also be available yourself to media calls and inquiries. Reporters value and use public information people who know their subject. But they also want to get their information from the horse's mouth. They want to talk to the people at the top who are responsible, and they want to talk to the people who put together and really know the data. When Cap Weinberger was secretary of then-HEW, he and his deputy, Frank Carlucci, put out directives telling HEW peple at every level: talk to reporters, answer reporters' phone calls. And they did. And HEW had a remarkably good press. Silence, reluctance to talk breeds suspicion. Openness builds trust.

In 25 years at the *Washington Post,* I had a lot of occasions to call the National Center for Health Statistics, sometimes getting to someone through public affairs but often phoning directly, and never failed to get immediate and cheerful help. The same goes for the statistics people at the National Cancer Institute, whose numbers were in the news a lot, and not always with good news about progress against cancer. I'm sure the same would be true of some other agencies. I've heard good things about the Census Bureau and the BLS. And not such good things about some others.

Finally, let me repeat: candor, accessibility build trust. I said at the outset that you are important. This means that we depend on you to tell us candidly how your conclusions, your data meet the tests of validity and reliability and the other statistical rules that you know and use. Tell us that in a thousand difficult situations --risks, apparent risks, dangerous diseases, population trends, educational progress or lack of it, the efficacy or lack of efficacy of welfare programs, the performance of many agencies and systems, yes, and in that controversial count, the census -- your conclusions, your data are not written in stone, but rather, "Here is the best we can say today, and we may know more tomorrow. The pursuit of knowledge is a continuing struggle."

About the census, now in contention. I can't give you a single formula for getting across the message that the best count in today's United States inevitably must include the best possible estimates, which includes sampling. Some of the opponents of sampling of course know that, which is why they oppose the best possible estimates.

I can only say: Educate, educate, educate, and, where appropriate to your job, politic too -- it's a respectable activity in Washington. Don't just wait to be called on to testify. Keep repeating your best opinions and the best possible facts in every possible way, and get them again and again to the political leaders and their staffs. Get them to the media. Do not give in. Obey, of course, what directives you must. But do not retreat on your beliefs. And eventually, truth may prevail.

The candor I have recommended on this and all issues -- including the admision at times of uncertainty -- flies in the face of the usual Washington argument for one course or another, which is "Here are the immutable facts, my friends, and if you follow my recommendations, everything is going to be wonderful." Honesty may make selling some ideas a bit harder, but it distinguishes an honest seeker after truth from a snake oil seller. It can build trust. It can prevent future disappointment and disillusionment and lack of trust.

A wise person once said, "If you would have public confidence, confide in the public." Any other course may work temporarily, then disastrously backfire and build the mistrust that is so common in America today.

* *Iowa State University Press, Revised Edition, 1994.*

Session 1

ELECTRONIC DISSEMINATION OF FEDERAL STATISTICS

Electronic Dissemination of Energy Data and Analysis

Mark Rodekohr, and Henry Weigel

Energy Information Administration

Abstract

This paper provides a discussion of the Energy Information Administration's (EIA) Electronic Dissemination program. This program is designed to deliver EIA energy information, statistics and analysis in the most timely fashion possible using the latest technologies to provide economies to both the provider (the U.S. government) and users of EIA information products. The EIA is responsible for producing the nation's statistics and analysis on energy production, consumption, imports and prices. These statistics are often available by month, year, region or other disaggregation. Just one of our databases (the Oil and Gas Resource Information Database) contains 50 MB of energy data. The delivery of these data in an efficient and timely manner is critical to EIA.

The paper is organized around several sections which describe:

- Electronic Dissemination Goals and Strategies
- Dissemination Techniques
- EIA's Experience with Electronic Dissemenation Methods
- Possibilities for the Future
- Electronic Dissemination Policy Issues

Strategy, techniques usage statistics and other policy related factors are discussed in some detail.

Goals and Strategy

The goal of electronic dissemination is to maximize the use and ease of use of EIA's energy data and analysis products. Minimizing the cost to both users and producers of these products goes a long way to achieving this objective. The rapid improvement in electronic publishing technologies is lowering costs to both users and producers but increases the costs and frustration of technology choice.

Electronic dissemination strategy involves matching appropriate technologies with users. For example some users do not have World Wide Web access but do have access to fax machines. This implies that products should be made available in fax formats where appropriate. In addition there are classes of users who again do not have Internet access but who do have modems, this suggests that some type of bulletin board system may be a useful addition to the electronic dissemination technologies in use at EIA.

The rapid improvement in electronic dissemination techniques has clearly changed information dissemination techniques. The old paradigm (excluding data collection and processing) involved:

- Writing reports in long hand,
- Having a secretary type the report,
- Using a desktop publishing system to format and produce print quality masters,
- Print reports,
- Mail reports,
- Answer questions on the phone and by hand fax out selected tables.

The new paradigm involves:

- Analysts type and largely format reports,
- Publish electronically,
- Publish traditional paper products where needed,
- Customize information reports using a variety of technologies to meet user needs.

Clearly the time involved with the old and new paradigms in meeting customer needs is significantly different. From user surveys conducted at the Energy Information Administration **the number one user concern continues to be timeliness.** Electronic dissemination strategies can and are being used to address this concern. Simply put, customers want more timely data more than ever and are consistently demanding that these needs be met.

Dissemination Techniques

In this section the advantages, disadvantages and primary customer for each electronic dissemination technique is discussed in turn.

Internet Servers	
Advantages	**Disadvantages**
Distribution is instantaneous	Users must be fairly high tech.
10 - 30 million users	Moving large amounts of data is time consuming
Customer base growing rapidly	Can be slow during peak hours
Fairly low marginal cost	

This description of the Internet clearly shows some of the reasons for its rapidly increasing popularity, the point and click ease of use and almost instantaneous distribution characteristics make it one of the more popular technologies.

Bulletin Board Server	
Advantages	**Disadvantages**
One of the first electronic dissemination technologies	Still requires a modem
Has an existing customer base of 1,800 users who have been using this system for years	Slower than the Internet for many users
Does not require an Internet connection	Cannot use some of the more user friendly Internet applications
Can be replaced with a dialup ftp server to reduce costs	

This system has been in use for several years and many customers have constructed custom scripts to retrieve and format data to meet their special needs. It is likely that the use of this technology will decline over time and be replaced with dialup access to the Internet ftp server.

Applications Server	
Advantages	**Disadvantages**
Meets specific customer needs by providing only those data needed	Development of new applications can be very expensive
Does not require the downloading and installation of large software packages	Somewhat higher marginal cost than some of the other Internet technologies
User costs low	

The Applications Server (or the Interactive Query System) allows users to specify by clicking on a menu exactly what data elements they desire. Once the menu has been specified the request is transmitted to the applications server which then, using a query system constructs tables of the requested data and then returns the data to the user. This technology allows users to view only those data that they are interested in. Most other database technologies require users to download entire databases and then using either downloadable run time engines, or client based software construct queries to produce the needed data aggregations. Clearly the applications server technology is less demanding from a users point of view and therefore superior to providing specific data needs. The downside to use of this technology is the cost to the producer organization. It is somewhat more costly but widely used by some Federal data providers.

Listserv	
Advantages	**Disadvantages**
Very low marginal cost	Can only be used for short products
Ability to provide interested readers with short report very quickly	
Allows interested customers to subscribe	

The listserv technology has turned out to be a very popular electronic feature. At the Energy Information Administration's Web site users are offered a menu of listserv products, such as press releases, data summaries etc., to which they can subscribe. Once they have subscribed the products are sent automatically to their mailboxes. They can subscribe to as many or as few products as they wish. At present there are about 18 separate listserv products with a total distribution of about 3000 mailings per month.

Fax on Demand/Broadcast Fax	
Advantage	**Disadvantages**
Only requires that users have a fax machine	User pays phone charges
Menu systems allows users to get a list of all products	Only a limited amount of information can be transmitted in paper formats
Good for the occasional user	Transmission time is slow

This technology is especially suited to users with more limited access to technology. Since fax machines are very commonly used throughout the world they enable users to receive small amounts of information in a fairly timely fashion. It should be noted that most broadcast fax systems work best in the late evenings when phone lines are not usually tied up however this produces greater delays in receiving information.

CD-ROM	
Advantages	**Disadvantages**
Reaches a broad customer group since over 10 million cd readers are in use	Relies on the mail system
Libraries are very attracted to this medium since it saves shelf space	Will not reach low tech users
Low production cost	Users still must print to get hard copy
Holds up to 650 MB of information	
Easy to search	

The search ability and high volume make this medium ideal for a variety of users. Libraries find that this is ideal medium since rows and rows of shelf space can be saved. At the same time it is ideal for archival purposes since it does not degrade over time. The ease of conducting searches makes it ideal for users who are interested in very specific subjects. It is also the only medium that can deal well with large databases since space is not a limiting factor. As stated earlier one EIA database contains about 50 MB of data which is too much to download over most web and ftp sites.

The Electronic Dissemination Experience at the Energy Information Administration

The experience of using electronic dissemination techniques at the Energy Information Administration has been very instructive. While many things have gone as expected there have been several surprises. The table shown below presents some aggregate statistics on the use of electronic dissemination products on a monthly basis starting in January 1996. While some of the products have been available before this date this is when the EIA product line was largely established.

Electronic Product Statistics			
Month	Web[1]	Ftp[2]	Listserv[3]
Jan	NA	NA	NA
Feb	NA	NA	NA
March	12,723	81,026	NA
April	15,732	101,440	2,500
May	19,711	113,987	3,050
June	14,280	93,118	3,750
July	12,161	96,385	NA
August	12.498	100,898	NA
September	16,657	122,981	6,100
October	22,892	175,399	NA

These data show that starting in January traffic on the Web and Ftp sites typical grew at a rate of 20 to 25 percent per month except during the summer. Several factors caused the decline in June, these included a large increase in May figures due to the rapid increase in gasoline prices which by June had largely abated. This increase in interest increased the figures for May dramatically but was a one time event. Other significant factors included the end of the school year and the summer vacation season which tends to reduce commercial

---

[1]Number of unique daily users per month.

[2]Number of file requests from Web, Ftp and EPUB per month.

[3]Number of individuals subscribing to E-mail products, not necessarily unique.

use of the Internet site. Starting in September the Internet traffic continued at a higher rate. In October it jumped to about 40 percent. It is interesting to note that the listserv traffic continues to increase by 15 to 25 percent per month over the period following the establishment of this product.

An examination of the detailed logs produced by the Internet server turns up a number of other interesting observations including:

- Over 60 percent of our users are from U.S. commercial firms and universities.

- About 15 percent of users are from other Federal agencies who tend to look at a larger number of files than commercial users.

- About 15 percent of users generate about 70 percent of the hits in other words the repeat customers are the heaviest users.

- About 10 percent of our users are from foreign sources.

Not shown in these statistics are the number of customers for the CD-ROM. It turns out that these users are much more limited in number than other users. By June 1996 about 1,000 CD's were being distributed every quarter. A large number of these users were represented by libraries who have expressed a great interest in this medium. Since library distributions can be looked at by a large number of users it is difficult to get an accurate measure of the number of total number of users of this technology.

The Future of Electronic Dissemination Techniques

Rapid changes in technology make the prediction of future dissemination techniques very difficult. However, there are some promising technologies that are being developed. One of the first that comes to mind is the use of a corporate database, structured using a database management system such as Oracle, which could permit a common user interface that would generate databases, tables and even permit the use of query systems that would operate across all data series maintained by a statistical agency. Another concerns client-server software that allows users to seamlessly get database updates and additions given the existence of a predefined database structure. The advantage of this approach would be that users would no longer have to download entire databases but only the database outline. Given this outline they could fill the data they need over the Internet, as desired. This could greatly reduce download time. In addition CD technologies already exist which allow users with Internet connections to get updates of CD databases over the Internet. Technologies such as Amber could allow users to download pdf files a page at a time thereby making the pdf file a much more usable technology and at the same time greatly reduce the need for the construction of time consuming html file formats.

Policy Issues

There are many policy issues that remain to be resolved with regard to electronic dissemination. A few include:

- Choice of file formats. Private sector companies are very sensitive to the choice of file formats because these tend to be proprietary in nature and can have a impact on future sales. Agencies must be aware of these impacts on production and usage costs. For example the choice of one spreadsheet format over another can greatly affect the kind of software users are required to purchase. Another example relates to the choice of document formats such as WordPerfect, pdf etc. which are tied to specific company products. Since no generic format exists these policy decisions could be difficult.

- Client-Server software. Again, this can impact company sales and user requirements and become a sensitive issue.

- Browser support. Most web sites will support many browsers but are developed to be easiest to use with one specific browser. Companies are already complaining of the choices of Federal agencies in this area and could be more vocal in the future. From a resource point of view developing web sites that work well across a wide range of browsers can be very costly.

- Fees. Some Federal Agencies are already charging users for access to specialized statistics. The pressure to expand or start this practice in the future could become greater. Given the long history of providing low cost data to users this transition could pose some very difficult policy choices.

COMMENTS

Robyn Levine
Congressional Research Service
Library of Congress

Note: The views contained herein are those of the author and not necessarily those of Congressional Research Service or the Library of Congress.

Session 1: Electronic Dissemination of Federal Statistics-- Census and Energy Data

Note: My remarks were directed to two papers I had received beforehand, although the Census Bureau did not present the paper it provided me and gave a hands-on demonstration of DADS (Data Access and Dissemination System) instead. Thus some of my remarks were not necessarily addressed by the Census speakers/presenters.

Obviously, there are myriad benefits a reference librarian gains from the electronic dissemination of agency reports and data, particularly if the data tables can be user-defined. We can get the information when and where we want and be assured that what we see on the screen is the latest, most up-to-date version of the report or data, reflecting the most recent revision. Both the Energy Information Data System and the Census DADS system take advantage of these inherent benefits of the Web environment.

My particular areas of specialization are transportation and the census, although at CRS librarians are expected to be generalists when we work in our reference centers in the House and Senate office buildings and in our reading rooms. Each is a mini-library with a basic core of reference sources and finding aids. Given the relatively small size of the facilities and the ever-present budget crunch, the Internet environment expands our collections to reflect the expertise and imagination of the librarian. The librarian, thus, has tools heretofore provided only for the specialist.

Because we have various possible places that can supply statistical answers to questions, depending on the question (e.g., via our Inquiry Section, to a member of a particular team or to an analyst in our research divisions, or in a reference center or reading room) by having the data online we can ensure that the same question will be answered with the same data.

While CRS functions much like a special library, we are not in a position to consider payment mechanisms such as user fees. We certainly can appreciate the tough budgetary choices Congress has imposed on the Census Bureau, but Congress has placed similar restraints on us. Thus, we would hope that the policy of waiving fees to Congress will extend to us, as Congress is our sole client.

Another area of concern for government librarians in general, and us at CRS in particular, is the depository library network and its Census collections. The mandate CRS received from Congress is to support Congress not only in its legislative needs, but also in its representational duties. Some of our work is in response to constituent inquiries, supplying information for Members to pass on to their constituents in their districts. In this respect, we view our work as much educational as informational. We in the Congressional Reference Division (CRD) are 43 professional librarians/researchers out of 750 people in the entire Service. CRD answers about 65 per cent of the 340,000 congressional requests that come into the Service annually. We rely on the depository network as a referral; we expect it to be able to provide certain government and congressional publications.

Following the 1990 Census, depositories received CD-ROMs whether or not they had the necessary equipment to make them available to the public. There were still paper products available. The Web environment opens up exciting possibilities for obtaining Census information, but it also opens a Pandora's box as to how depositories will handle the potential costs for any special products created by using the DADS system. The Library of Congress is wrestling with a similar question in enabling patrons to use fee-based electronic systems in the reading rooms. If this is a problem here, it is very likely to be a problem elsewhere. For some depository libraries, their ability to provide service could be limited by the electronic dissemination of publications by government agencies. Thus far, EIA is attempting to offer paper products as well as the electronic ones, including Faxed documents, and I urge them to remember the depository network when making future decisions concerning data dissemination.

I can see very positive benefits from the DADS system. After the 1990 CD-ROMs became available, it was obvious that the GO software that Census provided had some significant drawbacks, namely in the ability to manipulate the data. We spent considerable in-house resources to convert the CD-ROMs, with SAS-Assist, into a system where we would be able to rank and sort the data to our clients' specifications. Our system is very time-consuming to use and not very user-friendly. Thus, how easy DADS will be in producing custom-tailored tables is of great interest to us. The more user-friendly the better, since we have had numerous occasions to assist Congressional staff directly in their own use of the Internet.

That CRS has a vested interest in the Internet as a warehouse of government information can be seen in our Home Page, which is available to our own staff and to all Congressional offices. We make available our own databases and products, of course, and we enable our patrons to reach agency home pages with direct links. Staffers approach CRS staff in our reference centers for help in using the variety of Internet sites. In our prepared information, we include the URLs to government agencies and associations when relevant, and if we include any products obtained from the Internet, we provide the URL for further information.

In terms of government information, CRS may potentially be the primary intermediary outside the agency congressional liaison offices to assist

congressional users of government statistical databases. In terms of Census data, Congress's need for small-area data is going to be more demanding than in the past. Hopefully, DADS will be able to integrate various data sources by comparable geographic levels, especially smaller geographic entities like congressional districts.

With a devolution of authority for many government programs, federal involvement in data production may be more significant in statistical areas than ever before. Now that Congress has created devolution, it will have an oversight role with the shift in program responsibility to the states. We cannot rely solely on the states and local areas to produce data, as they are hard-pressed financially, and there is no overall standard for comparable data. Thus, for oversight, it will be imperative that Congress has comparable data to assess the impact of programs across state and local areas.

CRS has begun to examine resources and consider options for the information that will be required due to the impact of devolution. We have established a Federal-State Task Force since we believe devolution will be of significant congressional interest in the 105th Congress. We hope that other agencies are looking into how to present their data for very small geographic entities and that there will be interagency cooperation to enable this to happen in as consistent and unified a manner as possible. This likely will be a challenging and difficult situation for all of us, and the cost of our information gathering and production potentially will increase as Congress assumes this oversight role. In the end, the use of certain data may be very specialized and for Congress' eyes only (given inherent confidentiality problems).

While I used the EIA system in anticipation of this meeting, my familiarity with energy-related questions is not as keen as with the census. Thus I tested the EIA system against some typical questions one of my colleagues receives in terms of energy production, consumption, imports and prices, and found it to be easy to use. At the same time it allowed us to customize certain reports. Most of the data, however, is national in scope, with limited state information.

The rest of these comments will be addressed to questions for the Census Bureau that are active concerns of data users at CRS, although EIA may want to consider them also.

1. As a historical tool, how long will data be available electronically? Will this data be archived?

2. How accessible will raw data be on DADS or from EIA, and will it be available across state lines or just on a state-by-state basis?

3. How does the Census Bureau decide which of the printed reports are being discontinued or changed, and are users involved in the decision-making process?

4. Concerning congressional districts, as they are redrawn as a result of court decisions, will new data be issued for the reconfigured districts?

5. Congress wants more and more detailed data for smaller areas such as congressional districts, something the data user community at large is not

particularly interested in. Will the Bureau, and other agencies as well, be prepared to face the potential burden of this need?

6. One suggestion made in the paper I read concerned partnerships with non-governmental parties to conduct joint research with the Census Bureau. This has advantages in being able to trust data supplied by non-governmental sources if in sync with the Census Bureau (e.g., private marketing demographic systems and the formulas employed to updata decennial census data). Where this partnership will be used to find ways to combine various federal agency data sets has far-reaching applications, especially if small-area geography will be available across the board.

But if the information will not be made available for up to 5 years due to the contractual nature of the joint partnership, then the benefits discussed in the outset of these comments, namely the supply of current and timely data, has been negated.

Thank you for allowing me the opportunity to be here this morning.

Session 2

WHAT THE PUBLIC NEEDS TO KNOW ABOUT FEDERAL STATISTICS

Organizational Perspectives and the Agenda of Federal Statistical Agencies: How What We Know Reflects Who We Are

By Daniel Melnick[1]

National Science Foundation

Here is the dilemma federal statisticians face every day in their work.

> When we use numbers to enlighten public [and private] policy, particular perspectives are key to the meaningfulness of assertions. Agency mandates shape perspectives and in turn data collection, aggregation and analysis. The models of data analysis in common use assume these perspectives are or should be uniform. In fact, they are diverse. How can we create general purpose statistics if the goals we serve are particular to the questions put by our constituents? Yet, without this infrastructure, individual studies would be impossible.

Some Examples

Consider three examples of apparently flawed procedures that might be criticized from a statistical perspective but have considerable validity in the context of the data collectors' expectations and their relationships with respondents and constituents.

- A Member of Congress sends out a questionnaire to all of the people living in the district. It is a simple flyer addressed to occupant and it contains questions which express the Member's views as much as ask for opinions. 15,000 constituents or about five percent of the adults in the district return the form. The Member touts the results and seems to pay a great deal of attention to them.

- A group of mental health researchers designs a survey to measure the prevalence of mental disorders in the general population. Because they are concerned about the need to accurately measure the respondent's condition, they recruit psychiatry students and staff from five leading tertiary care teaching hospitals with strong departments of psychiatry. In each catchment area near these institutions, they draw a probability

---

[1] Daniel Melnick, Ph.D. is a Senior Advisor, Research Methods at the National Science Foundation. The views expressed in this paper are his own and do not represent an opinion of the National Science Foundation.

sample of the population and administer a detailed diagnostic protocol. Subsequently, they use demographic data from the decennial census to extrapolate national estimates. When the next census is published, they update the extrapolation and report on expected changes in the prevalence of mental and addictive disorders. These figures become part of the calculations for a major national health care initiative in an attempt to estimate the number of people who will demand mental health services under the new plan.

- A group of substance abuse researchers wants to estimate the number of people needing drug abuse treatment, the number receiving it and the treatment deficit. They are particularly concerned with measuring the changes that followed increased funding for treatment. Drawing on an annual survey of a probability sample of US households, they create a composite indicator of treatment need which the psychiatrists in the second example would probably not trust. They calibrate these results against administrative data to try to adjust for reporting biases. They compare the treatment deficits of population groups. Yearly change can be reported to satisfy OMB and the White House Drug Policy office.

It is easy to show why each of these procedures is wrong in the abstract:

- The Member of Congress obviously does not have a representative sample of the district and the questionnaire probably did not provide a reading of the "true" opinions of the respondents—i.e. the views they held before they received the Members mailing.
- The psychiatrists don't really have a nationally representative sample either. They may also be missing variation between places with and without tertiary care institutions. When the data are updated with new census findings, the researchers assume that there is no change in the relationship between the demographic characteristics of study participants and mental health conditions.
- The substance abuse researchers make the grand assumption that the limited data they have from their survey really measures the need for treatment. They also assume that respondents have been frank about very sensitive personal matters. They justify this by the need to measure the amount of change from one year to the next and between groups.

While its easy to use an abstract standard to judge these studies, its wrong to do it. Each of these procedures must be understood in terms of the underlying assumptions and expectations of the data

collectors and their mandate. As long as they remain within the limits of their own framework, the results have meaning. When they depart from it, they experience difficulties.

- The Member of Congress uses the questionnaire to provide constituents with a way to express their views. While information about a representative sample would be useful, from the Member's perspective the mailing provides a way to measure the number of people who are deeply concerned about issues. The questionnaire is a crude kind of behavioral test—those who return it tend to be more exercised about the issues included in the questionnaire than those who do not. For example, knowing that there are at least 300 people who are deeply troubled about the siting of a government office helps the Member fulfill his duty to represent his constituents.

- The psychiatrists believe mental disorders result from long lasting personality or biological factors which are not likely to vary much from place to place. Because they assume that the key to understanding the prevalence of these conditions lies in expert diagnosis, they believe its better to compromise on sampling and timeliness to insure accurate measurements. When they have to respond to demands for updated information, it seems reasonable to adjust decade old results with the new census data. A subsequent National survey using probability sampling techniques uses procedures that are calibrated to the original 5 site study.

- The substance abuse researchers focus on tracking change and geographic variation. Their constituents believe that drug abuse is the product of overt infusion of pernicious activity. Therefore, they demand indicators that can track its progress yearly [if not semiannually, quarterly or monthly] and report results for cities and states. The demands to justify expenditures in annually submitted budgets force them to estimate the effect of treatment initiatives. Criminal justice authorities also use this data to plan their strategies. They compare the reported prevalence to estimates of the supply of drugs which is equally if not more uncertain.

The psychiatrists and the substance abuse researchers both report on drug abuse, but their different perspectives lead them to different conclusions. Policy makers sometimes ask which of these is "correct", but in fact they both appropriately answer different questions and careful analysis shows they can be reconciled.

We may think that federal statistical agencies are shielded from these kinds of concerns. After all, many of them are organizationally distinct from mission agencies. But, each data collection and statistical reporting organization carries its own perspectives to the field, in the office and ultimately in its data tapes and reports. Each constituent evaluates data in terms of their particular inquiry--but statistical agencies cannot adopt a narrow view or they risk eroding their main role as providing the infrastructure for many particular analyses. This dilemma is at the heart of many disputes. For example, consider how BLS's wage statistics program must meet the needs of federal employee wage adjustments while still providing general wage statistics.

Organizations shape the content, quality and meaning of federal data series. Their expectations, relationships and mandates shape the view provided by seemingly unbiased factual presentations. To account for these factors, we need to go beyond the definition of uncertainty used by statisticians in which total error equals sampling error plus non-sampling error.

We can draw upon a large literature about organizations which has recently been enhanced by studies of the decision and cognitive sciences. These perspectives have been applied to the interactions between survey interviewers and respondents.[2] But, we need to assess their implications for the entire system of collecting and reporting statistics. These factors are not exclusively found in federal statistical agencies, they are common to every organization that collects and reports information. But, today we are focused on the federal effort.

Expectations Shape the Outcome

What we expect shapes what we report, because it limits what we look for and thereby what we find. Different approaches to compiling information grow out of the different questions policy makers and the public ask. It helps to think about the stages at which information is used, and what it is used for. Whatever else they do, statistical agencies provide a factual foundation for the "authoritative allocation" of resources.[3] Relationships with users and constituents provide the context for this work and shape its impact on how public policy is grounded in facts.

Decision makers and the public use facts to:

- **monitor conditions,** by reporting on trends, tracking changes, or comparing groups,
- **plan action or propose responses,** in legislation, government rules or corporate policies,
- **implement action strategies,** [For example, companies use them to decide how to shape their marketing messages, candidates use them to decide where to put scarce campaign resources and legislators use them to fashion allocation formulas for block grants.]

---

[2] For example see Norbert Schwarz and Seymour Sudman, editors, Answering questions : methodology for determining cognitive and communicative processes in survey research /.1st ed. San Francisco : Jossey-Bass Publishers, c1996. xi, 469 p. : ill. ; 24 cm. Includes bibliographical references (p. 403-441) and index.

[3] David Easton, A systems analysis of political life. New York, Wiley [1965] xvi, 507 p has an elaborate presentation of the implications of this phrase.

- **evaluate the results** of those strategies, by trying to judge what actually happened compared to what would have occurred if they had not acted, *and*
- **weigh evidence** during the adjudication of disputes arising from these actions. [For example, when experts report about the results of psychological tests or the number of children from different groups admitted to a school.]

There seems to be a temporal logic to these activities, but it is a constructed logic. In fact, frequently these activities occur out of order. Thus, their linear appearance is just that, a neat way to describe them, but far from what actually happens.

Modes of Data Collection

The relationship between statistical agencies and their sources might best be summarized by borrowing from a classical categorization that divides data into those coming from administrative records, censuses, periodic surveys, focused or single project surveys, and experiments.

- **Administrative records** are the result of corporate activity either in private companies, community organizations [like universities or schools] or government agencies [like the IRS];[4] For example, when goods sold in a store, loans paid to the bank, votes cast, passports issued, taxes paid or unpaid, tickets issued for a performance, books lent in a library, grant applicants processed, and payments made to mothers with children, transaction records are a part of the administrative process.
- **Censuses,** although specially organized, are generally so comprehensive that they use many of the same procedures and have many of the same characteristics as Administrative records except that censuses count the stock of people, organizations or property while administrative records track the flow of actions on these elements; [5]
- **Periodic** sample surveys can be much more carefully controlled. They generally employ a permanent and professionally trained survey crew, led by skilled statisticians and analysts to collect the same or very similar information each day, week , month or year.
- **Focused surveys** are specially organized to collect information about a specific topic, population or issue. They use many of the same methods as periodic sample surveys but have greater flexibility to target information needed to answer specific questions. And
- **Experiments** purposefully vary the lives of people or organizations to test the effect that this has on attainment of stated goals.

The match between the questions policy makers and other users want to answer and the particular type of data available is a key factor in the validity of reported results.

---

[4]These have been the stock in trade of researchers for hundreds of years. What differs now is the increasing mining of them for quantitative estimates that are integrated with data from more controlled studies.

[5]Censuses also have a long history, going back to Biblical times. See Anderson, Margo J., 1945-The American census : a social history / Margo J. Anderson. New Haven :Yale University Press, c1988. xiii, 257 p. : ill. ; 25 cm

Disputes about statistical results were focused by the impact that statistical work has on policy choices, limiting leaders options and setting the context for constituent support. Nothing gets the public's attention more than learning that a new report will shift resources or power, when new census figures shift representation from one state to another, new crop figures effect prices, or the latest Consumer Price Index shifts billions of dollars from employers to employees, and the government to its beneficiaries.[6]

This effort is important because:

- Federal data are the foundation for what we know about our country, its people, economy and society. A large part of the information generated in the private sector and social science research conducted by Universities relies on federal data for benchmarks, sampling frames or controls. *And*
- Federal data must suit multiple purposes: providing both the descriptive and analytical information and mechanisms to directly allocate resources. When policies are drafted, statistical results provide key information to calibrate impacts. The data is used to determine if proposals are feasible. Just consider the way the Congressional Budget Office uses federal statistics to cost out legislative proposals which are then scored against goals to see if they comply with the budget resolution. And when the statistical basis for allocations is ambiguous, it is difficult to get the provisions enacted.

Facts, Figures, and War

Statistical indicators are central to our public policy debates. 55 years ago, in October of 1941, President Roosevelt appointed the Librarian of Congress as the head of a Bureau of Facts and Figures.[7] This office which was ultimately absorbed into the Office of War Information, marshaled accurate information--the Facts and Figures-- to counter war rumors. The largest part of its staff was assigned to a statistical operation- the Bureau of Intelligence which monitored US public reactions. Its results enabled Archibald McLeash to have extraordinary influence in the early months of the war-- because he had charts and graphs to back up his assertions. The OFF also monitored and tried to gain control over the reports of US statistical agencies so as to provide

---

[6]Other disputes are summarized in The Decennial census : an analysis and review / prepared for the Subcommittee on Energy, Nuclear Proliferation and Federal Services of the Committee on Governmental Affairs, United States Senate : by the Congressional Research Service, Library of Congress. Washington : U.S. G.P.O. : For sale by the Supt. of Docs., U.S. G.P.O., 1980. ix, 465 p. : ill. ; 24 cm. OTHER NAMES: Melnick, Daniel. United States. Congress. Senate. Committee on Governmental Affairs. Subcommittee on Energy, Nuclear Proliferation and Federal Services. Library of Congress. Congressional Research Service. NOTES: At head of title: 96th Congress, 2d session. Committee print. Authors: Daniel Melnick and others."November 1980." And, Anderson, Margo J., 1945-The American census : a social history / Margo J. Anderson. New Haven :Yale University Press, c1988. xiii, 257 p. : ill. ; 25 cm. Also, Numerous issues have been reviewed in reports issued by the National Research Council's Committee on National Statistics.

[7]This account is based on papers found in boxes 52 and 53 of the Archibald McLeash Papers at the Library of Congress. See. History of the Office of Facts and Figures, typescript HFG 9/23/1943.

needed information to the public while denying intelligence to the enemy. In support of this effort, Library staff went on a 24 hour schedule, so that the information would always be available on short notice. The speed with which this was done and its importance illustrates the role that statistical results play in policy analysis, strategic planning and implementation. [8] By the summer of 1942, these activities were subsumed in the work of the Office of War Information. Analysts applied the results to pressing decisions. The importance of statistical data and the attention paid to it marks a real shift in our view of government action.

Of course, this library mobilization followed several decades during which a revolution in federal statistics was begun.[9] The period after the First World War saw the institutionalization of federal statistical efforts and the introduction [into the 1930's and 1940's] of modern sampling and statistical controls. At the same time, the changed role of the American government and the altered economic and social system brought on by the depression and World War II intensified demands for statistical information. For example, the rise of the radio and television created a whole new kind of demand for numbers to show the impact [and monetary worth] of totally new information products. At its base, each new private statistical service [like the Nielsen ratings] rested on federal statistics which provided its foundation.

Rhetoric and Facts

Most of us recognize that there are different versions of the facts depending on our points of view. Sometimes, we easily see that an apparently factual presentation is rhetoric subtly or not so subtly expressing the policy preferences of analysts or the decision makers who use their work. Yet, these purposeful attempts to present facts so as to make an argument are not the most difficult or troublesome instances of distortion. As long as the purposes are clearly set forth, these arguments are a legitimate part of the policy dialog. For given the constructed nature of facts, there is no escaping the impact of opinion on fact. [But of course we can try to present different views of the same circumstances in an attempt to balance the argument.:-- the sort of pro/con analysis developed into a high art at the Congressional Research Service.]

In fact, a conscious attempt to mold the public's view of the world by shaping the statistics we report is easier to deal with than the more subtle impact of organizational expectations. In our open society, the biases are often apparent or can be easily identified. Opposing analysts are quick to articulate them. Even without a planned attempt to shape statistical reports there are numerous disputes about the fairness and accuracy of reported results. If there were none, we would wonder if the reports were important. In fact, the very controversies generated by statistical reports show their importance for public and private decisions.

---

[8] Annual reports of the Librarian of congress detail this work. It is described in Jean M. Converse. Survey Research in the United States. Roots and Emergence 1890-1960. University of California Press. Berkeley. 1987 pp. 171-172.

[9] Duncan, Joseph W. Revolution in United States Government statistics, 1926-1976 / written by Joseph W. Duncan and William C. Shelton. [Washington] : U.S. Dept. of Commerce, Office of Federal Statistical Policy and Standards : for sale by the Supt. of Docs., U.S. Govt. Print. Off., 1978. ix, 257 p. Shelton, William Chastain, 1916- joint author. See also Converse

How we view statistical inference is one of the most important expectations shaping our assessment of these facts.

Organizational Structure and Inference

It is easy to think about inference as if the projects generating statistics were investigations conducted by a single researcher or a group focused on a defined set of issues. They formulate an hypothesis and design a procedure to test it including the data collection procedure [sample, questionnaire, interview, etc.] and a specific inferential plan. The statistical tests we use work best when part of this kind of process.

Statistical agencies, on the other hand, collect data for general purposes. Even statistical offices in mission oriented agencies frequently have mandates to report on conditions rather than test hypotheses. Analysts using this data impose post hoc hypotheses and try to test them. The agencies operate with a distinct mandate that shapes their expectations and defines the conceptual maps that guide their work. These cultures give purpose to the efforts of the agency by showing how its work fits into a broader pattern of action.

Before the growth of the counting industry, it was sufficient to address the problem of analysis by reflecting on the dual nature of facts: that they are defined by our ideas and refract our environment. What we know arises from our own expectations about what we might know. At the same time, what we know reflects learned definitions of what there is to know. Thus, facts are a product of individual innovation and learned behaviors. Language sets up the definitions and names the feelings that allow individuals to see and not to see aspects of the world. Without such filtering, the mind cannot comprehend the unlimited impressions thrown at it. Facts are therefore created and arise from the fictions inherent in the learned environment.

As anthropologists have recently shown[10], intelligence--the skill of manipulating information and relating it to activity--may have grown from the increasing dependence and communication between primates and ultimately men and women. Thus, the apparent innate skill of knowing arises from cooperation and is learned. It lies exposed as yet another cultural artifact-- perhaps the master artifact of society.

As we try to use modern procedures to collect, compile, analyze and report facts, the organizations we work in [such as the Census Bureau, the Income tax department, the social entitlement office, the polling firm, the market research department, the academic survey research center, or the statistical office of a large company] become dominant instruments of fact finding. It is as if our fictions [the meaning we gave to sense impressions that make them facts] have taken on a corporeal existence and are manifest in these Bureaus and Institutes. We should recognize that this continues a tradition of institutions which came before our enchantment with counting. Yet, the increasing quantitative expression of facts has radically altered the way we think.

---

[10]For example see Esther N. Goody ed. Social Intelligence and Interaction. Expressions and Implications of the Social Bias in Human Intelligence. Cambridge University Press. 1995.

To be sure, the distinction is a matter of degree. The family, guild, state, church, library and legislature surely also embody such domains of meaning. Yet, with the rise of the professional fact finders, we see a newly ordained fact-space in which dogma and values are not supposed to dominate. These new bureaucracies certify the veracity of stated realities while they also specify the limits to our confidence in them. Meanwhile, they legitimate what we know by giving it a quantitative face that adds an appearance of definiteness even as it allows analysts to indicate the probable degree of associated uncertainty. The fact finders treat events as mundane and emotionless, but yet, their product provides bold new banners--sentinels of modern life. And as such, they become the vortex for heated debates. [11]

Consider the question of accuracy: how do we know that our reports accurately and appropriately reflect the truth? For many of the most interesting questions, no gold standard of veracity can be found. We replace it with more relative yardsticks testing the adequacy of what we do know. Thus, truth is estimated by comparing the answers with the questions--and consistent responses often stand in for valid ones. Validation is predicated on the expectation that analysts pose problems, they observe phenomena focused on these issues, and the results approximate an answer. Statisticians step in to help us measure the limits of the outcome.

Yet, as we periodically collect large scale surveys and censuses or compile data from the records of bureaucratic files, frequently the hypotheses are only implicit-- and often the only way to collect data is to allow different participants to adhere to their own view of the purposes.

All of what we think we know about society, politics and its environment -- the most basic facts -- are refracted through these structures. This means that, we cannot know about the uncertainty of our results -- the variances from some absolute truth-- without understanding the effect of these organizational mechanisms on what we think we know.

To do this, we must look beyond the conventional statistical tests of variance and the impact of measurement as well, to the systematic effects of these social structures. The lead indicators that are the artifacts of our age drift off into this uncharted space. Analysts struggle to extract meaning from this constructed environment where bits of data are transformation into information which becomes fact.

Each bit of data ultimately relates to our perception of the subjective relationship. And each in turn is both observation and activity. By focusing attention on a particular aspect of social relations, we take a stand on the issues. Yet, we agree to suspend our acknowledgment of this ultimate subjectivity--attempting to cast what we want to say in objective fact.

---

[11] Herbert Simon. Administrative Behavior . The Free Press. New York 1945. See especially Chapter III Fact and Value in Decision-making.

Implications

Although it may seem convenient to treat the organizational perspectives discussed above as biases to be identified or controlled, they are as much a part of statistical work as questionnaires, forms, samples or formulaic tests. They provide the purpose and direction for much of the work. Inferences are best judged in light of their contribution to these purposes. There is no escaping that who we are greatly influences what we know.

IMMIGRATION STATISTICS: NO LONGER NEGLECTED, BUT STILL INADEQUATE*

by
Jeffrey S. Passel†
Urban Institute
2100 M St., NW
Washington, D.C. 20037

Introduction

Immigration remains a very contentious area of public policy, even after a total overhaul of U.S. immigration policy between 1980 and 1990. The 104th Congress passed three significant pieces of legislation relating to immigration and narrowly failed to pass a fourth. The Antiterrorism and Effective Death Penalty Act of 1996 contained a number of provisions relating to immigration, focussing mainly on easing removal of aliens from the United States. The law limited appeal and judicial review for persons seeking asylum and expanded criteria for deportation to include relatively minor crimes. Welfare reform legislation — the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 — included major modifications to immigrant and immigration policy. Two notable legal changes were severe limits on legal non-citizens' access to a range of public assistance programs and making the sponsor's affidavit of support legally enforceable by the government and the sponsored immigrant.

The final law passed, the Illegal Immigration Reform and Immigrant Responsibility Act of 1996, represents a significant alteration of immigration policy. The Act increased border enforcement by doubling the size of the Border Patrol and strengthening the physical barriers at the Mexican border. It also reinforced the limits to review, appeal, and access to public benefits in the previous two laws. The law also criminalizes many activities related to illegal immigration

---

* Presented at the Seminar on Statistical Methodology in the Public Service: A Conference Dedicated to the Memory of Maria Gonzalez, sponsored by the Council of Professional Associations on Federal Statistics, Bethesda, Maryland, November 12–13, 1996.

† Jeffrey S. Passel is Director of the Urban Institute's Program for Research on Immigration Policy, a program of public policy research supported principally by The Ford Foundation. The views expressed are attributable to the author and do not necessarily reflect those of other staff members, officers, trustees of the Urban Institute, The Ford Foundation, or any organizations that provide financial support to the Program.

(e.g., smuggling and document fraud) and sets up pilot programs to test employment verification systems.

In an environment such as this, with numerous policy changes and legislative considerations, one would assume that a great deal is known about immigration, immigrants, impacts, and behaviors. Unfortunately, many basic facts about immigration and immigrants remain unknown or, at best, based on estimates of unknown or questionable quality, often because of serious data deficiencies. Detailed studies of immigration impacts and immigrant characteristics face similar limitations. The result is often "dueling studies" in which even reputable researchers reach diametrically opposite conclusions. Partly as a consequence of data deficiencies, development of new immigration and immigrant policies has only occasionally been the result of measured analysis of immigration phenomena and considered deliberation of empirical research.

At the heart of these difficulties over alternative interpretations lies the inadequacy of data on immigration and immigrants, principally from federal sources. Two studies by the National Academy of Sciences (Levine et al. 1985; Edmonston 1996) have described various barriers to producing needed data and prescribed potential solutions. Both of these studies led to some improvements in immigration statistics. However, notwithstanding the significant progress of the last decade, immigration data are still woefully inadequate for addressing a number of key areas of both immigration and immigrant policy.

Organizational and conceptual obstacles continue to plague analysts and policymakers attempting to deal with immigration issues. No agency or office is in charge, overseeing the planning, collection, and production of data on immigration and immigrants. Statistics and statistical analysis at INS still do not have a central role in the agency's mission; they remain a stepchild of administrative processes. The same is true of many other agencies as well, including the Bureau of the Census. Nonetheless, a number of new data sources are now available, with yet more coming on line soon. However, we still cannot answer some very basic questions, such as: "How many people come to the United States each year?", "How many illegal aliens live in the country?" or, even more basic, "How many legal aliens and naturalized citizens live in the country (and each state)?" Finally, there are so many areas where experts' opinions are at odds in large

part because we have not developed a common framework for analysis and discussion, nor the appropriate data to address the pressing questions.

This paper covers four main areas of concern. First, what data do we need on immigration and immigrants? What would we like to know? Second, what data do we have? What is being planned and is likely to be available? What needs remain? Throughout these discussions the focus is on some basic measurement issues, largely dealing with the demography of immigration. Notably, how well are the concepts measured? What deficiencies and gaps are there? Finally, the paper closes with some proposals for reform and change.

Data Needs and Analytic Issues

In assessing our data needs, we must examine linkages among data, data users, policy formulation, and policy implementation. Specifically, what data and analyses are needed to *formulate* policy. We also need to examine how the policies are *implemented* and to *evaluate* both the implementation and the policy outcomes. For all of these purposes, we need to keep in mind that we must not only investigate *immigration policy* (i.e., the numbers and types of persons admitted to or excluded from the country), but also and *immigrant policy* (i.e., those policies and programs for dealing with immigrants after they arrive in the country — integration issues — and the impacts of the immigrants).

Impacts of immigration across a broad spectrum of realms are either poorly understood or remain subjects of contention. The bases for admission of immigrants into the country are principally: family unification, employment needs, and humanitarian concerns. The labor market impacts of not only the employment-based immigrants, but also other legal immigrants and illegal immigrants represent the major arena of competition between immigrants and natives; the purported labor market impacts are frequently cited in policy debates. The research in this area has been generally inconclusive. Case studies and anecdotal evidence indicate the possibility of significant displacement of low-skilled natives. However, macro studies do not find support for such displacement effects, nor do they find substantial wage effects (Fix and Passel 1994). Some researchers (Borjas 1996) have argued that such effects must exist, but that the data and

measurement techniques available are simply not sufficient to detect them. Clearly more and better data are needed.

While the issue of economic competition between immigrants and natives may eventually be settled, there are more subtle impacts of immigration that continue to be even more difficult to discern. The availability of immigrant labor may affect the mechanization and transformation of entire sectors, such as agriculture. Interactions between trade, capital flows, and the migration of labor also require better data and methods to be measured accurately.

The impact of immigrants on the public fisc is another area where more and better data are sorely needed. While some researchers have found an overall fiscal surplus from immigrants (Passel and Clark 1994), others found significantly more costs than revenues (Huddle 1993). There does, however, seem to be general agreement that negative fiscal impacts are felt at state and local areas, driven principally by education costs, while revenue streams flow to the federal level. Not only are data generally inadequate for careful and complete assessments of immigrant (and native) impacts on the balance of public costs and revenues, but the proper analytic framework has yet to be developed and agreed upon. A National Academy of Sciences panel is currently addressing these issues.

One of the difficulties in fully assessing the fiscal impacts of immigrants is the inadequacy of methods for measuring indirect economic benefits derived from immigration. Again, anecdotal evidence abounds across the country for the positive impacts of immigrants in revitalizing once deteriorating neighborhoods. Similarly, immigrant businesses are known to be the principal driving factor behind economic activity in some neighborhoods and to account for significant amounts of tax revenues. However, none of the available fiscal impact studies takes these factors into account.

While analytic and policy issues abound in the immigration field, even more fundamental measurement issues need to be addressed. We simply do not have accurate measures of either the stock or flow of immigrant populations. Such basic questions need better answers and, in many cases, the answers to the larger policy questions raised above must be predicated on having basic demographic information.

Immigrant Numbers

The passage of welfare reform in August 1996 has brought into focus some of the basic deficiencies of the nation's immigration data systems. This law limits access to a range of benefits for many categories of non-citizens, some of whom had formerly been eligible and receiving benefits. For example, legal permanent residents who have not worked 40 quarters in the United States and refugees who have been in the U.S. for more than 5 years are no longer eligible; naturalized citizens remain eligible. In addition, whereas eligibility rules governing aliens (legal and undocumented), had been the exclusive province of the federal government, the welfare reform law devolved responsibility for setting many of these eligibility rules, such as those covering undocumented aliens, to the states. Since the financial responsibility for providing services to the groups no longer eligible for federally-provided coverage was also devolved to states and localities, a number of new parties became interested in the numbers of naturalized citizens and aliens of various types, and in rates of naturalization; moreover, such data are needed for states and even smaller governmental units. Many of the population numbers needed are only available as rough estimates or simply do not exist.

There are five main immigrant populations of interest to most observers. In roughly decreasing order of size, they are:

Legal permanent residents;
Naturalized citizens;
Undocumented aliens;
Refugees, asylees, and parolees; and
Nonimmigrant residents.

In addition, for some purposes, many want data for subgroups of these major groups — family-sponsored immigrants, employment immigrants, foreign students (a category of non-immigrants) to name just a few. Data are needed on the numbers in each category, the inflows and outflows each year (or at least the annual net change), and characteristics of the particular individuals. Interestingly enough, the only one of the five groups for which there are "official" estimates of size and annual net change is the one most would characterize as the hardest to measure — undocumented aliens.

There are three major impediments to obtaining data (or estimates) on the sizes of the various immigrant populations. First, the data we actually do get from INS relate to the numbers of people attaining the legal status each year, not the numbers actually entering the country. Thus, many persons becoming legal permanent resident aliens each year are currently living in the United States, for many years in some cases. The difference between new entrants and the total number can be quite substantial in some years, with new entrants generally running about 60 percent the total. However, in some years the difference can be so large as to be totally misleading. For example, in fiscal year 1991, official INS figures indicated about 1.8 million persons attained legal permanent resident status. This figure was touted, in the press and by anti-immigration advocates, as being the largest annual immigration in the nation's history. About 1.2 million of these "new" immigrants, however, were formerly undocumented aliens, most of whom had been living in the country for more than 10 years. The number of non-legalization legal permanent residents that year was approximately 600,000 or the same number as the year before; the number of true new entrants was even smaller.

A second major deficiency is the lack of any hard data on departures from the United States. The most recent estimates available cover emigration of the foreign-born population during the 1980s (Ahmed and Robinson 1994). These data suffer from a number of known problems. There is little detail available on the emigrants — only age, sex, and country of birth. Information of the emigrants' legal status and socioeconomic characteristics is completely absent. The third significant impediment to obtaining good demographic data is that there is virtually no information on movement of immigrants between immigration categories or within the United States.

Data Sources

There have been major improvements in the amount of data on immigration and immigrants available in recent years. There are more sources, more topics covered, and more types of data. New data sources will be coming on-line soon. Nonetheless, there remain significant shortfalls and deficiencies. The major data gaps will continue to be a need for better, more up-to-date information on legal status and more longitudinal data. Government data sources remain the most extensive, but nongovernment surveys often provide richer data.

Immigration and Naturalization Service. The Immigration and Naturalization Service naturally remains the principal data source on in-flows of various types of immigrants. However, the INS is not organized as a statistical agency, even in its data-gathering divisions. One consequence of this organizational perspective is that administratively-based categories dominate data collection efforts. Counts of immigrants are based solely on official categories. Thus, it is very difficult to derive a figure for the demographically-based concept of *the number of people moving into the country each year* (Levine et al. 1985). The focus on administrative needs means that only limited information is available on the demographic and socioeconomic characteristics of the immigrants. Strangely, however, there is little linkage across data sets even though many administrative needs could be served by linked data sets. For example, little information is available on sponsors of immigrants.

The INS' focus on administrative data means that little analytic work related to immigration or integration policy is done in the agency. Even the analyses done in the agency tend to be idiosyncratic rather than institutionalized. The INS estimates of undocumented immigration serve as an example of both analytic work done at the agency and some of the institutional issues raised by these efforts. Generally, the analytic estimates of undocumented aliens released by the agency are of high quality (Warren 1994; INS 1997), use advanced methodological techniques, and respond to criticism and comments (e.g., Woodrow-Lafield no date; GAO 1995). However, a detailed description of the methodology employed, such as generally supplied by other statistical agencies, is lacking. The review and release procedures tend to be somewhat *ad hoc*. INS did convene an expert panel to review the latest estimates, but the release provides virtually no information on the statistical properties of the estimates. Furthermore, a lack of institutional commitment is quite apparent. The estimates are the product of a single researcher working largely without support staff and the principal data system required to measure visa overstayers has been allowed to fall into a state of disrepair and neglect.[1] In fact,

[1] Warren's estimates make significant use of matched arrival and departure forms from the Non-Immigrant Information System. These data may no longer be available on a regular basis. A similar event occurred in the early 1980s. Warren and Passel (1987) made extensive use of the Alien Registration system to develop estimates of legal residents and of undocumented aliens included in the 1980 Census. The collection of Alien Registration data was halted in 1982, shortly before Warren and Passel first released their estimates in 1983.

Warren's estimates became the "official" INS estimates largely because the new Commissioner (Meissner) knew of his work and started citing them, rather than though a concerted agency effort to measure this politically relevant, clandestine migration.

There are some other examples of analytic, policy-related research done at the INS, but the list is not long. Analyses of naturalizations, including cohort-based measures, are presented in recent *Statistical Yearbooks* (INS various dates). However, even this work does not go much beyond simply reporting rates of naturalization; there is little analysis of determinants of naturalization or attempts to correct the rates for emigration and mortality. Another recent example is an INS analysis of records of immigrant sponsors.

Current Population Survey. Within the last few years, there has been a substantial expansion of data on the foreign-born population available from national surveys. There is now more data available on socioeconomic characteristics of the foreign-born population for intercensal dates than ever before. Efforts of INS, NICHD, and a few other agencies have led to the creation of some new data sets. Substantial gaps remain, but the first steps have been made.

Beginning in 1994, the Current Population Survey (CPS), the monthly labor force survey conducted by the Bureau of the Census for the Bureau of Labor Statistics, has included questions on country of birth, year of immigration, citizenship, and country of birth of parents. These data are available monthly, but the annual March demographic supplement to the CPS provides a wealth of information on social and economic characteristics, income sources, and program usage. Although the sample is fairly large (approximately 50,000 households), its size does limit to some extent the amount of detail available on immigrants. The CPS is, however, the first significant source of data on second-generation Americans (i.e., the native-born children of immigrants) since the 1970 Census.

The new CPS data are extremely important and their continued collection needs to be guaranteed and institutionalized. The data, however, are not without problems. Some of these can be attributed directly to the lack of institutional support for immigration statistics. The data on the foreign-born population from the 1994 CPS were available only for a very limited list of about 22 countries of birth. Immigrants from other countries were grouped together into an "all

other" category. Thus, it was not possible to tabulate information on immigrants from Asia, Europe, or any aggregation other than the specific list of countries. Although this problem was fixed by 1995, it would not have occurred if the Census Bureau had a staff or branch devoted to immigration statistics. In 1995, another glitch occurred. The list of countries on the data collection instrument was expanded and all responses on country of birth were coded directly, permitting aggregations at any geographic level desired. However, data for 6 countries were lost; all new responses of South Korea, Taiwan, Thailand, Trinidad and Tobago, Vietnam, and Yugoslavia were treated as "unknown country of birth" and imputed to other specific countries. The result was serious shortfalls in the estimated numbers from these countries and substantial overstatements for other Asian countries including India and the Philippines. The Census Bureau has recently released corrected data some 15 months after the initial release, but again the lack of staff dedicated to immigration statistics undoubtedly delayed finding the error and resulted in release of erroneous information.

A yet more serious problem affects the CPS data for 1994 and 1995. A change in editing procedures for the race question led to inconsistencies between the CPS data for persons who were not white or black and the demographic estimates used as control totals for the survey. (See Passel 1996 for more details.) As a result, CPS estimates of Asian/Pacific Islanders and American Indians/Alaska Natives are understated by approximately 20 percent in 1994 and 30 percent in 1995. In March 1995, the shortfall amounts to more than 1.5 million Asians. Since about two-thirds of this group is foreign-born, the resulting data on the foreign-born population, its origins, and its geographic distribution within the United States are seriously distorted. By 1996, this weighting problem and the coding problems described above have all been fixed.

Two other issues illustrate some of the problems deriving from the lack of institutional support for immigration statistics. The Census Bureau has consistently downplayed the significance of the weighting problem and has no plans to issue reweighted data.[2] Thus, many researchers remain unaware of a deficiency that could have a substantial impact on research and

[2] The Urban Institute is planning to release new weights for the March 1994 and 1995 socioeconomic supplements which should permit immigration researchers to better approximate the size and characteristics of the foreign-born population. See Passel 1997.

have no official mechanism for correcting the data even if they are aware of it error. The citizenship data also fail to meet fully the needs of researchers. These data, like other information on immigration, are collected only at the initial interview of CPS respondents. Because of the rotation patterns for the CPS sample, some of the information released on naturalization may be 16 months old. This lag makes very little difference for invariant statistics such as country of birth, country of birth of parents, and year of immigration. However, when more than 1 million immigrants have naturalized in each of the last two fiscal years, data on naturalized citizens could be significantly affected. More fundamental, however, is the lack of information on immigration status (e.g., refugee, legal permanent resident, undocumented, student visa).

Other Government Surveys. Data on immigrants from other government surveys is more limited, but new information will be available in the foreseeable future. The Survey on Income and Program Participation (SIPP) is a semi-longitudinal survey, tracking respondents for two-and-one-half years, which provides information on immigrants. The sample size is smaller than the CPS so information on immigrants is more limited. However, the welfare reform legislation provides funds for a substantially expanded SIPP (called the Survey of Program Dynamics) that will follow respondents for longer periods and have a larger sample.[3]

The American Community Survey (ACS) has been proposed as an alternative to detailed collection of sample data in the decennial census. As currently envisioned, the ACS would consist of large, independent monthly samples amounting to perhaps 3 million households every year. As an alternative to decennial census data, the ACS would fill most needs of immigration researchers who currently use census data. Moreover, the ACS data would be more timely since they will be collected on an on-going basis rather than once every 10 years. The ACS also represents an opportunity to expand data related to immigration. More information on the second generation and more detailed information on citizenship and legal status, if collected, would make the ACS an exceedingly useful source of immigration data. At this time, plans and funding for the ACS remain uncertain at this time.

---

[3] The PSID, a longitudinal survey that has been tracking respondents for almost 30 years, does not have many immigrants since the sample selection predated the current wave of mass immigration and new immigrants have not been added in representative numbers.

Another new survey, if fully operationalized, will provide sorely needed longitudinal information on immigrants. A pilot test for the so-called "Green Card" survey is currently in the field, funded by NICHD. This survey will be a sample of aliens admitted for permanent residence (i.e., legal immigrants) during a year, augmented by a sample of certain nonimmigrants. The immigrants will be followed for a number of years at specific intervals. This survey will have good data on legal status and will provide information on adaptation of immigrants the U.S. society. The survey will be limited in that it is only a single cohort and limited only to legal immigrants. Nonetheless, it will be an unprecedented source of data when fully funded and operationalized for a number of years.

Administrative Data. There are a number of potentially useful administrative data sets available for the study of immigration, including data from Social Security, Supplemental Security Income, and Food Stamps. These data have been underutilized for immigration research, in part, because their scope is limited. (For example, a study of Food Stamp recipients needs a comparison group of non-recipients that requires another source of data.) Nativity and immigration status are often not routinely or systematically collected. New data from welfare reform may be a gold mine for immigration researchers, as certain data collection activities are mandated in the law. To be fully utilized, researchers will need to ensure that needed data, especially on nativity and citizenship are collected, and that there are some parallel data on no-welfare users.

Data Deficiencies

The preceding discussion has covered a number of issues and problems affecting immigration data. While there are a number of new sources available now and in the near future, the major deficiencies can be summarized by four key points:

INS Administrative Data. The main purpose of data collected by INS is administrative; this limits severely its utility for analytic work. Only data for administrative actions are routinely collected, even when other information could be collected from immigrants at minimal cost, in terms of either money or respondent burden. The concepts and definitions used can be quite confusing and are often irrelevant for analysis or policy implementation. In addition, significant

changes in collection methods, coding, and data processing have occurred, making comparisons over time difficult.

Legal Status Information. The legal status of immigrants is a key characteristic for many new programs and policies as well as for many existing ones, with the most important statuses being legal permanent resident, naturalized citizen, refugee (on admission), undocumented alien, and nonimmigrant. These characteristics are not routinely collected in censuses and surveys, nor even in some administrative data sets. A number of technical issues must be addressed before such data can be systematically collected. The impact of legal status questions on response rates is unknown, but one would expect response rates to be reduced overall by such questions; undocumented aliens, in particular, may be dissuaded from participating in government surveys. In addition, many respondents may not be able to provide accurate information on their own legal status, nor for other household members. Children who derive U.S. citizenship when their parents naturalize may be an especially problematic group.[4] INS record systems may, in fact, need to be upgraded to provide accurate and timely information on the legal status of individuals.

Demographic Flows. As noted, we do not have even basic information on the numbers of immigrants living in the United States in various legal statuses; the only official estimates are for undocumented immigrants. Estimates of resident legal permanent residents and naturalized citizens are critical for assessing welfare reform and its impacts. Yet, available estimates are little more than "back of the envelope" approximations. Nor do we have information on the flows of immigrants into and out of the country each year. The in-flows can be closely approximated by manipulation of various INS data bases. However, estimates of out-flows are almost completely the product of informed guesses.

Even more serious for the future of research and impact assessment is that no one in either INS or the Census Bureau is tasked with providing such estimates. To the extent that estimates are made, they are often by-products of other work.

---

[4] Even INS does not have complete information on this group.

Longitudinal Data. Many of the significant issues concerning immigrants, their impacts, and integration require data on changes over time in the number of immigrants and their characteristics. There is virtually no such data covering significant numbers of immigrants. Synthetic cohorts, put together from successive data collections (i.e., decennial censuses, monthly CPSs) can fill some of the gaps, but selective outmigration hinders most serious analyses. Eventually, if systems currently in testing come to full fruition, we will have the required longitudinal data. Unfortunately, it takes time to collect such data.

Conclusions and Recommendations

Policies get made and implemented even if no data, or only deficient data, are available. Furthermore, some policies are instituted even in the face of data and analyses. For example, a number of studies showed that discrimination against foreign-looking and sounding persons legally in the country resulted from IRCA's employer sanctions. The law called for sunsetting of these provisions in such circumstances, but the sanctions remain in place today, largely because of the politics surrounding the issue. Nonetheless, it is generally thought that data and research aid the policy formulation process. Immigration and immigrant policies should be no different. More plentiful, more targeted, and more accurate research is clearly a desideratum. However, better and more plentiful data are a necessity. What can be done to improve upon the current situation? The following four main points, if implemented, will clearly lead to a better understanding of immigration and its impacts on the United States.

Centralize Federal Control of Immigration Statistics. No agency or office has official responsibility for immigration data collection (and analysis) activities. Consequently, there are many important activities which are either not being done or are being done in a very *ad hoc* manner. In some cases, there is no institutional support for essential data systems.

Some active oversight is needed for immigration statistics. There must be someone or some group defining data needs, setting data standards, and allocating resources so that essential tasks do not get overlooked. The lack of such oversight and planning is very apparent — in a recent book assessing the entire Federal

statistical system (Norwood 1996), immigration statistics are not even mentioned. Such an oversight or policy office could be located in the Office of Management and Budget or at INS, but it any case it needs cross-agency powers for allocating resources and defining data needs.

Strengthen INS Data Production & Analysis Functions. INS data production activities are clearly not an agency priority. There are some resource allocation problems, although the agency is flush with dollars and actually faces more problems of finding sufficient staff than budget problems. The administrative and policing mindset of the agency is more of an impediment to collection of useful, high quality data. These priorities need to be changed.

Analytic capacity in the area of immigration and immigrant policy needs to be strengthened within the federal government generally. INS' policy office has begun to expand in this area. It should be given higher priority and expanded. However, the ideal locus is not immediately apparent — INS may not be the best location. This issue is the type that could best be addressed by a centralized immigration statistics office. Some important projects could also be done outside the government. Wherever the research is done (either inside or outside the INS or federal government), analysts need to have access to essential data systems.

Improve Data Collection Systems. Progress has been made in this area, but a great deal more could be and needs to be done. One deficiency affecting virtually all current data systems is the lack of data on legal status. With the new welfare regimes, particular attention needs to be given to acquiring quality data on legal status, especially naturalized citizenship. Integration of data from disparate sources could be a particularly valuable method of filling gaps in the data.

Develop a Consensus on an Analytic Framework. A common language and reference system facilitates communication and policy formulation. In the area of immigration, such a common analytic framework may be extremely difficult, if not impossible, to achieve, given the variety of views on the subject. Nonetheless, a general, if not

complete, consensus will still be valuable. Several National Academy of Sciences' panels are beginning a significant effort in this area, as they assess existing studies of fiscal impacts and conduct their own. These efforts need to encouraged and strengthened. "Dueling" experts are, in part, a consequence of this shortfall, as many studies simply "talk past" one another. However, even a consensus framework will not be helpful if we only have bad data.

REFERENCES

Ahmed, Bashir and J. Gregory Robinson. 1994. "Estimates of Emigration of the Foreign-born Population: 1980-1990." Population Division Technical Working Paper No. 9. Washington, DC: U.S. Bureau of the Census. December.

Edmonston, Barry (editor). 1996. *Statistics on U.S. Immigration: An Assessment of Data Needs for Future Research*. Washington, DC: National Academy Press.

Fix, Michael and Jeffrey S. Passel. 1994. *Immigration and Immigrants: Setting the Record Straight*. Washington, DC: Urban Institute.

General Accounting Office. 1995. *Illegal Immigration: INS Overstay Estimation Methods Need Improvement*. Letter Report, GAO/PEMD-95-20. September 26.

Huddle, Donald. 1993. "The Costs of Immigration." Washington, DC: Carrying Capacity Network Research Series, Updated July 20.

Immigration and Naturalization Services. 1997. Press release on estimates of undocumented immigration. February 5.

Immigration and Naturalization Services. Various dates. *Statistical Yearbook of the INS*. Washington, DC: U.S. Government Printing Office.

Levine, Daniel B., Kenneth Hill, and Robert Warren (editors). 1985. *Immigration Statistics: A Story of Neglect*. Washington, DC: National Academy Press.

Norwood, Janet L. 1995. *Organizing to Count: Change in the Federal Statistical System*. Washington, DC: Urban Institute Press.

Passel, Jeffrey S. 1997. "New Weights for the March 1994 and March 1995 Current Population Surveys." Washington, DC: Urban Institute. Memorandum, forthcoming.

Passel, Jeffrey S. 1996. "Problem with March 1994 and 1995 CPS Weighting." Washington, DC: Urban Institute. Memorandum, November 12.

Passel, Jeffrey S. and Rebecca L. Clark. 1994. "How Much Do Immigrants Really Cost? A Reappraisal of Huddle's 'The Costs of Immigrants.'" Washington, DC: Program for Research on Immigration Policy, Urban Institute. Photocopy, February.

Warren, Robert. 1994. "Estimates of the Unauthorized Immigrant Population Residing in the United States, by Country of Origin and State of Residence: October 1992." Washington, DC: Statistics Division, Immigration and Naturalization Service.

Warren, Robert and Jeffrey S. Passel. 1987. "A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 Census." *Demography* 24 (3, August): 375–393.

Woodrow-Lafield, Karen A. No date. "A Sociology of Official Statistics on Undocumented Immigrants." Photocopy.

Seminar on Statistical Methodology in the Public Service

November 12-13, 1996

Session 2: What the Public Needs to Know About Federal Statistics

Discussion

Constance F. Citro

Committee on National Statistics

[Note: The views expressed are those of the author and not of CNSTAT.]

When I was asked to be a discussant at this session, my first question was: What do papers on immigration statistics and organizational perspectives have to do with the topic of the session? My second question was: What is the topic of the session anyway? After all, the title of the session, "What the Public Needs to Know About Federal Statistics," can be interpreted in several ways simply by changing a word or two:

(1) One interpretation is: "What the Public Needs to Know FROM Federal Statistics." That is, what data does the public need to have from the statistical system to inform policy debates and serve other important purposes? I interpret "the public" in broad terms as the data-using community of policy makers, analysts, planners, researchers, the media, and the general public; and I stress the need for the statistical system to be relevant to policy concerns. Jeff Passel's paper on immigration data relates directly to this interpretation of the topic; Dan Melnick's paper on organizational perspectives also bears on it, although indirectly.

(2) A second interpretation is: "What the Public Needs to Know TO FIND OUT About Federal Statistics." This has to do with one-stop shopping, which is the topic of another session, but it is relevant to the first interpretation of this session's topic because data of which users are ignorant cannot be used to inform policy or serve other public purposes and most likely represent a waste of resources.

(3) A third interpretation is: "What the Public Needs to Know About THE QUALITY AND USEFULNESS OF Federal Statistics." Both papers relate to this very important topic: bad or irrelevant data will almost always be worse than no data at all.

(4) A fourth interpretation is: "What the Public Needs to Know About Federal STATISTICS-PRODUCING AGENCIES." I had not thought of this one before seeing Dan Melnick's very interesting paper. His thesis is that how different agencies in the system go about providing what the public needs to know is shaped by their organizational perspectives, which, in turn, are shaped by their mission, history, experience, and expertise. The public needs to understand the different organizational perspectives in order to be able to assess the accuracy and appropriateness of the different outputs from the system.

Interpretation (1): What the Public Needs to Know FROM Federal Statistics

Users, if asked, will say they need to know everything about everything, but this expectation is not reasonable. Given resource constraints (and resources are always constrained, more so

now than ever), priorities must be set and choices must be made. What the data-using community *can* rightly expect is that the federal statistical system will continually monitor current and emerging policy issues and other important user needs and seek to determine the data collection vehicles that can best serve those needs-- rather than that the system will look first to the care and feeding of existing data collection vehicles and only secondarily consider whether they continue to be relevant to user needs.

Yet, from Jeff Passel's paper, it is clear that currently available immigration statistics only partly address today's policy needs in what is an increasingly contentious area of public concern. Agencies have made improvements in specific immigration databases, but no agency has a mandate to improve immigration statistics overall. The Immigration and Naturalization Service, with the perspective of an administrative agency, does not want this mandate. The Census Bureau, with the perspective of a statistical agency, could have but has not sought this mandate either, most likely because it already has a lot of other things on its plate. But the result for users is that the system of immigration statistics does not provide even basic numbers, such as how many people leave the United States each year, let alone more detailed information for policy purposes.

What Jeff Passel says for immigration statistics could be said for many other areas of current and emerging policy debate. Here are some examples from work at the Committee on National Statistics.

(1) Data on Children Historically, it has been difficult to study trends in the well-being of children--a topic of increasing public concern. One reason is that, until recently, longitudinal surveys have almost always followed adult members of households and not children, even

though many children these days move in and out of different family situations. Also, statistical agencies that collect data about children have tended to look only at the domain of direct relevance to their mission--school, health, work, and so on.

(2) Data on Transportation Historically, transportation data have been collected by "mode," that is, in terms of highways, rail, airlines, etc. Consequently, while there is a great deal of information about each mode, basic data on the total movement of freight or people in the country are lacking, and there are not comparable data across modes with which to address important public concerns about the safety, cost, and efficiency of the transportation system.

(3) Data on Retirement Income Security A growing policy debate that may be all-consuming in future years is how the United States will provide for the baby boomers' retirement. Employer-provided pension and health benefits are a key component of retirement income security, and employer behavior is one of the critical things to know about when considering alternative policies. There are many employer surveys conducted by such agencies as the Census Bureau and Bureau of Labor Statistics (BLS); the Department of Labor has an administrative database on pensions; and the Department of Health and Human Services is starting up a large new employer health insurance survey. But all of these data, while serving many useful purposes, do not in fact provide what the public needs to know to address important policy questions about employer-provided benefits and how employers might respond to different policy incentives.

In each of these areas, the statistical system recognizes the need to inject a cross-cutting, system-wide perspective, and is taking action. There is a recently-established federal interagency forum that is working hard to improve data on children (CNSTAT had a workshop on the topic, from which came a 1995 report on *Integrating Federal Statistics for Children*). In transportation, the 1991 Intermodal Surface Transportation Efficiency Act (ISTEA) established the newest federal statistical agency, the Bureau of Transportation Statistics (BTS). In its short life, BTS has made significant efforts to fill gaps in intermodal data and make accessible the data that have been buried away within USDOT. In immigration, there have been some efforts to bring agencies together, but Jeff Passel suggests that more needs to be done--specifically, that a single agency should be made the center of responsibility for improving immigration data. In the area of retirement income security policy, a CNSTAT study panel has just issued a report, *Assessing Policies for Retirement Income: Needs for Data, Research and Models*, that recommends an interagency working group, led by BLS and the Pension and Welfare Benefits Administration, to address how to provide the employer data that are needed to inform the policy debate.

These developments are all promising--or potentially promising. However, the U.S.'s highly decentralized statistical system--the most decentralized in the developed world--can make it hard to keep cooperative efforts going and, indeed, can make it hard to identify areas in which cooperative efforts are needed in the first place. Speaking in the context of Dan Melnick's paper, the individual agencies in the system bring valuable perspectives that ensure that it does not get locked into one approach or viewpoint, but without cross-cutting perspectives, the system is in danger of not providing what the public needs to know in many areas of policy concern. All of the participants in the system--the agencies, the OMB

Statistical Policy Division, of which Maria Gonzalez was a leading light for so many years, and such outside review groups as CNSTAT, need to work continually to find creative--and low-cost--ways to inject the necessary cross-cutting perspectives.

Here are some ideas for consideration by OMB and the system as a whole:

(1) Follow up Jeff Passel's recommendation and designate a central or lead agency for immigration statistics and other cross-cutting areas. A lead agency would not try to take over all of the data collection in an area, but it would have responsibility to identify data gaps and work with other agencies to develop strategies to fill them.

(2) Mandate that interagency forums and working groups establish specific action goals and timetables so that they do not lose momentum, which can all too easily happen as staff are pulled back to the perspectives of their own agency.

(3) Periodically charge an interagency task force to look around at what is happening in the U.S. society, economy, physical environment, and polity to identify emerging areas of policy concern that may require new or improved data. (Emerging issues will almost always require a cross-cutting perspective and consideration of how to revamp existing data systems.)

(4) Establish outside advisory groups in cross-cutting areas--that is, in addition to committees of outside experts that advise particular agencies or data programs, establish committees that advise in such areas as data for children or retirement income security policy needs.

(5) Interpret the Government Performance and Results Act (GPRA) to require statistical agencies not only to establish performance goals for their own programs and

services, but also to set goals for contributing to the effectiveness of the federal statistical system as a whole in providing what the public needs to know.

(2) What the Public Needs to Know TO FIND OUT About Federal Statistics

This topic is addressed in the session on one-stop shopping. I would simply make a point about the potential of the Internet and World Wide Web to help the user community. Statistical agencies and OMB are already making innovative use of the Internet to provide large amounts of data to users; the agencies and OMB could use the flexibility of the Internet to even greater advantage to serve users' cross-cutting data needs. For instance, the White House Economic and Social Statistics Briefing Rooms, which provide key statistics and links to the agencies, are likely to be a popular route of access for users. The briefing rooms could be made even more useful by such simple means as adding cross-cutting categories--for example, statistics for children, the elderly, immigration. Such expansion in the number of categories could help users find needed data and also help the federal statistical system identify cross-cutting areas in which more or different data are needed to serve user needs.

In the move toward electronic dissemination of data, which the agencies are absolutely right to pursue, I would raise one caution. The agencies hope not only to serve users more effectively via the Internet, but also to save costs. However, such cost savings should not target agencies' in-house analysis capabilities. In-house statistical and subject-matter specialists are needed to work with their agency's data in order to provide benchmarks for users and valuable insights for their agency about needed improvements in data quality

and relevance.

(3) What the Public Needs to Know About THE QUALITY AND USEFULNESS OF Federal Statistics

This is a very important topic on which there could be an entire session. I would simply make a point again about the role of the Internet. Data dissemination via the Internet is exciting in that it is expanding the user community, but it is also scary in that it is expanding the opportunities for users to misapply data because they do not understand their quality and limitations. The federal statistical system has an obligation to provide information to Internet users about data quality and to make the issue of quality standards a priority concern. Perhaps the Federal Committee on Statistical Methodology could work to develop standards for documentation of data on the Internet. Also, OMB could enforce such standards by including statistics on the White House Economic and Social Briefing Rooms only for those agencies that follow the standards. It is true that many users will not read the footnotes, but they cannot heed them or have the opportunity to be educated if information on quality does not accompany the data.

(4) What the Public Needs to Know about Federal STATISTICS-PRODUCING AGENCIES

It is unrealistic to expect that the public, or data-using community, will seriously study the sociology of the federal statistical system. However, it is important that federal statistical agencies and OMB periodically remind themselves of the insights from Dan Melnick's paper about the role of organizational perspectives in shaping how the system responds to data needs for policy and other public purposes. Each agency needs to be cognizant of and periodically reassess its own perspective. Each agency and OMB also needs to consider how best to inject the cross-cutting perspectives that are so necessary to enable the system to provide what the public needs to know.

Council of Professional Associations on Federal Statistics

Seminar on Statistical Methodology in the Public Service
November 12 - 13, 1996

WHAT THE PUBLIC NEEDS TO KNOW ABOUT FEDERAL STATISTICS

Summary of Remarks by
TerriAnn Lowenthal, Discussant

I am pleased to serve as a discussant this morning on the topic of, "What the Public Needs to Know about Federal Statistics." I will focus my remarks on two specific groups of users of Federal statistics: the general public, and Congress as the elected representatives of the general public. It is important to bear in mind that Congress both **directs** Federal statistical policy through agency authorizations, funding decisions, and data requirements, and **uses** (or misuses, as the case may be) Federal statistics to support policy and program development.

I want to address two basic issues that I think will help us get to the heart of the topic at hand. First, how do the public and Congress form their opinions about Federal statistics? And second, what is your responsibility, as experts in the field, to help these two large, generally non-expert, groups of users understand the value of the work you do?

Let's start by looking at what the public and Congress know about Federal statistics. Actually, it is much easier to look at what they **don't** know. Generally, you can assume that the level of knowledge about Federal statistics among Members of Congress and the people they represent isn't all that much different: very low. There appears to be a vast 'disconnect' between the prevalence and use of statistics and any understanding of how and why they are produced and what they represent.

It seems to me that, intuitively, people ought to have some level of faith in the numbers that assault them daily through the media and public officials. After all, basic math (to which most of the population has had some exposure) is taught as an exact science. There is usually a right answer and a wrong answer, a right number and a wrong number. So why are people so skeptical or distrustful of Federal statistics?

It's important to understand that Federal statistics probably suffer from the company they keep (although usually not by choice). It is unlikely that the public distinguishes between Federal statistics and statistics produced by a wide range of private sources. Jeff Passel aptly referred to the phenomenon of "dueling" figures, when reporters of official statistics have a need to find a competing statistic in order to present what they believe is a balanced view of an issue. Ironically, I think it's likely that the

public would tend to trust statistics produced by a Federal agency above all others. But it is difficult (and, I suspect, probably not worth the effort to non-statisticians) to sort through the barrage of statistics they may be confronted with every day through the news, advertising, politicians, and other sources.

The fact is that non-Federal (or otherwise unofficial) statistics are often produced for a narrow purpose, to support a specific policy outcome or cause. Some may be useful and produced through reliable means, and some may not be. But the reporting of these numbers is unlikely to explore such technical issues. And if the public, as the recipient of this information, is skeptical of the cause that the numbers support or the organization that produced the data, then it is likely to be distrustful of the statistics themselves, and any other numbers that happen to be reported at the same time.

Another, similar cause of this environment of mistrust may be found in the persona of the bearer of the news. Federal statistics often are released publicly not by the statistical agency itself but by a senior spokesperson of the parent department, often in conjunction with a policy initiative or political (small "p") statement. Therefore, the perception of objectivity that might otherwise attach to the statistics may be undermined if the underlying policy context is politically (capital "P") charged.

I also think that the media, a primary source of statistics for the general public and Congress, also contributes to the confusion surrounding the meaning of statistics in several ways. First, as I mentioned earlier, in the name of balance or objectivity, the media often finds a statistic to counter every statistic it reports, thereby diminishing the meaning of any given statistic and contributing to the perception that numbers only mean what the people producing them want them to mean.

Second, the media often doesn't take the time (or doesn't have the time) to understand the full meaning of statistics, how and why they were collected, and what they tell us and don't tell us.

And third, members of the media also can be skeptics, further undermining public confidence in the numbers to a significant degree. I have the perfect example, which I've been waiting to share with an audience like this for quite awhile.

As 5,000 statisticians were meeting in the Windy City this past August for the annual meetings of the American Statistical Association, an Op-Ed on welfare reform (which had just been passed by Congress) by columnist Mike Royko appeared in the Chicago Tribune. Mr. Royko started his column by conjuring up an image of the often-quoted statistic that 1.2 million American children would be thrown into poverty if the welfare reform bill being considered by Congress were enacted. This statistic, said Royko, was produced by "alleged experts." (This, by the way, is probably better than being called a "know nothing.") Heartless, if true, Royko said. But then he asks: How can anyone be certain about something like this? "How can these experts be sure it is

true? How do they know it won't be 1.4 million. Or 900,000. Or 2,365. Or 317," Royko asked.

Well, for one thing, Mr. Royko, even a non-statistician couldn't live with that large of a margin of error.

Mr. Royko then goes on to say, with unveiled sarcasm, that he doesn't doubt the findings because they came out of "some kind of think tank, where bright people ponder stacks of stats, charts, and other data, then make significant pronouncements." That sounds like statisticians aren't too far above lawyers in the public perception department! You've got tough skins, I know. The trouble is, if memory serves me correctly, that the study in question actually came out of the Department of Health and Human Services. But let's not let the facts get in the way of a good argument.

I suppose we should be comforted by Mr. Royko's statement two-thirds of the way through his column: "Since I'm not an expert and statistics give me a migraine, my guess is that nobody knows what the effect of the new welfare law will be." Mr. Royko, if you don't know what you're talking about, how come you get to write an influential column in a major newspaper?

Right after informing us (rather proudly) that he's not an expert, though, Mr. Royko proffers the following statement: "The biggest of all welfare problems is unmarried, uneducated, unemployable young women having illegitimate kids." Hmmmm. I wonder if Mr. Royko has any data to back up that statement? Or maybe it's just another guess?

Now, aside from the fact that this column really made me angry, I have a larger point to make. It seems to me that because the media is responsible for a great deal of the public's exposure to statistics, both Federal and non-Federal, it ought to be a little more responsible in how it portrays and reports numbers. And Federal statisticians have a responsibility here, too. Don't underestimate the value of the advice about dealing with the press that Victor Cohn gave you in this morning's keynote address. It may be tempting to dismiss reporters as hopeless when it comes to reporting fairly or fully on scientific matters, but please bear in mind that the media is the primary avenue of communicating the work you do to most of the nation, so you might as well strive to be friends, if not partners and allies.

Now that we've established that the public and Congress have a weak understanding of Federal statistics, we probably need to determine what they know about the Federal statistical system. The answer is "not much," I'm afraid. In fact, not many more Members of Congress than the people the represent even know that there is a Federal statistical system. There are several reasons for this.

First, oversight and funding responsibility among congressional committees is dispersed too widely for all but a handful of members to see the big picture (or even

know that there is a big picture). There are many different committees and subcommittees in both the House and Senate overseeing the work of the many Federal statistical agencies, and it is rare that any one committee or subcommittee is more than mildly aware of the activities of other panels with respect to oversight or funding of statistical agencies within an entirely different Federal department.

Second, the activities of individual Federal statistical agencies often are overlooked or lost within debates about larger agency or department programs. The decennial census is probably the one notable exception, but few Members of Congress can name any other activity of the Census Bureau. Even the relatively noteworthy monthly release of the employment and unemployment figures is generally viewed in a vacuum by most legislators. The statistics are eagerly anticipated and much analyzed both within and outside the halls of Congress, but rarely is there an effort to understand how the statistics were compiled and produced.

Speaking of the decennial census, don't look for any deep understanding among Members of Congress of that process, either. For example, while the Committees on Appropriations have been busy directing a reduction in number of questions on the census form, the rest of Congress was **adding** a question -- on children who are being raised by their grandparents -- to the 2000 census through the welfare reform bill. And to demonstrate the level of sophistication in Congress on matters statistical, the welfare reform provision requiring this new data directs the Census Bureau to collect "statistically significant data." And that's in connection with the decennial **and the mid-decade census**. You remember that much anticipated data-collection vehicle. If you blinked in 1995, you missed it!

The bottom line is that Congress doesn't even have a system for organizing its own activities affecting Federal statistics, never mind understanding the organization of the Federal statistical system!

And if you think Members of Congress have a minimal understanding of the Federal statistical system, generally, just think about their ability to grasp methodological issues. Many of us are familiar with the ongoing debate about the use of sampling and statistical techniques in the decennial census. It's often painful to watch, as Members of Congress make well-intentioned but usually uninformed pronouncements about the accuracy of sampling methods.

But perhaps the most egregious example of the failure to understand the meaning of Federal statistics, how they are produced, the constraints in producing them, and how Congress itself may have contributed to the limitations of the data it now rails against, is the debate (if one can call it that) over the accuracy of the Consumer Price Index. I never thought I would see the day when Congress would actually consider setting a statistic **legislatively**. I think this is outrageous; manipulation of data at its worst!

To me, this effort to change the CPI symbolizes nothing more than how palpable the fear over our most significant policy questions has become. Some of our most respected, thoughtful legislators see a legislative change in one of our most important statistics as the key to resolving some of the most fundamental, difficult, and titanic problems of social and economic policy -- namely how to control the burgeoning costs of entitlement programs, such as Social Security, that threaten our ability to balance the Federal budget. **You ought not to let that happen! The damage to the Federal statistical system may be irreparable if it does.** The non-Federal statistical community must speak up soon, loudly, and often, before this misguided effort goes too far.

I will close my remarks this morning by asking: What do you need to do to raise the level of public understanding of Federal statistics, as well as the level of public confidence in the ability of the Federal statistical system to provide the information that the country needs to develop and implement sound policy?

First, you need to establish useful working relationships with legislators and their staff, apart from the parent agency or department within which you operate. I know this may be difficult from an organizational or protocol standpoint, but you need to find a way to do it. Confidence in the reliability and objectivity of your products will increase in direct proportion to your ability to establish an independent identity as a statistical agency.

Second, rather than hiding from policy debates, you should strive to understand fully the policy context within which your statistical products will be used, now and in the future. You don't have to take sides in a debate, but you must be able to understand how your statistics will be used by people with differing goals and agendas, so that you can present and explain your product in a way that is most meaningful to the user.

Third, it is important to build meaningful and enduring relationships with the media. Try to understand the constraints under which reporters operate (such as the 5:00 p.m. deadline) and then develop innovative ways to increase the media's ability to use your statistics properly within those constraints.

And, finally, reach beyond your traditional audiences and professional relationships to establish meaningful and mutually beneficial relationships with public constituencies that are the end users and beneficiaries of the numbers you produce. Talk to the NAACP and the National Council of La Raza, to the U.S. Conference of Mayors and the National Association of Community Action Agencies, to senior corporate executives (as well as corporation statisticians or demographers) and editorial writers (as well as beat reporters). Reach out not just to professors of statistics, but to professors of history and political science and government.

When you hold conferences and seminars, such as today's session, include individuals and organizations outside of your traditional professional circles. Try to understand their concerns and perspectives, the worlds in which they operate, the people

they represent, and the objectives they are trying to achieve. These individuals and organizations are your window on the broader audience which you hope will understand what you do, and that understanding must be mutual in order to be effective.

Federal statistical agencies should go beyond their official advisory committees and communicate directly to the widest range of users and stakeholders possible, on their turf, in their environment. You will find, I think, that this kind of communication will help you move beyond the walls that we naturally build around our sciences, walls that often serve as barriers to a full understanding of what you do and why.

It is these kind of relationships that will help you build a strong and lasting foundation for the acceptance of, appreciation of, understanding of, and confidence in Federal statistics, and instill true meaning in the work you do for your government and country.

Session 3

TRAINING FEDERAL STATISTICIANS

SURVEY AND STATISTICAL TRAINING AND TRAINING STATISTICIANS: FEDERAL COMMITTEE ON STATISTICAL METHODOLOGY (FCSM) SUBCOMMITTEE INTERIM REPORT

Cynthia Z.F. Clark, Bureau of the Census

1. Introduction

This paper presents an overview of the progress of the FCSM Subcommittee on Survey and Statistical Training and Training Statisticians since its inception in November 1995. It also introduces planned activities of the subcommittee to further its study of this topic in Federal statistical agencies. The subcommittee plans to prepare a final working paper in the coming year.

The second section presents the charter of the subcommittee: its background, charge, stakeholders, membership, and scope. The third section briefly reviews selected agency survey and statistical training programs. The fourth section discusses some subcommittee discoveries. The last section summarizes activities of the subcommittee to date.

2. Charter of the Subcommittee

2.1 Background

The topic of survey and statistical training received by statisticians was proposed to the FCSM by Monroe Sirkin and David Williamson for study by a working group. A somewhat larger group met to clarify the topic for subcommittee study. Maria Gonzalez asked me to chair the subcommittee. Subcommittee members were identified through the FCSM and the committee was first convened in November 1995, meeting every 4 to 6 weeks since then.

2.2 Charge to the Subcommittee

The subcommittee was charged to document and compare survey and statistical training programs of Federal agencies. They were asked to provide baseline measures of these programs and to assess the strengths and weaknesses of these programs for statisticians. The subcommittee was directed to provide guidelines for agency self-improvements of these programs and for interagency coordination and collaboration in providing these programs. There was an expectation that the subcommittee would discover ideas that were worth sharing and identify areas of future need or improvement.

The subcommittee was asked to look toward the future by defining future needs, resources to meet those needs, and potential for collaborations between agencies. The subcommittee was also asked to identify areas where the Joint Program in Survey Methodology (JPSM) might enhance its contributions to the Federal statistical agencies. The subcommittee was directed to prepare a final report documenting its findings and making recommendations to improve survey and statistical training for statisticians.

2.3 Stakeholders and Customers

There are several important organizational stakeholders for the subcommittee -- the Federal Committee on Statistical Methodology chartered by the Office of Management and Budget (OMB), the Statistical Policy Office at the Office of Management and Budget, and the Committee on National Statistics of the National Academy of Sciences. These organizations are interested in the status of statistical training, in assessing the skills of the Federal statistical workforce, and in developing strategies to meet current and emerging training needs of that workforce.

The Federal statistical agencies are viewed as having a particular interest in this topic as survey and statistical training is relevant to a large proportion of their workforce. The agencies of particular focus are those represented on the Council of Statistical Agency Administrators chaired by Katherine Wallman, and those represented by members of the Federal Committee on Statistical Methodology. However, there might be other Federal agencies with a smaller contingent of statisticians who would have interest in the topic. The management of these agencies were considered to be the ultimate customers for the products of the group.

The Joint Program in Survey Methodology and other academic institutions and professional societies, particularly in the Washington, DC, area should be interested in the results of the subcommittee's efforts. The information about survey and statistical training courses and unmet training needs will help academic institutions and training providers plan their curriculum and courses.

2.4 Membership

The membership of the subcommittee consists of representatives from seven agencies and the Joint Program on Survey Methodology. The members are from the Bureau of Labor Statistics (BLS) (Nathan Howard), Bureau of the Census (Romeo Munoz, Nancy Bates), Centers for Disease Control and Prevention (David Williamson, Russ Roegner, George Ryan), Energy Information Administration (EIA) (Carol French, Renee Miller), National Agricultural Statistics Service (NASS) (Linda Raudenbush), National Center for Education Statistics (NCES) (Samuel Peng), National Center for Health Statistics (NCHS) (Monroe Sirkin and Joyce Crossman), and the Joint Program in Survey Methodology (Nancy Mathiowetz). I serve as chair of the subcommittee. During the first seven months of the subcommittee's existence, Denise Myers served as secretary while she was on detail to the OMB Statistical Policy Office from NASS. We have a combination of agency managers, practicing statisticians, agency training officers, and academic statisticians as committee members.

2.5 Scope

The first dilemma the subcommittee faces was to define its scope. The subcommittee was directed to address training received by statisticians employed by Federal agencies. Several questions immediately arose. What training was relevant to statisticians? Who are statisticians? What Federal agencies were interested in training received by statisticians?

In order to address these topics we decided to review training programs at several agencies. We hoped that a review of this training would help the subcommittee to identify those areas that would be of particular interest for training statisticians and for the customer Federal agencies.

3. Review of Agency Programs

We began our investigation with reports from each agency represented on the subcommittee. We discovered that agencies fairly universally offered training in new software (word processing, spreadsheets), general office skills (writing, presentations, team work), supervision and management, and personal development. After some discussion, we decided not to focus on these types of training opportunities because the needs are not different for statisticians than for other non-quantitative professionals in the workforce. We also discussed whether we should include training in statistical computing. We decided to include those courses where the statistical content was an important factor in the course material. Highlights of the agency review included the following findings.

3.1 Bureau of Labor Statistics. BLS developed a training plan for mathematical statisticians based on six technical Knowledge, Skills, and Abilities (KSAs). Additionally, BLS identified 3 KSAs for supervisory level positions and 3 more for management level. Training was also identified for the supervisory and management KSAs. BLS set priorities for different levels of training. Training needed to perform the current job had first priority; training that was expected to have an impact on how the current job was done had second priority; training expected to have an impact on future jobs had third priority. Priorities will be considered in determining training eligibility.

BLS provides in-house training, and also supports academic training. A particular example of in-house training was a six-month series of courses on quality management. BLS supports employees attendance in JPSM courses and degree programs as well as other academic course training.

3.2 Census Bureau. The Census Bureau supports academic training for staff on an individual course basis and for students at JPSM on a half-time basis. It also occasionally sponsors in-house statistical courses in topics such as variance estimation, time series, and categorical analysis taught by Census Bureau staff who are experts in these topics. Three years ago a mathematical statistician career development program was initiated.

In 1986, the Census Bureau developed a several-day orientation program and a six week course entitled Professional Skills Development. All professional employees took the course during their first year of employment at the Census Bureau. During the course the employees designed and conducted a survey, giving them hands-on experience in all aspects of a survey. These courses have not been held in the past 3 years due to an insufficient number of entry level employees. There are plans to revise these courses to meet expected future needs.

3.3 Centers for Disease Control and Prevention. The Applied Statistics Training Institute sponsors short-term (2-1/2 day) training courses across the country to focus on data issues related to current public health concerns. The CDC offers courses specific to its program area (e.g. Introductory Biostatistics, Epidemiology for the Non-Epidemiologist, Introduction to Methods for Public Health Program Evaluation, Utilization of Data by the Public Health Manager, Marketing Information to Policymakers: How Statisticians can produce what Politicians Want). They also offer more standard survey and statistical courses (e.g. Basics of Survey Research, Introduction to Survey Sampling, Small Area Analysis). They have also recently developed a Quantitative Methods Career Enhancement Program for their statisticians.

3.4. Energy Information Administration. The professional workforce at EIA includes industry specialists, operations research analysts, economists, survey statisticians, mathematical statisticians, computer specialists, and others. EIA participates in formal classroom training. It also has special training provided by its Office of Statistical Standards. These courses are specific to needs of individuals working in the energy industry (e.g. Determinants of Long-Run Energy Demand, Intermediate Econometrics, Commodity Pricing of Natural Gas, FEDWORLD Internet System).

3.5 National Agricultural Statistics Service. The NASS is very conscious of career development and training its professional statisticians. All employees have Individual Development Plans (IDPs). IDPs are standardized for each professional series with the opportunity to provide individual training opportunities. The agency has developed a formal week-long orientation program and a series of agricultural survey and estimation training program for all its statisticians. These courses cover specifics of agricultural survey design, data collection, and processing at several experience levels. The NASS has long supported a program of full-time academic training at the graduate level for mathematical statisticians, computer scientists, and survey methodologists.

3.6 National Center for Education Statistics. The NCES has a training program for staff to provide skills in statistical design, analysis, and project management. These courses are either taught by agency staff with a particular expertise or outside experts. The NCES has a unique program of training for external data users to promote effective and correct use of NCES data. Data users often are also data providers, so the training assists in improving data quality. Instructors are internal experts or known experts in a field.

3.7 National Center for Health Statistics. The NCHS supports academic programs for its staff, including participation in the JPSM courses and degree program. The NCHS also conducts in-house training. The NCHS brings in vendors to teach technical courses. The agency has developed a training database and has collected training costs systematically since 1995.

4. Discoveries

The review of agency training programs helped the subcommittee to focus the task of the group. We discovered that the training that was most relevant for the topic included both survey and statistical training relevant to the collection and publication of official statistics. We also discovered that this training was relevant to a broad group of quantitative professionals and one series of support staff working at statistical agencies. Thus, we decided to include as "statisticians" individuals classified in a number of series: mathematical statisticians (GS-1529), survey statisticians (agricultural, economic, demographic, health, education -- GS-1530), survey methodologists (GS-1530), quantitative social scientists (economists, sociologists, psychologists, anthropologists, demographers and other researchers), health scientists and biostatisticians, program analysts, operations researchers (GS-1515), and the support series of statistical assistants (GS-1531).

We discovered, however, that survey and statistical training provided to others who are not employees but who have some connection with the statistical or survey operations of a Federal agency was also of interest. The individuals receiving the survey or statistical training could include several types: interviewers, data providers, data users, or collaborators (clients).

The subcommittee review revealed several agency career development programs for statisticians, and two that were specifically designed for mathematical statisticians. The subcommittee felt that other agencies might benefit through knowledge of these career development programs. They each had aspects that had proved to be very effective and might well be adapted to other agencies. The subcommittee decided to provide information on these career development programs in the working paper.

The subcommittee discovered many good ideas that need to get broader visibility because they are applicable to other organizations. To do this, the subcommittee decided to include case studies of selected statistical agency training programs in the working paper. Greater knowledge of current agency training programs has given the subcommittee some synergistic ideas for collaborating in the area of training. The subcommittee plans to develop these ideas further in the working paper.

The review also demonstrated the need to have a common data set to make comparisons between agencies. The subcommittee felt that it would be desirable to have similar information on the scope and cost of agency survey and statistical training for employees, on the number of agency participants, and some information on survey and statistical training for non-agency employees. This information would help agencies gauge their performance in relation to other organizations. The subcommittee plans to collect this information from all the "customer Federal statistical agencies" previously mentioned.

5. Activities of the Subcommittee

Early in the tenure of the subcommittee, I discussed subcommittee plans with the chartering parent committee -- the Federal Committee on Statistical Methodology. I also met with the Council of Federal Statistical Agency Administrators to solicit support, additional subcommittee members, participation in later data collection, and interest in the end product.

The subcommittee collected written information on the workforce training of statisticians. We did a literature search using resources of the subcommittee members and the Internet. We contacted statistical agencies in other countries, receiving, in particular the training and development handbook for methdologists at Statistics Canada. We also communicated with the American Statistical Association Committee on Statistical Education to let them know of our project and to receive information they had on workforce statistical training. Nancy Mathiowetz worked with a JPSM student to prepare an annotated bibliography of the papers and documents that we discovered. This annotated bibliography will be in our working paper.

We obtained a copy of the Washington Area Alliance for Education in Survey Methods Consolidated List of Graduate Course Offerings for 1996-97. This includes information for American University, George Mason University, Georgetown University, George Washington University, University of the District of Columbia, JPSM at the University of Maryland, and the U.S. Department of Agriculture Graduate School. We plan to highlight this information in the working paper.

As was previously mentioned, the subcommittee reviewed agency training programs to determine the scope of project. As we were doing this we also reviewed agency training databases to determine what information was available. We identified software and databases that were more desirable and will highlight those in the working paper. We also identified information that was particularly relevant for comparisons between agencies --average training costs and average number of training opportunities per

employee, amounts and kinds of training provided and to whom, and total cost and the cost as a percent of program budgets. The subcommittee developed a questionnaire that will go to the agencies represented on the FCSM and the Statistical Agency Heads to request data on training costs, courses, and numbers of attendees for both employees and nonemployees. We will also get a distribution of classification types of employees attending specified courses and grade level of employees.

The subcommittee also recognized that we would not be able to obtain information from agency training databases on employee satisfaction with training opportunities for present work assignment, for keeping up with technology, and for career development. We were aware of an opportunity to collect information on employee perception of various aspects of their organization on the 1996-97 JPSM Practicum Survey of Organizational Climate being conducted at 10 of the Federal statistical agencies. The subcommittee proposed questions for the Practicum Survey that would provide insight into employee satisfaction with training.

The subcommittee is presently in the process of defining the content of the working paper report. We expect it to include statistical comparisons of survey and statistical training at Federal statistical agencies, case studies of survey and statistical training programs for employees, descriptions of career development programs for statisticians, descriptions of survey and statistical training provided by Federal statistical agencies for nonemployees, recommendations to improve training opportunities, identification of areas of collaboration across the statistical system to address future needs, and an annotated bibliography of workforce survey and statistical training.

The subcommittee played an active role in organizing this session at the COPAFS Conference. We wanted to share with the conference attendees some of the ideas that we have gleaned from our efforts thus far. We hoped that this would stimulate your thinking and provide for an increased exchange of ideas and information. In particular, we wanted to present information on the exciting career development programs for statisticians uncovered in our review of agency training to encourage other agencies to consider such programs. We also wanted to stimulate our thinking about future skill needs for statisticians in our agencies and to begin to identify those training needs.

FEDERAL STATISTICAL CAREER DEVELOPMENT PROGRAMS

1. National Agricultural Statistics Service -- Fred S. Barrett

1.1 Abstract

The National Agricultural Statistics Service (NASS) recruits and trains entry level professionals mostly in its 45 State Statistical Offices (SSO). NASS's career development and training program is designed to progress entry level statisticians (GS grades 5-7-9) to Senior SSO Statisticians (GS-12) in a noncompetitive environment. (See Attachment A for Statistician/ADP Career Training Paths). This paper describes the core training program and the competitive training programs available for employees seeking the GS-13 career level and above.

1.2. Agency Training Program

The National Agricultural Statistics Service is the primary statistical agency in the Department of Agriculture. The agency needs employees that have broad agricultural experience with special skills in survey design and administration, knowledge of data analysis and estimation procedures, and computer data processing. NASS's training program is designed to develop and improve the individual's knowledge, skills and abilities while enhancing the overall agency performance. All professional employees participate in a broad-based training and work program that introduces them to several disciplines and possible career paths. NASS expands this broad-based training with a number of competitive formal training opportunities designed to fill highly technical and specialized positions which are critical to the organization.

NASS believes a successful training program must be tailored to the individuals recruited and the career path opportunities made available to them. NASS recruits are hired mostly as GS-7's with Bachelors degrees or GS-9's with Masters degrees. They are generally classified into one of three disciplines, agricultural statisticians, mathematical statisticians, or computer specialists. All recruits must meet the minimum requirements of a Bachelor of Science degree. Agricultural statisticians must have at least 15 semester credits of mathematics and statistics, of which 6 credits must be statistics, plus 9 additional credits in other physical or social sciences. Experience in agriculture is very desirable. Mathematical statisticians must have at least 24 semester credits in mathematics and statistics, of which 12 must be mathematics and 6 statistics. A Masters degree in mathematics or statistics is preferred. Computer specialists must have 30 semester credits in computer science and mathematics.

Each new recruit will have a career path that is noncompetitive to the GS-12 journeyman level. The length of the training from entry to journeyman is about 6 years. To progress beyond GS-11 requires a second assignment in another State Office. Once the GS-12 journeyman level is reached, the statistician is expected to have a working knowledge of agriculture, an understanding of statistical concepts and applications, the ability to conduct surveys, be skilled in the use of basic computer software, and be able to operate in a LAN environment. They are also expected to have the ability to write and speak effectively, be able to plan assignments, and delegate work. During this training period each person will be offered the opportunity to cross-train in either of the other two disciplines.

1.3. Noncompetitive Training Program

The following are the chronological steps of noncompetitive career development and training for new professionals at NASS.

Office Orientation: The basic orientation is intense during the first two weeks and then continues for several months. The employees study materials on the agency mission and its history. They review agency and office policies and administrative procedures and they are trained to use their computer workstation as well as getting acquainted with the LAN operations. They are given their first work assignments and their performance elements and standards on which they will be evaluated.

Individual Development Plan (IDP): Each individual starts with a generic IDP that prescribes all the basic elements required of everyone to reach GS-12 along with the career goals and aspirations of the individual. In addition, the supervisor and employee are to specify training and development needs that meet the employee's objectives and are in accord with the agency goals and staffing needs.

Headquarters Training and Orientation: Groups of new employees come to Headquarters for a week of training. They receive an overview on all aspects of NASS survey and estimation procedures, and participate in an Agricultural Statistics Board simulation. They are also given a briefing on current research activities and computer operations. They become acquainted with the Headquarters environment and meet the Headquarters staff, as well as meeting with top management in a question-and-answer session.

On The Job: Learning while working is the most important element of NASS's training program. Opportunities are provided to travel with the State Office managers and senior statisticians to agricultural meetings, field days and commodity meetings. These meetings help increase their knowledge of agriculture and acquaint them with the agricultural industry. They are given assignments requiring them to conduct survey interviews and do crop observations. Their workloads and responsibilities are gradually increased in accordance with their performance and promotions. They will generally work in their first State Office for at least four years, and during this time they are expected to have different assignments in at least two of the three major functional areas of responsibility which are surveys, estimates, and systems services.

Basic Concepts Training: All new statisticians attend formal training sessions on NASS survey procedures, estimates and analysis, and yield measurement. These are usually four-day training sessions conducted by the Headquarters Survey Training Group. Basic concepts are taught and everyone is expected to know and understand these basics regardless of their current assignments.

Advanced Survey and Estimation Training: Statisticians who have completed the basic concepts and are assigned major responsibilities for either surveys or estimates are provided formal training on specific topics. These are usually four-day sessions conducted by the Survey Training Group. This training is directed toward specific actions and programs that are designed to give the participants the knowledge and skills to perform these activities at the full performance level.

Special Survey Training: Statisticians assigned to special and more complex surveys are provided training specific to that survey. This training covers all topics involved in conducting the survey including list building, sampling, questionnaire design, training of enumerators, data collection, editing, data analysis, summarization, and publication.

Senior Statistician Workshops: When a statistician reaches the journeyman level, they are often designated as the technical leader for one of the operational groups. When this occurs, they will periodically attend workshops along with their counterparts from other State Offices. These workshops emphasize project planning, coordination of office activities, and overall project management. This training involves sharing of ideas and interaction among participants and Headquarters technical leaders.

Professional Training: Statisticians are encouraged to engage in professional training opportunities such as college courses, seminars, toastmasters, and self-development training provided by local institutions or the NASS resource library. NASS pays for all such training, provided the training is related to the overall mission of the agency. The IDP is used to identify specific employee training needs and indicate appropriate professional training.

Pre-supervisory Training: A specially designed course has been developed by the USDA Academy at Texas A&M University to meet the unique needs of NASS statisticians and computer specialists. Training topics include values clarification, workplace diversity, stress management, effective meetings, presentation techniques, team building, communication, change, and ethics. In addition, NASS requires all of its statisticians and computer specialists to attend at least 80 hours of supervisory/management training.

Mathematical Agricultural Career Enhancement (MACE): The MACE program is a combination of "on-the-job" and formal educational program designed to permit agricultural statisticians to become cross-qualified as mathematical statisticians and mathematical statisticians to become cross-qualified as agricultural statisticians. Applicants accepted into MACE will complete the portion of the IDP's for both the agricultural statistician and mathematical statistician required for classification in the respective series.

Computer/Agricultural Career Enhancement (CACE): The CACE program is designed to permit computer specialists to become agricultural statisticians and to permit agricultural statisticians to become computer specialists. Applicants accepted into the CACE program complete the portion of the IDP's for both the agricultural statistician and computer specialist required for classification in the respective series.

1.4. Competitive Training Programs

When NASS professionals have completed their first year and are making satisfactory progress on their IDP, they have the opportunity to apply for certain competitive training programs. These programs include:

Full-Time Graduate Education Program: To be eligible, employees must attain the GS-9 level with at least one year of experience and be performing in a superior manner with satisfactory progress on their IDP. The full-time training programs provide at least one year of graduate level academic training. Agricultural statisticians, mathematical statisticians, and computer specialists are competitively selected for these training programs and, upon successful completion of the training, are placed noncompetitively in GS-13 headquarters positions. Selected candidates are given a new IDP which include any "warm-up" courses required. They are generally relocated to an SSO near a university with a NASS-approved graduate program. They must meet the selected educational institutions qualifications for admission to graduate school.

The full-time graduate level training programs are:

(1) Mathematical Statistician: This program is designed to provide education for agricultural and mathematical statisticians in advanced statistics and statistical theory to become highly-trained mathematical statisticians.

(2) Information Technology: This program is primarily designed for computer specialists to provide training in software engineering, telecommunications, or management information systems. However, the program is open to agricultural and mathematical statisticians who have a strong interest and background in computer systems and information technology.

(3) Survey Methodology: This program is designed for agricultural statisticians and mathematical statisticians to receive advanced training in survey methodology. Participants attend the Joint Program for Survey Methodology at the University of Maryland.

Career Development Intern Program (CDIP): The CDIP program is designed to provide accelerated training and career enhancing experiences for agricultural statisticians in State Offices. The training program is designed to prepare statisticians for specific assignments in Headquarters at the GS-13 level. Agricultural statisticians can apply as GS-11's when they are expecting a relocation to their second State Office assignment. They will be expected to maintain a full workload assignment in the SSO and complete all the IDP requirements for the GS-13 position targeted.

1.5. Summary

NASS's training program up to the journeyman level is designed to provide each professional employee with a broad base training in agriculture, statistics, surveys, and computer science. This gives the employee the opportunity to choose the career path most suited to their skills and abilities, but also offers them the opportunity to switch career paths. Everyone receives similar training and career development opportunities that allows them to compete for competitive technical positions at the GS-13 level in Headquarters and supervisory and management positions after a Headquarters assignment. This program has been very successful in providing NASS with a highly trained staff of agricultural statisticians while at the same time providing a source of specialized mathematical statisticians and computer specialists who have State Office experience.

2. Bureau of the Census -- Charles P. Pautler, Jr.

2.1 Introduction

The Census Bureau highly values its employees and works hard to have staff developmental programs that help both to train and retain our most valuable asset. Two such programs are the Joint Program in Survey Methodology (JPSM) at the University of Maryland and the Census Bureau's Mathematical Statistician Intern Program. This paper focuses mainly on the Intern Program since most readers are very familiar with the Joint Program.

2.2 Census Bureau's Commitment to JPSM

The Census Bureau makes a strong commitment to the JPSM since it is such an important component in our approach to training tomorrow's statisticians. When the first class of JPSM students was formed in September 1993, the Census Bureau saw the opportunity to have our staff trained specifically in the ways of the Federal Statistical System, and to have the classes focused on the methodologies that in some ways are unique to the Government methodologist. There is a commitment and burden on the organization to have a valued employee away from work about half time for three years--and to pay their full salary during this time along with tuition, books, and local travel. There was much discussion as to whether or not the Census Bureau could afford such an investment and to what extent, but the realization was that we could not afford not to be involved. Since the beginning, the Census Bureau has competitively selected 6 employees each year to start the program, and today has 19 students enrolled full-time in JPSM. In addition, several staff are supported in taking one course a semester, plus the Bureau has actively participated in the numerous short courses offered by JPSM. We are also very pleased to have our first 5 graduates from the program. Two of them followed the math stat track and three the social analyst track. This Spring, we are expecting 7 more graduates with 3 from the math stat track and 4 from the social scientist track. Our attrition from the program has been due to personal circumstances such as maternity, one person went to another Federal agency, and only one person decided that the program was not a good fit for her circumstances.

As statisticians, we are all into evaluation and measurement of results. However, in the case of the JPSM, it is probably too early to say if we have made a good investment and, in some cases, we may never know. Some may say that the fact that the staff we are selecting for the program are graduating and staying with the Census Bureau is proof of results. However, all students sign a commitment to Federal employment equal to 3 times the amount of time released to take courses so, at this time, no student has repaid that commitment. Others may say that the graduates will have to contribute significantly for many years before we can declare success, and how we measure significant contributions is a problem within itself. The students are pleased with the quality of their education and enthusiastically support the program. The students report that they come back to their jobs with new insights and techniques to apply to their work. It is also very clear that they are effectively networking amongst themselves and with students from other agencies. These staff are eagerly sought after by Census Bureau managers to fill vacancies and to accept positions of further responsibility.

2.3 Mathematical Statistician Intern Program

The Census Bureau started the program three years ago at the same time the JPSM was coming online. The two programs were seen as complimentary and yet appealing to two different universes of staff, and fulfilling two different missions. In general, the participants in the Intern Program already have a Masters Degree plus, in many cases, several more graduate level courses. The general profile of the participants has been staff who have had 5 years or more experience at the Census Bureau, who had worked in only one division, who were about 30 years old, and who were generally recognized as the best in their peer group. Competition for one of the four intern positions selected each year has been intense.

The Intern Program was established with 5 objectives:

- Identify staff for the fast-track to the GS-13 level and perhaps later management assignments.

- Provide exposure to each of the Census Bureau program areas--economic, demographic, decennial census, and statistical research.

- Provide opportunity for statistical assignments that require different areas of knowledge.

- Provide opportunity for professional growth through formal paper preparation and presentation in a professional forum.

- Provide enhanced training opportunities to meet career goals.

Implementation of the program has had the following features:

- Competitive selection which has involved intensive group interviewing by the Associate Director for Methodology and Standards, the senior mathematical statistician from each of the four program areas, and a division chief selected each year from each of the subject areas.

- One-year assignments in each program area where the intern has not had experience. Thus, with the four program areas previously listed, the Intern Program normally lasts three years for most interns.

- Presentation/participation at the annual ASA meetings. This is an important benefit since competition to attend ASA meetings is very intense amongst other staff. The interns automatically are expected to prepare a paper and go to the meetings.

- Each intern is assigned one of the four mathematical statistician division chiefs as a mentor. Regular meetings are held between the intern and the division chief, and individual development plans are prepared.

- Increased exposure to senior staff. Quarterly meetings for all interns with the Associate Director for Methodology and Standards and the four mathematical statistician division chiefs are held. Usually a member of the Executive Staff is invited to come and talk about their program area.

- Increased training opportunities. With the crunch on training funds the past few years, this has proven to be a valuable benefit of the program, as interns have been given priority for training money. Numerous JPSM short courses have been taken with these training funds, along with personal development type courses.

- When rotated to another area, the interns have been given priority for assignments that can be completed in a year and lead to an ASA paper.

- Experience in working as a group on a broader management or organizational problem. For example, the interns recently worked together to prepare a proposal for reorganizing the Census Bureau mathematical statisticians.

Intern Program Results/Positives: Overwhelmingly, the interns have been very pleased with the program and enthusiastically support it. On a personal level, it has been very beneficial for them with the implementation features I have just listed, and the program continues to meet the objectives that management set when the program was established. In addition, we are seeing the benefits of increased communication across the Bureau as the interns ban together for numerous networking opportunities, and take back to their respective branches news from across the Bureau. For example, they have regular luncheons without senior management involvement, and the group assignment to propose a Census Bureau mathematical statisticians' reorganization provided an excellent bonding experience.

Although senior management originally discussed giving the interns their first opportunities to obtain supervisory experience through the rotational assignments, this has not materialized for several reasons. However, the interns are seeing and taking notice of the various management styles they are being exposed to, and these differing styles are discussed and compared during their informal meetings.

Another positive from management's viewpoint is the increased pool of highly qualified staff to fill our future vacancies. We have already seen this positive as we have filled three GS-13 level positions with former interns.

Intern Program Concerns/Issues: Even though senior management has been generally pleased with the progress of the Intern Program, there have been valid issues and concerns raised by managers across the Bureau. These include:

- When a division has a person selected for the Intern Program and that person leaves the division to start the rotational assignments, the programs of the division are affected because the divisions have not been able to back-fill the vacated position. With many staffs down to only a few people, one person leaving is a significant loss of resources.

- Although most managers support the objectives of the Intern Program, several believe that we have created twelve positions that are not providing prime value to the programs of the Census Bureau. It has become a question of whether or not we can support the "luxury" of this program.

- Some managers believe that we are sending a negative message to other staff members in the units that the interns are assigned to by giving the interns priority in assignments. These managers make the argument that there are other deserving employees in the units who should be given these assignments.

- Some managers express the concern that we have created a caste system. Given that we are, in effect, creating four new GS-13s every year, will there be any promotion opportunities for the GS-12s who choose to dedicate themselves to becoming expert in one of the more complicated surveys of the Bureau or who choose a rotational and development program of their own?

- In the view of some managers, we have put more focus on the people than on the programs that are the primary purpose of the Census Bureau.

- And from our personnel division comes the concern that when the interns complete the program, there will not be a permanent GS-13 position available for them to fill.

None of these concerns/issues are trivial and, in some cases, the emotions run deep. With the first interns completing their three years this fall, we had no problem in finding permanent positions for them to fill. In fact, there were more positions than interns. It can also be argued that we have taken every step possible to rotate the interns to positions of the greatest need throughout the Census Bureau, but that is of little solace to the manger who ends up with one less resource. Of course, it is true that there is always a learning curve when a new person enters any position and, with the one-year assignments, there is the constant overhead of the learning curve, but senior management does not believe that this price is too high an investment in the future of the Census Bureau.

The issue of opportunity for those who are not a part of the Intern Program is a little harder to measure and difficult to demonstrate. In the past year, there have been job postings at the GS-13 level that all could apply for and were not filled by an intern. It is sure to happen in the future that there will be positions that an intern will be placed in and the manager to which this happens will surely feel that he/she was not given the opportunity to select who they wanted for the position. There will be some deserving employee who was not a part of the Intern Program who could have possibly done quite well in that position. All employees know about the Intern Program and all have the opportunity to apply for it and to be selected in the competitive process. It is the belief of senior management that through the Intern Program there will be a better pool of candidates who, through broadening work experiences, will be better equipped to fill the vacancies of the future.

2.4 Conclusion

The Census Bureau recognizes that its highly trained and specialized staff is its most valuable asset. Creating staff development programs to train and retain these staff is in the best interest of the Census Bureau. To that end, the Census Bureau has significantly supported and participated in the Joint Program in Survey Methodology at the University of Maryland by selecting six staff members each year to begin the three-year program on a full-time basis. All expenses, plus a full salary, are paid while releasing the student half-time to attend classes. In addition, the Census Bureau has created the Mathematical Statistician Intern Program. In this program, four people are competitively selected each year to begin three one-year rotations to program areas they have not previously worked in. Special training, assignments, and a mentor are significant parts of the Intern Program.

Both programs are a current burden on the organization, and issues and concerns have arisen about these investments in the future. However, both staff development programs are viewed positively and seen as vital components in preparing and retaining a highly technical pool of staff to provide both the technical and managerial leadership necessary for the Census Bureau of the next millennium.

3. Centers for Disease Control and Prevention -- G. David Williamson and Donald R. Betts

3.1 Abstract

We have developed the Quantitative Methods Enhancement Program (QMEP) at the Centers for Disease Control and Prevention (CDC) in response to the recent emphasis on reinventing government and to a need to provide alternative career development training for statisticians. In addition, the program is designed to

sustain and enhance statistical capacity within CDC. The QMEP is an innovative career enhancement alternative for CDC statisticians and other scientists who have a strong career interest in statistical and other quantitative methods. The program provides employees with a unique opportunity to move temporarily to another group within CDC to acquire new skills and understanding of specific analytic methods from CDC experts in such areas as Geographic Information Systems (GIS), longitudinal data analysis, meta analysis, risk assessment, sample survey analysis, and small area methods. We describe the joint development and management of QMEP by CDC's Statistical Advisory Group, the Human Resources Management Office, and the Statistics and Epidemiology Branch of the Epidemiology Program Office. We also discuss the goals, assets, and process of the program, and our plans to expand it for interagency participation.

3.2 Introduction

In a January, 1989 memorandum, the Associate Director for Science, Centers for Disease Control and Prevention (CDC), established CDC's Statistical Advisory Group (SAG) in recognition of the increasingly important role statistics and statisticians play in fulfilling the agency's mission. The SAG was asked to act in an advisory role to CDC's Office of the Director on statistical issues, to oversee and coordinate CDC-wide statistical activities, and encourage communication among statisticians and other scientists at CDC. In 1991 the SAG cosponsored CDC's Planning Retreat for Epidemiologic and Statistical Methods in Public Health to produce a plan for maintaining and developing expertise in statistical and epidemiologic methods essential to preserving CDC's national leadership role in assessment of health status and in public health practice. One of the high priority recommendations from the retreat was enhanced recruitment and retention of statisticians and other data analysts with expertise in methods to analyze public health data. This recommendation, coupled with the reinvention/reengineering environment in government fostered by the 1993 National Performance Review, became the impetus for us to consider innovative ways to provide positive reinforcement for deserving CDC employees who have a strong career interest in analytic methods.

In December, 1993 the SAG convened a focus group comprised of CDC statisticians, management analysts, and personnel experts to discuss and lay the foundations for an internal rotation program which would identify outstanding employees who demonstrate interest and promise in analyzing public health data and temporarily reassign them to another group within

CDC to acquire and develop new statistical skills. During the next year, the focus group and others in CDC's Epidemiology Program Office (EPO), the group which provides personnel to coordinate and support much of the SAG activities, discussed and revised the original proposal for the methods rotation program. The resulting proposal was one which provides alternative career development training for statisticians and, at the same time, sustains and enhances the statistical capacity within CDC. Now, with approval and support from SAG, the Statistics and Epidemiology Branch of EPO, and CDC's Human Resources Management Office, we introduce the Quantitative Methods Enhancement Program (QMEP).

3.3 Program Description

The purpose of QMEP is to provide an innovative career enhancement opportunity for CDC and Agency

for Toxic Substances and Disease Registry (ATSDR)[1] scientists. The program facilitates professional growth and development for statisticians and other data analysts, assists in maintaining and strengthening CDC's capacity in analytic methods expertise, and promotes retention of CDC scientists.

The QMEP provides CDC employees with a unique opportunity to temporarily be assigned to another group at CDC to acquire new skills in specific analytic methods from CDC experts on current statistical methods (for example, generalized estimating equations, meta analysis, neural networks, risk assessment, sample survey analysis, and small area estimation). The program consists of 1) a competitive application process that is used to match an applicant with a mentor, 2) an internship training period, and 3) an evaluation of the program experience by the intern, mentor, and sponsoring Center, Institute, or Office (CIO) of CDC. It is anticipated that the intern will be released from all job duties during the time of participation in the program.

The QMEP is open to health and mathematical statisticians and to other scientists who have a strong career interest in statistical and epidemiologic analytic methods. Applicants must be permanent employees of CDC with a minimum of two years service in the agency, and must have secured approval from supervisors to participate in the program. Applicants should be at the GS-11/12/13 (or CO-04/05 level for Commissioned Corps employees), and have received a rating of "Excellent" (or "D" for Commission Corps employees) or higher on their most recent end-of-year personnel evaluation.

Each applicant must submit to HRMO a current position description, including job series and grade, CIO, and location; curriculum vita; name, address, and phone number of immediate supervisor; and one page memo that addresses the following topics:

- Reason for applying to program
- Specific methods area(s) in which to learn new skills or knowledge
- Primary learning objective(s)
- Description of how assignment will benefit career goals
- Description of how new or enhanced skills will benefit sponsoring CIO.

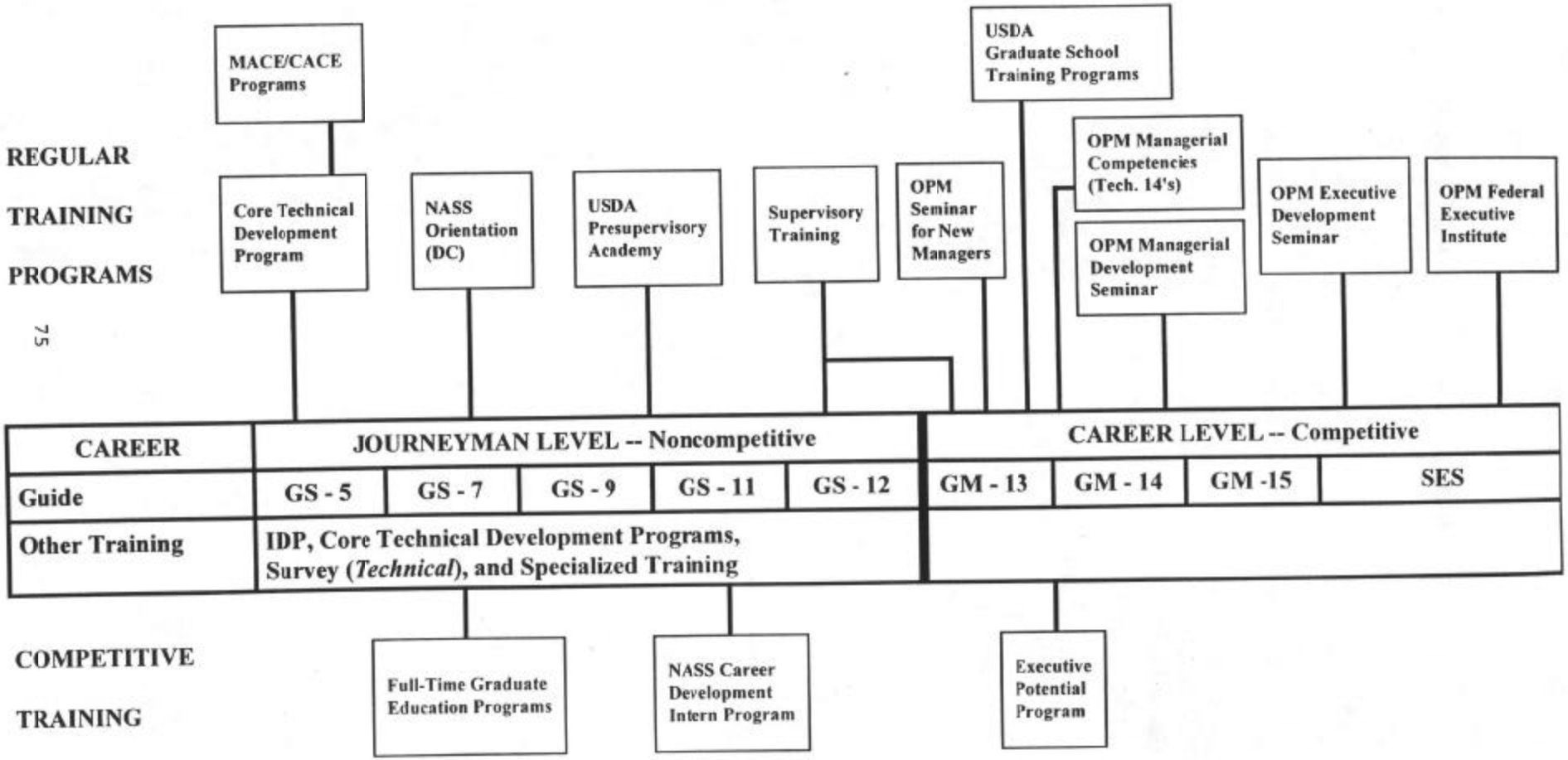
After a SAG subcommittee screens applicants, prospective applicants will receive a listing of available projects/methods areas and associated mentors for the program. Mentors will be located throughout CDC, including locations other than Atlanta, where CDC is headquartered. Subsequently, mentors and applicants will interview each other and rank their choices, and a matching process will be used to team selected program participants with mentors. A maximum of three applicants will be selected for the initial year of the program, depending on qualifications and availability of interns and mentors.

The duration of the training is variable, depending on the length of projects. Generally, enhancement assignments will be four months to one year. The QMEP, modelled after CDC's long-term training program, calls for the applicant's sponsoring CIO to provide the FTE and salary support throughout the training period, but there is flexibility in this arrangement and exceptions to this model should be mutually agreed upon by CIOs. The intern will return to their CIO upon completion of training.

---

[1] Future reference to CDC includes ATSDR because the QMEP applies to both agencies and allows participation between those agencies.

NASS STATISTICIAN/ADP CAREER TRAINING PATHS



STATISTICAL SKILL REQUIREMENTS AND TRAINING NEEDS FOR THE 21ST CENTURY

Panel Discussion

1. National Agricultural Statistics Service -- Comments by Rich Allen

When Cynthia Clark first asked me to be on this panel, I interpreted the topic very broadly as "what knowledge would we want our staff to have?". I was reminded of the search for knowledge in *The Hitchhiker's Guide To The Galaxy* by Douglas Adams. What the characters in the book were seeking was the answer to: **LIFE, THE UNIVERSE, AND EVERYTHING**

While we can't assume our staff will have all of that knowledge, my first point is that well-rounded staff members will be more valuable than narrower, more technically trained individuals.

As an aside, *The Hitchhiker's Guide* had words of encouragement for supervisors. On the cover--in friendly letters--was the message: **"DON'T PANIC"**

1.1 Basic Assumptions

Everyone in 21st century statistical offices will need good computer literacy skills. Those must extend to the ability to efficiently link files across different software packages. I do mean everyone since the support staff role in the future will involve applying the same software as do statisticians. Organizations must decide what software packages will be supported with training and upgrades. Staff members will then be able to write extensions and improvements which can be shared throughout the organization.

Sampling knowledge will become even more important in the future, particularly for organizations which historically focused mainly on census-type techniques.

1.2 Specialized Skills

Speaking from an economic statistics viewpoint, I see three related skill requirements as being very important to good quality national and small area agricultural statistics. The first is knowledge of techniques for handling large data bases. It will not be possible, nor wise stewardship of public resources, to collect new data for all applications. Previous survey and administrative data will be linked with newly collected data for many statistical efforts.

Whenever data files are linked or individual firms contacted on a repetitive basis, there will be data relationships which seem to be invalid. These may be due to changes in respondents for the same unit, corrections of earlier responses, new respondent-interviewer nonsampling errors, or a multitude of other reasons. Being able to interpret and adjust for these "MESSY DATA" situations is the second requirement for publishing consistent statistics.

The third requirement is the ability to link data together and develop statistical profiles. An operation profile, based on previous information, can quickly identify current reporting problems, can speed up edit

and summary operations, and can sometimes minimize the amount of new data to be collected. NASS has had great success with statistical profiles for all Federally-inspected slaughter plants. Based on a plant's own data, the profile "knows" what livestock species are handled, what weight ranges are usually processed, the general volume level per week, and even what days of the week the plant operates. By using the profile in an on-line edit, it is now possible to concentrate data collection on the few large plants and, by processing whatever small plants do come in quickly, to use the profile for all missing plants to estimate the actual weekly slaughter one week earlier than using only conventional methods. This type of profiling will be valuable in reducing future respondent burden.

1.3 Most Important Considerations

In giving the knowledge topic more consideration, I came to the conclusion that there are two very important emphasis for the future.

The first of these I like to call **Statistical Thinking**. The American Statistical Association, the Washington Statistical Society, and others have done a lot to promote Quantitative Literacy (QL). QL focuses on the importance of working with mathematical relationships. Statistical thinking is the proper application of statistical tools and techniques. For instance, there will be continued advances in computer software and it will be possible to automatically run many test and analyses--even if they are not appropriate. Thus, it will be vital to analyze all factors and properly direct our staff's efforts. Improvement in statistical thinking comes from experience, from practicum type course work, and from listening to other people's experiences at technical presentations and symposia such as this one.

My final topic is not new. However, as telecommunication speed increases, as data users have more processing power, and as businesses consolidate and focus on international issues, **Communication** will become even more important in order to provide accurate, timely information. Particularly for statistical organizations relying on voluntary cooperation, it will be vital to communicate with respondents--in their "language"--and build close relationships. As we use more technology, we must develop better communication within units, across statistical organizations, and with respondents and data users.

2. Bureau of Labor Statistics -- Comments by Cathryn Dippo

When contemplating the topic of statistical skill requirements for the 21st century, I began with the obvious--everyone will have access to a computer. This train of thought led to a list of issues, including:

What skills are needed to use a computer and existing software?

What skills are needed to facilitate the adoption of new environment?

How can the work environment be changed to enhance/promote more efficient use of computers?

What skills are needed to promote innovation or the design of even better tools for increasing productivity and job satisfaction?

Presumably, few people would quibble over whether or not these questions are issues or not; more than likely, the natural inclination would be to just say the list is incomplete.

But what if I change the wording only slightly:

> What skills are needed to use **statistics** and existing **statistical** software?
>
> What skills are needed to facilitate adoption of new **statistical** environment?
> (One would hope that the adoption of various TQM and continuous quality improvement philosophies would result in a more statistically-based production environment.)
>
> How can the work environment be changed to enhance/promote more efficient use of **statistics**?
>
> What skills are needed to promote innovation or the design of even better **statistical** tools for increasing productivity and job satisfaction?

The questions are still just as relevant, if not as obvious. This leads me to two observations:

> The issues related to statistical skill requirements for the 21st century are not really different from those for any other skill.
>
> In some ways, statistical skills cannot be separated from other skills. This may be obvious with respect to computer skills, but it is also true of communication skills.

If I had to pick one statistical skill requirement that deserves special attention, it would be an ability to communicate about statistics. In these remarks I focus on just two of many aspects of communication about statistics--fundamental concepts and metadata. By fundamental concepts of statistics, I mean the ideas of uncertainty, probability, variation, bias, and relevance. These notions are not well understood by the population at large, including the vast majority of people working in statistical agencies. And those who have a good grasp of the concepts are often poor at communicating them to others, myself included. For example, when someone asks me what the unemployment rate was last month, I do not automatically say 5.2 percent plus or minus .2, or that I'm 90 percent confident that the unemployment rate is somewhere between 5.0 and 5.4 percent. By metadata, I mean all the background information on concepts, question wording, sample design, data collection procedures, estimation methods, etc. that give a number meaning and that should be used in determining a particular statistic's fitness for a specific use.

Communication about statistics is one of our most important tasks. Whether you are an interviewer trying to gain the cooperation of respondents, a methodologist discussing alternative procedures, or an analyst conveying survey results to the public, it permeates your work.

And, if we look to the 21st century, in the near term, what one factor is most likely to have a major effect on communication in our work lives? The Internet. The creation of the Internet is often likened to the invention of the printing press. The printing press provided a means for communicating with more people than was possible previously. The same is even more true of the Internet. In little over a year, the number of downloads from the BLS web site has increased to more than 800,000 per month, and we can expect the number of users of statistics to continue to grow. Many of these new users will be less knowledgeable than

current users. For example, the Federal One-Stop Statistics web site will point people to agencies and statistics that may be unknown to them.

If the staff at Federal statistical agencies are to cope well with the expanding customer base for statistics, they must be prepared to communicate both the fundamental concepts of statistics and statistical metadata to the public. Already, those responding to telephone inquiries at BLS note that there are more questions about concepts and survey methods (metadata) and fewer requests for just a number. While it is important for these information providers to be knowledgeable about the statistics (both the numbers and the metadata needed to give the numbers meaning), it is imperative that these people recognize that the statistics are just estimates, that they are surrounded by uncertainty, and that they communicate the uncertainty, along with the estimates, to the user.

We also need to do more than just **hope** that Internet statistics users will look or call for the metadata needed to interpret the statistics. We need to be pro-active in designing effective means of communicating both metadata about the statistics and the basic concepts of statistics within the web sites themselves.

While I am confident about what basic skills are needed, I have no answers on how to achieve the goal of a statistically literate staff in a statistical agency, much less society. I do believe we need seminars and classes in basic concepts; classes without formulae that do include exercises in writing and speaking about statistics. Such classes would be better if we had a manual or book like Victor Cohn's *News and Numbers*, but written for the statistics provider. Or, maybe we need a prescriptive counterpoint to *How to Lie with Statistics*. Whatever the mechanism, there is a clear need for making sure staff have a firm grasp of the fundamental concepts of statistics and the skills necessary for communicating these concepts to others.

I also believe that one very effective means of learning is to put yourself in someone else's shoes. Try conducting research on a topic via Internet. One quickly learns that a whole new set of skills are needed on how to find pertinent information, how to use some not very user-friendly interfaces, how to assess the quality of information from unknown sources, etc. Everyone in a statistical agency should have these skills. Interviewers should be encouraged to have a usernet group for discussing how to convert refusals and to learn what's new in other field organizations. Methodologists and computer scientists should be able to access Statistics Sweden's current best methods for improving response rates or the functional specifications for the Australian Bureau of Statistics data warehouse. Analysts should know how to assess the quality of a new statistic vis-à-vis a similar one from several years ago or in another survey or census, either in the U.S. or from another country.

When using the Internet for research, one quickly begins to recognize deficiencies. Methodological metadata is in very short supply. This brings us to the ubiquitous chicken and the egg problem of statistical skills. Which comes first: methodological metadata or the skills necessary to prepare and/or use metadata? If we do not provide users with response rates and estimates of variance, how can we expect them to use them? If users do not know how to use response rates and estimates of variance to assess a statistic's quality, why should we expect them to want us to provide them?

I would like to be a firm believer in the basic concept of the *Field of Dreams*, that is, "Build it and they will come," but I have my doubts when it comes to statistical metadata. What if no one knew anything about baseball, and someone built a field and placed some balls and sticks on it? (That is analogous to providing statistics without metadata.) Eventually, people would start playing a game, but would it be

baseball? Now, what if, in addition, a book of rules were provided in hieroglyphics, along with the balls and sticks? (That is similar to providing metadata to people who do not know how to read or use it.) Again, without the Rosetta Stone, the probability is not very high that we will get the chance to see baseball as we know it. In our case, the statistical skills associated with knowing how to read and use metadata are the Rosetta Stone. Building metadata repositories, incorporating the production of quality measures into survey production systems, etc. are insufficient in and of themselves. If we want data users, be they inside or outside agency walls, to demand and make use of metadata, we need to begin defining and transferring the skills needed to understand and use survey metadata.

To sum up, statistical communication skills are the key to giving statistics meaning. In the 20th century, we have developed considerable expertise in producing statistics. If we are to make major advances in the 21st century, we need to develop similar expertise in communicating statistical information--both the fundamental concepts of statistics and survey metadata.

3. National Science Foundation -- Written comments made by Jeanne Griffith were not available

4. Bureau of the Census -- Comments by Paula Schneider

4.1 Skill Assessment

Last December, CB formed a council to assess this topic (Strategic Planning Council on Organizational Resources, SPCOOR) comprised of union representatives, managers from all program areas, Regional Offices and Data Preparation Division in Jeffersonville)

SPCOOR mission: design an action plan to improve training and development programs

To accomplish this:

Evaluate and inventory current training programs

Identify corporate skill needs

Examine competencies required by jobs

Bottom Line: We must identify skills needed in the future AND skills no longer needed.

4.2 What are we finding?

Advanced-level technical skills in statistics and survey methodology will continued to be valued:

Masters-level preparation is needed for many jobs in both areas.

Hiring at the Master's/Phd level and/or additional training through the JPSM are both ways of obtaining employees with these skill levels.

Technical skill needs will be driven by new technologies. Smaller budgets and changing customer expectations will force the Census Bureau to embrace new technologies for collecting, processing and disseminating data. These include, for example, the computer assisted information technologies of computer assisted telephone interviewing (CATI), computer assisted personal interviewing (CAPI), Internet (using computer assisted self administered questionnaires -CSAQ, or for data dissemination), and touchtone data entry (TDE).

The way in which work is organized will change radically:

Jobs giving way to roles (employees will be assigned work not on the basis of job description but rather on basis of need for skill sets and role expertise. Employees will become more like independent contractors).

Fewer employees will be needed, but they must possess higher skill levels and a broader range of skills.

Future training and development needs MUST include non-technical as well as technical. Non-technical skills are becoming increasingly important as statisticians find themselves working in organizations shaped by following trends:

A more diverse work force,

Greater use of teams, partnerships, and networks,

Increasing premium on ability to meet customers' needs for tailored products/services.

4.3 What does this mean for future training/skill requirements?

More cross-training (Subject matter specialists <---> Statisticians),

Broader array of skills (e.g., programming skills AND statistics skills),

Workforce of generalists rather than specialists (e.g., questionnaire designers who can also program the automated instrument they "design"),

Increased computer-literacy in all areas (e.g., internet HTML programming, data analyses, expanded use of administrative records requires new computer programming skills to match and unduplicate records),

Training must reflect new census methodologies (e.g., new uses of sampling for nonresponse follow-up, continuous measurement, etc., requires knowledge of variance estimation software like VPLX, CPLX, SUDAAN (not just SAS anymore).

Technical expertise in topics for corporate research (e.g., time-series, population projections, small area estimation).

4.4 Two examples that we're heading in the right direction:

Joint Program in Survey Methodology (JPSM -- University of Maryland, University of Michigan, Westat)

Offers broad-based training in all aspects of survey data collection...result=more cross-training.

For mathematical statisticians: Increases knowledge of social survey methodologies (e.g., questionnaire design, pretesting methods -- cognitive interviews, behavior coding, focus groups, mode effects, non-sampling error,

For survey statisticians and social scientists: Increases knowledge of statistical survey methodologies (e.g., sampling techniques, statistical data analyses).

JPSM short-courses: offers concentrated training to all agency employees in specific areas (e.g. questionnaire design, variance estimation, context effects).

Mathematical Statistician Internship Program - Competitively selects the "best" and then trains them in each of the Bureau's program areas.

4.5 Where do we go from here?

Use SPCOOR report to plan/design future training,

Continue programs that "groom" future managers and achieve cross-training such as the JPSM, and the intern program,

Continue to promote idea of corporate hiring across program areas.

Statistical Skill Requirements for the 21st Century

Cathryn S. Dippo, Bureau of Labor Statistics

When contemplating the topic of statistical skill requirements for the 21st Century, I began with the obvious--everyone will have access to a computer. This train of thought led to a list of issues, including:

- What skills are needed to use a computer and existing software?
- What skills are needed to facilitate the adoption of new environment?
- How can the work environment be changed to enhance/promote more efficient use of computers?
- What skills are needed to promote innovation or the design of even better tools for increasing productivity and job satisfaction?

Presumably, few people would quibble over whether or not these questions are issues or not; more than likely, the natural inclination would be to just say the list is incomplete.

But what if I change the wording only slightly:

- What skills are needed to use statistics and existing statistical software?
- What skills are needed to facilitate adoption of new statistical environment?
  (One would hope that the adoption of various TQM and continuous quality improvement philosophies would result in a more statistically-based production environment.)
- How can the work environment be changed to enhance/promote more efficient use of statistics?
- What skills are needed to promote innovation or the design of even better statistical tools for increasing productivity and job satisfaction?

The questions are still just as relevant, if not as obvious. This leads me to two observations:

- The issues related to statistical skill requirements for the 21st century are not really different from those for any other skill.
- In some ways, statistical skills cannot be separated from other skills. This may be obvious with respect to computer skills, but it is also true of communication skills.

If I had to pick one statistical skill requirement that deserves special attention, it would be an ability to communicate about statistics. In these remarks I focus on just two of many aspects of communication about statistics--fundamental concepts and metadata. By fundamental concepts of statistics, I mean the ideas of uncertainty, probability, variation, bias, and relevance. These notions are not well understood by the population at large, including the vast majority of people working in statistical agencies. And those who have a good grasp of the concepts are often poor at communicating them to others, myself included. For example, when someone asks me what the unemployment rate was last month, I do not automatically say 5.2 percent plus or minus .2 or that I'm 90 percent confident that the unemployment rate is somewhere between 5.0 and 5.4 percent. By metadata, I mean all the background information on concepts, question wording, sample

design, data collection procedures, estimation methods, etc. that give a number meaning and that should be used in determining a particular statistic's fitness for a specific use.

Communication about statistics is one of our most important tasks. Whether you are an interviewer trying to gain the cooperation of respondents, a methodologist discussing alternative procedures, or an analyst conveying survey results to the public, it permeates your work.

And, if we look to the 21st century, in the near term, what one factor is most likely to have a major effect on communication in our work lives? The Internet. The creation of the Internet is often likened to the invention of the printing press. The printing press provided a means for communicating with more people than was possible previously. The same is even more true of the Internet. In little over a year, the number of downloads from the BLS web site has increased to more than 800,000 per month, and we can expect the number of users of statistics to continue to grow. Many of these new users will be less knowledgeable than current users. For example, the Federal One-Stop Statistics web site will point people to agencies and statistics that may be unknown to them.

If the staff at Federal statistical agencies are to cope well with the expanding customer base for statistics, they must be prepared to communicate both the fundamental concepts of statistics and statistical metadata to the public. Already, those responding to telephone inquiries at BLS note that there are more questions about concepts and survey methods (metadata) and fewer requests for just a number. While it is important for these information providers to be knowledgeable about the statistics (both the numbers and the metadata needed to give the numbers meaning), it is imperative that these people recognize that the statistics are just estimates, that they are surrounded by uncertainty, and that they communicate the uncertainty, along with the estimates, to the user.

We also need to do more than just **hope** that Internet statistics users will look or call for the metadata needed to interpret the statistics. We need to be proactive in designing effective means of communicating both metadata about the statistics and the basic concepts of statistics within the web sites themselves.

While I am confident about what basic skills are needed, I have no answers on how to achieve the goal of a statistically literate staff in a statistical agency much less society. I do believe we need seminars and classes in basic concepts; classes without formulae that do include exercises in writing and speaking about statistics. Such classes would be better if we had a manual or book like Victor Cohn's *News and Numbers*, but written for the statistics provider. Or, maybe we need a prescriptive counterpoint to *How to Lie with Statistics*. Whatever the mechanism, there is a clear need for making sure staff have a firm grasp of the fundamental concepts of statistics and the skills necessary for communicating these concepts to others.

I also believe that one very effective means of learning is to put yourself in someone else's shoes. Try conducting research on a topic via Internet. One quickly learns that a whole

new set of skills are needed on how to find pertinent information, how to use some not very user-friendly interfaces, how to assess the quality of information from unknown sources, etc. Everyone in a statistical agency should have these skills. Interviewers should be encouraged to have a usenet group for discussing how to convert refusals and to learn what's new in other field organizations. Methodologists and computer scientists should be able to access Statistics Sweden's current best methods for improving response rates or the functional specifications for the Australian Bureau of Statistics data warehouse. Analysts should know how to assess the quality of a new statistic vis-à-vis a similar one from several years ago or in another survey or census, either in the U.S. or from another country.

When using the Internet for research, one quickly begins to recognize deficiencies. Methodological metadata is in very short supply. This brings us to the ubiquitous chicken and the egg problem of statistical skills. Which comes first: methodological metadata or the skills necessary to prepare and/or use metadata? If we do not provide users with response rates and estimates of variance, how can we expect them to use them? If users do not know how to use response rates and estimates of variance to assess a statistic's quality, why should we expect them to want us to provide them?

I would like to be a firm believer in the basic concept of the *Field of Dreams*, that is, "Build it and they will come," but I have my doubts when it comes to statistical metadata. What if no one knew anything about baseball, and someone built a field and placed some balls and sticks on it? (That is analogous to providing statistics without metadata.) Eventually, people would start playing a game, but would it be baseball? Now, what if, in addition, a book of rules were provided in hieroglyphics, along with the balls and sticks? (That is similar to providing metadata to people who do not know how to read or use it.) Again, without the Rosetta Stone, the probability is not very high that we will get the chance to see baseball as we know it. In our case, the statistical skills associated with knowing how to read and use metadata are the Rosetta Stone. Building metadata repositories, incorporating the production of quality measures into survey production systems, etc. are insufficient in and of themselves. If we want data users, be they inside or outside agency walls, to demand and make use of metadata, we need to begin defining and transferring the skills needed to understand and use survey metadata.

To sum up, statistical communication skills are the key to giving statistics meaning. In the 20th century, we have developed considerable expertise in producing statistics. If we are to make major advances in the 21st century, we need to develop similar expertise in communicating statistical information--both the fundamental concepts of statistics and survey metadata.

Statistical Skill Requirements and Training Needs for the 21st Century

Panelist: Paula J. Schneider

- Last December, Census Bureau formed a council to assess this topic (Strategic Planning Council on Org. Resources, SPCOOR)

(comprised of union representatives, managers from all program areas, Regional Offices and Data Preparation Division)

- SPCOOR mission: design an action plan to improve training and development programs

- To accomplish this:

  1) evaluate and inventory current training programs

  2) identify corporate skill needs

  3) examine competencies required by jobs

  Bottom Line: We must identify skills needed in the future AND skills no longer needed.

What are we finding?

1) Advanced-level technical skills in statistics and survey methodology will continue to be valued:

   - Masters-level preparation is needed for many jobs in both areas.

   - Hiring at the Master's/Phd level and/or additional training through the JPSM are both ways of obtaining employees with these skill levels.

2) Technical skill needs will be driven by new technologies. Smaller budgets and changing customer expectations will force Census Bureau to embrace new technologies for collecting, processing and disseminating data.

   Example: Computer Assisted Info. (CAI) technologies:

   - CATI
   - CAPI
   - Internet (CSAQ surveys, data dissemination)
   - TDE

3) The way in which work is organized will change radically:

   - jobs giving way to roles (employees will be assigned work not on the basis of job description but rather on basis of need for skill sets and role expertise.

Employees will become more like independent contractors).

- fewer employees will be needed, but they must possess higher skill levels and a broader range of skills.

4) Future training and development needs MUST include non-technical as well as technical. Non-technical skills are becoming increasingly important as statisticians find themselves working in organizations shaped by following trends:

   - a more diverse work force
   - greater use of teams, partnerships, and networks
   - increasing premium on ability to meet customers' needs for tailored products/services

What does this mean for future training/skill requirements?

- More cross-training:

  (Subject matter specialists <---> Statisticians)

- Broader array of skills

  (e.g., programming skills AND statistics skills)

- Workforce of generalists rather than specialists

  (e.g., Questionnaire designers who can also program the automated instrument they "design")

- Increased computer-literacy in all areas

  (internet HTML programming, data analyses)

- Training must reflect new census methodologies

  (e.g., new uses of sampling for NRFU, Continuous Measurement, etc., requires knowledge of variance estimation software like VPLX, CPLX, SUDAAN (not just SAS anymore).

  (e.g., expanded use of administrative records requires new computer programming skills to match and unduplicate records)

- Technical expertise in topics for corporate research

  (e.g., time-series, population projections, small area estimation)

Two examples that we're heading in the right direction:

1) **Joint Program in Survey Methodology**
   (JPSM University of Maryland)

- Offers broad-based training in all aspects of survey data collection...result = more cross-training.

- For Math Stats: Increases knowledge of social survey methodologies, e.g.,

  - Q'nnaire Design
  - Pretesting Methods (cog. interviews, behavior coding, focus groups)
  - Mode Effects
  - Non-sampling error

- For Survey Stats, social scientists: Increase statistical survey knowledge, e.g.,

  - sampling techniques
  - statistical data analyses

- JPSM short-courses: offers concentrated training to all employees in specific areas (q'nnaire design, variance estimation, context effects).

2) **Math-Stat Internship Program**

- Competitively selects the "best" and then trains them in each of the Bureau's program areas.

Where do we go from here?

- Use SPCOOR report to plan/design future training
- Continue programs that "groom" future managers and achieve cross-training:
  - Math Stat interns
  - JPSM cohorts
- Continue to promote idea of Corporate Hiring across program areas

Session 4

MEASURING CUSTOMER SATISFACTION

SEMINAR ON STATISTICAL METHODOLOGY IN THE PUBLIC SERVICE

Sponsored by:
Council of Professional Associations on Federal Statistics

TITLE of PRESENTATION:

Customer Satisfaction Surveys in the Federal Government

PRESENTER:

Sam Rives
National Agricultural Statistics Service
U.S. Department of Agriculture

November 1996

This presentation focuses on two primary objectives related to customer service surveys: 1) How and Why the NASS established the Customer Service Consultative Working Group (CUSCO Working Group) and 2) What NASS has done as an Agency toward customer satisfaction measurement.

1. CUSCO WORKING GROUP

President Clinton issued Executive Order No. 12862 in September 1993. In this Order, titled "Setting Customer Service Standards,' the President affirmed strong commitment to customer service and putting people first. This Executive Order directed Federal agencies to survey customers and then continually reform our management practices and operations to provide service to the public that matches or exceeds the best service available in the private sector.

At its core, customer service is the act of listening to customers about their needs and satisfaction and then meeting those needs and ensuring a high level of satisfaction. This is as important for those providing government services or programs as it is for those in business delivering goods and service. To provide excellent customer service, we need to take the time to listen to our customers. Our 16th President provided a perfect role model. Even during the Civil War, President Lincoln threw open the doors of his office to the public twice a week. To many of his advisors, this seemed like badly wasted time. Not to Lincoln:

> "No hours of my day are better employed," Lincoln asserted. People moving only in an official circle are apt to become merely official; not to say arbitrary in their ideas. Now this is all wrong. I call these receptions my public opinion polls. Though they may not be pleasant in all their particulars the effect as a whole is renovating and invigorating."

In response to Executive Order 12862, NASS proposed to provide statistical and survey services to other agencies of the USDA to help them identify and survey their customers. In order to provide these services, NASS established a Customer Service Consultative (CUSCO) Team or Working Group.

Background and reasons NASS proposed CUSCO:

The National Agricultural Statistics Service (NASS) is the statistical data collection agency of USDA. It has a staff of statisticians trained in survey, sample, questionnaire design, data collection, and analysis procedures. Additionally, the Service has resources available that could be used in a cost effective way to conduct surveys for USDA agencies. The Service maintains both

a geographic area frame and a list frame of farm and ranch operators. NASS has the expertise to select representative samples from these frames. NASS also has an on-going contract with the National Association of State Departments of Agriculture (NASDA) that provides a core of enumerators who can conduct personal and telephone survey interviews. As part of its on-going operation, the NASS has a staff who design mail, personal, and computerized survey data collection instruments. The Service has a computer processing system designed for analysis and tabulation of statistical data.

The NASS anticipated that several agencies in the Department would identify farmers and ranchers as their customers and would want to conduct surveys of this target population. The Service desires to coordinate any USDA surveys for this population. This effort would minimally provide nonoverlapping sample designs reducing multiple farmer contacts. It would also permit the NASS to lessen the impact of these surveys on its on-going agricultural survey programs. Additionally, where information is desired by several USDA agencies has the same target population and is compatible, NASS could design a single integrated survey that meets multiple objectives. The viability of this approach would depend on the number of cooperating agencies and the scope of the agency--specific questions. A Team USDA questionnaire could include some global service questions common across several USDA agencies. This approach would provide each agency the opportunity to meet its individual needs, but also benefit from the USDA core questions.

The NASS is also prepared to provide survey and statistical consultation services and/or to design and conduct surveys of other customer populations. The NASS has a staff with expertise in all aspects of survey design and operations.

The NASS is prepared to offer survey coordination and statistical consulting services to USDA agencies. This service is available for any target customer population of USDA agencies. However, for surveys of farm and ranch operators, the NASS desires to select nonoverlapping samples and design coordinated surveys to the extent possible. This effort enhances the USDA image by reducing cost, burden, and multiple contacts.

Nass Cusco Plan

- The USDA directed all agencies to survey their customers to establish baseline measures of customer service.

- NASS, functioning as internal consultants, will help USDA'S agencies meet their responsibilities to establish baseline measures of customer service.

- NASS will provide 20 hours of professional consultation to each agency at no charge. Additional services will be available on a reimbursable basis.

- Agencies contact NASS when ready to discuss their customer service plans. At these meetings, NASS discusses and shares information about:

    - How to identify customers.
    - How to identify customer services to be measured.
    - Procedures to be followed to collect data.
    - Resources NASS can provide to support their customer survey effort.
    - Other resources internal and external to USDA that might be available to support the customer survey effort.

- NASS develops and maintains a process to ensure that each customer is not surveyed more than once by anyone in USDA.

- NASS reviews and/or recommends survey methodology to ensure sound statistical procedures.

In addition, NASS, working with USDA, developed a process that dramatically decreased the time in obtaining clearance for survey instruments and packets which must have OMB approval.

The Office of Information and Resource Management (OIRM)(now renamed Policy and Analysis Coordination Center) serves as the liaison between USDA and OMB. Using a form developed by OIRM and NASS, customer service survey information requests from USDA agencies are sent to OIRM via NASS and forwarded to OMB. NASS indicates on the form if plans have been discussed and if it is a customer service survey. OMB response could be expected within approximately five (5) days.

The following tables show some of the agencies and projects which NASS assisted with customer service surveys. In some cases NASS provided consultative services, while in others the agencies contracted to NASS to conduct the entire survey. Attachments 1, 2, and 3 are types of aids used in obtaining information, initiating discussions, or answering questions.

AGENCIES AND CUSTOMER SERVICE SURVEYS WHICH NASS PROVIDED ASSISTANCE

AGENCY	PROJECT	CUSTOMER
APHIS	Quality of laboratory services including billing procedures. New APHIS unit refining its mission.	State and University Labs, Animal importers and exporters. Researchers/practitioners and heads of organizations with roles in biological control.
FS	Reinvention of Forest Service.	General public. Telephoned 5,000 respondents to obtain 500 responses.
CFSA	Measure customer satisfaction for six ASCS farm programs.	Individuals currently participating in the six ASCS farm programs.
INFOSHARE	Field Office of the Future	Focus groups among the general public.
FmHA	Quality Assurance Survey Counter Card Survey	FmHA/RDA large borrower groups (farmers, municipalities, waste treatment plants). Applicants and borrowers at 1,680 county and 250 district offices.
MAP-OD	Organizational Performance Assessment Survey. Focus Groups.	Federal employees grouped by grade levels.
OC	Readership Survey	1,100 broadcast newsletter subscribers.
FNS	NASS review of RFP responses	
ES	Employee Perception Survey Climate Survey	Program leaders at Land Grant Universities ES Employees

AGENCY	PROJECT	CUSTOMER
FAS/ICD	Customer Service Survey Survey to develop leads for exporting U.S. products. This project on hold.	A census of domestic and foreign customers.
USGS	User Needs Survey of USGS's primary data users.	Professional organizations, associations, magazine and newsletter subscribers.
FCIC	Survey to determine why some farmers are using their products and others are not. This project on hold.	
NAL	Survey to obtain information for setting standards for customer service. This project on hold.	
NASS	Press Service Survey	Press Service

Problems Encountered

After providing numerous hours of consultative discussions with many different agencies, it is interesting to note how the same problems and difficulties dealing with customer service are so similar.

The major difficulties agencies had and the areas where most help is needed are:

- Identifying their customers,
- knowing what they want to find out from customers,
- how to get representative responses, and
- the importance of having an **UPDATED** list of the population to be surveyed.

I will expand on the first problem, identifying the customer since this appears to cause the most disagreements within the different agencies.

Defining the Customer

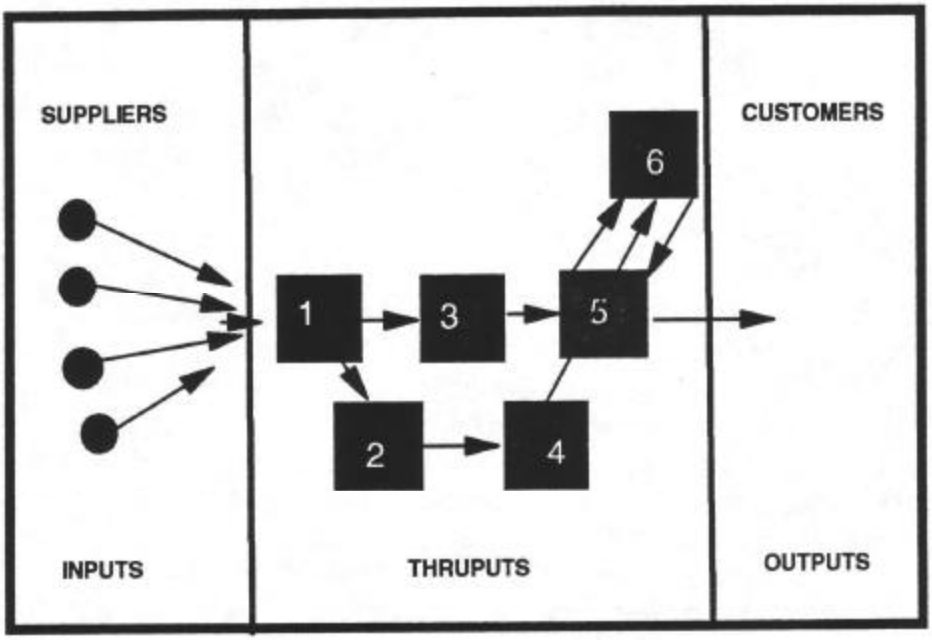
This is the topic that seems to perplex agencies more than any other. Usually during the consultative session, the discussion starts with the question "who is the customer and what business are you in?" The answers most often given are: "we have multiple customers, " or " The U.S. citizen is our customer. Usually Federal agencies do not have multiple customers but usually do have multiple players and multiple layers of interaction required to satisfy their customers. This statement cannot be used as a cop out for not defining your target. Also, saying that the U.S. citizen or the American public is your customer is simply too broad. The key is to establish a clear chain of events performed throughout the organization which satisfies the customer. In other words, before trying to determine who the customer is, look at the core process your agency performs. Once the core processes are identified, then determine who the recipients, i.e. customers, are. A "customer" is one who receives the results of the process.

Defining the Process

A process is a series of interrelated activities or work tasks that transforms an input and add value to it to create a product or service that fulfills a customer's need.

DEFINING A PROCESS

- ***A group of interrelated actiivities that begins with an input, adds value (transformation), and results in an output.***



A process starts with a customer need and is completed when the customer is satisfied.

Other Lessons Learned

- A factor in determining whether they are a customer depends on the relationship. They could be partner one time and a customer the next.
- A customer is at the end of the chain and helps accomplish the mission
- We usually measure what we are good at and then keep measuring it because it makes us look good.
- If you don't measure it, then you can't improve it.
- Increasing costs seldom increases quality. But, increasing quality often reduces cost.
- Make employees part of the solution.
- We do many wrong things right.

2. CUSTOMER SATISFACTION MEASUREMENT PROCESS

NASS focuses its current customer satisfaction review efforts on seven types of customers. These are: (a) individuals who contact NASS through its 1-800 customer hotline, (b) individuals who contact State Statistical Offices (SSO's) by mail or telephone for information, (c) individuals who contact the Agency through electronic mail, (d) data analysts who are often described as "power users" of the Agency's data series, (e) news organizations that participate in the immediate release of NASS statistical reports, (f) farm organizations, private companies, and government agencies which arrange to have visitors participate in briefings on lockup security and release procedures for major reports, and (g) State commissioners and directors of agriculture. Many of these efforts were already taking place prior to the National Performance Review[1]

For categories (a) and (b) above, prototype questions were designed for small scale tests. A formal survey has not been conducted for category (c), since we already get direct feedback. Much of the category (d) input was received from a data users meeting, plus other telephone and personal contact during the year. For categories (e), (f), and (g) customer responses come from on-going meetings and communications.

---

[1]Rich Allen: Implementing Customer-Drive Quality in Federal Statistical Agencies Panel, ASA Winter meetings, January 8, 1994

One unique feature of NASS is that it operates under cooperative agreements with State Departments of Agriculture and land grant universities or both in every State. NASS is both the Federal and State statistical organization for agriculture in each State. NASS State Statisticians work very closely with customers in carrying out their State responsibilities.

Definition of Customer

NASS places an emphasis on strengthening communications among **internal** customers to improve delivery and quality of products. These communication efforts have particularly been successful in expanding capabilities for analysis of survey data and survey factors which may have affected indication levels. Those efforts improve quality as they lead to the need for fewer subsequent revisions and more consistent data series.

NASS serves a wide variety of **external** customers--from the legislative branch of government to individual farmers. Among the most active data users are general farm organizations, producer and marketing associations, farmer advisory services, farm and general purpose media, State government offices, agribusinesses, public interest groups, university and other researchers, and other Federal Government agencies.

Contacts with Customers

NASS is in contact with customers every day through mail, telephone, and personal interview surveys and through a variety of written and telephoned requests for data or clarification of estimates. These data contacts occasionally plant a seed for changes in data products or services, but most improvements have come about from more specialized contacts.

Data Users Meetings - Over the past 8 years, NASS has conducted annual meetings with data users around the country. NASS takes the lead in organizing and scheduling but the other "economic" agencies of USDA, the Economic Research Service and the World Agricultural Outlook Board, also take part and the Agricultural Marketing Service of USDA is often included.

Before 1985, NASS held "listening" meetings where data users were invited to come to a location to comment on economic and statistical reports and on data needs. Those meetings often exposed considerable confusion about which agencies conducted what surveys, what reports were issued by specific agencies, and methods and procedures used by statistical agencies. The fact that some of the input received at these meetings was misguided, led to the concept of data

users meetings which are instructional in nature.

Headquarters Briefings - NASS hosts groups of visitors in conjunction with the release of major reports nearly every month. The monthly Crop Production report for which extra security procedures are used which include securing the staff working on the report in a lockup area until its signing by the Secretary of Agriculture and release at 3:00 p.m., is extremely popular with visitors. The visitors receive briefings on NASS sampling, survey, and estimation procedures in the morning and have the opportunity after lunch to witness the security procedures first hand and to receive a detailed briefing on the procedures and specifics of the reports being issued that day.

These briefings have been well received by data users such as farm organizations, agribusinesses, farm publication reporters, and analysts. They also provide an on-going series of contacts for the Headquarters staff and monthly questions about procedures and reports. While fewer suggestions for changes are received during the briefings than at the data users meetings, ideas are occasionally raised that can be acted upon.

Commodity Organization Meetings - Many organizations of the producers of specific commodities have statistics committees within their structure. NASS meets regularly with some of these organizations and has scheduled meetings with a variety of other organizations when the opportunity presented itself.

Statistics committees often have very specific questions about data series, timing, and data definitions. Changes made in response to these meetings with commodity organizations are highlighted below under discussion of format and timing changes.

Other Communication Efforts - Another major communications effort with customers involves attendance at national farm and commodity group meetings. Displays of NASS products and specifically prepared information brochures open discussions with data users and providers. NASS also participates in various types of outlook meetings where analysts discuss current production issues and share their forecasts on upcoming trends. NASS data form much of the underpinnings for their analyses and the analysts often offer suggestions for improving data series.

New Distribution Procedures Case Study

One of the best illustrations of NASS working with customers for unique solutions comes from its reports distribution contract. When government agencies went to

a pay-for-publication policy many subscribers were disappointed with service and choices available through the Government Printing Office (GPO). NASS had about 40 individual subscription choices in order to allow customers to select only those reports they desired. However, under GPO regulations, each series required a separate payment each year unless the organization did a large volume of GPO business and set up a drawing account (which still required considerable paperwork).

The main topic at listening meetings at that time was criticism of the GPO arrangement and suggestions of needed features. NASS, working with the Economic Research Service and the Assistant Secretary for Economics of USDA, was able to get specific legislation included in the next farm bill to allow NASS to distribute reports and keep the proceeds. An outside contract was issued for a vendor to handle reports of the economic agencies. This ERS-NASS service allows subscriptions of up to 3 years and many different titles can be ordered with one payment. The service has a 1-800 telephone number for convenience and it now handles historic electronic products and facsimile transmissions. Reports are still printed by GPO, but ERS-NASS distributes them the same day or next day when printed.

Two different survey approaches were tested for direct followup customers. The Kansas SSO performed a test using short mail questionnaires. Telephone followup was used for customers who requested data through the Headquarters 1-800 hotline. In both cases, the focus was on communication as well as whether the person's data needs were answered. Since the State office test would not have human interaction, it was designed as seven questions to be rated on a 5 point (poor to excellent) scale. The telephone followup had three questions with a similar scale, other "yes, no" questions, and a request for additional narrative comments.

In the Kansas test, about 80 individuals were chosen both in January and in April who had recently received information. The response rate was about 50 percent which was very encouraging for a mailout-mailback approach. The sample size for the Headquarters specific 1-800 hotline was 50 people who had called the previous month. This sample was in addition to making a few follow-up calls on a regular basis, particularly to callers who were transferred to other agencies.

In the case of data analysts, a formal Data Users meeting was held in October 1995. The meeting specifically involved discussion of a proposed new Agency estimation program which would change timing and content of a number of reports.

During the year, NASS implemented and finalized new reporter access procedures for major reports. Two meetings were held with all participating news organizations. (All 15 news services participate in one release each month and most participate at least weekly.) Considerable informal feedback was also received during the year and some changes in procedures were made because of those questions and suggestions.

Written responses were received from a number of organizations which participated in lockup briefings. NASS did not send a specific inquiry but one organization polls its members with an evaluation form after each visit.

Input on our service to State commissioners and directors is received throughout the year by NASS State Statisticians. In addition, the managers of our Field Operations Division and the Agency Administrator receive feedback through a series of national and regional meetings conducted each year by the National Association of Departments of Agriculture.

NASS customer service responses to the Kansas and Headquarters inquiries were very encouraging. In Kansas, using a 5-point scale, with 5 a excellent and 3 as good, all seven questions received at least a 4.4 average. The highest average numeric reading of 4.9 was received for both courtesy and willingness to respond to questions.

In the Headquarters followups, which use a 4-point scale, almost all responses were good to excellent to the three rating questions. In fact, the only ratings below "good" were from individuals who objected to the fact that we took time to ask for their name and address so that we could conduct the follow-up survey. They felt their question was so simple that a follow-up would not be necessary. Almost all answers were "excellent" to the "was your contact pleasant" question.

One sidelight to the telephone hotline was that many callers objected to our telephone system's intermittent beeping sound when they were placed on hold while the operator verified that a commodity specialist was available. Arrangements were made to connect a radio playing classical music instead, which makes the waiting time seem shorter.

We also learned there are people who desire quicker access to our data who do not currently have access to Internet. To address these needs, we recently installed an Autofax system to activate another means of timely information delivery to the public.

Considerable input was received from electronic data users through the comment

feature on the NASS Home Page, other electronic mail responses, and telephone inquiries. A number of changes have already been made based on feedback. Some customers said it was confusing to determine how to order products, so a "hot button" approach is being added for the next Home Page update. Customers also indicated that, because of the nearly 400 reports per year, it took time to get to the report they need. A separate "Today's Reports" feature is now being added to our Home Page. We made improvements for our current CD-Rom of *Agricultural Statistics* to make it easier to download individual data tables but some people wanted to download an entire chapter so we made the necessary changes.

Internal Customer Measurement

Climate Survey

The 1,330 employees at NASS as of November 1, 1993, (1,094 of them Federal and 236 State employees) were given a self-administered, anonymous survey. To encourage employees to complete the survey, its cover memorandum was signed by the Agency Administrator. It asked employees to "provide an honest appraisal of the present working climate." It also stressed that results would help in identifying problem areas in the Agency. One copy of the survey form was distributed to each employee (including State employees). Individual employees who reported a lost or misplaced copy of the survey questionnaire received replacement copies. Reminders to respond appeared in the Agency's monthly Staff Letter, and cc:Mail bulletin board messages.

Employees returned 835 questionnaires (63 percent of the 1,330 distributed). Exluding State employees, who may have been less inclined to respond, this rate was 739 of 1,094, or 68 percent. (The 1990 rates were 66 percent overall and 70 percent, excluding State employees,) Survey repondents generally were similar to all NASS Headquarters and SSO employees, based on their job series, work location (Headquarters or SSO), and number of years at NASS.

Informal Customer Service Feedback

NASS uses its survey-specific evaluation forms to help measure how well Headquarters delivered service to the field office statisticians and to the enumerator staff. The survey evaluation forms are designed to provide an overview of survey operations. Included are comments on response rates, counts of quality control contacts, and survey comments and recommendations relating to specific subjects. These forms help describe any problems with survey materials, including quantity received, sampling concerns, edit limits and other

pertinent information. The feedback from NASS field office statisticians and enumerator staff are circulated to each Headquarters Branch and Section responsible for the subject. These suggestions and recommendations for changes to any aspect of the survey are given full consideration and are a very important tool that NASS uses to improve quality and service.

Technical Reviews

NASS appoints Technical Review Teams (TRT) for technical review of the 45 State Statistical Offices (SSO's) and Headquarters (Washington, D.C.) activities. This team examines the methods and procedures used to carry out theAgency's statistical programs. The Agency's goal is to review all SSO's within a 5-year period. This review provides for a way to obtain feedback, recommendations and suggestions on technical procedures and methods which only NASS employees could provide. It also allows for ideas and improved methods to be shared and passed on to other employees.

Customer Service Stories

Our favorite story of the year is the hearing aid call. It demonstrates positive success, it truly helped an individual, and it illustrated how we want our employees to pursue difficult questions. We received a 1-800 hotline call from an individual who said his father, who farms, needed a hearing aid and he understood that USDA had a program which could help. That was not familiar to our person staffing the telephone (nor would any of our trained staff members have known the answer--then). Our person took the caller's name and telephone number and started on a search within USDA. She had no success with the first agencies she contacted, and even got "you must be wrong" type responses. However, she then found someone who had an inkling of some special Extension Service Programs. That was the right answer; there is a program to help farmers with disabilities obtain appropriate accommodations. We were able to inform the caller about the program within 3 hours.

Lessons Learned

We have learned we are on the right track both with efforts such as our 1-800 hotline and with the training program established for staff members who take those calls.

We have also learned that people calling for information do not want to take extra time to give their name and telephone number--if we can give them an answer while they are on the telephone. One reason might be that people may have been

bounced from one agency to another without getting answers and are skeptical.

Another conclusion is that we need to do any followup contacts soon after we have helped a customer. Calling people back even a month later meant that some people did not remember they had called us. (We assume that means we did such a good job that the contact seemed natural and didn't leave a lasting impression.) We also need to shorten the number of questions on follow-up telephone calls. Even very satisfied customers did not want to take much time. More questions can be asked on a mail out, fax out, or e-mail out questionnaire since the respondent can see the whole survey at one time.

We were reminded that the wording can mean different things to different individuals. We improved our annual *Agricultural Statistics* publication so much with more current data that we named it 1995-96 rather than 1995. We recently heard from a data user who incorrectly assumed that meant we will now only issue it every 2 years.

The final lesson was to realize that efforts to provide good service of one type can cause problems elsewhere. One example of this comes from working with reporters who have advance access to reports in lockup so that they can be prepared when telephone are turned on at release time. We provide copies of the relevant NASS reports, the World Agricultural Outlook Board (WAOB) report, and electronic diskettes. We felt having one diskette with all the reports would be the best customer service. However, that added one extra last minute step and created stress between NASS and WAOB in accomplishing the goal. The two agencies worked together and found the best solution was for each agency to distribute diskettes containing their report(s).

Attachments

CUSTOMER SERVICE ASSESSMENT and INFORMATION COLLECTION PLANS

Agency _____ Date _____

Division/Unit _____

Contact person(s)_____Phone_____ FAX_____

_____Phone_____ FAX_____

CUSTOMER SERVICE NEEDS ASSESSMENT:

Please answer the following questions. Be specific.

1. What are you Agency's products or Services? _____

_____

2. Who are the customers of your Agency? _____

_____

3. Has your Agency developed a customer service plan? □ Yes □ No

4. When do you expect to begin collecting baseline information of the services your Agency offers? _____

INFORMATION COLLECTION PLANS TO EVALUATE CUSTOMER SERVICE:

5. Do you have specific survey plans at this time? ..... □ Yes..... □ No

   a) Which of your Agency's products or services are to be addressed in this survey?

   _____

   _____

   b) In this survey, do you want to know about: (check one)

   □ your entire customer base?

   □ selected subgroups only? (Specify: _____)

   c) What exactly do you want to know from this survey?_____

   _____

   _____

   d) How do you expect to measure it? _____

   _____

   e) How will this information be used in your program or activity to improve customer satisfaction?

   _____

   _____

INFORMATION COLLECTION RESOURCE MANAGEMENT:

What types of descriptive or identifying information do your records contain about your customer? (Check all that apply)

- ☐ Name (person)
- ☐ Telephone number
- ☐ Organization Name
- ☐ Address
- ☐ Social Security No.
- ☐ County
- ☐ Other (specify: _____)

a) Where are your customer records maintained? (Check all that apply)

- ☐ Headquarters
- ☐ Local Offices

b) Are these records: (Check all that apply)

- ☐ Stored on mainframe?
- ☐ Stored on PC, microcomputer, or LAN?
- ☐ Not stored in machine media?

7. In what areas would you want NASS to assist you? (Check all that apply)

	Advice, consultation, and/or review	Conducting the task
Sampling	☐	☐
Questionnaire design	☐	☐
Data cleaning and processing	☐	☐
Estimation and summary	☐	☐
Data analysis	☐	☐
Other (Specify: _____)	☐	☐

Please return this form to Sam Rives at USDA-NASS, Room 4162 South Building, or call at 720-2248 or FAX to 720-6396.

Thank you.

ATTACHMENT 2

The Steps in a Customer Survey

Step 1	Determining the Scope of the Survey Measurement Program
Step 2	Identifying the Factors and Characteristics that Underlie Customer Satisfaction
Step 3	Identifying the Target Customer Population for the Survey
Step 4	Developing a Sampling Frame of the Target Customers
Step 5	Choosing a Data Collection Method Best Suited to Your Customer Survey
Step 6	Choosing Who Will Collect the Survey Data from Customers
Step 7	Developing and Pretesting the Survey Questionnaire
Step 8	Constructing the Statistical Design of the Sample of Customers
Step 9	Designing Procedures to Achieve High Response Rates in the Customer Survey
Step 10	Ensuring Quality While the Survey Data are being Collected
Step 11	Processing the Survey Data and Preparing them for Analysis
Step 12	Analyzing the Data, Summarizing the Results, and Pretesting the Findings

DATA COLLECTION COSTS RULES OF THUMB

Following are estimated cost ranges per attempted contact for the different modes of data collection. Also listed, are some factors that could vary considerably and will impact the overall survey costs.

Mode	Cost per Contact	Factors to Consider
Mail	$1 - $5	◆ Postage (rates vary by class) ◆ Size of envelope & weight ◆ Address correction requested ◆ Number & type of reminders ◆ Non-response follow-up ◆ Pre-survey Notification Letter
Telephone	$8 - $15	◆ Length of Questionnaire ◆ Salary Rates (Interviewers & Supervisors ◆ Long Distance vs. Local call charges ◆ Time of day calls are made ◆ Quality of phone numbers ◆ Target Respondent ◆ Training ◆ Quality Control ◆ Pre-Survey Notification Letter ◆ Number of Call Backs ◆ Call Management System ◆ Auto Dialing Capability
Face-to-Face	$25 plus	◆ Salary Rates (Interviewers & Supervisors ◆ Mileage & Per Diem ◆ Training ◆ Interview Length/Length of Questionnaire ◆ Call Back Plan/Scheduling ◆ Distribution of sample relative to interviewers ◆ Quality Control ◆ Field Data Purview ◆ Field Supervison ◆ Time of Day Visit Made ◆ Sample Size per Interviewer ◆ Target Respondent ◆ Quality of Addresses ◆ Pre-Survey Notification Letter

Measuring Customer Satisfaction with Census Bureau Products

Presented by
John Kavaliunas
Assistant Chief, Marketing Services Office
U.S. Census Bureau

for the
Seminar on Statistical Methodology in the Public Service
Bethesda, Maryland
November 12, 1996

Since 1993 the Census Bureau has conducted 11 surveys to measure customer satisfaction with specific products and product lines. Our purpose in doing so is to continually upgrade and improve our products, find out more about how customers are using these products, and to develop ongoing measures of customer satisfaction.

Through these surveys we've done we've learned a lot about taking customer surveys, as well as customer likes and dislikes. We've gathered some market intelligence, and we've been able to monitor trends in data access and user preferences.

Some Background

We usually focus on a particular product or product line and survey identified customers of that product. Our universe ranges from several hundred to several thousand customers. In addition to paying customers, we generally include all our state data center lead agencies and a sample of affiliates. We also include a sample of federal depository libraries--usually 1 in 10-- so that no one library or State data center affiliate has to answer a questionnaire more than once every three or four years. We have benefitted from having a customer data base and good working relations with the Government Printing Office. We have not attempted to reach any additional end users that would not be covered directly through the survey, although we found that users share copies of Census Bureau reports with, on average, 10 colleagues.

We try to keep the questionnaires simple with no more than 4 pages, but often just 2, focusing on issues that are specific to the product. However, we have a standard set of questions that are asked on all surveys and that include overall product satisfaction and future media preferences. We have been able to construct a satisfaction matrix and to track acceptance of the Internet as a data dissemination vehicle. Occasionally we've also included questions of a marketing nature, such as how did users find out about particular reports or products. We always leave blank space for write in responses and dry run questionnaires with selected data users beforehand.

We have received a generic clearance from the OMB for customer surveys, as I'm sure many of you have. Basically, we send the questionnaire and a letter to the OMB, specifying our intention to conduct a customer survey. The letter describes the survey, the customer universe and response burden hours anticipated. If OMB doesn't respond within a week, we proceed with the survey operations.

Our response rates range between 25 and 30 percent. I understand from professionals in the private sector world that anything over 10 percent is considered excellent. On several occasions we tried to increase our response rate with telephone follow-up and through remailing the questionnaire with some success. But the limited number of additional returned questionnaires may not be worth the additional efforts, costs and time required to carry out the follow up.

We prepare a fairly lengthy analysis of the survey for the program area. We also do a short 2-page *Marketing Brief* which provides some of the highlights of the survey. We send the brief to survey respondents so that we close the loop and customers can see that something was done with the information they provided. We also distribute the brief throughout the Bureau. An important feature of the document is identifying the product changes that will take place as a result of our customers' input.

Some Specifics

We have surveyed our Internet customers each spring for the last three years. We conducted our most recent survey entirely electronically and surveyed persons who e-mailed us during the month of April or who downloaded files and for whom we could determine an e-mail address. From this survey we discovered that about 20 percent of our Internet users access our site once a week or more. Thirty-eight percent of respondents indicated that they found our site through a search mechanism. And 14 percent indicated that they had purchased a Census Bureau product as a result of finding out about it on the Internet.

Since we began surveying customers about 4 years ago, we've asked a question on future media preferences. In our first survey, in the pre-Internet days of 1993, only 5 percent of respondents indicated any preference for online dissemination. We've seen that percentage grow and in 2 surveys this spring we saw preferences for Internet surpass CD-ROM as the data dissemination media of choice. In our recent Internet survey --a high tech, Internet-friendly universe-- 74 percent said Internet was the preferred media for accessing Census Bureau information.

We also discovered through two surveys this spring --one of a print product and one of a CD-ROM product--that at least 4 out of 5 respondents had access to the Internet, but only half of these had actually accessed the corresponding files on the Census Bureau Internet site.

We have also asked a question on general satisfaction and satisfaction with various aspects (such as timeliness, reliability and value) of our products and have constructed a user satisfaction matrix. Overall satisfaction has run from 72 percent on the first *Statistical Abstract* on CD-ROM to 94 percent for users of our *Current Population Reports*. (The satisfaction rating for subsequent *Statistical Abstracts* on CD-ROM has gone to 77 percent.)

Some Practical Results

It does little good to solicit customer opinions unless you do something with the information. The program areas at the Census Bureau have been very accepting of user suggestions and have instituted product changes and enhancements as a result.

A couple of examples:

- Users did not like the software we had originally included on the *1993 Statistical Abstract* on CD-ROM. We added Adobe Acrobat -readable files to later versions.

- Users wanted more timely delivery of our foreign trade data. We offered expedited order fulfillment with pick-up on the same day as the press release.

- Users of our printed *Current Population Reports* told us they were unaware of similar information on the Internet. We will be adding references to Internet addresses in the printed reports and press releases.

- Users of our Economic Census data told us that they used information from the various economic sectors, rather than just one sector. With the 1997 Censuses, we will introduce a new report series with cross-sector data.

Some Plans for the Future

We plan to continue to survey users of specific products or product lines with three targeted customer groups this coming year. We also plan to survey Internet users in the spring of 1997.

An interdivisional team is currently reviewing our customer service standards. Once approved and publicized, we will gather some benchmark data through an independent customer survey on how we're measuring up to these standards.

We are also in the planning stage of developing a bureau-wide customer comments data base so that write-in comments, letters, and other feedback from our customers can be recorded, characterized, and communicated back to the divisions that produce

the data and throughout the Bureau.

Some Lessons

Here are a few observations, based on our work, that can help you with your customer surveys.

- Keep it simple. Keep it focused. Sure it would be interesting to know some things about your users, but if you can't use the information for product improvement or marketing, don't ask it.

- Beta test your questionnaire with persons outside the organization. What seems obvious to someone inside your organization may not be so clear to users on the outside. And outside beta testers may identify issues that you hadn't even considered.

- Don't send unsolicited questionnaires through the Internet unless you let potential respondents know beforehand that the questionnaire is coming.

- Ask the same questions on different surveys and at different times. This will enable you to compare across products and measure improvements.

- Give feedback to the rest of your organization and to respondents. Customers who know that their suggestions and opinions are taken seriously are more likely to respond to future inquiries.

In Summary

We've made substantial progress at the Census Bureau in last several years in knowing something about our customers. We have good information on how they find out about Census Bureau products, how they access Census Bureau data currently and how they would like to do so in the future.

We have used the customer survey process to measure customer satisfaction with specific products and to make changes to those products to better meet the needs of our customers.

We will continue to use this process, as well as other mechanisms, to gather customer input, so that the Census Bureau can provide products and services that are of value to our users.

WHEN TO LISTEN AND WHEN TO MEASURE

comments by Jerry Coffey

These two papers have a lot to say about the learning process most agencies have been through in the last year or two. Sam Rives' paper emphasized the importance of accurate measurement, while John Kavaliunas mentioned a number of activities with response rates so low they couldn't measure much of anything. I should point out that statisticians don't have a copyright on the term "response rate" -- those "professionals" who bragged about a 10% response rate were undoubtedly professional marketers. Many years ago, at the U. S. Postal Service, I got a lot of exposure to the mail marketing community, and they define "response rate" as the number of sales divided by the size of the mailing.

"Customer Surveys" include a wide range of activities from the rigorous measurement processes that Sam described to simple feedback processes where "response rate" is not a consideration. When the Office of Management and Budget released its "Resource Manual for Customer Surveys" in 1993, one of the things we emphasized to agencies trying to improve customer service was that "listening to your customers" does not necessarily mean sitting at your desk drafting questionnaires.

Statistical agencies have useful experience to offer in all of these activities. Meetings with data users and other forms of communication developed by statistical agencies have become effective tools for listening to customers. This kind of feedback provides critical insights into those attributes of your products or services that customers value. At that point, the statistical agency's methodological expertise can be effectively applied to measure the right things and measure them accurately.

To use this experience and expertise effectively, the statistician must first understand the task and then match the method to that task.

WHO IS THE CUSTOMER FOR YOUR CUSTOMER SURVEY?

1) Suppose the information collected is exclusively for the use of front-line manager or employees.

Such a survey may be nothing more than a simple feedback device, a tool to enhance communications, one of the many ways to "listen to customers." The object of such an exercise is usually to make it easier for customers to communicate their complaints, suggestions or other ideas to those who immediately serve them. Such tools do not *measure* anything, and thus need not be designed to support a measurement process.

2) Back away one step and suppose that the "customer" for the survey needs data collection that supports some sort of comparative measurements.

Now we are supposing a real measurement process that must be repeatable (measure the same thing at different times and for different groups of respondents) and reasonably complete (does not miss significant bodies of opinion or activity). The "customer" for these kind of results is typically a manager trying to *track the performance* of his own organization or subordinates (who may themselves be managers of other units within the organization). These more stringent requirements demand more *control of the data collection process* (e.g., a rigorously designed and implemented sampling process). If comparison over time is the only need, then there is still flexibility in choosing the measurement method (it only needs to be consistent over time). If comparisons *across* units are needed, then there is another constraint -- comparable measurement methods must be used in all the units to be compared.

3) Back away one more step and suppose that the "customer" is outside the agency bureaucracy.

This kind of use is implied by the Government Performance and Results Act and by some parts of the Customer Service Executive Order. The "customer" here is generally the Executive Office of the President (e.g., OMB) or top agency management or the Congress, who must have measures that support comparability across Departments or agencies. This case requires all the rigor of the second category above plus the use of measures that are consistent across all the units that are to be compared -- leaving very little room for flexibility sought in the first category.

Session 5

ONE-STOP SHOPPING FOR FEDERAL STATISTICS

One-stop Shopping for Federal Statistics

Alan R. Tupek

National Science Foundation

Background

It is difficult for the general public, and even frequent data users such as social science researchers, to know about and to access the extensive amount of statistical information produced by the decentralized U.S. federal statistical system. Publications, such as the *Statistical Abstract of the U.S.*, have provided access to a limited amount of statistics from the federal statistical system. With the growing acceptance of the World Wide Web (WWW), there is an opportunity to provide access to the broad range of statistics of interest to the public. Federal agencies have, for the most part, made tremendous progress in developing easy access to their data and databases through the WWW. Data users accessing information from one Federal agency may even find out about related statistics available from other agencies. However, these efforts have been very limited to date. The technological barriers no longer exist that once made it difficult to have a coordinated effort to allow users to locate the statistics that are most appropriate for their uses.

In May 1996, the Office of Science and Technology Policy (OSTP) and the Office of Management and Budget (OMB), along with representatives of the major statistical agencies released the Economic Statistics and Social Statistics Briefing Rooms that can be accessed through the White House WWW Home Page. This initiative provides the general public easy access to key current economic and social statistics and to the approximately 15 agencies that provide these statistics.

A Task Force on One-stop Shopping for Federal Statistics was formed under the auspices of the Interagency Council on Statistical Policy (ICSP) in November 1995. The Task Force is chaired by the National Science Foundation, with representatives from the Bureau of the Census, the Bureau of Labor Statistics, the Bureau of Economic Analysis, the Economic and Statistics Administration, the Energy Information Administration, and ex *officio*, the Office of Management and Budget.

The Task Force was formed as a culmination of several activities. These included the early development stages of the Briefing Rooms, a November 1995 OMB report on Electronic Dissemination of Statistical Data, meetings of an ad hoc interagency group on dissemination of statistics on the Internet, as well as initiatives of the

Office of Statistical Policy at OMB. The Briefing Rooms would provide access to a limited amount of Federal statistics, but would also provide an avenue to access the larger array of statistical information provided by the Federal Government. The report on Electronic Dissemination of Statistical Data provided a guide to survey managers on issues to consider when expanding the modes of data dissemination. The report highlighted the broad range of dissemination modes used by Federal agencies and drew attention to the need for better coordination of dissemination practices. The ad hoc interagency group on dissemination of statistics on the Internet had been monitoring the various practices agencies were developing for Internet dissemination. This group was especially concerned with dissemination of metadata and also saw the need for interagency cooperation to provide easier access to all Federal statistics. Finally, but certainly not least, monthly meetings of the Heads of Statistical Agencies chaired by the Statistical Policy Office at OMB had just been sanctioned as the Interagency Council on Statistical Policy (ICSP). As one of its first official acts, the ICSP formed the Task Force on One-stop Shopping for Federal Statistics.

Task Force Charge

The initial charge to the Task Force was to examine the Federal Government's various systems for disseminating statistical data, and to provide recommendations to the OMB on the structure and implementation of new one-stop shopping mechanisms for users of Federal statistics. The charge to the Task Force includes: (1) identifying the major statistical series to be incorporated into the One-stop Shopping mechanism; (2) identifying and developing new concepts, structures, and methodologies to facilitate the dissemination of Federal statistics; (3) developing and testing a mechanism to provide one-stop shopping for users of Federal statistics; (4) planning and implementation of the new one-stop shopping mechanism; and (5) ensuring that there is ample opportunity for widespread public participation of stakeholders in the development process.

As one approach to its charge, the Task Force should oversee the design, development, implementation, and maintenance that would provide easy access to Federal statistics through the WWW. The Task Force will develop and monitor contracts or other mechanisms to develop, implement, and maintain the agreed upon approach.

The Task Force should adopt processes that ensure ample opportunity for public participation. These processes should involve all stakeholders, including the range of Federal data users, both government and private, as well as data collectors and data providers. The Task Force should consider forming a consultation group, composed of Federal agencies not represented on the Task Force. Such a group would meet on a flow basis, as necessary, to provide input to the work of the Task Force. Notice of the Task Force's work should be widespread and should be

published in the Federal Register for all interested public and private parties. Interested parties will be given the opportunity to be included on a mailing list.

Since its inception, the Task Force has solicited opinions from various communities on its prototype development. Task Force members have made presentations on the One-stop development work in a variety of forums, including meetings of the Association of Public Data Users, the Council on Professional Associations on Federal Statistics, the Federal Web Consortium, and the American Statistical Association. In September 1996, the Task Force asked each of the agencies on the ICSP to appoint a liaison to the Task Force. The liaisons have been asked to provide input to the prototype development of the One-stop site.

Features of the Prototype Web Site

The purpose of the WWW site is to permit easy access via an initial point of entry to the wide array of Federal statistics of interest to the public. The site is intended to supplement the excellent WWW sites that individual agencies have developed for disseminating Federal statistics. Other currently available WWW services do not adequately meet the objectives of One-stop Shopping for Federal Statistics. For example, the Economic Statistics and Social Statistics Briefing Rooms provide access to about 100 key statistics from a dozen or so Federal agencies. STAT-USA does not cover the full array of Federal statistics and its fees may limit the potential audience. A few academic sites provide access to some of the Federal statistics that are produced by the larger statistical agencies. It is also possible to find a limited amount of Federal statistics through Government and commercial WWW search services, such as GovBot, Excite, Yahoo, and Infoseek. However, using these services to find Federal statistics often results in many more links to non-relevant information than to links to Federal statistics. It is especially hard to find Federal statistics from those agencies whose entire mission is not statistics. The primary objective of the WWW site for One-stop Shopping for Federal Statistics is to help users find the information they need without having to know and understand how the decentralized Federal statistical system works. The WWW site for One-stop Shopping is intended to help guide users to statistics from the 70 or so agencies that provide statistics of interest to the public. It is expected that most of the centralized information WWW sites mentioned above will be able to use One-stop Shopping for Federal Statistics to improve their services.

The prototype site includes the following features --

Subjects A to Z - A table of contents approach to the wide-range of statistics available from Federal agencies. The list of subjects allows users to see the various sources of statistics for each subject area. Links are provided to the statistics on agency WWW sites.

Regional Statistics - A collection of agency WWW sites that provide easy access to state, metropolitan area, and other sub-U.S. geographical statistical information. Most agencies provide a clickable map approach to access these data.

Agencies that Provide Statistics - A list of all agencies that provide statistics with links to the agencies WWW Home Page and to the statistics by subject area for each agency.

Statistical Programs of the U.S. Government - Adapted from the OMB report with the same name, this feature provides links to statistics available from federal agencies within 10 broad topical areas.

Subject Matter Contacts - A collection of agency WWW sites that provides contact names, telephone numbers, and e-mail addresses for questions about statistics.

Statistical Press Releases - A collection of agency WWW sites that includes the latest statistical news releases.

Statistical Policy - Includes links to Federal budget documents, statistical policy working papers, and selected federal register notices.

Site Maintenance

The members of the ICSP recently agreed to continue the development and maintenance of a WWW site for One-stop Shopping for Federal Statistics. The ongoing development and maintenance of the site will be funded through an annual, renewable interagency agreement between the agencies represented on the ICSP and the Bureau of the Census. Listed below are the responsibilities of the organizations involved in this agreement.

Responsibilities of the Bureau of the Census (BOC)

The BOC is responsible for the continuing development and ongoing maintenance of the WWW site for One-stop Shopping for Federal Statistics. Enhancements to the site will be made as directed by the Task Force on One-stop Shopping, operating under the aegis of the Interagency Council on Statistical Policy. The BOC will also be responsible for hardware and software maintenance and upgrades. The BOC will provide –

a separate server for the WWW site with the domain name **http://www.fedstats.gov;**

staff to maintain the server and up-to-date links to Federal statistics and to enhance the site as directed;

usage statistics to allow the Task Force to determine traffic to and from the site and other information needed to assess the usefulness of the site;

a limited feedback mechanism to report broken links and to suggest improvements. Users will be directed to contact the appropriate agency for replies to all other questions.

Responsibilities of the Task Force on One-stop Shopping for Federal Statistics

The Task Force will direct the maintenance and enhancements to the site and maintain contact with agency liaisons to determine the need for maintenance and enhancements.

The Task Force will monitor the usage of the site and develop and implement a public information plan for the site in consultation with the ICSP.

The Task Force will provide periodic status reports to the ICSP.

The Task Force will provide an assessment report to the ICSP at the end of FY 1997. The assessment report will include a summary of the usage of the site, the resources expended to maintain and enhance the site, the enhancements made to the site during FY 1997, the amount of and type of feedback received, and recommendations for FY 1998, including funding options.

Responsibilities of the ICSP Agencies

Each agency will appoint an agency liaison to the Task Force on One-stop Shopping for Federal Statistics. The liaison will provide guidance on changes in an agency's WWW site that would need to be updated on the One-stop site. The liaison will also be responsible for suggesting improvements to the site.

Each agency will provide a link to One-stop Shopping for Federal Statistics on its home page.

Each agency will provide funding to support the ongoing development and maintenance of the One-stop site. The five largest Federal statistical agencies (based on FY 1997 statistical program budgets) will contribute $20,000 each to the Bureau of the Census on an annual basis through an interagency agreement. These agencies are the Bureau of the Census, the Bureau of Labor Statistics, the National Agricultural Statistics Service, the National Center of Education Statistics, and the National Center for Health Statistics. The contribution from the Bureau of the Census will be to provide staff resources in addition to those funded through the interagency agreements. The other agencies represented on the ICSP will contribute $10,000 each to the Bureau of the Census on an annual basis through an interagency agreement.

The funds will be used by the BOC to support one GS-13 (100%), one GS-05/07 (50%).

In addition, the Bureau of Labor Statistics has agreed to conduct usability testing of the site with library science researchers. Results of this work will be implemented, as directed by the Task Force on One-stop Shopping for Federal Statistics.

Also, the National Science Foundation has agreed to purchase the initial hardware and software needed for the site at a cost of approximately $23,000.

FY 1997 Plans and Activities

The site is expected to be publicly available in February 1997. During FY 1997, the Task Force on One-stop Shopping for Federal Statistics plans to direct the following activities --

set up hardware, install software, procure upgraded hardware and software as needed;

through interactions with agency liaisons, identify and implement appropriate links for the one-stop site tools, "Federal Agencies That Provide Statistics," and "Subjects A to Z;"

provide access to agencies' data maps (regional statistics) and press releases;

design and develop usage statistics software and reports;

conduct routine site maintenance through interactions with agency liaisons and evaluation of reported problems;

procure, install, test, and release automated searching software which will allow users to enter their own keywords for searching for Federal statistics across agencies,

provide links to publicly available information about statistical research efforts; and

develop access tools for data maps, press releases, and statistical research that permit cross-agency searching for information.

Future Enhancements

Future enhancements will depend on an analysis of the site usage, user feedback, and the results of the usability testing of the site. Possible new features include: a data dictionary or glossary that describes differences between various terms, thematic search tools, statistical hypernews that allows users to participate in dialogues on statistical issues, and interagency coordination on database tabulation and extraction tools and user interfaces.

Acknowledgments

The members of the Task Force on One-stop Shopping for Federal Statistics are: Paul Bugg (Office of Management and Budget), Connie DiCesare (Bureau of Labor Statistics), Valerie Gregg (Bureau of the Census), Daniel Larkins (Bureau of Economic Analysis), Ken Rogers (STAT-USA), Rachael Taylor (Bureau of the Census), Al Tupek (National Science Foundation), and John Weiner (Energy Information Administration). Others who have made significant contributions to the work of the Task Force include Cherie Vaden and Marion Brady of the Bureau of the Census, Michael Levi and Deborah Klein of the Bureau of Labor Statistics, and John Wells of the Bureau of Economic Analysis. The development of the One-stop site would not have been possible without the leadership of Katherine Wallman, Office of Management and Budget and Chair of the Interagency Council on Statistical Policy.

One-Stop Shopping for Federal Statistics

Discussant: Ann Gray, Cornell University

I am here as the Association of Public Data Users (APDU) representative in place of Joe Salvo who could not attend today. At APDU's recent Annual Conference we discussed this initiative along with many other issues related to electronic delivery of federal statistics. My remarks will draw on the APDU conference. They are intended to be comments, not criticisms. This Task Force has a very ambitious charge and very little resources. They are going to need all the help they can get.

APDU members are very concerned about a group of persons that can be called the "general public" -- persons who lack knowledge of the U.S. federal statistical agencies. We believe that this group should be able to obtain the best measures they need without undue effort. Toward that end, One Stop Shopping should incorporate features that allow subject or topical access to federal data using simplified but accurate terminology. This would include features such as cross references and indexes. In fact, such products already exist both in print and in electronic formats. County and City Databook and Statistical Abstracts of the United States are provento be useful and, in the case of Statistical Abstracts, fairly complete compendia of national level statistics.

I would urge Census to allow free access to Statistical Abstracts, even though they also market it on CD-ROM and charge for it via STAT-USA. Free, on-line use of the product may encourage the purchase of the CD-ROM as more people learn of its existence and the scope of the product.

Another APDU concern deals with the mis-use of data. This concern is not expressed in the notion that access should be limited, but rather that we need to recognize that people do make mistakes. One Stop Shopping should make sure that if mistakes are made, it is done by the end-user and not by the agency or provider. There is a standard disclaimer regarding data distribution that includes statements like "we bear no responsibility for the uses or interpretations or inferences you might make," but what is needed here is clear and ample information about the origin and history of the statistics that are presented.

APDU also encourages the inclusion of multiple levels of user support, including names, telephone numbers or electronic mail addresses of a real live person. This is useful for both novice and sophisticated data users. APDU also believes that the system should include the names and telephone numbers of local experts. Often people want local information that may not be readily available from federal agencies. There are people that are willing to be included in this list. The search of local experts could begin with those associated with the Census Bureau's State Data Center/Business and Industry Data Center program and the Depository Library Program. There may be other existing networks, such as the agricultural extension agencies, that could be brought into the system. I would also add that the media is an important intermediary in providing statistics to the public. They should be remembered in the design of the system.

Next, we need to consider the project for what it is and what it is not. It is not the democratization of information. The Depository Library Program, where everyone has equal access to government information distributed to the depository library, is democratization of information. Not everyone has equal access to the Internet, there are still problems with connectivity and we may never have universal access. This is an extension of the Depository Library Program and, in fact, not every community has a depository library.

This project is an attempt to create a single or virtual federal statistical agency. It should be possible, but there are massive intellectual problems. The success of this project depends upon solving those intellectual problems. It is not a technological problem, no matter how well technology may serve as a network or means of bringing things together. That is to say, the challenge is to organize the information is a way that makes sense.

The work of libraries is to organize information for retrieval. In order to locate information, control must be established. This control is made possible by a set of rules. We call them standards. These standards work in libraries and they work in technology also. In looking at other inter-agency or international efforts we see standards at work.

For example, the International Monetary Fund has set up a bulletin board -- it is really a web implementation -- dealing with national accounts and other economic data from various nations. There is a Special Data Dissemination Standard (copies are available from IMF: Telephone (202) 623 4415, FAX (202) 623 6165, Web: www.imf.org) that covers content, access, integrity, and quality. IMF created a voluntary standard where IMF acts as a centralized source for information about a nation's statistical products, including release dates, media, cost, and other issues dealing with availability. At this time about forty countries have subscribed to the standard. IMF enforces a consistent set of information about the statistics included in the bulletin board. Of course, it must be noted that IMF has a certain amount of influence that the Task Force on One Stop Shopping may not have and that IMF also limits the scope of the coverage for the bulletin board.

In another international effort, the Council of European Social Science Data Archives (CESSDA) has created an Integrated Data Catalogue that is available via the web (www.nad.uib.no/Cessda). Each data archive maintains its own catalog of data holdings and these catalog records include descriptive information about the contents of the data. Integration occurs through the use of field based WAIS indexing. WAIS incorporates a communication standard called Z39.50 that is widely used by libraries to integrate database functions.

Another effort to normalize information about data is being undertaken by a group headed by Merrill Shanks and Richard Rockwell. It is based on the use of SGML to allow automated systems to manipulate tagged text. SGML is also a standard and is, not incidentally, Z39.50 compliant.

Why are standards important? Standards are important because they allow for growth. The web itself is based on a standard and because there was a standard, the web has been able to grow into what we see today. The wide-spread availability of CD-ROM readers and CD-ROM products was made possible by a standard. The Federal Information Processing standards, especially the FIPS Codes for geographic and governmental units makes it possible to find a common identity for creating compendia of statistics about those areas from different sources. Likewise, there is a government standard for identifying government information: the Government Information Location System or GILS. If there is something wrong with GILS, fix it. Don't just ignore it.

Finally I'd like to offer some personal advice on the use of so-called "advanced" technologies and software systems based on the creation of "knowledge bases." It would be a costly waste of time to construct an enormous knowledge based system that will be obsolete by the time it is completed. What is needed to make this system successful is human intelligence and human intervention.

Council of Professional Associations on Federal Statistics
Seminar on Statistical Methodology in the Public Service

Session 5, "One-Stop Shopping for Federal Statistics"
Discussion of presentations by Alan Tupek (National Science Foundation)
and Valerie Gregg (U.S. Bureau of the Census)

Lauris Olson, Van Pelt Library, University of Pennsylvania

As I am merely a reference librarian at a large urban research university, I will restrict my remarks to the actual Internet resource (and its various parts) being presented today, leaving policy concerns to more qualified parties.

How useful is the "One-Stop Shopping" web site? Consider what's probably the federal government's single most popular and widely-used reference work, *Statistical Abstract of the United States*. The reader opens *Stat Abs* and is immediately confronted with numbers: statistical tables grouped in broad subject chapters. It never fails to surprise me that so many of our readers are satisfied with the tables they find in *Stat Abs*, or, if they're not fully satisfied, accept them with a grudging, "They'll do".

Of course, *Stat Abs* is an "abstract" in the bibliographic sense, its tables serving to indicate the contents of more expansive or detailed statistical publications through the source notes appended to each table. And *Stat Abs*'s wonderful appendix, "Guide to Sources of Statistics", serves as a subject-arranged bibliography for the whole volume by listing important U.S. statistical information resources.

An Internet counterpart to the statistical tables of *Stat Abs* has been available for some time as the "Federal Statistical Briefing Rooms" available through the White House's WWW site. As presently configured, FSBR provides current economic indicators and social statistics, sometimes with charts depicting indicators through time. Each indicator includes an hypertext source note, adding value to the presentation by allowing users to link from FSBR to the producing agency's relevant WWW page.

The "One-Stop Shopping" web site presented today seems most successful when viewed as analogous to *Stat Abs*'s "back of the book". That is, the web site is intended to provide an integrated access point to the wide range of federally-produced statistical programs and related products. Thanks to hypertext, the web site transcends *Stat Abs* by linking users immediately to data resources and documentation.

But to continue the *Stat Abs* / FSBR-"One-Stop Shopping" comparison is to show the latter's weakness. *Stat Abs* presents statistics gathered by non-federal sources; its "Guide to Sources of Statistics" includes federal and "other" statistical publications. The "One-Stop Shopping" Task Force's charge reasonably excluded these non-government data resources. But data dissemination partnerships within the federal government and between federal producers and nongovernmental distributors are proliferating. It's reasonable to link to "1990 Census Lookup" and other decennial census resources developed at Lawrence Berkeley National Laboratory, or to Cornell University's USDA statistical site. ICPSR now provides public access to the National Archive of Criminal Justice Data; CIESIN's Ulysses server pioneered decennial census microdata access; Right-To-Know Net has constructed interfaces for HMDA and other data sets. I worry that tracking these, let alone the indicators recently devolved from the Census Bureau to the Conference Board, are beyond the scope of the Task Force's charge.

The "One-Stop Shopping" web site will become one among several starting points for identifying and locating federal information. Most of these web sites are barren directory lists of agency web sites. "One-Stop Shopping" will join the handful of web sites providing multiple access tools. An instructive comparison can be made between "One-Stop Shopping" and the U.S. Government Printing Office's "GPO Pathway Services", which strikes me as possessing the best developed toolkit for finding federal information on the web.

Both "One-Stop Shopping" and "GPO Pathway Services" offer links to resources through subject arrangements. Of course, the GPO's web site attempts coverage of all federal government information Internet resources. But the two sites display significant functional differences. "One-Stop Shopping"'s "Subjects A to Z" links are selected and indexed by data-producing agencies. The collaborative nature of the project suggests that coverage will be more detailed and more comprehensive than GPO's offering. On the other hand, GPO's reliance upon the 200 or so durable subject headings of the venerable *GPO Subject Bibliographies* as applied by experienced indexers, makes its "Browse by Topic" links much easier to use, as does the skillful use of abstracts describing agency focus and web site content. I was stymied by the arcane jargon substituting for subject headings in "Subjects A to Z": "OASDI"? "Contingent workers"? "Displaced workers"? [Some alarming omissions, viz. "Immigration", probably reflect the current web site's prototypical nature, restricted to Interagency Council on Statistical Policy agencies.] Perhaps "One-Stop Shopping" should revert to the broad topics used by *Stat Abs*'s "Guide to Sources of Statistics".

Keyword searching of the content of federal information WWW pages is provided through the "GPO Pathway Indexer", whose GOVBOT-related Harvest software scans 22,000 links on more than 780 Internet servers in the "gov" and "mil" domains daily. It's safe to assume that a similar web searcher

implemented by "One-Stop Shopping" would be restricted to scanning a small number of federal Internet servers known to hold statistical information. Perhaps the collaborative nature of the "One-Stop Shopping" program will standardize the use of HTML indexing tags on relevant web pages produced by participating agencies to increase search result relevance. But these refinements won't matter: given the present interest in locking statistical information behind passwords or within WAIS servers, I suspect that keyword searching for specific pieces of federal statistical information will be useless. A similar effect has blighted the GILS initiative, albeit from a different cause, that no one server holds a copy of every GILS record makes searching that genre of fielded record impossible.

The "One-Stop Shopping" web pages on statistical programs reflect changes in public perceptions of federal statistical data and their availability that relate to the acceleration in electronic dissemination beginning in the late 1980s with the Census Bureau's CD-ROM data releases. At our reference desk, it's not just a matter of readers asking for a summary tape file or a printed report; we now meet readers asking for the Current Population Survey to reproduce analyses published in *Current Population Reports*. I hope that the web site's interest in program information moves participating agencies to make available online documentation including data dictionaries or queriable databases of program questionnaire data items. [A similar "reference shelf" approach seems to have been adopted in "One-Stop Shopping"'s "Federal Statistics Policy Documents" web pages, which link to Federal Committee on Statistical Methodology working papers and relevant *Federal Register* notices. Why not expand this to include, for instance, the NAS/NSF decennial census reports?]

The importance of collaboration in making "One-Stop Shopping" a success is most apparent in the underdeveloped web pages on regional statistics. At present, links are provided to agency web pages that serve as directories of agency regional offices. But our reference desk readers want regional, local, and small area statistical information, regardless of its bureaucratic birthplace. The Census Bureau acknowledges this universal demand with the American Community Survey and its recently advertised county data CD-ROM compendia. I hope that "One-Stop Shopping" will be able to provide more refined access to regional statistical information, perhaps by linking individual regional office web pages through geographic hierarchies.

I've examined the principal features of "One-Stop Shopping for Federal Statistics". There's much that's praiseworthy: I would certainly use it in my daily work assisting researchers, teachers, students, and the general public to identify and locate statistical information. But it doesn't matter which feature of "One-Stop Shopping" is examined: the key to the web site's success is through collaboration among its contributing agencies.

LO

Session 6

REVIEWING AND REPORTING QUALITY IN SURVEY DATA

Quality Declarations at Statistics Sweden - Principles and Practises

Claes Andersson, Håkan L. Lindström and Lars Lyberg

Statistics Sweden

Abstract: The general principles for quality definition and quality declaration at Statistics Sweden are presented. Their development over the last two decades is discussed in the light of an increasing concern for the users of statistics. Some background is given to explain how variation in ambitions, techniques and resources have changed the possibilities to study and measure quality. For the major quality component accuracy the statements of the quality level of the subcomponents are presented one by one. We mention general approaches to promote measurement and presentation of product quality. Finally we give some examples of products with good quality declarations.

Key words: Quality declaration, accuracy.

The Development of the Quality Concept

The production of high quality statistics has always been a concern at Statistics Sweden (SCB). Quality, as we see it, is a vector of components. Together these components describe the quality but they cannot all be measured in a quantitative way and therefore they cannot be added to each other to form a measure of the total quality.

Different sets of principles have been applied for different groups of products. In 1979 some general principles were formulated for quality definition and quality declaration at Statistics Sweden. There was a strong recommendation from the chief statistician that these should be followed by all surveys. At that time the concept focused on the effects on estimates of those procedures that influence the mean squared error, i.e., strategy, data collection method, and measurement, nonresponse, data entry, editing, coding and estimation models. The general recommendation was supported by more detailed definitions and recommendations in specific areas. Special decisions were made regarding mandatory reporting of nonresponse, variance calculations and questionnaire testing, i.e., three specific areas. The basic principles were quite general and should be applied on censuses and sample surveys based on administrative records or survey data.

But other aspects of quality were not totally forgotten. Also the content of the survey, the comparability with other surveys and time aspects were included in the general principles. The main reason why these aspects were less elaborated upon is obviously that the work with the quality declarations was led by statisticians while subject matter specialists were less interested in quality issues at that time. Another reason may be that the production of

official statistics at that time was highly concentrated to Statistics Sweden. Most of our statistics were then centrally financed by the government. Other clients were less important. Ideas on user or customer satisfaction were virtually non-existent.

Eventually Statistics Sweden has become more and more dependent on the market. The responsibility for official statistics has been transferred to about two dozen agencies since 1994 and Statistics Sweden is one of them, albeit the largest. Now only about half the volume of our work is funded by the government and the other half by other clients. In most cases we are competing for the contracts. This has forced us to turn to a broader definition of quality, highly influenced by ideas from total quality management. Current important principles are:

(i) **The focus is on the user**. A product's quality is determined by the user's opinion of the product and its usefulness. The user's opinion should direct the approach of the development work.

(ii) Quality encompasses **all characteristics of a product** (commodity or service) influencing how well the product satisfies the user's needs and expectations.

With this definition quality has a **descriptive meaning from the producer's point of view**. This is why we use the concept **content** and do not speak about relevance. The **user makes the judgement** and decides if the quality of the product is good or bad in relation to his/her intended use of it. It is essential for the producer to be aware of the quality judgements of existing and potential users, since these judgements provide a basis for development leading to higher quality, which in turn leads to increased user satisfaction.

One consequence of this point of view is that the user have to take more responsibility for the level of quality and that resources are set aside to reach it. Before 1994 the producer felt this responsibility and could argue with the government about the resources needed. The user could expect the producer to provide good accuracy within the limit of given resources. One can hope that the new situation will promote more intense discussions.

The development of the present quality concept was preceded by long discussions within SCB in order to make it both useful and generally accepted. Quality can be seen as a vector consisting of four main components: **Content**, **Accuracy**, **Accessibility,** and **Timeliness and periodicity**, which are broken down into 23 subcomponents. There was a fairly common acceptance of the subcomponents *per se* but opinions varied on the best grouping into main components. Especially the subcomponents on time and comparability were under intense discussion. The subcomponents are given below:

Content:	**Timeliness and periodicity:**
Statistical entity: • Type of unit and population • Variables • Types of statistical measure • Study domains Comparability with other statistics	• Time of reference • Length of production time • Punctuality • Periodicity • Comparability over time
Accuracy:	**Accessibility:**
Overall accuracy Sources of uncertainty: • Coverage • Strategy (sampling and estimation) • Measurement • Nonresponse • Processing • Model assumptions Presentation of uncertainty measures	• Forms of dissemination • Presentation • Documentation • Access to microdata • Information services

All producers of official statistics now have to follow these principles.

It is obvious that the quality information will be different for different subcomponents. Especially it is difficult to produce information on the accuracy subcomponents in the way users want. The users want to know the quality of the product and are less interested in the process leading to the product. Still we often have to compromise since process information is what the producer can offer and also has to offer. The information on accuracy that is offered can be classified in the following levels:

1. Quantified quality of the product, like evaluation results, variance calculations, response variation.
2. Quantified process indicators, like nonresponse and editing rates.
3. Generalized knowledge on error tendencies from "comparable" surveys.
4. Process descriptions like coding or editing rules.
5. Common sense conclusions/Vague knowledge - for example about the presence of a black market economy, car accidents not accounted for, etc.
6. No knowledge whatsoever of the quality.

Other than saying that level 1 is what everyone should strive for and level 6 is unacceptable, there is no absolute ordering between the other levels. In practise we get more information about the process quality that is useful

mainly for the producer than we get information about the product quality needed by the user. The computerization of survey processes will easily provide us with more and more process information. This is much cheaper and easier to come by compared to information on product quality.

The widening of the quality concept mirrors how the users are given more and more influence on official statistics in Sweden and, as a consequence, how "new" quality components increase in importance. Accessibility is a component of great importance for many users. They tend to prefer to produce their own statistics and ask mainly for edited files or easily accessible data in databases. The pressure on fast production is also growing and so is the demand for decreased costs. Demand for a high-level quality regarding one component is often in conflict with the possibility to maintain quality regarding other components. The demand for fast production will make it harder to obtain an acceptable response rate.

One important issue is how to use of the quality declaration. Statistics Sweden is no longer the only producer of official statistics from the collected data and is consequently no longer in a position to advise on its use. A quality declaration cannot be written to suit all situations since data files and statistics will be disseminated in many ways and also further processed by different users. We have to provide the potential users with information that makes it possible for them to derive the quality of their own statistical products. It must also be made obvious that it is their responsibility to declare the quality of their products.

In the near future we anticipate an increasing demand for statistics to be standardized in order to make comparisons and mergings with other statistics possible, especially with those produced by international organizations. International comparability may for example be in conflict with the "the most relevant" content in a survey or in a country.

Organizational and Business Aspects

Statistics Sweden has become a hybrid. We are an agency responsible for our share of statistics funded by government together with half a dozen other agencies with identical responsibilities. Most of the surveys that Statistics Sweden is responsible for concern areas where there is no obvious subject matter agency. One example is the national accounts, another is the labour force survey. An example of a responsibility shift is the following: the responsibility for environment statistics has been transferred to our agency for environmental protection. The shift in responsibility means that the agency gets the funding for statistics production and can do the work itself, or let an outside firm do it. Statistics Sweden could be one of those firms.

Thus Statistics Sweden is also a statistical firm. Most of the production work has so far gone back to our agency following extensive negotiations regarding cost and quality. But we are in a very competitive environment. It is very tempting for all these agencies to try to do some work themselves or perhaps let several firms, including Statistics Sweden, do different parts of the work. On top of this Statistics Sweden shall oversee this new system. Basically, one

specific department at Statistics Sweden, R&D, where the authors work, has this task. We are supposed to report to government via our Director General how all the agencies, including our own, perform. We are also responsible for general methodological development that should benefit the entire Swedish statistical system. This also means that we should train and consult with these other agencies, as we always have done within our own agency.

The organizational structure and the different roles of Statistics Sweden might seem very complicated. To some extent that is true but so far the transition has been really smooth. We can notice an increased interest in official statistics among agencies and we can definitely notice an increased interest in quality and methodology, which must be good for the status of statistics in Sweden.

So how does all this reflect on product quality? We do not know yet. There is little money for evaluations and methodological studies. We believe that product quality must be achieved through improved and stable processes. As has been pointed out there is no shortage of process data and the collection of such data and trying to develop standardized procedures will help improve quality. Insofar these other agencies choose to let Statistics Sweden do the production work this approach is valid also for them. If they work on their own it is our job to see to it that the products are up to par when it comes to quality declarations, publishing, and proper use of accepted methodology. But we are in no position to tell them how to do things.

The Development of Quality Assurance - Variation in Resources

The view on how to tackle quality problems has varied at Statistics Sweden during the last decades, due to variation in financial and methodological resources. In the beginning of the period finances were in good shape and a number of evaluation studies took place. For a number of years, special funds were available for quality studies. There were both smaller studies shedding light on specific error sources in specific surveys, and larger ones, most notably the evaluations of the population censuses. These studies led to improvements in the methodology used but rarely to profound process improvements. The improvements had a tendency to stay within the surveys evaluated. For instance, evaluations of the coding process in the censuses led first to the use of independent verification and then to automated coding but very little of these achievements spilled over to other surveys.

Then there was a period of redesign of surveys. Typically, a group of methodologists studied the design and came up with quite remarkable suggestions on how to change the design so that it became more efficient. In some cases the improvements led to significant reductions in sample size with reductions in costs and respondent burden as consequences. This approach was very demanding, since each attempt drew a lot of methodological resources. Perhaps it was possible to conduct one or two such efforts per year and knowing that the number of surveys is around 100, it is

easy to see that the approach does not seem too efficient. It was basically abandoned since it exhausted the methodological staff.

There was a feeling that procedures that were common to many surveys, like questionnaire development, coding, editing, nonresponse reduction and adjustment, estimation, and analysis should be done similarly across the organization. Therefore a lot of current best practices were developed during the 80s, most notably in the areas of questionnaire development, editing, automated coding, nonresponse, and estimation. These practices, however, had difficulties becoming known across the organization. Partly this state of affairs was due to a lack of financial pressure and a lack of competition. Meanwhile, of course, a lot of developments and improvements took place within the individual surveys but the common slogan that visualizing good examples would do the trick simply was not true. There was no systematic benchmarking within our organization.

During this period two general measures were developed to keep track of quality: the nonresponse barometer and the yearly quality report based on self-evaluations performed by the survey managers. Both of these efforts are described below.

Statistics Sweden bought into the total quality management concept in 1993. The reason was that our position had become more vulnerable in connection with the creation of the new statistical system. The customer became more visible and we realized that improvements must involve those who work on the processes. All of a sudden Statistics Sweden had to compete for work which called for some changes. One such change is to create current best methods of the kind just mentioned but in such a way that these methods are readily accepted by those involved in everyday survey work. There is a great need for such standardizations since, if they are applied consistently across the organization, they will reduce variation and save resources. Currently two such standardization projects are underway, one on nonresponse reduction and one on editing. Next year two new ones will start, one on questionnaire development and one on time series analysis. The project groups are set up such that implementation becomes more or less "automatic." Implementation will be assured through management follow-up and the fact that process owners have participated in the development.

Also, we are developing a system for quality assurance based on process thinking. Product quality is generated through process quality by means of checklists. For a number of survey operations, there are checklists that each survey manager should use to make sure that all process steps have been taken.

We are fairly convinced that there will be little room for large quality studies in the future. We believe that the route to take is working on processes, standardize them, measure key process variables on a continuing basis and use checklists to assure product quality.

The Accuracy Components

In many surveys at Statistics Sweden the presentation of accuracy can be made in a rather standardized way. This is due to the existence of a number of administrative registers which have been transformed to sampling frames. Whether we make a complete enumeration or a sample survey by using these registers as sampling frames, they give us possibilities and set limits to what we can do.

In the following paragraphs we present the subcomponents of the major quality component Accuracy.

Overall accuracy

The final goal for the quality declaration is to present the overall accuracy. This is, however, seldom accomplished. The existence of frames makes it almost always possible to check that estimates from the survey agree with known parameters computed from the register, though.

Some surveys present overviews of their current knowledge of errors referring to a series of experiments, observations and analysis.

Coverage

Most surveys can easily express coverage rates in relation to the sampling frame when there are sampling units like individuals, organizations, farms, enterprises, etc. From contacts with the authorities producing administrative records and from experiences of earlier surveys one often has a very good understanding of the number of units that has not been included in the frame and those who have not been excluded in time. Typically there are only a few percent over- and undercoverage in the surveys conducted by Statistics Sweden.

For some surveys the demand for very quick presentation is strong. Sometimes preliminary results have to be published before all the data have been collected. If this is to be classified as undercoverage or nonresponse error depends on the " time cut-off" rules.

A more important problem often appears when the units are events like road accidents, crimes and some types of economic activities. Nonobservation or underreporting results in some estimates being much too low. It is very rare that we have a good knowledge of the size of this error.

Strategy

The concept strategy includes sampling plan, sample size and estimation plan. Most surveys at Statistics Sweden use some stratification by register variables. The sample is usually simple random or systematic within each stratum. The sample is more often than not allocated to promote good quality in study domains. Allocation to reach best possible precision in population

estimates is rare. The chief statistician has declared that the calculation of precision must not be neglected in any sample survey.

Formulas for estimation and variance estimation are usually straightforward. Some surveys have found it useful to develop generalized variance functions to reduce the amount of calculation.

At Statistics Sweden several computer programs for the estimation in sample surveys have been developed. One of the recent programs, CLAN, is designed to estimate several different rational functions $f(\mathbf{t})$ of different totals, $\mathbf{t}$, (for instance means, ratios, ratios of ratios, etc.) and their standard errors in the same run. Since CLAN was written in the SAS macro language it works on PCs as well as on mainframe computers.

A large number of estimators, including the use of auxiliary information and calibration can be handled. The user may combine the choice of estimators with the specification of complex sets of domains in a very flexible manner.

So far four strategies have been implemented in CLAN. The strategies imply stratification of elements and clusters and the sample selection with SRSWOR. The majority of surveys conducted at Statistics Sweden, including a number of surveys that use pps-sampling, various types of network sampling and two-phase sampling schemes for stratification, can in different ways be brought back to these four strategies.

Measurement and data collection

Usually only process information or vague knowledge is available. Sometimes the producer will mention variables that are hard to measure, sometimes statements about the direction and size of the bias are made. A low occurrence of item nonresponse or absence of complaints by respondents (or the opposite) is sometimes mentioned as an indication of good measurement quality. Comparisons with other statistics are used as indicators of reasonable results. A number of surveys have conducted reinterview studies or other evaluations of their questionnaires.

A growing number of surveys have had their questionnaire pretested - usually by the Measurement, Evaluation and Development Laboratory (ML) of Statistics Sweden. Even if we can be rather confident that pretesting means an improvement of the questionnaire it is not designed to measure errors.

Processing

The data collection process may cause problems of different kinds. For instance, when data are collected from different administrative registers, these data are not always very well adapted to the needs we have. Actions have to be taken in order to edit and control the data.

Traditional data entry is becoming less frequent at Statistics Sweden depending on the use of computerized collection techniques, CAPI, CATI, TDE, etc. and scanning of questionnaires.

Automated coding of occupation and education is done on a regular basis.

Each moment in the data processing should have some effect on the final quality of both the micro data and the estimates computed from these data.

Nonresponse

A great effort has been made to develop and standardize the presentation of nonresponse rates. Since 1985 there is an overview, "The Nonresponse Barometer" that presents time series on response rates in all sample surveys and some censuses at Statistics Sweden. At first only a few important surveys on individuals and households were included and their rates had to be accepted without standardization. The reports covering the last few years include almost all surveys. The main features of their sampling and data collection plans are mentioned. Design changes that may have an influence on the response rate are mentioned.

Model assumptions

Statistical results sometimes rely on complex calculation schemes. These schemes may presume a model relationship among the input statistics if the calculated results are to be valid. This is the case for a lot of statistics on the environment and on the public economy. Errors in the model specifications may generate important errors. The models must be explained and robustness to specification errors explained.

Presentation of uncertainty measures.

Uncertainty of statistical estimates must be reported for all Swedish official statistics according to a set of recommendations by Statistics Sweden.

Evaluations

The results of censuses and sample surveys suffer from a number of errors. Editing, control of coding, etc. reduce the errors but cannot eliminate them. It is important that users of statistics have a possibility to judge how statistics can be used and what conclusions can be drawn from the published information. Also the producer should have an interest in knowing the actual quality of the published figures.

One possible way to get knowledge of the sizes of the errors in the final estimates is to carry out evaluation studies. These studies are not carried out on a regular basis and not in every statistical product at Statistics Sweden. The reason is of course limited financial and human resources as previously mentioned. An evaluation study is often considered complicated and it is often thought that available resources can be of better use in the main survey.

However, evaluation studies are carried out in the most important surveys and censuses as part of the quality control and as a basis for development of improved methods of data collection, editing, coding, etc.

Traditionally, evaluation studies at Statistics Sweden have been mostly *producer-* oriented rather than *user-* oriented and it is the major component *accuracy* and its subcomponent *measurement* that have been of interest. Perhaps this is natural since, in general, qualified statisticians have been responsible for these studies.

Random and systematic errors can occur both in estimated parameters and in the measurement of variable values. Systematic errors in variable values might eliminate each other when aggregated (if you are fortunate) and make the net bias in the estimates due to measurement error small, while the random part of measurement errors in general will increase the random errors in the estimates.

Often not only the net effect of the systematic errors are of interest, but also the gross effect. The reason for this is that in a survey, as well as in a census, the collected data are to be used not only to estimate the parameters that have been studied in the evaluation study but also to estimate other parameters that perhaps no one thought of at the time of the evaluation, for example estimates in totally different domains of study. Another reason is that the results from a census is usually used as a sampling frame where the variables are used to define different strata. Serious errors in the stratification might then ruin a sample survey. Further, when independent researchers outside Statistics Sweden use the data for different kinds of analysis, often by sophisticated methods, the researcher has to know the quality of the data he/she is using in order to draw valid conclusions.

Evaluation of the component accuracy is usually concentrated on the size of errors in the statistical estimates, for instance in estimated totals, ratios or mean values. In cases where the aim of the studies is to measure the systematic error in different estimates, evaluations can be conducted in different ways depending on level of ambition and available resources.

Crude measures and indicators of systematic errors can be obtained by comparisons between estimates from different surveys where related parameters are estimated, or by the study of correlated background variables whose values are known for the whole population. These types of measurements can often be made without further data collection but they give limited information about the measurement process.

Only in exceptional cases can good estimates of measurement errors be obtained without collecting additional data about the units in the main survey. If the aim of the study is to measure the reliability, you "only" need to repeat the main measurement process under conditions that are as identical as possible with the main survey. Such evaluation studies have been conducted rather frequently at Statistics Sweden. They are not to expensive and usually they give valuable information to the producer of the statistics. The results are seldom disseminated outside the agency but rather published in internal memos.

When you want to know something about the bias in published estimates, you need data with "true" values for at least a subsample of the units. The word "true" is used in an operational way here (in some cases there might not even be a true value). In practice it means that we are using a measurement process that is considered significantly better than the ordinary one.

"True" values are most often determined by matching and reconciliation. The statements given in the main survey are compared with the corresponding statements given in the subsample, where the questions or the wording of the questions need not be identical. If there is no discrepancy between the statements they are considered true. In other cases the respondent is asked to confirm the "true" statement, (the original, the new, or perhaps a third one).

This technique is rather expensive and is mainly used to evaluate the large registers and censuses.

Some discussions have concerned how to present the result from the evaluations. Let $\hat{\theta}$ be the estimate of a parameter based on the units in the evaluation sample and on the ordinary statements and let $\tilde{\theta}$ be the estimate of the same parameter based on the "true" measurements in the evaluation sample. Sometimes (1), $100 \times (\hat{\theta} - \tilde{\theta}) / \tilde{\theta}$ and sometimes (2), $100 \times (\tilde{\theta} - \hat{\theta}) / \hat{\theta}$, are used as measurement of relative bias. In (1) the deviation is shown as % of "true" (unbiased) value, while (2) shows the deviation as % of the "official" estimate. Of course you can always get (1) from (2) and vice versa but (1) is the *producer's* measurement as it tells the producer the deviation from the goal while (2) tells the *user* something about the error in the published figures. Measurement (1) was for example used in the evaluation of the Register of Employment while (2) has been used in the evaluations of the population censuses.

An example: The evaluation of the 1990 Census of Population and Housing

The evaluation program contained the following studies:

Evaluation of

- household data for dwelling households
- housing data for occupied dwellings
- employment data
- education data

The evaluation covered about 17 000 units which were sampled from persons in Sweden who were in the ages of 16 to 74 years and registered in Sweden on the 1st of November, 1990.

"True" values were determined by matching and reconciliation in the same way as described above. The statements given in the census were compered with corresponding statements given in the Labour Force Survey of November 1990. For about 16% of the units no "true" value could be

determined depending partly on missing values in one or both of the two and partly on discrepancies that could not be reconciliated.

The results from the evaluation study are published in the official statistical series. Both net errors and gross errors are estimated. Estimates are given for many different combinations of sex, age, region, type of dwelling, type of household, etc.

The Nonresponse Barometer

Since 1985 a yearly report on the nonresponse in some surveys at Statistics Sweden has been published. The aim of the report is,

- to show the amount of nonresponse in a number of surveys at Statistics Sweden,
- to give a picture of the "response climate" (i.e., do individuals, businesses, and other institutions become more or less willing to answer survey questions?),
- to be one (of several) instruments to compare different statistical products over time.

The aim of the barometer is not to describe the quality component named nonresponse, but rather to give a description of the size of the nonresponse for different surveys and over time. The *effects* of the nonresponse, for example nonresponse bias, are not handled here. An estimate of these effects is given in the yearly quality survey "The Quality Report" described below.

Only unit nonresponse is treated. Item nonresponse, that is where a unit has participated in the survey but has not given answers to every question, is not.

Measurement of the nonresponse rates are given both as weighted and unweighted figures. Weighting is done in many different ways depending on which measure that is considered best adapted to its purpose. In surveys related to businesses, the estimated number of employees or the turn-around (according to the information in the register) in the nonresponse businesses is used as well as the estimated number of nonresponse businesses in the population. When the units are individuals the weighting means that the estimated number of nonrespondents in the population is used.

The total nonresponse is classified according to cause (*refusals*, *no contact* and *other*), where this is possible (mostly in surveys of individuals).

The response climate is measured by asking the person responsible for the survey to give an estimate of the changes in response climate from last year and the change from five years ago. Four alternatives are possible: "*Better*, *Neither better nor worse*, *Worse*, *No opinion/irrelevant.*"

Another measure of the response climate is obtained by asking the interviewers' supervisors about their opinions. Then we get more information and it is also possible to say something about regional differences.

Until now the presentations of the results have been made by the different departments at Statistics Sweden. In the future the presentation of the nonresponse results will probably be more difficult depending on the new distribution of the responsibility for official statistics.

The Quality Report

The Quality Report is produced in order to provide a basis for an analysis of the development of the quality in the statistical products at Statistics Sweden. The report has been published yearly since 1988.

A questionnaire is administrated to every person responsible for a statistical product at Statistics Sweden. The questionnaire consists of three parts. In part 1, the respondent is asked to state general external factors of importance for the quality changes of his/her product, also the measures (if any) carried out to handle them are to be reported. In part 2 the respondent is asked to give estimates of how the quality level of their product has changed since last year. Estimates are to be given for each of the 23 quality components described above. Changes are measured by a five degree scale, *"much worse," "slightly worse," "unchanged," "slightly better,"* and *"much better,"* or *"not relevant."* In part 3 the respondent is asked to report remaining quality problems that are judged to be of special importance. Also planned improvements of importance for the quality shall be reported.

The questionnaire is sometimes filled in by a team, involving several persons working with the product. The questionnaire is examined and approved by the respondent's manager. In each department the responsible statistician examines all the questionnaires, make necessary completions and the data processing.

The respondents are instructed to fill in the form with the intended user's perspective in mind but it is important to note that the users normally have no possibility to share their opinions since they are not explicitly ask to do so.

The measurement process should be continually improved. Especially in part 2 there are currently substantial possibilities for subjective judgement when a change is reported. However, in order to improve the quality of the responses, the respondent is asked to give a motivation and a comment whenever a change is reported.

Needless to say, generally the alternative "unchanged" dominates the statements in part 2, (70%-90%). When a change is reported the alternatives "slightly better" and "much better" dominate. Future plans include a redesign of the entire process.

Endnote

A number of factors affect quality and the possibilities to declare quality. Cuts in funding generally means that there is less room for nonresponse follow-up. Money is also the reason why there are fewer evaluation studies these days. Demands for faster production also contributes to less time for nonresponse follow-up but also to less planning time in general. The new technology means that it is easier to measure process variables which to some degree can compensate for the problems mentioned.

A continuing quality problem is that so much of statistics are based on the use of administrative registers. Normally the producer of statistics has very little influence on the collection of such data and normally does not know much about coverage and measurement error problems unless special studies can be designed.

International organizations get more and more influence on contents and methods. It is more and more common that specific surveys deliver to systems of different kinds, like index systems or accounts systems. Concepts such as comparability, additivity, and completeness might get new meanings and country comparisons, for instance, might take precedence before local needs. Thus, international cooperation seems necessary.

Data Quality at the Energy Information Administration: The Quest for a Summary Measure

Renee Miller
Energy Information Administration

At one point not too long ago, I found myself in a conversation with a relative who, like many taxpayers, was not convinced he was always getting his money's worth. He was familiar with the Energy Information Administration's (EIA) data and said that the data were useful. I thought I was off the hook, but then he said something like, "How do you know if the data are any good? Do you have some kind of a measure?"

Skeptical relatives have not been our only questioners. "How do we know if the data are any good" is a question with which we at EIA have been grappling for years. It recently emerged during our Business Re-engineering efforts. The Business Re-engineering team[1] was chartered to rethink three core business processes: data operations, data integration, and product preparation and dissemination.

At several points the team discussed making EIA data more timely to better meet the needs of our customers. During these discussions someone would raise the concern about balancing timeliness and quality. We thought a summary measure of data quality would be helpful in this situation and we discussed what that measure might be. The discussion of a summary measure of data quality led to thinking about how information on data quality is presented to the public. The issues of how we ensure data quality, whether there is a summary measure of data quality, and how we report information on data quality to the public are intertwined.

This paper begins with a few words about EIA, then discusses our attempts to measure data quality. It continues with a discussion of additional activities to ensure data quality. It then presents a proposal for a summary measure of data quality and goes on to describe some recent developments. The paper ends with some thoughts on where we go from here.

EIA in a Nutshell

EIA is almost twenty years old. Congress established the agency in 1977 to be an independent source of energy information. It combined data gathering functions formerly performed by the Bureau of Mines, the Federal Energy Administration, and the Federal Power Commission. Besides combining data gathering functions, the new agency was also a combination of people. Some came from the three predecessor agencies. Some of us are from other statistical agencies such as the Bureau of the Census, Bureau of Labor Statistics and National Center for Health Statistics. Others came from academia and many other places.

---

[1]Members of the Business Re-engineering Team included: Project Director, Chuck Heath; Core group members: Ray Boyer, Clyde Boykins, Ann Ducca, Sue Harris, Mike Lehr, Dorine Andrews, and Lori Gillespie; Champions: Chuck Allen, George Baker, Noel Balthasar, Yvonne Bishop, Ken Brown, Bill Dorsey, Lamar Gowland, Mike Lehr, Nancy Leach, Bob Manicke, Renee Miller, Ken Vagts, Howard Walton, and John Weiner.

The activities described in this paper reflect traditions carried over from many of these agencies, plus new ones that we developed as we strive to become "team EIA."

There will be references to EIA's re-engineering efforts. These efforts stemmed from the realization that with declining budget and staff levels EIA could not improve (or even maintain) its level of customer service by doing "business as usual." In August of 1995 EIA's Quality Council chartered a team, and in April of 1996 the team delivered a blueprint for a re-engineered EIA. We are in the process of implementing parts of the blueprint.

Attempts to Measure Data Quality

During the development of the Business Re-engineering blueprint, the team developed measures for various processes. The idea was that instead of having a long line of staff checking and rechecking work, we would have measures that would indicate whether the processes were functioning effectively. One measure that eluded us, but that we kept coming back to, was a summary measure of data quality. In searching for a measure, we reviewed some approaches tried or suggested previously.

These approaches included validation studies, data comparisons, supply/disposition balances, and elements of data quality. Revision error, response rates and sampling error were also revisited as described below.

Validation Studies

In its early days, EIA conducted validation studies. Reflecting their extensiveness, they were called "cradle to grave" examinations of the data. They included a search for deficiencies in the universe list, an audit of company records to determine if they corresponded to what was reported, a check for transcription errors by comparing hardcopy to the automated data file, and many other activities. Reference [1] provides an example of a study pertaining to data collections on coal production.

As a result there was some information pertaining to each source of nonsampling error (coverage, measurement, nonresponse, and processing). Sometimes the information was quantitative. What was lacking was a way of adding it all up to get total survey error, because sometimes the errors were offsetting.

Overall, the studies were not popular with either the survey respondents or with the survey managers. Furthermore, they were expensive. They were stopped when our budget was reduced in the early 1980's.

Data Comparisons

EIA staff also compared the data series of interest called the reference series with other data collected by EIA or other organizations. In the early days comparative sources were plentiful. At the aggregate level, we computed each comparative series as a percentage of the reference series.

An early study focussed on the data on imports of crude oil. There were three comparative series

with data for three years, which we considered as nine independent estimates. We had nine ratios of the comparative series to the reference series and computed the mean, the standard deviation of the mean and a 95 percent confidence interval. The 95 percent confidence interval was 99.2 - 100.8 for imports of crude oil based on data for 1977, 1978 and 1979. We then concluded that the EIA reference estimate was accurate to within 1 percent[2].

As might be expected, these estimates of accuracy were not well received. Sometimes the comparative estimates had well-known problems. There was often no indication that they had been validated and most of the time there was no documentation on how the comparative series were obtained. Although we continue to perform comparisons and present the results to the public, we stopped coming to conclusions about data quality based on them.

In the eighties, we presented the results of data comparisons in a series of reports that became known as the "State-of-the-Data" reports [2]. Staff in the Office of Statistical Standards prepared these reports with input from survey staff. These comparisons were performed at the respondent level as well as at the aggregate level.

Currently survey staff members prepare annual feature articles comparing EIA data with other sources. The articles appear in EIA's *Petroleum Supply Monthly* and *Petroleum Marketing Monthly*, which have a wider distribution than the earlier "State-of-the-Data" reports. Examples of comparisons the user can find in these reports include data on imports of crude oil and petroleum products from EIA and the Bureau of the Census and prices of petroleum products from EIA and the Bureau of Labor Statistics.

These articles provide a vehicle to let users know that some observed differences in the data series stem from the different definitions or universes used in the data collections. In addition to the feature articles, results of comparisons have been presented at conferences such as the annual meetings of the American Statistical Association and Washington Statistical Society meetings. Some EIA programs, such as end-use consumption and electric power, routinely include comparisons in Appendices to their data publications [3, 4, 5, and 6].

Comparisons often raise more questions than they resolve. While it is comforting when data collected from different sources correspond well, there is still the possibility that they are not correct. Comparisons have been most useful when data are collected from the same respondents and we can match records. In these situations we can identify the individual respondents with differing responses and follow up to find out why. We have gained information about how respondents are interpreting our definitions and instructions through these follow-ups.

Supply/Disposition Balances

In addition to data comparisons, EIA staff members look for symptoms of problems in the published data by examining supply/disposition balances. The expectation is that supply should equal disposition. Both components, in turn, consist of several parts. Production, imports, and

[2] An Assessment of the Accuracy of Principal Data Series of the Energy Information Administration, DOE/EIA-0292, June 1981, page 23.

withdrawals from storage make up supply. Consumption, exports, and additions to storage make up disposition. Because we obtain the data that comprise the supply/disposition figures from different surveys, a balancing item is needed for supply to equal disposition.

EIA staff members have been using the balancing item as a warning signal. If, for example, the balancing item increased sharply from one year to the next, it could be an indication of errors in one or more of the components. A small balancing item, however, does not conclusively show that all the figures are of good quality. There could be offsetting problems; e.g., production, a component of supply, could be overstated and imports, another component of supply, understated. Therefore, it is difficult to use the balancing item as a measure of data quality.

While looking for symptoms of problems in the data through comparisons and balances is likely to identify major problem areas, it does not give us a systematic way of quantifying errors in the published data.

Elements of Quality

In August 1991, then EIA Administrator, Dr. Calvin Kent took a different approach to assessing data quality. In a presentation at the annual meeting of the American Statistical Association entitled, "Quality of Energy Data," he discussed four elements of quality: timeliness, consistency, continuity, and customer satisfaction. In the past few years we have made progress in measuring two of the elements: timeliness and customer satisfaction.

We measured timeliness as the number of days between the last day of the reference period and the "released for printing date" shown inside the front cover of the publication [7]. We have compiled data on timeliness for annual publications for the years 1990 through 1994 and for monthly and quarterly publications for 1993 to 1995. In the future we will use the date the publication returns from the printer to better reflect the date the customer receives it.

To measure customer satisfaction, for the past two years we have been surveying EIA's telephone customers. In February of 1996, EIA volunteers surveyed 264 telephone customers. The volunteers asked customers about their satisfaction in two broad areas: customer service and information quality. The first area included: ease of access, courtesy, familiarity with the information, understanding the customer request, and promptness in responding. The second area included: availability, relevance, accuracy, comprehensiveness, and timeliness.

About 73 percent of the respondents said they were either satisfied or very satisfied with the timeliness of the information. By contrast, 90 percent of respondents said they were satisfied or very satisfied with the accuracy of EIA data. Several interviewers noted, however, that a few respondents said they had no way of knowing whether the data were accurate or not. These respondents, nevertheless gave us a high rating because they said they had no reason to believe the data were not accurate.

With respect to timeliness, during 1995 there were widespread efforts to make data available earlier through electronic dissemination. However, in 1996 EIA received about the same overall rating on timeliness as in 1995. We think that one reason that the ratings did not change was that the

customer survey was conducted with telephone customers, a group that may not be fully aware of the electronic data.

EIA does not have measures for the two remaining elements: consistency and continuity. By consistency we meant--how do EIA data compare with other similar series and are our data internally consistent? (An example of a data inconsistency would be more domestic electricity sold than generated). For continuity our questions were: Do we measure the same thing over time? How does EIA handle revisions to the data? How does EIA handle breaks in its time series resulting from industry changes and modifications to our survey forms?

The four measures, in contrast to the work performed in the validation studies, do not attempt to measure total survey error. Rather, they are related to the fitness for use of our statistical products.

The Usual Suspects: Revision Error, Response Rates, and Sampling Error

Revision error, the difference between preliminary and final estimates, has been suggested as a measure of data quality. This measure has been criticized, however, because it does not address the issue of the quality of the final estimates. In addition, suppose there are no revisions. Does that mean there is no error?

Nevertheless, we have found the computation of revision error to be useful in improving our preliminary estimates. We present information on revision error to the public in annual feature articles to EIA's *Petroleum Supply Monthly* and *Natural Gas Monthly*. Other program areas, such as electric power and petroleum marketing, include the information in appendices to their publications [6, 8]. We show the preliminary estimate, the final estimate, and the percent difference. In the feature articles, we provide explanations of the differences, if available.

In addition, we have been tracking revision error as part of our organizational performance measurement system. EIA developed the system while participating as a pilot project under the Government Performance and Results Act of 1993 [7].

We also compute and publish information on response rates and sampling error, generally in the explanatory notes section of our publications. While these are important measures, they do not tell the whole story with respect to data quality.

Other Activities to Ensure Quality

Besides the activities just described, there are other activities performed throughout the agency to ensure the quality of the data. While we have found these activities useful, we also found that they did not lend themselves to measuring data quality. The activities include editing of the data and the development of statistical standards. In addition, we conducted audits to check compliance with standards. Furthermore, we have performed site visits with selected respondents. This section summarizes these activities.

Edits

Prior to publication, survey staff members edit the data using consistency checks, comparisons with previous reported values, and other more complex methods [9, 10]. They follow up by phone with respondents who have reported seemingly anomalous values. Because of the wealth of historical data for the weekly and monthly surveys, time series methods have been used to predict the current value and to construct tolerance limits for the new data. Examples of these methods are featured in Statistical Policy Working Paper 18, "Data Editing in Federal Statistical Agencies" [11].

Edits tell us about the quality of the reported data to some extent. We have found, however, that there are errors that edits cannot detect such as a respondent consistently reporting residential deliveries as commercial deliveries. Therefore, we have not been able to translate information from edits to a measure of quality for the published data.

Standards and Audits

EIA has developed a manual that contains copies of the agency's statistical standards [12]. In the foreword to the manual, we state that standards "help ensure data quality, remove ambiguities, avoid duplication of effort, and improve responsiveness to our data users." The standards cover both data collection and processing, and data presentation.

Using an analogy from the health area, standards are much like the advice to maintain a low-fat diet and to exercise regularly. While a low-fat diet and exercise purport to contribute to our long-term well-being, they do not ensure that we are disease free on a daily basis. There is a similar situation for statistical standards which is why adherence to standards has not been accepted as a measure of data quality.

As noted in the paper, "Quality in Federal Surveys: Do Standards Really Matter?" [13], the relationship between standards and data quality is tenuous. Nevertheless the paper notes that standards were helpful in establishing the credibility of EIA data, along with rigorous programs of enforcement, evaluation, and education.

EIA has conducted audits as part of its enforcement program. The initial round was broad in scope and mainly concentrated on standards and documentation review along with data processing issues. These audits checked each system for compliance with each standard [14].

The next round focussed on the quality control activities to determine whether they were adequate to control nonsampling error [15]. For each source of nonsampling error, we developed checklists of activities that could control these errors. The EIA standards manual and Statistical Policy Working Paper 15, "Quality in Establishment Surveys" [16] were used to develop the checklists. In addition, results of data evaluations were used to determine whether any identified anomalies resulted from a failure in a quality control procedure.

Site Visits

In the early 1990's, we started a program of site visits. During the visits we spoke to respondents about how they are interpreting selected items on the forms. In addition, we asked whether their records correspond with the items we are requesting. Unlike the validation studies of prior years we have not asked respondents to produce records for verification. We visited about a dozen respondents, covering coal and natural gas production, consumption and distribution data. While we have obtained useful data from these visits, the sample size has been too small to draw inferences. Furthermore, we did not ask all respondents the same questions.

Proposal Developed During Business Re-engineering to Measure Data Quality

During our re-engineering efforts the issue of a summary measure of data quality arose several times. And several times we concluded it could not be done. With the activities previously described as background, following is a rating scheme we tried to develop.[3]

The Business Re-engineering team ultimately decided it would not be workable because it would require much time and judgement. It is being presented because we learned something from the experience. The team was a diverse group consisting of managers, statisticians, analysts, computer specialists and interdisciplinary staff. We all had different reference points. The attempt to develop a summary measure proved helpful in making us all realize what was involved.

Dimensions of Data Quality

We began by listing some of the dimensions of data quality:

- sampling error
- measurement error (the difference between the value collected during the survey and the true value. It includes both reporting error and specification error)
- coverage
- nonresponse
- methodological consistency (this is the same as "continuity" in Dr. Kent's scheme described earlier. It pertains to breaks in the data series and whether the changes and their impact on the data are documented)

These dimensions differ in the ease with which we can quantify them. Sampling error, on the one hand, can be computed directly from the survey data. Methodological consistency, on the other hand, cannot be directly computed. Ideally a series should be stable over time; i.e., not have any

[3]Dwight French, Office of Energy Markets and End Use, participated as a subject matter expert in the re-engineering effort and worked with this author on the rating scheme for data quality.

breaks. Sometimes due to changes in the industry, it is inevitable that a data collection is modified. Does that series get penalized for having breaks?

Some other dimensions sound like they should be easily quantified, such as measurement error. EIA does not have information for each survey on an ongoing basis. This is the type of information that we obtained from the validation studies which have been discontinued.

Since we had some information for each survey, we thought we could gather it together. We would then rate each survey on each dimension using a 1 to 5 scale (where 5 is very satisfactory and 1 is very unsatisfactory). We would have 2 scales: level of knowledge about the category and level of quality. In this way we would obtain information on how much we know about data quality as well as information about the quality of the data.

Using nonresponse as an example, a survey might get a score of 5 on knowledge if documentation was available on the response rate, on our follow-up and imputation procedures, and key information was presented in the publications. A survey might get a 5 on the quality scale for nonresponse if the response rate was 98 percent in terms of both number of respondents and volumes reported.

We would then combine component scores into an overall score for a survey, program, or EIA as a whole. To ensure consistency we started to develop guidelines on what represents a "5" versus a "4" and so on. It got very complicated quickly. Furthermore, for measurement error and methodological consistency, we found it difficult to develop "quality" measures; therefore, we only had "knowledge" measures.

Perceived Complications

Because of the perceived complications and other issues, the Business Re-engineering team decided not to pursue this procedure. One issue was who would do the ratings. Another concern was that we could not really ensure consistency. Furthermore, there was the perception that a lot of time would be involved in performing the ratings. The general feeling was that even if we could be precise enough to ensure consistency, we would not be giving the user much more information than is provided in the explanatory notes section of our publications.

This comment raised the issue of whether the approach we should take should focus on the descriptive explanatory material, perhaps standardizing it. All of our data publications contain explanatory material. We have a standard on publication of energy statistics which is based on a directive in the *Statistical Policy Handbook* [17]. The EIA publication standard specifies that we describe the survey design and provides a checklist of activities to include. It also specifies that we point out the limitations of the data. The detail we provide on the limitations of the data and on features of the survey design varies across EIA.

Recent Developments

There are two recent developments at EIA that could affect the approach we take to presenting information on data quality. One is quite specific, the development of a quality profile for the Residential Energy Consumption Survey. The other is more global, electronic dissemination.

Quality Profiles

Last spring, EIA published its first quality profile, an extensive profile of the Residential Energy Consumption Survey. It was prepared by Thomas B. Jabine [18] in a joint effort between the Offices of Energy Markets and End Use and Statistical Standards. As described in the report the purpose of the *Residential Energy Consumption Survey Quality Profile* (*Quality Profile*) is "to present, in a convenient form, a report on what has been learned about the quality of RECS data since the survey began."

The report provides an overview of the survey and presents information about three major sources of nonsampling error: coverage error, nonresponse, and measurement error. It also discusses the contributions to nonsampling error of data processing and imputation procedures. In addition, it looks at the effects of estimation procedures on data quality. Furthermore, the report presents results of studies that have compared RECS data with data from other surveys, and describes relevant research currently in progress.

The *Quality Profile* has been very well-received. Several members of our energy advisory committee said it was a good model of how we should document our surveys [19]. They pointed out that customer satisfaction depends on data quality and that a quality profile would give users all the information they would need to determine data quality. They suggested that we do more profiles. Unfortunately, due to our budget constraints and reduction in staff levels, that does not appear likely.

Electronic Data Dissemination

As mentioned earlier, there was a concentrated effort to make data available electronically at EIA. Electronic dissemination has produced new possibilities. One is that the user would click on a data value and see a standardized description that explained it [19].

We have taken a couple of steps in that direction. EIA has developed an Electronic Styles and Standards Manual [20]. It requires that when a publication is released electronically that it is released in its entirety so that the explanatory material is included. For products released as files, we are required to provide data sources and caveats concerning the data.

Another step is the development of a succinct set of notes for data from the Commercial Building Energy Consumption Survey that will be released on the internet [21].Topics in the notes include: survey methodology, target population, sample design, changes in the survey from the previous

cycle, sampling rates, data collection procedures, response rates, minimizing nonresponse to the survey, and a general discussion of sampling and nonsampling errors.

Where Do We Go from Here?

The issue of a summary measure of data quality does not appear to be going away. EIA is moving toward a performance-based budget. In addition, implementation of the business re-engineering blueprint includes a pilot test to integrate survey operation activities. As part of this effort, staff members are developing measures to monitor the process overall. Ideally we would like to include a measure of data quality, apart from revision error.

We realize it is not likely that we will find the perfect measure. While we have not found a summary measure of data quality, there is agreement that providing users information on what we know about the quality of the data is crucial.

We have been giving our users explanatory material for years. Yet during our customer satisfaction survey, some have told us that they have no way of judging the quality of our data. They think EIA has good quality data, but they say they do not know for sure. Something seems amiss here.

Perhaps the future direction should be to make the information on data quality easily accessible and understandable. We could cover in a concise way the dimensions of data quality that we identified: sampling error, measurement error, coverage, nonresponse and methodological consistency. The notes developed for commercial consumption data that will be released on the internet are in the direction of this goal. They were based on work done at the National Center for Education Statistics.

Building on each other's work perhaps we can attain information on data quality that is so clear and accessible that users, themselves, will be able to answer the question, "How do we know if the data are any good?"

References

1. Management Engineers, Incorporated. "Coal Production Data Systems Validation: Final Report." Reston, Virginia, September, 1982.

2. Energy Information Administration. *An Assessment of the Quality of Selected EIA Data Series.* DOE/EIA-0292 (83), (85), (86), and (87). Washington D.C.

3. Energy Information Administration. *Household Energy Consumption and Expenditures 1993.* DOE/EIA-0321(93). Washington D.C. October 1995.

4. Energy Information Administration. *Commercial Buildings Energy Consumption and Expenditures 1992.* DOE/EIA-0318(92). Washington D.C. April 1995.

5. Energy Information Administration. *Manufacturing Consumption of Energy 1991.* DOE/EIA-0512(91). Washington, D.C. December 1994.

6. Energy Information Administration. *Electric Power Monthly March 1996 With Data for December 1995.* DOE/EIA-0226(96/03). Washington D.C. March 18, 1996.

7. Kirkendall, N.J. "Organizational Performance Measurement in the Energy Information Administration." Presented at Bureau of the Census Annual Research Conference, 1996.

8. Energy Information Administration. *Petroleum Marketing Monthly April 1996 With Data for January 1996.* DOE/EIA-0380(96/04). Washington D.C. April 2, 1996.

9. Swann, T.C. "Electronic Data Collection in the Petroleum Supply Reporting System." Presented at the meeting of the American Statistical Association Committee on Energy Statistics. Washington, D.C. April 28-29, 1988.

10. Weir, P., Emery, R., Walker, J. "The Graphical Editing Analysis Query System." Presented at Conference of European Statisticians, Work Session on Statistical Data Editing. Statistical Commission and Economic Commission for Europe. Athens, Greece. November 6-9, 1995.

11. Subcommittee on Data Editing in Federal Statistical Agencies, Federal Committee on Statistical Methodology. *Data Editing in Federal Statistical Agencies, Statistical Policy Working Paper 18,* page 35. Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget. Washington D.C. May 1990.

12. Energy Information Administration. *The Energy Information Administration Standards Manual.* DOE/EIA-0521.

13. Freedman, S.R. "Quality in Federal Surveys: Do Standards Really Matter?" Presented at the Annual Meeting of the American Statistical Association, 1990.

14. Dandekar, R. "Future Data Quality Audits." Presented at the meeting of the American Statistical Association Committee on Energy Statistics, Washington D.C. November 2-3, 1989.

15. Energy Information Administration, Office of Statistical Standards. "Electric Power Data Evaluation." Unpublished report, 1994.

16. Subcommittee on Measurement of Quality in Establishment Surveys. Federal Committee on Statistical Methodology. *Quality in Establishment Surveys. Statistical Policy Working Paper 15.* Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget. Washington D.C. July 1988.

17. U.S. Department of Commerce. Office of Federal Statistical Policy and Standards. *Statistical Policy Handbook.* Washington, D.C. May 1978.

18. Energy Information Administration. *Energy Consumption Series: Residential Energy Consumption Survey Quality Profile.* DOE/EIA-0555(96)/1. Washington, D.C. March 1996.

19. American Statistical Association Committee on Energy Statistics. Transcript of the Meeting. April 26, 1996, pages 93 - 118.

20. Energy Information Administration. "EIA Style and Standards for Electronic Products." Unpublished report. Washington D.C. January 1996.

21. Energy Information Administration, Office of Energy Markets and End Use. "Methodology and Technical Notes." Unpublished notes, prepared for Commercial Buildings Energy Consumption Survey data to be released on the internet.

REVIEWING AND REPORTING QUALITY

JAY WAITE, DISCUSSANT

I really appreciated the opportunity to discuss these two papers. It was especially refreshing that the two papers were actually on the same subject and that the papers presented were faithful to the papers as written.

Both papers were broadly similar and both discussed work of two statistical agencies in searching for what I fear is an unreachable goal. They are both striving for a linear measure of the quality of a survey.

In fact, quality like beauty, is most often in the eye of the beholder.

In the final analysis, quality is defined by the customer. Even if massive resources are expended to precisely measure the primary components of quality, different customers with different uses in mind will want to weight the components differently. To some, timing is critical and without it nothing else matters. To others, accuracy is the last and final determinant of a quality data product.

It seems that as statistical organizations, either public or private, we are fundamentally in the business of producing and delivering information. In this context, we need to be clear about what we are producing, for whom we are producing it, and what are our customers' key determinants of quality. This is harder than it appears, especially for public- sector agencies.

Before trying to get a producer's view of a linear measure of quality, we should look toward our customers.

Who are our customers, or at least who would we like our customers to be? What do these people want from our data? (This may be quite different from what we think they should want.)

How can we get them what they want with the aspects of quality that they value most?

What will it cost us to give them that?

What are they willing to pay us for it?

The Swedish paper talks about the four faces of quality. They defined quality in the context of four areas.

1. Statistical Entity
2. Accuracy
3. Timeliness
4. Accessibility

These aspects of quality are often conflicting and while with greater or lesser degrees of success, we may be able to measure each of them. It doesn't really make sense to me to just add the up somehow either weighted or unweighted to get a total quality measure.

Consider an example of the restaurant business. Suppose we wanted to compare two restaurants on the same quality scale. Let's begin with the four categories defined in the Swedish paper. First, we would need to make a transformation of language in order to talk about restaurants. Let us define the four areas of a quality restaurant as follows:

Statistical entity	------->	Menu
Accuracy	------->	Taste/pleasing to the palate
Timeliness	------->	Quick service
Accessibility	------->	Location

Now consider two restaurants.

Restaurant one is a Five Star restaurant with food to die for. It is a the Greenbriar resort in West Virginia.

The second restaurant is a wonderful little McDonalds restaurant just four blocks from my house.

Now let's try to measure these two restaurants on quality.

First: Menu

The Greenbriar has a vast menu and if they don't have it, they will get it for you. It is fair to say that on the menu scale, the Greenbriar is a clear 10.

The local McDonalds? Well, if you want hamburgers, chicken, or fish and french fries cooked in a tub of grease, this is the place for you. A rack of lamb with some fine wine might be a problem though.

Second: Taste/ Pleasing to the palate

Here again, the Greenbriar is at the top of the chart. If you have ever eaten there, you know what I mean.

The local McDonalds tastes good too, but in a different way and to a decidedly different clientele.

Third: Quick Service

Well, here the Greenbriar is not so good. While it is true that you can get literally anything you want at the Greenbriar, it is not so clear that you can get it quickly.

The local McDonalds, on the other hand, specializes in getting you its admittedly limited menu quickly. In fact, its name is synonymous with fast food.

Fourth: Location

Well, here there is no comparison. The Greenbriar is four hours away by train. The McDonalds--well you know the McDonalds.

Which of these two restaurants should score the highest on the summary measure of quality scale? Clearly it depends on who is voting, and what they are looking for in a restaurant experience. It obviously doesn't make any sense to try to get a numeric score for these two restaurants and then compare them somehow. They are both quality restaurants, but in a different way. The same is true of surveys.

This ambiguity does not mean that we shouldn't try to measure quality for both internal management improvement decisions and for advertising purposes.

We clearly must measure all we can and seek to improve all we can, but we should not fall prey to the fiction that somehow we can put a universal metric of quality on all surveys.

THE ENERGY PAPER

On this paper as well, many measures of quality are proposed--both internal and external. Clearly, resources available to measure quality are shrinking. Does this imply that interest in quality is shrinking or only that the measures that we are producing are not what our customers feel they need?

If it's the former, then we are in trouble. If it is the latter, we may be able to get the needed resources if we measure and improve the aspects of quality that our customers and sponsors think are important.

One statement in the energy paper struck me. The authors state that users don't seem to know very much about the quality of your surveys now and yet they are basically happy with the existing quality. I found myself wondering what would be accomplished if they succeed in answering the questions about the quality and their customers find that their happiness has been misplaced.

Maybe a better question for statisticians to attempt to answer is "Are our data good enough that they are the data of choice for most of our important users? How can we improve them for this purpose?"

All this gets us back to the three really important questions we should be asking ourselves about quality.

1. Who are our customers?
2. What do they want?
3. What are they willing to pay for it?

My final advice:

Seek to understand as much about the quality of your survey as you can.
Seek to understand your customers needs and wants.
Make sure that you are measuring improvement in the areas of quality that are most important to your customers.
And finally, don't waste your time searching for the Holy Grail.

DISCUSSION:
Reviewing and Reporting Quality in Survey Data

Richard A. Kulka
Research Triangle Institute

Obviously, there have been many papers presented in the past few years on the quality of survey data, including the preparation of extensive "quality profiles" for selected major surveys in U.S. Statistical agencies, such as the Survey of Income and Program Participation (SIPP) at the Census Bureau, the Schools and Staffing Survey (SASS) at NCES, and the Residential Energy Consumption Survey (RECS) at EIA. There is also a large volume of methodological work on various aspects or dimensions of data quality, both as addenda and footnotes to major reports and as full-blown publications in their own right. While both of the statistical agencies described in these papers have contributed substantially to this literature and tradition, these two papers and the new philosophies to approaching data quality that they describe and represent are both remarkably congruent with one another and substantially different than the current mainstream in a few major respects.

In essence, these papers might be jointly entitled "Approaches to Survey Data Quality in an Era of Declining Resources for Statistical Data Collection and in the Face of Increased Competition for the Collection and Analysis of Statistical Information." Both address the meaning of data quality in the face of dwindling resources for statistical data, and both explicitly recognize competition for those resources and the need to focus more directly on quality from a customer, client, or user's perspective, as contrasted with that of the developer or producer of statistical information and products.

In effect, both papers argue, at least implicitly, that the market for statistical information is such that we must increasingly view data quality largely from a customer's or user's perspective. That is, the ultimate standard for quality is in the eye of the user. While this perspective is in evidence to some degree throughout the U.S. statistical system, and a recent focus on customer satisfaction and reinventing government (e.g., GRPA and other legislation) essentially requires that all government agencies give at least some serious attention to this perspective, both EIA and Statistics Sweden have been specially zealous in pursuing these concepts, in both cases apparently due to significant changes in their business environments or climates (i.e., budget reductions and the need to compete with other agencies in providing their services and products). More specifically, the approaches being taken by each agency derive from a "Total Quality Management" paradigm. In one form or another, this perspective has drawn significant attention throughout the statistical establishment, and, assuming that such attention will not diminish significantly over the next few years, it is important to explore more thoroughly some of the key *implications* of this particular view of data quality. Even if one does not fully agree with this particular point of view, it is still important that we recognize some of the important implications of taking its basic tenets seriously.

First, consider two of the basic distinctions made in these two particular papers. One, already alluded to, is an emphasis on quality as defined by the customer or data user rather than a

producer-centered view of data quality. As the paper by our colleagues at Statistics Sweden notes, the vast majority of approaches we have typically taken to define and assess data quality are decidedly producer-focused rather than client-focused in their basic orientations. A second dimension highlighted in this session is an emphasis on process versus product quality, a distinction also emphasized in particular in the paper developed by Statistics Sweden.

The distinction between a user-based versus producer-based definition of statistical data quality is stated most clearly by Claes Andersson and his colleagues at Statistics Canada:

> A product's quality is determined by the user's opinion of the product and its usefulness. The user's opinion should direct the approach of the development work. . . . Quality encompasses **all characteristics of a product** (commodity or service) influencing how well the product satisfies the user's needs and expectations. . . . The **user makes the judgement** and decides if the quality of the product is good or bad in relation to his/her intended use of it.

Thus, as noted by Anderson et al., it is essential for the producer to be fully aware of the quality judgments of both existing and potential users, since these provide the very basis for efforts to improve quality .

But that is not all. Another consequence of this way of defining quality is that users must then take greater responsibility for the level of quality and for ensuring that sufficient resources are available to reach this level, a responsibility previously vested in most instances with the producers of statistical data. As a result, this point of view also entails yet another important responsibility implied by the EIA paper--a responsibility to provide the information necessary for users to make such judgments in an informed and effective manner. However, this clear and obvious need also presents a potential dilemma. As noted by Renee Miller, EIA has:

> been providing our users with explanatory material for years. Yet during our customer satisfaction survey, some have told us that they have no way of judging the quality of our data. They think that EIA has good quality data, but they say they do not know for sure. Something seems amiss here.

Indeed. And, is it not in effect a major new responsibility of producers of statistical data and products to ensure that these customers and users have adequate information to make fully informed judgments of quality in relation to their critical uses and needs?

In turn, this need for information touches on the other basic distinction alluded to earlier--process versus product quality. As our colleagues at Statistics Sweden point out, users are fundamentally most interested in the quality of the product and less so in the process leading to the product. However, process information is often the only thing producers have to offer. Thus, there is a natural tension between the information generally available and what the user ideally wants or needs. Ironically, the major new surge throughout our industry to computerize survey and related information collection processes offers considerably more and higher quality

information on process quality (of substantial utility to producers of statistical data) but very little more information on product quality (of greatest interest to users), except by inference.

There is a general faith that product quality can be and is achieved through increased process quality. Thus, we work hard on improving processes (e.g., by developing standardized procedures and checklists, by continuously measuring key process variables, etc.) to achieve product quality. That process quality will automatically result in product quality cannot always be assumed, however. For example, in her paper, Miller correctly notes that adherence to standards is not synonymous with achieving data quality, just maintaining a low-fat diet and exercising regularly does not ensure that one will be disease free. Thus, the relationship between standards and standard processes and data quality is not an exact one, although standards are most clearly quite helpful from another perspective—for establishing credibility among our users. However, as the EIA customer satisfaction data described by Miller suggest, the relationship between standards and credibility may still be a very tenuous one.

Perhaps the key implication then of this new perspective on data quality—a user or customer-based perspective—is that one must take quite seriously a fundamental responsibility to provide the information necessary for users to make well-informed judgments on the quality of our statistical services and products. And, these two papers provide several good examples of how difficult it may be in practice to shift our basic paradigms in this direction. First, consider the EIA observation that, in spite of their having provided users with detailed explanatory information (derived from a producer-based persepective) for several years, customer surveys indicated that users still felt that they had no good way of judging the quality of EIA data.

Similarly, Andersson and his colleagues highlight two additional examples from Statistics Sweden in describing their evaluation studies, which have mostly been producer rather than user-oriented. For example, they note that while statisticians undertaking such studies (producers) typically focus on the *net* effects of systematic bias, from a broader, user-based perspective *gross* effects are also of legitimate interest for a number reasons. In the same context, they note that measures of relative bias derived from evaluations can be presented either as deviations between observed and "true" values as: (1) % of "true" (unbiased) value (of interest to *producers* as deviation from goal), or (2) % of the "official" estimate (of greater interest to *users* as an indication of error in published figures).

Another example is evident in the description by Andersson et al. of The Quality Report published annually since 1988 by Statistics Sweden to provide a basis for an analysis of the development of quality in their statistical products. These reports are based on questionnaires administered to every person responsible for a statistical product at the SCB, who are asked to respond from the intended user's perspective. However, users themselves are not routinely asked to share their opinions directly.

A final example that best illustrates perhaps how far we may still need to go to fully meet the demands and responsibilities associated with a user-based orientation to data quality is derived from Miller's discussion of EIA's *Quality Profile* for the Residential Energy Consumption Survey (RECS):

The *Quality Profile* has been very well-received. Several members of our energy advisory committee thought it a good model of how we should document our surveys. They pointed out that customer satisfaction is a function of data quality and that a quality profile would give users all the information they would need to determine data quality. It was suggested that we do more profiles. Unfortunately due to our budget constarints and reduction in staff levels that does not appear likely.

In an era where we purport to be adopting and take quite seriously a true user or client-based approach to data quality, can we really afford *not* to provide such information both routinely and in considerable depth?

In closing, I wish to thank the authors from SCB and EIA for presenting these very stimulating papers. Read in combination, they are extremely informative and thought-provoking and may well serve as precisely the right type of stimulus to ensure that we clearly recognize the full potential, implications and responsibilities that accompany these new and still somewhat controversial ways of conceiving of data quality.

Session 7

PERFORMANCE MEASUREMENT IN STATISTICAL AGENCIES

Performance Based Management:
Using the Measures
Nancy Kirkendall and Paul Staller

Today, and in the foreseeable future, government agencies will need to operate with decreasing resources. Concurrently, there is a rising level of expectation concerning the service quality provided by government agencies. These two trends present a challenge to government managers and staff. Another current operating today is the ever-increasing focus on the outputs and outcomes of government agencies' operations and policies.

Traditional government management has been focused on the preservation of resources (inputs) as opposed to the results of programs (outputs and outcomes.) Increasingly, citizens are asking the government and Congress, "What am I getting for my tax dollar?" The Energy Information Administration (EIA) has already received feedback on its latest budget submission to the Department of Energy (DOE) and the Office of Management and Budget (OMB) asking for specific EIA goals and objectives. It is anticipated that the Congress will ask the same questions during the next budget cycle. In short, EIA and the rest of the Federal Government are being asked to describe what we provide for the resources we are given.

Over the past three years, Congress has codified these trends principally in two Acts. In 1993 Congress enacted the Government Performance and Results Act (GPRA), and in the following year enacted the Government Management Reform Act (GMRA). In enacting these two laws, Congress has directed the Federal Government to manage itself using performance measurement (to include the establishment of performance goals and objectives), provide for reasonable managerial flexibility while ensuring managerial accountability, and provide for the financial stewardship of the funds and other assets entrusted to its care according to established government-wide standards.

Over this same period, EIA has been a leader within the DOE Headquarters in implementing not only quality practices as part of its Quality Program, but also best business practices in its line operations. Additionally, EIA has been a leader in the implementation of the GPRA, in which EIA has been participating as a pilot project (in fact, the only statistical agency to participate). As part of this pilot project, EIA has developed a set of agency-wide performance measures and collected some of the necessary data to support these measures. The results of this data collection effort were included for the first time in the Fiscal Year 1998 Budget Submission to the Congress.

From 1994 through mid 1996 the EIA identified a set of performance measures to monitor progress toward its strategic plan and started collecting the data to support them. EIA's efforts in the development of performance measures is described in Kirkendall (1996). A more complete description of the background for EIA's work in the development of performance measures isavailable among the case studies assembled by the American Society for Public Administration's Task Force on Government Accomplishment and Accountability Task Force

(Reference 2). In retrospect, this first part of the process, the development of performance measures and collection of data, seems relatively straight forward.

The next step is to implement performance based management, which is defined as "the strategic application of information generated by performance plans, measurement and evaluation to strategic planning and budget formulation[1]". To achieve performance based management, a major change is needed in how the organization is managed. To achieve the change, managers must accept the measures, the targets set for those measures, and must use them to guide their planning and resource allocation. Additionally, staff at all levels of the organization need to agree that the measures and their targets are reasonable, doable, and constitute a challenge for the future. This paper describes EIA's approach to the implementation of Performance Based Management. This report documents work in progress. We anticipate a successful outcome.

Background

Through the summer of 1996, the performance measurement effort in EIA concentrated on several steps:

1. The Strategic Plan
2. The input/output chart
3. Deciding what to measure
4. Collecting the data

EIA's Senior managers developed their first strategic plan in the Spring of 1994. In their annual strategic planning sessions since then, they have reviewed the strategic plan and made minor revisions. EIA's mission, vision and strategic goals are shown in Attachment 1.

Shortly thereafter, the Performance Measurement Development Team developed an input/output chart for the EIA. Using the input/output chart, and the EIA strategic plan, the team identified 14 performance measurement categories. The input/output chart, the 14 measurement categories, and the measurement types are illustrated in Attachment 2.

We believe that this information is particularly relevant to other statistical agencies. While we all do things a bit differently, we have in common the collection and processing of information, the analysis of information, and dissemination. EIA's strategic plan and input/output chart should be similar to those of other statistical agencies, and many of EIA's measures are likely to be of importance to other statistical agencies as well.

EIA concentrated on collecting the data to support the computation of the measures during 1995 and early 1996. As statistical agencies, we are all experienced in data collection and know how to do it. However, data collection is a major undertaking, and requires the commitment of resources by managers and staff. Though EIA's performance measurement data systems are by

---

1 *Guidelines for Performance Measurement*, U.S. Department of Energy, June 1996.

no means complete, we believe that the time has come to implement performance based management and to use the measures .

The Next Step - Where We Are Going

Part of the process of implementing a system of performance measurement is convincing everyone that the measures are important and useful. If measures are not used, staff and managers alike will believe that there is no reason to spend resources to collect and maintain them.

Attachment 3 demonstrates that Performance Based Management is a cycle providing input to the organization (at the bottom of the chart). This cycle involves strategic planning, collection of measures which demonstrate how well the organization is performing in response to strategic planning initiatives, and the evaluation of results and measures. The evaluation is fed into the next cycle of strategic planning.

In EIA we have had a strategic planning process since 1994. We have had some performance measurement since 1995. The next challenge is to insert the evaluation of results and measures into the process and to assure that measures and results are used by the strategic planners.

To achieve this, in 1996 the Performance Measurement Team drafted a performance agreement, listing measures for each strategic goal, along with a specific targets for the year 2002. This performance plan is based on the measures for which we have data. This plan was submitted to senior staff and circulated for comment throughout the organization. The plan was revised based on input, and was ultimately adopted by the senior managers and the quality council.

This is the first step in the implementation of performance based management. Adoption of the measures and targets by senior managers will help to institutionalize the process. It now requires follow through. Managers and staff are expected to review and use the information to evaluate their progress toward targets. Managers are provided flexibility to allocate resources to achieve the agreed-to targets. The Administrator of EIA has said that he would like to see performance measurement information quarterly. Ultimately the process will help managers and staff communicate about what is important and how success will be measured. It also provides both managers and staff the information they need to communicate with internal and external customers and stakeholders, including the OMB and Congress.

The EIA Performance Agreement

In previous years EIA has labored through a process that had been designed to allocate the resources available to EIA amongst EIA's programs. In the current environment, this process showed numerous weakness and provided limited value to the management of EIA. The process does have the advantage of providing a forum for the Administrator and Deputy Administrator to provide input to the direction of the EIA programs in the coming year.

With the direction provided by the Congress and as implemented by the OMB and DOE, it was decided that now was the time to drastically revise the processes used by EIA. EIA also had the opportunity to continue with it's tradition of leadership. For Fiscal Year 1997 the present resource allocation process was scrapped. Program direction and input will continue to be provided during less formal discussions between the EIA's Office Directors and either the Administrator or the Deputy Administrator or both as appropriate. EIA shifted the focus of it's process from inputs to corporate outputs/outcomes by conducting a one day session designed to establish performance objectives for each of the strategic goals and the associated performance measures. The actors in this process were EIA Senior managers and representatives from the EIA Performance Measures Team and the resource management office. The tangible output of this session was a performance agreement for the Energy Information Administration that is loosely modeled after the *Performance Agreement between The President of the United States and The Secretary of Energy for Fiscal Year 1996.*

The intent of the performance agreement is to establish a set of measurable short-term and long-term objectives for the agency, as envisioned by the GPRA, and to base these upon the established performance measures and EIA's existing Strategic Plan. Features of the performance agreement are:

- The agreement is for the agency as a whole.

- The five goals in the EIA Strategic Plan will be used as the basis of the agreement. EIA's established set of performance measures are linked to these goals. This combination provides a solid foundation for the agency to measure the continuing success of its operations.

- The agreement establishes performance objectives for the EIA strategic goals for the year 2002, as required under the provisions of the GPRA.

- Managers are expected to manage towards these objectives by allocating resources to meet them and, where necessary, redesigning processes under their control.

- EIA's Annual Report to Congress will become EIA's performance report documenting progress toward its established objectives and the fulfillment of its performance agreement.

- The 1997 agreement is a mock agreement and used in-house only. The 1998 agreement is expected to be the formal performance planning document required under GPRA.

Using the Measures

A concern that emerged from initial discussions with the EIA Performance Measures Team, senior managers and selected staff, on this process was the linkage between the objectives/targets to be established in the performance agreement and the allocation of resources. If EIA meets it's

objectives do we get more or less money? That's not the objective of performance measurement. The objective is to improve the performance of the EIA, not necessarily to increase the size of EIA's budget or any portion of EIA's budget. The measures will need to be examined as a whole, and there will be opportunities to explain why performance did or did not meet the objectives. One possible result could be that the objectives are unachievable.

This year is a pilot year, the information will not be used for resource allocation. This year the performance objectives/targets will be established, and the process is viewed as being more akin to strategic planning. The actual performance information will be collected and then analyzed by the EIA Performance Measures Team. The results of the analysis will then be presented to the EIA senior managers in time for the next round of Strategic Planning, that is now scheduled to begin in February 1997.

In conducting this analysis, the EIA Performance Measures Team will need to keep in mind that there are two types of performance measures: efficiency measures or "doing things right", and effectiveness measures or "doing the right things". EIA needs to have measures supporting both aspects of performance, and each measure should have a target. The measures and their objectives will enable each program to be described more intelligently, and managers will be in a better position to make informed decisions. At this point in the development of performance based management at EIA, it is difficult to say *exactly* how performance measurement information will feed into budget decisions. What most likely will happen is the inclusion into the budget submissions of the description(s) of how the performance measures information and the resulting analysis were used to make decisions concerning the agency's programs.

Steps Followed

Development of the performance agreement was a collaborative process that involved senior management, line and staff personnel and a cross-cutting commttee. The development process consists of six major steps. The six step process has proceeded in the following manner:

Step 1: EIA's resource management office provided Senior Staff with the resource allocations for Fiscal Year 1997.

Step 2: Draft performance agreement proposing corporate objectives circulated for review and comment. Draft performance agreement was developed by the EIA Performance Measures Team.

Step 3: EIA offices respond with comments on the draft performance agreement.

Step 4: Performance Measures Team consolidated comments on the draft performance agreement from EIA offices and highlights areas of agreement and disagreement. Consolidated comments and recommended objectives were reported back to EIA senior managers. A copy of the proposed Performance Agreement, as reported back to the senior managers is included at the end of this paper.

Step 5: EIA senior managers adopt the Fiscal Year 1997 Performance Agreement.

Step 6: Fiscal Year 1997 Performance Agreement distributed to all EIA employees.

Summary

In summary, this is a drastic revision to EIA's processes that entails moving the focus from the resources to be used and towards the outputs and outcomes of the use of those resources. In addition, the process will provide for increased managerial flexibility while instituting some limited managerial accountability. All of this is consistent with the expressed desires of the Congress as expressed in the GPRA and the GMRA. Other additional benefits that will accrue to the EIA from adopting this approach are a reduction in the "Us versus Them" behaviors associated with resource allocations, clear direction for management and staff implementation and some increased credibility with EIA's stakeholders.

References

1. Kirkendall, Nancy, "Organizational Performance Measurement in the Energy Information Administration," Proceedings of the 1996 Annual Research Conference, Bureau of the Census, U.S. Department of Commerce, August 1996.

2. "Use and Development of Performance Measures: Department of Energy, Energy Information Administration", American Society for Public Administration, Government Accomplishment and Accountability Task Force, July, 1996."[2]

---

[2] Abstracts from ASPA's Government Accomplishment and Accountabiltiy Task Force are available on the World Wide Web at: http://globe.lmi.org/aspa/taskfrc.htm. Documents can be ordered from American Society for Public Administration, 1120 G St. NW Suite 700, Washington DC 20005-3885.

Attachment 1

Energy Information Administration
Mission, Vision and Goals

Mission

The Energy Information Administration is a leader in providing high, quality, policy-independent energy information to meet the requirements of Government, industry, and the public in a manner that promotes sound policymaking, efficient markets, and public understanding.

Vision

- EIA is a unified team committed to excellence and customer satisfaction
- EIA leaders recognize employee's potential and together create a workplace where team work and innovation are encouraged, supported and realized.
- Everyone in EIA develops their technical and analytical capabilities to keep abreast of new technologies and changes. This enables our employees to reach their full potential and enables us to rely more on our in-house capabilities.
- EIA expands its customer base and becomes nationally and internationally recognized as the premier source of energy information.
- EIA rengineers and standardizes core business systems.
- EIA improves productivity and supports the delivery of customer-oriented products and services.
- The EIA Strategic Plan is a road map for a EIA decisions and is used as the basis for alignment of human and financial resources.
- EIA works in partnership with the National Treasury Employees Union to accomplish our mission and reach our vision.

Goals

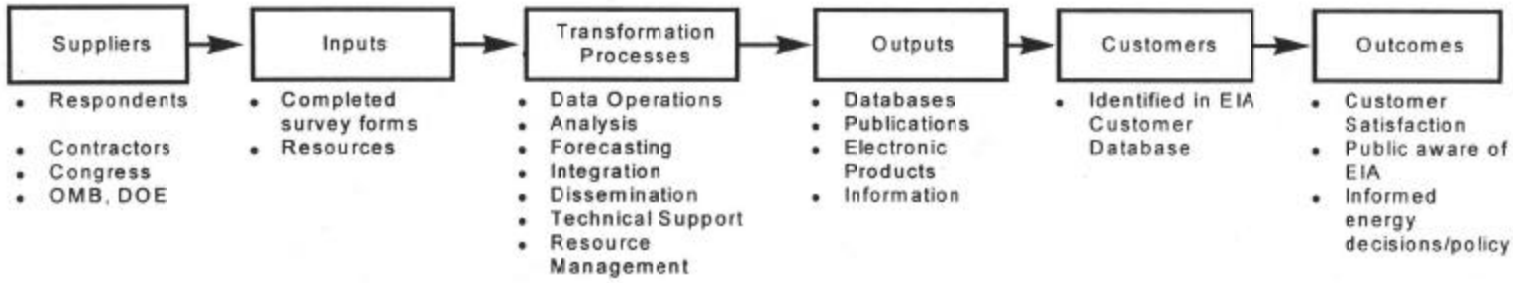
- We will work together to achieve the full potential of a diverse workforce through team work and employee development.

- EIA will assure its data and analyses are of the highest quality and relevant to the needs of its customers.
- EIA will provide its customers fast and easy access to public energy information.
- We will make resource and program decisions based upon customer input and conduct our business in an efficient and cost-effective manner.
- EIA will be an objective partner in fulfilling the mission of the Department of Energy.

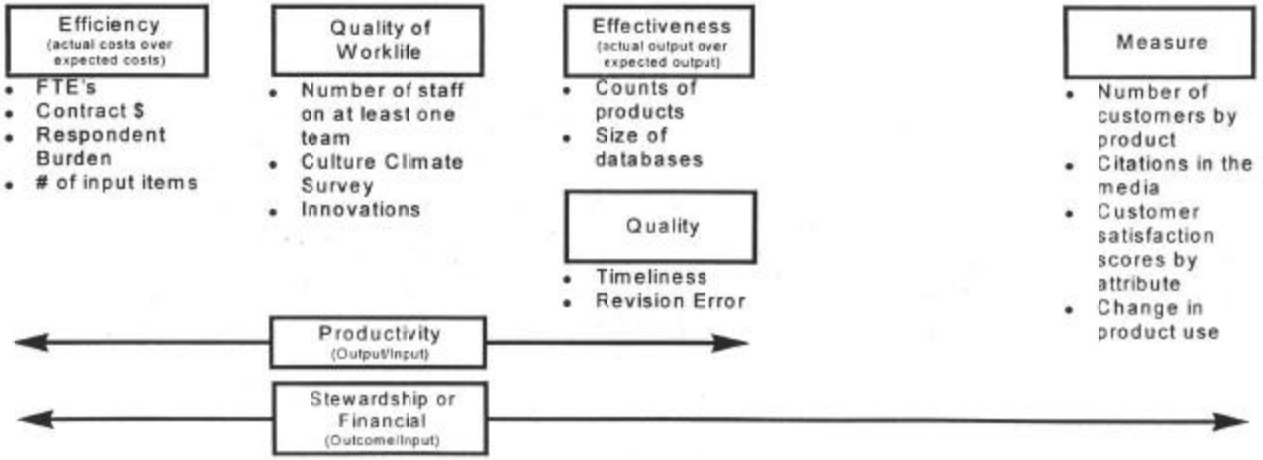
Attachment 2

EIA's Input/Output Chart and Performance Measures

Simplified Input/Output Chart:



Measures:



Attachment 3

EIA's Performance Based Management Cycle

Proposed EIA Performance Agreement

Goal 1: We will work together to achieve the full potential of a diverse workforce through teamwork and employee development.

Success will be measured by:

1.1 By the year 2002, the average score for the following questions on the culture climate survey will increase to the benchmark noted.

1.1.1. "There is a high spirit of teamwork among my co-workers." The benchmark is 5.69. In the 1994 culture climate survey, EIA's score on this question was 4.30 and in 1995 the score was 4.64.

1.1.2. "My supervisor ensures that I get job related training when needed." The benchmark is 4.99. In the 1994 culture climate survey, EIA's score on this question was 4.49 and in 1995 the score was 4.63.

Goal 2: EIA will assure its data and analyses are of the highest quality and relevant to the needs of its customers.

Success will be measured by:

2.1 During the period between 1996 and 2002, accuracy will remain stable, or improve over time, as the EIA improves the timeliness of it's products.

2.1.1. The accuracy of data will be measured by percent sampling error, percent revision error, and unaccounted for balances.

2.1.2. The accuracy of forecasts will be measured by:

2.1.2.1. The percent difference between actual and forecast for STEO.

2.1.2.2. Compare AEO forecasts of key variables with historical data and provide a qualitative discussion of factors that led to differences.

2.1.2.3. Compare IEO forecasts of total world energy consumption and world consumption by fuel with historical data in 5 year increments beginning with the availability of 1995 international data. Provide

a qualitative discussion of factors that led to differences.

2.2 By the year 2002, the EIA will increase the number of customers who are very satisfied with accuracy to 60 percent. In the 1995 and 1996 EIA Customer Surveys, the results indicated that 51% and 52%, respectively, of the customers surveyed were very satisfied with the accuracy of EIA's products.

2.3 By the year 2002, the EIA will increase the number of customers who are very satisfied with relevance to 70 percent. In the 1995 and 1996 EIA Customer Surveys, the results indicated that 60% and 58%, respectively, of the customers surveyed were very satisfied with the relevance of EIA's products.

2.4 During the period between 1996 and 2002, citations of energy information attributed to EIA in the media:

2.4.1 Overall growth in media citations will increase by 10% per year.

2.4.2 Citations in major media will increase by 40% per year.

2.4.3 EIA's share of a market basket of energy citations will increase.

2.5. Growth of customer base:

2.5.1. During the period between 1996 and 2002, the number of unique daily users of EIA's Internet site will increase by 25% per year.

2.5.2. During the period between 1996 and 2002, the distribution of published copies of data reports, analysis reports, and feature articles combined with the downloads of the electronic file versions will increase by 25 percent per year.

2.5.3 During the period between 1996 and 2002, the number of Energy InfoDiscs sold will increase by 5% per year.

2.5.4. By the year 2002, the Energy InfoDisc annual subscription renewal rate will be 50%.

Goal 3: EIA will provide its customers fast and easy access to public energy information.

Success will be measured by:

3.1. By the year 2002, the EIA will increase the percent of customers who are satisfied or very satisfied with timeliness to 80 percent. In the 1995 and 1996 EIA

Customer Surveys, the results indicated that 72% and 73%, respectively, of the customers surveyed were satisfied or very satisfied with the timeliness of EIA's products.

3.2. By the year 2002, the EIA will increase the percent of customers who are very satisfied with ease of access to 70 percent. In the 1995 and 1996 EIA Customer Surveys, the results indicated that 64% and 54%, respectively, of the customers surveyed were very satisfied with the ease of access to EIA's products.

3.3. By the year 2002, the EIA will improve the timeliness of it's products to:

3.3.1 The median for all EIA annual publications will be 180 days after the close of the reference period. The median for electronic release of EIA annual publications will be 165 days after the close of the reference period. In 1993 and 1994 the median for all EIA annual publications was 342 and 321 days, respectively, after the close of the reference period.

3.3.2 The median for all EIA quarterly publications will be 90 days after the close of the reference period. The median for electronic release of EIA quarterly publications will be 75 days after the close of the reference period. In 1994 and 1995 the median for all EIA quarterly publications was 146 and 144 days, respectively, after the close of the reference period.

3.3.3 The median for all EIA monthly publications will be 30 days after the close of the reference period. The median for electronic release of EIA monthly publications will be 20 days after the close of the reference period. In 1994 and 1995 the median for all EIA monthly publications was 74 and 71 days, respectively, after the close of the reference period.

Goal 4: We will make resource and program decisions based on customer input and conduct our business in an efficient and cost-effective manner.

Success will be measured by:

4.1 By the year 2002, the EIA will increase the percent of customers very satisfied with overall service to 80 percent. In the 1995 and 1996 EIA Customer Surveys, the results indicated that 68% and 69%, respectively, of the customers surveyed were very satisfied with the overall service provided by EIA.

Goal 5: EIA will be an objective partner in fulfilling the mission of the Department of Energy.

Performance measures for Goal 5 will be discussed during the next EIA Strategic Planning Session.

An Overview of

The Strategic Management Initiative

A Commitment to Customer Satisfaction, Employee Satisfaction, and Productivity

Office of Technology and Survey Processing

Section 1

1.0 Executive Summary

The Office of Technology and Survey Processing (OTSP) is an organization of approximately 650 people within the US Department of Labor, Bureau of Labor Statistics (BLS). OTSP delivers roughly 114 million units of 550 unique information technology products and services. We are in the midst of implementing the Strategic Management Initiative (SMI).

The SMI is a business management process for defining and objectively measuring our success. It links strategic planning concepts to results through measurement and accountability. It is also a structured way for our customers and employees to set our priorities and hold us accountable for measurable results. We expect the SMI to do for us what the business plan, profit/loss, and balance sheet does for the private sector -- help us articulate organizational goals and priorities, define success in measurable terms, and stimulate self-correcting behavior.

The SMI yields three key measures or indicators --

- Customer Satisfaction with Our Products
- Employee Satisfaction with Job Factors
- Productivity of Our Processes

We believe that these indicators are easy to understand, promote action, get to the heart of the "value adding" mission of any organization, and align the best interests of our customers, employees and the taxpayers. The key challenge will be to improve results in all three of these indicators simultaneously.

A thorough review of this guide will acquaint you with our process for --

- Developing an inventory of products, processes, and customers
- Setting measurable baseline and target scores for customer satisfaction with products, employee satisfaction with job factors and productivity of processes
- Establishing strategic enterprise-wide goals linked to products, processes and/or job factors
- Deploying improvement/innovation work down to the line staff level
- Re-measuring and holding managers and line staff accountable for results.

Section 2

2.0 Introduction To The SMI

2.1 What is the SMI?

2.2 Why are we doing it?

2.3 Will the SMI work?

2.4 Who is involved?

2.5 What's in it for . . .?

2.6 What will it cost?

2.7 What is the process?

2.8 When will results happen?

2.9 How do we define success?

2.10 Sample of SMI Products

2.1 What Is The Strategic Management Initiative (SMI)?

The SMI is a business management process for defining and objectively measuring our success. It links strategic planning concepts to results through measurement and accountability. It is also a structured way for our customers and employees to set our priorities and hold us accountable for measurable results. We expect the SMI to do for us what the business plan, profit/loss, and balance sheet does for the private sector -- help us articulate organizational goals and priorities, define success in measurable terms, and stimulate self-correcting behavior.

Because OTSP exists to deliver valued technology-based products to our customers, customer satisfaction is a prime component of the SMI. Because our employees operate the processes that result in product delivery, employee satisfaction is another key SMI component. Since the Federal Government must be accountable to the taxpayers, productivity is the third component of this business management process.

Therefore, the SMI is designed to yield three key measures or indicators --

- Customer Satisfaction with Our Products
- Employee Satisfaction with Job Factors
- Productivity of Our Processes

We believe that these indicators are easy to understand, promote action, and get to the heart of the "value adding" mission of OTSP. The SMI attempts to align the best interests of our customers, employees, and the taxpayers. The key challenge will be to improve results in all three of these indicators simultaneously.

We want the SMI to help create on-going incentives for all of our employees to work in concert across the Office towards enterprise level success goals. Business process reengineering and quality gurus call this "organizational alignment."

Through the SMI, we will have the ability to --

- Survey customers & employees for baseline data on OTSP performance
- Set measurable long & short term goals for future performance
- Deploy & implement goals across and at all levels of OTSP
- Re-survey customers and employees for evaluation of our results.

In short, the SMI is designed to help us improve:

1. OTSP accountability to its customers,
2. front line staff accountability to managers, and
3. manager accountability to front line staff.

2.2 Why are we doing it?

OTSP, an office of approximately 650 people, delivers hundreds of information technology related products to our BLS customers every year. These products fall in five major categories: *software*, including all of our computer systems; *system outputs*, including tables, data files and paper listings; *customer support products*, including training, help-line services, software and hardware maintenance services and manuals; *technology management products*, including research reports on new technology and delivery orders; and *management and administrative products*. They touch every BLS employee and are cornerstones of the agency's ability to deliver its external statistical products. OTSP delivers roughly 114 million units of 550 unique products that fit into these five categories.

While OTSP has enjoyed much success, we currently face many complex challenges.

- The Commissioner and Secretary of Labor challenge us to promote action on customer outreach, employee involvement, and customer service standards initiatives.
- Our customers challenge us to deliver more, better, and less costly products faster.
- The administration challenges us to cut costs, increase productivity, and align with the National Performance Review & Reinvention; the Government Performance and Results Act of 1993 (mandates strategic planning and results measurement for agencies); and Executive Order 12862 on Setting Customer Service Standards (mandates customer service standards, plans for customer surveys and identifying other organizations that will be used to benchmark performance.)
- We challenge ourselves to infuse rapidly changing technology to boost the performance of our production systems, and to maintain a highly skilled and motivated staff as we change technology, cut budgets, and increase workloads.

While these challenges represent added workload, they also offer opportunities for greater success. The SMI represents a carefully planned integrated effort to help us absorb the workload and realize the success opportunities. As we face these and other challenges, we must not forget that OTSP exists to deliver valued technology-based products to our BLS customers. The SMI will measure the success of our product line in the eyes of our customers.

2.3 Will the SMI work?

The SMI is a business management approach modeled after what works in the private sector. As mentioned above, the private sector has powerful financial and survey based measurement tools to clearly and forcefully articulate organizational goals and priorities, define success in measurable terms, and stimulate self-correcting behavior. These tools are effective because they enforce accountability. The SMI is designed to be our counterpart to proven private sector tools.

2.4 Who is involved?

Everybody! The SMI represents a cooperative team effort between our managers, our front line staff, and our customers to prioritize and focus OTSP resources where results are most needed.

As stated before, the SMI revolves around three key measures -- Customer Satisfaction, Employee Satisfaction, and Internal Productivity.

To develop **Customer Satisfaction** scores, our employees will identify the products we deliver to each customer. Then, customer involvement becomes the cornerstone of our strategy to measure and improve customer satisfaction. Through a survey, our customers will prioritize the product list and score their satisfaction with individual products. Our customers fall into two categories -- Sponsors and Users. Sponsors pay for OTSP products. Users take delivery and make direct use of our products. We created these two customer categories to help us deal with the conflicting demands and priorities that these different customers sometimes place on us. The classic example of this occurs when a program office sponsor customer pays for a product that is not delivered to that program manager. Rather, the product is delivered to a regional office user customer.

We will measure the satisfaction level of both customer groups with our products. For sponsors, we will survey BLS program managers. The user survey process is not as direct. We have hundreds of users throughout the national office, regional offices, and state agencies. For a few products, we have general public users. To simplify the user survey process, BLS cost center managers and branch chiefs will serve as our focal point for gathering user satisfaction scores. To combine user and sponsor scores, we will apply weights: users, 35%; sponsors, 65%.

To develop **Employee Satisfaction** scores, we will survey all OTSP employees for their satisfaction with nine specific job factors.

To develop **Internal Productivity** scores for OTSP, our managers and front line staff will work together to calculate the unit cost of operating selected processes.

Our managers, front line staff, and customers must all work together if we are to realize our goal of simultaneous improvement of all three scores.

2.5 What's in it for . . . ?

For any undertaking or system of work to sustain itself as a successful ongoing enterprise, it must deliver products that customers sufficiently value and yield benefits that sufficiently motivate the producers. Customers sufficiently value a product when they are satisfied to pay the necessary "price" for the product. Producers are sufficiently motivated when the monetary and non-monetary benefits associated with producing and delivering the product satisfy them.

The SMI is a system of work that yields information based products. Currently, OTSP is the producer and BLS managers and line staff are the primary customers of SMI products. We expect customers and producers to obtain the following benefits from these products:

OTSP Customers

We will be more accountable to our customers. We will give our customers:

1. an inventory of the products that we produce for them,
2. a structured way to set product priorities,
3. a structured way to communicate their satisfaction with our products, and
4. a meaningful set of measures that will allow them to observe the results of our efforts to increase customer satisfaction and reduce targeted unit costs.

The distribution of baseline and target scores to our customers should generate incentives for OTSP employees to improve customer satisfaction with their products and to improve the productivity of their processes.

OTSP Managers

Our front line staff will be more accountable to our managers and will be more productive through empowerment. Products and processes targeted for specific measurable customer satisfaction and productivity improvements will be clearly linked to front line individuals or teams.

Our managers will get more productivity from an empowered front line staff; they will gain a structured tool to help them prioritize improvement efforts; and they will have a clear focused feedback mechanism. The distribution of baseline and target improvement scores to top management and customers should generate incentives for our managers to take beneficial risks to achieve the target scores.

OTSP Front Line Staff

Our managers will be more accountable to our front line staff. Front line employees will have more job satisfaction through empowerment. All employees will get a survey that allows them to declare priorities among job satisfaction factors and a set of division level baseline and target employee satisfaction scores. The distribution of these scores within each OTSP division should generate incentives for managers to improve employee satisfaction.

Our front line staff will gain empowerment. They will be empowered to increase customer satisfaction with the products they produce and to increase productivity within the processes that they operate. They will gain a clear understanding of customer needs and priorities; and like managers, they will have a clear focused feedback mechanism. The distribution of baseline and target improvement scores to managers and customers for specific products and processes should generate incentives for our front line staff to take beneficial risks to achieve the target scores.

BLS Top Management

Top BLS management will reap the human and financial rewards that may come from all of the above -- more satisfied BLS customers, reduced rework, more efficient processes, and higher employee morale. In addition, they will have a mechanism in place for responding to ever increasing external demands for performance based information, and results that satisfy DOL and administration efforts to reengineer and improve governmental operations.

2.6 What will it cost?

We estimate that OTSP will expend between .75% and 1.75% of its internal human resources on planning and measurement related work for the SMI. We believe that this is very much in line with private sector costs for these kinds of activities. In addition private sector managers view SMI-like activities as an absolutely essential factor in achieving success and productivity gains.

We estimate that our BLS customers -- program managers, cost center managers, and some staff members -- will each spend on average between 1.5 and 4.0 hours per year providing structured feedback on our products. Bureau-wide about 850 people will expend a total of 1275 to 2500 annual person hours.

2.7 What is the process?

The following table provides a brief overview of the methodology that we have developed for measuring Customer Satisfaction (CS), Employee Satisfaction (ES), and Internal Productivity (IP).

An Overview of the SMI Model – Table 2.1

What Outcome Do We Want?	How Do We Get It?	What Needs To Be Measured?	How Do We Measure It?
Increased Customer Satisfaction (CS)	Deliver Better Products (Goods & Services)	Customer Satisfaction with Product Factors -- Features/Completeness Easy to Obtain/Use Timely Delivery Quality (Errors/Defects) Cost Courteous Treatment	• Inventory Products • Map Products to Customers • Measure Baseline CS Scores • Select Products to Improve • Set Target CS Scores • Innovate/Improve Products • Re-Measure at Year End -- Compare Results to Baseline & Target Scores
Increased Employee Satisfaction (ES)	Design Better Jobs & Align The Incentive System	Job Satisfaction Factors -- Job Duties Quality of Supervision Training/Skills Dev. Mgmt. Leadership Communications Work Rule Flexibility Job Growth/Promotions Awards & Recognition Workforce Diversity & Fairness	• Measure Baseline ES • Select Factors to Improve • Calculate Baseline for Factors • Set Target ES Scores • Take Action to Improve Factor Satisfaction • Re-Measure at Year End -- Compare Results to Baseline & Target Scores
Increased Internal Productivity (IP)	Redesign & Reengineer Processes	Product/Process Unit Cost	• Map Processes to Each Product • Select Product/Process Pair to Improve • Baseline IP -- Calculate Unit Costs • Set Target IP Scores • Innovate/Improve Processes • Re-Measure at Year End -- Compare Results to Baseline & Target Scores

The activities described in Table 2.1 will be implemented and deployed across nine events:

Event 1 - Create Product/Customer Inventory

Event 2 - Measure Baseline Customer & Employee Satisfaction

Event 3 - Analyze Baseline CS & ES Scores

Event 4 - Establish OTSP-Wide Strategic Goals

Event 5 - Set Measurable Division Level Goals for Year

Event 6 - Finalize 5-Year & 1-Year Goals

Event 7 - Innovate/Improve Products, Processes & Job Factors

Event 8 - Measure Change From Baseline Scores -- Reward for Success

Event 9 - Improve SMI Process & Repeat Events

2.8 When will results happen?

While the SMI draws resources from part-time volunteers only, we have made major strides in developing and deploying the SMI. Key accomplishments are summarized below:

1990	customer demands for faster product delivery sparks total quality management (TQM) & systems development live cycle (SDLC) research
1992	published article on CS, ES, IP organizational success factors and held workshops on business process reengineering (BPR)
	decision made to launch results oriented strategic planning project
1993	completed pilot of product, process, and customer inventory
	completed version 1 of SMI measurement methodology, analyzed its cost/benefits, and began developing systems and procedures for collecting data and calculating the measures
1994	installed version 1 of SMI client/server based system onto OTSP PC's and launched SMI pilot to measure customer satisfaction, internal productivity and the success of strategic outcomes

1995	pilot customers surveyed for CS, delivered baseline scores, declared strategic outcomes, set CS & IP target scores
	deployed managers & staff to innovate & improve to reach target scores
1996	completed SMI pilot by resurveying customers, calculating actual CS and IP scores for comparison to targets, and evaluating strategic outcomes.
1997	delivered CS and IP score results to customers.

The SMI process is based on a continuous recurring cycle, at the end of which we will have measurable results regarding our efforts to improve Customer Satisfaction (CS), Employee Satisfaction (ES), and Internal Productivity (IP). To date, we have not piloted the ES component of the SMI. Throughout the pilot we have solicited feedback from customers and OTSP participants, and we continue to listen for and make needed improvements to the SMI process, so that we can proceed smoothly with full implementation.

2.9 How do we define success?

One mark of success for the SMI will occur when OTSP front line staff and managers use SMI tools to help them --

- Define measurable success goals
- Set work and resource priorities
- Generate plans of action to improve products, processes and job factors
- Deploy resources
- Maintain accountability by measuring and re-measuring results

Another success milestone for the SMI will be the distribution to our customers and employees of the CS, ES, IP scores.

To summarize, our short term SMI goal is to focus our improvement efforts on customer and employee priorities. Our long term goal is simultaneous measurable improvement to **Customer Satisfaction**, **Employee Satisfaction** and **Internal Productivity**.

2.10 Sample of SMI Products

A wide array of useful measures and information products will be generated by the SMI for use by our customers, managers and front line staff. The tables on the following page provides a small sample of high level SMI measures. The data contained in these tables are for example purposes only.

OTSP Level FYXX Targets

Vision Category	FYXX Targets
Continuously Increase Customer Satisfaction With Our Products	OTSP will improve its Customer Satisfaction score from 49.5 to 54.4.
Continuously Increase Employee Satisfaction With Job Factors	For Targeted Job Factors OTSP will improve its Employee Satisfaction score from 47.3 to 52.3.
Continuously Increase Productivity Within Our Processes	Currently, OTSP is measuring productivity changes for 0 % of its total budget dollars. Within the next year, OTSP will measure productivity changes for **10%** of its total budget dollars. Over the next year, for its measured resource base, OTSP will increase productivity at an annual rate of **1.5 %**.

Summary of FYXX Scores - Division Level

Producer Division: Division of International Price Systems (DIPS)

Column 1: Vision Category	Column 2: FYXX Division Level Baseline Scores	Column 3: FYXX Division Level Target Scores	Column 4: FYXX Division Level Target Improvement	Column 5: FYXX Division Level Actual Score	Column 6: FYXX Percentage of Division Level Targets Achieved
Customer Satisfaction	42.7	47.1	10.3 %	48.5	103%
Employee Satisfaction	43.4	47.4	9.2 %	49.6	105%
Internal Productivity Unit Cost	$12,844	$12,529	2.5 %	$12,651	61.3%
Internal Productivity % of Dollars Measured	5%	11%	120%	9.5%	86%

TOOL CS1A: OTSP Customer Satisfaction Survey

FY XX

Customer/Contact: Jack Galvin, Program 203, Cost Center 230

OTSP Producer: Business Establishment Surveys **Bob Carlson (606-7300)**	Product Priority (1-10) 1=low, 10=high	Satisfaction Score (1-10) 1=low, 10=high
SYSTEM OUTPUTS - DATA/FORMS		
1. UDB - Universe Database Data		

OTSP Producer: Producer Price Systems **Phil Kirsch (606-7500)**	Product Priority (1-10) 1=low, 10=high	Satisfaction Score (1-10) 1=low, 10=high
SOFTWARE/HARDWARE MAINT/OPERATION SERVICES		
1. Apprise Maintenance Service		
SYSTEM OUTPUTS - DATA/FORMS		
2. Assignment Listing		
3. Compressed Print Files		
4. DIP LIST		
5. Frame Listings		
6. IIQMS Listings		
7. Index Analysis Listings		
8. Industry synopsis		
9. LABSTAT Updates		
10. Product Checklists		
11. Pub Tables		
12. Refined Sample (Listing)		
13. Re-pricing Data		
14. Re-pricing Forms		
15. SSR Listings		
16. Weekly Collection Listings		
NEW, REDESIGNED OR ENHANCED SOFTWARE		
17. Imaging system		
18. Maintenance Service for APPRISE		
CONSULTING/TECHNICAL ASSISTANCE/HELP SERVICES		
19. Misc. Estimation Requests		
MANUALS/MEMOS		
20. Sampling Maintenance Service		
21. ARTS		
22. Downsized define and enter system		
23. EDI Feasibility Study		
24. FAXing Feasibility Study		
25. Pen Based Data Collection Test		
26. Revised PPI Seasonal Adjustment System		

OTSP Producer: Management Information Systems **Jarred Coram (606-7547)**	Product Priority (1-10) 1=low, 10=high	Satisfaction Score (1-10) 1=low, 10=high
SYSTEM OUTPUTS - DATA/FORMS		
1. BLS Financial Profile Reports		
2. SF-52 Processing System		

OTSP Producer: Systems Design **Gwen Harllee (606-7572)**	Product Priority (1-10) 1 = low, 10=high	Satisfaction Score (1-10) 1=low, 10=high
BLS-WIDE ADMINISTRATIVE MANAGEMENT PRODUCTS		
1. Memos or Letters on contract compliance		
TRAINING SERVICES		
2. Training Schedules		

OTSP Producer: Systems Modernization **Rich Fecher (606-7552)**	Product Priority (1-10) 1 = low, 10=high	Satisfaction Score (1-10) 1=low, 10=high
NEW, REDESIGNED OR ENHANCED SOFTWARE		
1. LABSTAT Info Module (IKON)		
2. LABSTAT Microdata Trans. Module (MTS)		

OTSP Producer: Technology & Network Management **Tom Zuromskis (606-5950)**	Product Priority (1-10) 1=low, 10=high	Satisfaction Score (1-10) 1=low, 10=high
CONSULTING/TECHNICAL ASSISTANCE/HELP SERVICES		
1. Answer/Response to PC/LAN Help Request		
2. Document on How to Use LAN Service		
3. Functional Central LAN Services		
SOFTWARE/HARDWARE MAINTENANCE/OPERATION		
4. IBM 3800 Laser Mainframe Print-Out		
5. LAN Hardware Item Ordered, Delivered & Installed		
6. Mainframe Computer Account Financial Report		
7. Mainframe Computer Manual Delivery Service		
SOFTWARE/HARDWARE DELIVERY &/OR INSTALLATION		
8. Mainframe Impact Printer Print Job		
9. New/Updated Mainframe User Account		
10. PC Print-Out (Via Central LAN Printer)		
11. Repaired PC, Server, Printer, etc.		
12. Upgraded LAN Server		

TOOL CS9: Customer Satisfaction Score - Division Level

FY XX

Purpose:	To determine a Division CS score across all customers and products.
Tool Repetitions:	Create 1 CS9 per OTSP Division, sort by Col2/Col3 in descending order
Tool User:	Division management and staff
Operations:	TRANSFER CALCULATE INPUT

Producer Division: DPPS TRANSFER From SMI DATABASE for each OTSP Division

Column 1: List of All Division Products by Product Category	Column 2: Normalized Product Weight for the Division	Column 3: FYXX Product Level Baseline CS Score	Column 4: Division - Weighted Product Satisfaction Score
TRANSFER From SMI Database list all Status 1 & 2 products (with priority weight and score) for this producer division. Sort order for Product list is by Product Category (in ascending order of category number), and within each category, products in descending order of ratio Col2/Col3.	TRANSFER From the CS7B (Baseline) for the product listed in this row, insert Col3 For product category row, shade this column	TRANSFER From the CS8 form for the product listed in this row, insert Col4 Total For product category row, shade this column	CALCULATE Col2 x Col3 For product category row, shade this column
Weekly Collection Review Processing	0.034	5.0	0.17
Sampling Maintenance Service	0.028	14.5	0.41
Weekly Collected Data Listing	0.033	20.4	0.68
IIQMS Listings	0.027	22.0	0.60
NIH Move	0.036	60.0	2.17
Automated Regional Tracking System (ARTS)	0.029	48.6	1.42

SSR Listings	0.009	38.4	0.33
	CALCULATE Column sum (should equal 1.0)		CALCULATE FYXX Division Level Baseline CS Score Column sum
Totals	1.0		56.0

TOOL CS17: Target Customer Satisfaction Scores - Division Level

FY XX

Purpose:	To determine the overall target CS score and improvement percentage for each OSTP Division
Tool Repetitions:	Create 1 CS17 per OTSP Division
Tool User:	Division Management/Staff
Operations:	TRANSFER CALCULATE INPUT

Producer Division: DPPS

TRANSFER from SMI DATABASE for each OTSP Division

Column 1: List of All Division Products by Product Category	Column 2: Normalized Product Weight for Division	Column 3: FYXX Product Level Baseline CS Score	Column 4: FYXX Product Level Target CS Score	Column 5: Division - Weighted Product Satisfaction Score	Column 6: Division-Weighted Target Satisfaction Score	Column 7: FYXX Product Level Target CS Improvement
TRANSFER From SMI DATABASE list all Status 1 & 2 products (with at least priority weight) for this producer division. Sort order for Product list is by Product Category (in ascending order of category number), and within each category, products in descending order of ratio Col2/Col3.	TRANSFER From CS7B (Target) for the Division above and product in this row, insert Col3 For product category row, shade this column	TRANSFER From CS8 for the Division above and product in this row, insert Col4Tot. If no CS8 exists, leave blank For product category row, shade this column	TRANSFER If product in this row is listed on CS16, insert CS16 Col4. All other products copy CS17 Col3 entry For product category row, shade this column	TRANSFER From CS9 for division above and product in this row insert Col4. If product does not appear on CS9, leave blank For product category row, shade this column	CALCULATE Col2 x Col4 If Col4 blank, leave blank For product category row, shade this column	CALCULATE ((Col6 - Col5)/ Col5) x 100 If Col5 or Col6 blank, leave blank For product category row, shade this column
	CALCULATE Column sum (should equal 1.0)			TRANSFER FYXX Division Level Baseline CS Score From CS9 Col4 Total	CALCULATE FYXX Division Level Target CS Score Column sum	CALCULATE FYXX Division Level Target CS Improvement ((Col6Total-Col5Total)/ Col5Total) x 100
Totals	1.00					

TOOL ES1 - OTSP Employee Satisfaction Survey

FY XX

Division:	DIPS	DPPS	DCPCS	DDCPS	DFSMS	DBES	DSHS
	DCCT	DSD	DSM	DTNM	DMIS	ACOMM	

Instructions

Please circle the name of your OTSP Division from the list above.

Please enter **only one importance factor** and **one satisfaction score** for each of the nine factors. The bulleted questions are there to further clarify the overall factor.

Please enter **one value** in column 2 that tells us how important each factor is to you. This will help us prioritize our efforts to improve employee satisfaction. The importance factor can be any value from 0 to 100. Use the following guidelines:

0 - 20	Very Low Importance
21 - 40	Low Importance
41 - 60	Medium Importance
61 - 80	High Importance
81 - 100	Very High Importance

Please enter **one score** in column 3 to indicate your current satisfaction with each factor. The score can be any value from 0 to 100. Use the following guidelines:

0 - 20	Very Dissatisfied
21 - 40	Dissatisfied
41 - 60	Neutral
61 - 80	Satisfied
81 - 100	Very Satisfied

When OTSP or OTSP management is mentioned, please include in your assessment your Division's management team as well as the Directors and Assistant Commissioner. Please feel free to add any additional comments in the space provided at the end of the form.

Column 1 Satisfaction Factor	Column 2 Importance Factor 0 - 100	Column 3 Factor Satisfaction Score 0 - 100
1. Job Duties and Responsibilities How satisfied are you with your work assignments? • Is your work challenging? • Do you have the right amount of work?		
2. Quality of Supervision How well does your supervisor ensure that you are a productive and effective member of your work unit? • Does he/she provide clear guidance and feedback? • Does he/she delegate authority and work effectively? • How open or willing is your supervisor to discuss job-related problems? • Does your supervisor capitalize on your individual skills and talents?		
3. Training & Skills Development How well trained are you to perform your job? • Do you receive training in a timely fashion? • Is the training that you receive of high quality? • Is job-related training easily available to you? • Do you have access to training (not directly job-related) to foster your growth and development?		
4. Management Leadership How well does OTSP management set direction and provide support that you need to achieve organizational goals? • Do they solicit and respond to your ideas? • Is the process for setting priorities and making decisions clear? • How effective is OTSP management in resolving problems? • Are the organization's plans and priorities clearly articulated?		
5. Communications How satisfied are you with the timeliness and effectiveness of communications within OTSP? • How effectively are communication channels utilized (e-mail, memos, meetings, etc.)? • Are you promptly informed of changes in policy that affect your work? • Does OTSP encourage communication?		
6. Work Rule Flexibility How satisfied are you with the flexibility of work rules within OTSP? • How supportive is OTSP management with regard to alternative working arrangements such as flexitime, flexiplace, job sharing, etc.? • How accommodating is OTSP management when dealing with individual situations? • How satisfied are you with the amount of flexibility in the work rules?		

Column 1 Satisfaction Factor	Column 2 Importance Factor 0 - 100	Column 3 Factor Satisfaction Score 0 - 100
7. Job Growth & Promotion Potential How satisfied are you with your advancement within OTSP? • Has your rate of advancement within OTSP met your expectations? • Are opportunities provided to you to prepare for career advancement?		
8. Awards and Recognition How satisfied are you with the process used to select and recognize recipients for awards within OTSP? • Monetary Awards Process • Non-monetary Awards Process		
9. Workforce Diversity and Fairness How satisfied are you with the fairness of your treatment in OTSP? • Are you treated with respect? • Are you considered for all task and team assignments fairly? • Are you satisfied with the diversity in the work groups to which you are assigned?		

Comments:

TOOL IP3: Internal Productivity Scores For Target Product/Process Pairs

FY XX

Purpose:	To calculate unit cost baselines for the product/process pairs selected by the Divisions for scoring in the current year
Tool Repetitions:	Create 1 IP3 form for each OTSP Division
Tool User:	Division Chiefs
Operations:	TRANSFER CALCULATE INPUT

Producer Division: DIPS TRANSFER from SMI DATABASE for each OTSP Division

Column 1: FYXX Product/Process Pairs Targeted For Productivity Improvement	Column 2: FYXX-1 Actual Dollars Consumed By Targeted Product/Process Pair	Column 3: FYXX-1 Product Units Produced	Column 4: FYXX Baseline Product/ Process Unit Cost
TRANSFER From IP1 for the division above, insert all product/process pairs with a check for FYXX in Col3	TRANSFER From FYXX-1 IP16 Col3Tot for the product/process pair in this row. If no FYXX-1 IP16 exists, leave blank.	TRANSFER From SMI DATABASE for the Division above & Product named in this row, enter total FYXX-1 annual units produced If no FYXX-1 units exist, leave blank	CALCULATE Col 2/Col 3 If Col2 or Col3 blank, leave blank
IPP Reporter Tracking System User Manual /Develop Document			
Monthly Index Listings /Production Run			
Monthly Index Listings /Review Listings			
	CALCULATE FYXX-1 Division Total Dollars Consumed by Targeted Product/Process Pairs Column sum	CALCULATE FYXX-1 Division Total Product Units Produced Column sum	CALCULATE FYXX Division Baseline Average Product/Process Unit Cost Col 2Tot/Col 3Tot
Totals			

TOOL IP7: Target IP Improvement Scores - Division Level
FY XX

Purpose:	To determine the overall target IP costs and improvement percentage for each OTSP Division
Tool Repetitions:	Create 1 IP7 per OTSP Division
Tool User:	Division Management and Staff
Operations:	TRANSFER CALCULATE INPUT

Producer Division: DIPS

TRANSFER from SMI DATABASE for each OTSP Division

Column 1: FYXX Product/Process Pairs Targeted for Productivity Improvement	Column 2: FYXX Product/ Process Baseline Unit Cost	Column 3: FYXX-1 Product Units Produced	Column 4: FYXX Product/Process Target Unit Cost	Column 5: FYXX Projected Total Cost For Baseline Units	Column 6: Division Target IP Improvement
TRANSFER From IP1 for the division above, insert all product/process pairs with a check for FYXX in Col3	TRANSFER From IP3 for the Division listed above, and for the product/process pair in this row, insert Col4	TRANSFER From IP3 for the Division listed above and product named in this row, insert Col3	TRANSFER From IP6 for the Division listed above, and for the product/process pair in this row, insert Col4	CALCULATE Col3 x Col4	CALCULATE ((Col 2 - Col4)/Col 2) x 100
	TRANSFER FYXX Division Baseline Average Product/Process Unit Cost From IP3 for the Division above, insert Col4Total	TRANSFER FYXX-1 Division Total Product Units Produced From IP3 for the Division above, insert Col3Total	CALCULATE FYXX Division Projected Average Unit Cost Col5Total/ Col3Total	CALCULATE FYXX Division Projected Total Cost for Baseline Units Column sum	CALCULATE Division Level Target IP Improvement ((Col2Total - Col4Total)/ Col2Total) x 100
Totals					

TOOL ALL6: Year 1 Targets - Division Level

FY XX

Purpose:	To summarize a Division's Year 1 SMI targets
Tool Repetitions:	Create one ALL6 for each Division
Tool User:	All Division Staff
Operations:	TRANSFER CALCULATE INPUT

DIPS

Vision Category	Year 1 Targets
(PRE-PRINT)	TRANSFER For CS row, from CS17 insert Col5Total and Col6Total For ES row, from ES10 insert Col5Total and Col6Total For IP Row Transfer 1 from IP14 Col4 Last Year for the Division above Transfer 2 from IP14 Col4 This Year for the Division above Transfer 3 from IP7 Col6Tot This Year for the Division above
Continuously Increase Customer Satisfaction With Our Products	DIPS will improve its Customer Satisfaction score from 42.7 (TRANSFER) to 47.1. (TRANSFER)
Continuously Increase Employee Satisfaction With Job Factors	For Targeted Job Factors DIPS will improve its Employee Satisfaction score from 43.4 (TRANSFER) to 47.4. (TRANSFER)
Continuously Increase Productivity Within Our Processes	Currently, DIPS is measuring productivity changes for 0 % (TRANSFER 1) of its total budget dollars. Within the next year, DIPS will measure productivity changes for 4% (TRANSFER 2) of its total budget dollars. Over the next year, for its measured resource base, DIPS will increase productivity at an annual rate of 1.5 %. (TRANSFER 3)

TOOL ALL12: Summary of Scores - Division Level

FY XX-1

Purpose:	To measure each OTSP Division's annual performance relative to their SMI targets
Tool Repetitions:	Create one ALL12 form for each OTSP Division
Tool User:	Division Management/Staff
Operations:	TRANSFER CALCULATE INPUT

Producer Division: DIPS TRANSFER from SMI DATABASE for each OTSP Division

Column 1: Vision Category	Column 2: FYXX-1 Division Level Baseline Scores	Column 3: FYXX-1 Division Level Target Scores	Column 4: FYXX-1 Division Level Target Improvements	Column 5: FYXX-1 Division Level Actual Scores	Column 6: FYXX-1 Percentage of Division Level Targets Achieved
(PRE-PRINT)	TRANSFER For CS row, from CS9 Col4Tot (FYXX-1) for the above division For ES Row ... For IP row, from IP3Col4Tot (FYXX-1) for the above division	TRANSFER For CS Row, from CS17 Col6Tot (FYXX-1) for the above division For ES Row ... For IP row, from IP7Col4Tot (FYXX-1) for the above division	TRANSFER For CS row, from CS17 Col7Tot (FYXX-1) for the above division For ES Row ... For IP row, from IP7Col6Tot (FYXX-1) for the above division	TRANSFER For CS Row, from CS9 Col4Tot (FYXX) for the above division For ES Row ... For IP row, from IP16A Col4Tot (FYXX-1) for the above division	TRANSFER For CS row, from CS43 (FYXX-1) Col5 For ES Row ... For IP row, from IP17Col6Tot
Customer Satisfaction	42.7	47.1	10.3 %	48.5	103%
Employee Satisfaction	43.4	47.4	9.2 %	49.6	105%
Internal Productivity Unit Cost	$12,844	$12,529	2.5 %	$12,651	61.3%
Internal Productivity % of Dollars Measured	5%	11%	120%	9.5%	86%

Comments made as a discussant in the session on Performance Measurement in Statistical Agencies at the Seminar on Statistical Methodology in the Public Service sponsored by the Council of Professional Associations on Federal Statistics, November 13, 1996.

Robert Wise, Ph.D.
Corporate and Government Consulting, Inc.
4027 Brandywine Street NW, Suite 200
Washington DC 20016
202-237-2600

I appreciate this opportunity to contribute to this discussion on a topic that I care very much about. I work full time consulting with public and private organizations to help them improve their work systems through the use of quantitative methods, but my special interest is in organizational performance measurement. I am also the humble moderator of a local Study Group on Measuring Organizational Performance which is sponsored by the Northern Virginia Chapter of the American Society for Quality Control. We call ourselves the MOP Group, MOP referring to "Measuring Organizational Performance." We are almost three years old and have about 30 members. We meet once a month to study and discuss theory and practice related to measuring organizational performance and we operate a MOP Clinic that provides free advice to organizations that want help in solving a performance measurement problem. You are all invited to contact me if your organization would like some free help in developing performance measures.

The co-authors of the first papers presented in this session, Nancy Kirkendall and Paul Staller, are members of the MOP Group. And, in fact, as Nancy mentioned in her presentation, their inspiration for initiating a project to develop organizational performance measures for their agency came from a meeting of the MOP Group. At that meeting, another federal agency was describing how it developed a performance measurement system and, as she listened, Nancy said to herself, "We can do that!" The results of that inspiration, these several years later, have been recognized as an exemplary model for how to design and implement an organizational performance measurement system in a federal agency. With all due respect to Nancy and Paul, the MOP Group wishes to claim all possible credit for their accomplishments!

In my allotted time, I would like to address my comments to the question of how to begin to develop a system for measuring organizational performance and draw on the papers presented here for illustration and reinforcement.

The challenge of measuring organizational performance is not fundamentally a technical matter. Many performance measures involve nothing more than counting. Of the "fundamental four" measures of process performance--quantity, quality, timeliness, and cost--quantity, time, and cost are pretty easy to count in most situations; quality can get a

bit tricky to measure in some settings, but it often involves counting errors or defects. I do not mean to say that there are not great challenges in the development and implementation of a system for measuring organizational performance, only that these challenges are, for the most part, not technical from a measurement point of view. Now, it may take some effort to define what constitutes an error or a defect or other failure to meet a quality standard, or even to define what a measure of some particular quantity is, but once these definitional tasks are completed by consensus, the actual measurement and analysis of data is often fairly straightforward.

It has been my experience that the challenge of measuring organizational performance is more a matter of corporate will, that is, a sincere desire followed by commitment to improve performance in the delivery of product or service. It involves a change in management philosophy and a learning curve to develop new management habits. Building a system to measure your organization's performance takes time. Note in the papers by Nancy and Paul that their measurement development project has been going on for two years and is still not finished. They began design work in September of 1994, began collecting baseline measurement data in September of 1995 and reported in April of 1996. In George's project, the decision to begin was made in November of 1992, the measurement methodology was established by September of 1993 and the first baseline measures were completed in August of 1994, a little less than two years.

Developing a system to measure your organization's performance goes through a number of stages. But we usually hear about what I call the "advanced" stages such as benchmarking, vertical alignment across organizational levels, the balanced scorecard, and statistical process control. These are all valuable techniques and principles that can contribute significantly to organizational performance improvement. But they are examples of more advanced stages in the development and implementation of a system of performance measures. For example, benchmarking involves exchanging your measurement methods, standards, and results with other organizations for mutual benefit. Benchmarking therefore requires that you already have a well-developed performance measurement system which you are using and have confidence in. It is not, as is sometimes misunderstood, something you do at the beginning of the start of developing measures to find out what other organizations are doing. I note that the papers presented in this session do not focus on these advanced topics and so are especially helpful to those interested in beginning the task of developing performance measures.

Based on my experience, I want to communicate one message to you today. If you are taking your first steps on the journey to developing and using measures in the management of your organization, start at the beginning and keep it simple at first. All three presenters provide good stories of how they approached their tasks. Nancy describes how they first described their work processes before addressing the measurement question. Her remarks illustrate the Wise Theorem of Sequence which states that "In the development of organizational performance measures, process thinking precedes quantitative thinking." In his remarks, Paul indicated that their agency is currently reviewing their measures and trying to reduce their number and improve their

accuracy. His remarks highlights the Wise Theorem of Action which states that "It is better to improve your data on the move than to design the perfect measurement system before moving." George pointed out in his presentation that his agency selected three measures of performance--just three; this illustrates the Wise Theorem of the Vital Few which states that "A few measures are all you need, if they are the right ones."

If I were asked what three pieces of advice I would give to an agency that is starting out to develop and implement measures of organizational performance, I would offer the following four:

1. Process thinking comes before quantitative thinking in the development of measures. In other words, you can't measure a process, program, or purpose that you have not described. So begin by describing your program. Identify the core work processes in your agency, what value they deliver, and to whom. Based on this articulation of your work, you can identify what the strategic objectives of your agency or program are. This may demand new thinking, because you need to view your organization as a set of related processes, not as lines and boxes on an organization chart. The EIA experience described by Nancy is an excellent illustration of the use of an Input-Process-Output model to develop a process view of an organization. Measurement of strategic priorities begins after you have described those priorities.

2. Establish baselines for the key measures you select. You need to find out how well you are currently delivering your value as early as you can. All three presenters made this point either in their remarks or in their papers. Without a baseline, you will have no basis on which to interpret later measures to see if your performance is changing. Comparison of data is the basis for interpreting performance measures and although an organization can use external standards (such as industry standards) to interpret their performance over time, most organizations like to interpret performance by comparing data to their own past performance.

I would like to register here an opinion that is somewhat in opposition to one particular notion that appears in several of the presenters' papers. I do not mean to say that they are advocates of this principle, but they refer to the notion of "targets," and I hear it often advised that you write your program objectives in the form of numerical "targets." I have seen organizations try to do this even before they had a good understanding of their current operations. I worked with one organization that had already set a target of a 50 % reduction in the amount of time it took to complete a certain process. When I met with them, I learned that they had not analyzed the process and had no data on how much time the process was taking. Not only did they not have a number to calculate 50% of, but they had no "before" data to document the improvement between "before" and "after." I think that setting numerical targets as a basis for measuring performance is great if your organization is at a certain stage, but to do it well, to do it accurately enough to make a commitment to a specific percentage increase, you had better have your finger on the pulse of your process capability. As an illustration, I note in Paul's paper that in his agency's plans, the intention is to establish numerical targets for the year 2002. A

numerical target it is not the only way to measure change or results. An alternative is to track a trend on a line chart. For the beginning measurer of organizational performance, simply increasing quality or reducing time can be worthy objectives, and these performance changes are eminently measurable with the help of baselines and trend charts. With time and experience, you can move to a stage where you can understand and measure your process capability well enough to commit to numerical targets in future program objectives.

3. Don't build the whole system at once. Build a piece of it, say for one process or one program, and get the system working and in use. Develop the right management habits that will sustain your effort and will truly make use of the measurement data that you generate.

4. Involve members from all relevant parts of your organization in describing the process or program and in the development of performance measures. Distribute the measurement plans widely. Everyone needs to understand the purpose of the measures and to cooperate in data collection and interpretation. It would be best if the measurement reports, in the form of charts, are made public and visible so that everyone can see how the organization is performing. Organizational performance is the sum of the efforts of all its members.

Reports Available in the
Statistical Policy Working Paper Series

1. ***Report on Statistics for Allocation of Funds***, 1978 (NTIS PB86-211521/AS)
2. ***Report on Statistical Disclosure and Disclosure-Avoidance Techniques***, 1978 (NTIS PB86-2115/AS)
3. ***An Error Profile: Employment as Measured by the Current Population Survey***, 1978 (NTIS PB86-214269/AS)
4. ***Glossary of Nonsampling Error Terms: An Example of a Semantic Problem in Statistics***, 1978 (NTIS PB86-211547/AS)
5. ***Report on Exact and Statistical matching Techniques***, 1980 (NTIS PB86-215829/AS)
6. ***Report on Statistical Uses of Administrative Records***, 1980 (NTIS PB86-214285/AS)
7. ***An Interagency Review of Time-Series Revision Policies***, 1982 (NTIS PB86-232451/AS)
8. ***Statistical Interagency Agreements***, 1982 (NTIS PB86-230570/AS)
9. ***Contracting for Surveys***, 1983. (NTIS PB83-233148)
10. ***Approaches to Developing Questionnaires***, 1983 (NTIS PB84-105055/AS)
11. ***A Review of Industry Coding Systems***, 1984 (NTIS PB84-135276)
12. ***The Role of Telephone Data Collection in Federal Statistics***, 1984 (NTIS PB85-105971)
13. ***Federal Longitudinal Surveys***, 1986 (NTIS PB86-139730)
14. ***Workshop on Statistical Uses of Microcomputers in*** Federal Agencies, 1987 (NTIS PB87-166393)
15. ***Quality in Establishment Surveys***, 1988 (NTIS PB88-232921)
16. ***A Comparative Study of Reporting Units in Selected Employer Data Systems***, 1990 (NTIS PB90-205238)
17. ***Survey Coverage***, 1990 (NTIS PB90-205246)
18. ***Data Editing in Federal Statistical Agencies***, 1990 (NTIS PB90-205253)
19. ***Computer Assisted Survey Information Collection***, 1990 (NTIS PB90-205261)
20. ***Seminar on Quality of Federal Data***, 1991 (NTIS PB91-142414)
21. ***Indirect Estimators in Federal Programs***, 1993 (NTIS PB93-209294)
22. ***Report on Statistical Disclosure Limitation Methodology***, 1994 (NTIS PB94-165305)
23. ***Seminar on New Directions in Statistical Methodology***, 1995 (NTIS PB95-182978)
24. ***Report on Electronic Dissemination of Statistical Data***, 1995 (NTIS PB96-121629)
25. ***Data Editing Workshop and Exposition***, 1996 (NTIS PB97-104624)
26. ***Seminar on Statistical Methodology in the Public Service***, 1997 (NTIS PB97-162580)

Copies of these working papers may be ordered from NTIS Document Sales, 5285 Port Royal Road, Springfield, VA 22161; telephone: (703) 487-4650. The Statistical Policy Working Paper series is also available electronically through the Bureau of Transportation Statistics World Wide Web home page (http://www.bts.gov).