December 1985

WASHINGTON STATISTICAL SOCIETY

NEWSLETTER

CALENDAR

December 3	Tuesday	Revenue Estimating by Treasury: The Art and Science
December 5	Thursday	Statistical Applications in Auditing - Some Problems and Challenges
December 6	Friday	Sample Size Determination for Some Common Nonparametric Tests
December 10	Tuesday	Alternative Definitions of Longitudinal Households: Implications for Annual Statistics
December 10	Tuesday	Data Linkage at IRS
December 11	Wednesday	Identifying Measurement Error in A Consumer Unit's Report of Expenditures
December 11	Wednesday	FEE EVENT: WSS Holiday Party
December 19	Thursday	FEE EVENT: Short Course: A 'Hands-on' Review of Statistical Software for Micro-Computers

****

EDITORS NOTE: This is the first of a series of articles on the statistical activities of various Federal agencies. This article summarizes how the U.S. General Accounting Offices (GAO) and the Congressional Budget Office (CBO) use statistical analysis to perform their mandated functions. Harry Conley is a sampling statistician at the GAO; Virginia deWolf is a general statistician at the National Highway Traffic Safety Administration, formerly at the GAO; and Nancy Gordon is an economist at the CBO.

U.S. GENERAL ACCOUNTING OFFICE--by Harry Conley and Virginia deWolf

In order to explain the role of statisticians and statistical analysts at the U.S. General Accounting Office (GAO) we decided to focus on three issues: how GAO gets its assignments, how GAO employs statistical analysts, and how GAO gets its data. We assumed that most readers are familiar with GAO products (such as the "blue book" reports) but not with the process whereby GAO accomplishes its work.

The following comments reflect the opinions of the authors and are not those of the agency. Since the authors do not speak for GAO, the article does not address such issues as the employment possibilities for statistical analysts at GAO.

(CONTINUED)

Background or how does GAO get its assignments? The GAO functions as the investigative arm of Congress. As such, a Senator, Representative, or Congressional Committee writes a letter to the Comptroller General, Mr. Charles Bowsher, and requests that GAO audit or evaluate a particular program. Congressional requests are subsequently forwarded to the appropriate GAO division depending on the particular issue. Divisions have different "issue areas" which parallel functional activities of the government (such as: law enforcement, and tax collection). There are four large divisions and three smaller technical divisions, as well as regional offices and international offices. In addition to Congressional requests, GAO also does self-initiated work which is tied to interests on the Hill.

How are statistical analysts used in GAO jobs? Once a Congressional Request reaches the appropriate division, audit staff (the majority of GAO staff members carry the position title of "evaluator") begin to "scope" out the job. In scoping, evaluators address such issues as follows: has this topic been studied by other Congressional agencies (such as CBO), and what are some alternative approaches which could be used to accomplish the needs of this request?

Statistical analysts in GAO work in a consulting role to teams of evaluators. One major function is to assist in designing jobs during the planning phase of assignments. In this phase analysts help develop the methodology with particular attention to the time frame within which the assignment must be completed.

Another major function of statistical analysts is sampling. Throughout the agency, the use of proper statistical sampling techniques is stressed. GAO sampling statisticians are involved in many different assignments and assist GAO audit teams in developing the appropriate sampling plan.

A third major function of statisticial analysts is to analyze the data received from GAO's data collection efforts. This includes projections from sample data, correlation and regression analysis, discriminant function analysis, and other techniques. We discuss how GAO gets its data in greater detail in the third section of this article.

In addition, GAO developed a course which teaches statistical auditing to evaluators and helps audit staff understand some of the basic statistical sampling and analytic techniques. Statistical analysts teach this course.

The official position titles of statistical analysts in GAO vary and include not only statisticians, but also operations research analysts, social science analysts, specialist evaluators (such as social science evaluators), and psychologists. Including GAO's regional offices, there are approximately 110 statistical analysts at GAO - or about 2 percent of the organization.

The analytic staff in GAO is decentralized. That is, each of the divisions and regional offices has a "Technical Assistance Group" or "Design Group" which houses its technical staff. There are a few exceptions, however, with a core group of sampling statisticians and actuaries located in one technical division.

Since WSS has economists as well as statisticians among its membership, it might be of interest for WSS members to know that GAO also employs about 80 economists. They are also decentralized except for a core group in GAO's Office of the Chief Economist.

(Continued)

U.S. GENERAL ACCOUNTING OFFICE (CONTINUED)

How does GAO get its data? Generally, GAO does primary data collection rather than secondary analysis. Questionnaires and other data collection instruments (such as instruments used to record telephone interviews, person-to-person interviews, or detailed source document reviews) are most frequently used to collect our data. GAO's response rate for questionnaires is usually at least 80 percent. Questionnaire development staff (another large segment of GAO's "technical" staff) work with GAO evaluators to develop these instruments.

In addition to preparing this data for analysis (i.e., editing, checking the reliability of the file after the instruments have been keypunched) and analyzing the file, statistical analysts also write technical sections about the analysis for inclusion in the final GAO report.

Lastly, GAO occasionally obtains computerized data from agencies, such as administrative records or the National Crime Information Center database. Statistical analysts work with audit staff to detemine the availability, accuracy, and reliability of such computerized files prior to using these data for analysis.

CONGRESSIONAL BUDGET OFFICE--by Nancy Gordon

The Congressional Budget Office (CBO) provides the Congress with information relating to the U.S. economy, the federal budget, and federal programs. It does not make recommendations, but identifies issues of concern and presents options for the Congress to consider.

Organization and Responsibilities of the CBO

The 222 employees of CBO are organized in seven divisions that reflect its major activities. The Fiscal Analysis Division is responsible for economic forecasting and special studies on the economy. The Budget Analysis Division projects spending in all Federal programs, estimates the budgetary effects of options being considered by the Congress, and follows final spending actions to see whether they adhere to the targets of the budget resolutions. The Tax Analysis Division projects revenues from various sources and examines the likely effects of modifying the tax code. In addition, three "program" divisions--covering National Security, Natural Resources and Commerce, and Human Resources and Community Development--analyze alternative federal policies with budgetary implications. These studies focus on the impacts policy alternatives would have on beneficiaries such as individuals, states, and firms, as well as the resulting increase or reduction in federal spending. Finally, the Office of Intergovernmental Relations coordinates CBO's activities with Members of Congress and the public, and analyze issues related to the management of the federal government.

The CBO is often closely involved in Congressional decisionmaking as a provider of objective information. Members of Congress are called upon to make difficult choices, which often involve value judgments about the appropriate federal role. These decisions are facilitated, however, by high-quality analyses. Analysts at the CBO primarily use available data--we do not collect data ourselves--to project what would occur if current law remains unchanged and how those outcomes would differ under various policy options.

(CONTINUED)

The information produced by CBO analysts is provided to the Congress in many different ways. Formal products--such as published papers and special analyses--are transmitted by the Director to the Chair or Ranking Minority Member of the committee or subcommittee requesting it. The Director and other members of the senior staff often testify before Congressional committees, generally reporting the results of recently completed studies. In addition, the analysts have frequent, informal contact with staffs of the committees--in meetings, by telephone, and through staff-to-staff memoranda.

Because of the variety of products the CBO produces, analysts need many different skills. Essentially all of them have strong quantitative backgrounds, generally in economics or public policy, although many other disciplines are represented. Also important is the ability to write quickly and well and to convey technical material in meetings with people from nontechnical backgrounds.

Some discussion of work done in the Human Resources and Community Development (HRCD) Division should help to make this general discussion of CBO's role more concrete.

The Human Resources and Community Development Division

HRCD is responsible for policy analyses in several areas--health, income security, employment, education, housing, and community development. Thus, the division deals with federal programs such as Social Security, Medicare, Unemployment Insurance, Guaranteed Student loans, Section 8 subsidized housing, and the Community Development Block Grant.

HRCD uses many different types of data, but most are "micro data" from surveys. Generally, individuals are the unit of observation, but occasionally it is institutions, such as hospitals. Adminstrative (or programmatic) data collected by federal agencies and, occasionally, data collected by private surveys also provide the basis for some analyses. The single most widely used source is the Current Population Survey (CPS), because it provides information on the overlap of participation in different programs and is the best source of information about incomes that is available for many past years. Other examples are the Annual Housing Survey, the Survey of Income and Program Participation (SIPP), the Medicare hospital cost reports, the Medicare Continuous History File (which provides programmatic information on individual beneficiaries' use of health care services), a file of Pell grant applicants, and data collected by the American Hospital Association from its members.

A variety of techniques are used in HRCD studies. In many cases, simulation models are used to examine how individuals and families, physicians, and hospitals would have been affected had different federal policies been in place in an earlier year, such as 1984. At other times, regression analyses are used to estimate relationships such as those between the proportion of low-income patients and a hospital's average cost of treatment. The results from these techniques are used by committee staff and Members of Congress in assessing the desirability of various options. In addition, the regression analyses sometimes contribute to setting technical parameters in legislation.

(CONTINUED)

Many projects within HRCD focus on the impacts of policy options on people in different income groups. Some specific examples of recent analyses based on simulations are:

- o reducing or eliminating cost-of-living adjustments (COLAs) in programs such as Social Security, including versions that attempt to protect those with low incomes;
- o modifying Medicare, for example, by requiring beneficiaries to pay increased premiums or increased cost sharing;
- o changing federal financial aid programs for higher education in ways such as those proposed by the Administration; and
- o altering tax policy, for example, by including a portion of the insurance value of Medicare or more of Social Security benefits in taxable income.

Other simulation studies have focused on the likely impacts of policy changes on hospitals with different characteristics, such as public versus private, large versus small, and urban verus rural. Particular alternatives would modify Medicare's prospective payment system for reimbursing hospitals, for example, by slowing or stopping the phase-in of national payment rates or changing reimbursements for medical education expenses. In contrast, regression analyses played a crucial role in the CBO study of how the Medicare reimbursement system might be adjusted to account for that some hospitals serve disproportionately large numbers of low-income patients who tend to be more severely ill, and more costly to treat, than higher-income patients.

PROGRAM ABSTRACTS

TOPIC: Revenue Estimating by Treasury: The Art and Science
SPEAKER: Howard Nestor, Office of Tax Analyis, U.S. Treasury
CHAIR: Lynda T. Carlson, Energy Information Administration
DISCUSSANT: Rosemary Marcus, Congressional Budget Office
DATE AND TIME: Tuesday, December 3, 1985; 12:30-2:00 p.m.
LOCATION: Forrestal Building, Room 1E-245, 1000 Independence Ave., S.W.
ABSTRACT: The Treasury Department is responsible for developing estimates of revenue impacts of various budget and tax proposals. This year with the various tax reform proposals, Treasury's estimates have come under scrutiny. Alternative revenue estimates for the same tax proposals have been developed by various groups such as the Congressional Budget Office. Mr. Nestor will describe how Treasury developes its estimates.

****

TOPIC: Statistical Applications in Auditing---Some Problems and Challenges
SPEAKER: John Neter, University of Georgia
CHAIR: Richard Bolstein, George Mason University
DATE AND TIME: Thursday, December 5, 1985; 6:30-7:00 p.m.-Wine & cheese reception, 7:00 p.m.-Lecture; 8:15 p.m.-Dinner at the Magic Gourd, 528 23rd Street, N.W. [For Reservations, call Mary Mulry-Liggan at 763-7140]
LOCATION: Marvin Center, Room 406, George Washington University, 800 21st St., N.W.
ABSTRACT: Auditors frequently use statistical sampling procedures to make inferences about the total amount of error in an accounting population. Some of the problems encountered in these statistical applications will be considered and several methods for overcoming these problems will be discussed. The interactions between auditors and statisticians in the development of this area of applications will also be considered.

****

(CONTINUED)

TOPIC: Sample Size Determination for Some Common Nonparametric Tests
SPEAKER: Gottfried Noether, University of Connecticut
DATE AND TIME: Friday, December 6, 1985; 2:00-3:00 p.m.
LOCATION: Room 2736, GAO Building, 441 G Street, N.W. (Contact Sandra West at 523-1874 to arrange building entry)
ABSTRACT: The problem of determining sample size for hypothesis testing purposes can be stated as follows. We want a test with significance level α to reject the hypothesis being tested with probability at least $1 - \beta$, whenever an alternative is sufficiently far from the null hypothesis. In a nonparametric setting, it is usually not at all obvious how to interpret the requirement "sufficiently far from the null hypothesis." We show that for the sign and Wilcoxon tests of the one-sample problem, for Wilcoxon two-sample test, and for Kendall's test of independence, there exist certain natural "parameters" that are not only easily interpreted, but also permit easy sample size determination.

****

TOPIC: Alternative Definitions of Longitudinal Households: Implications for Annual Statistics
SPEAKER: Constance F. Citro
CHAIR: Donald J. Hernandez, Bureau of the Census
DATE AND TIME: Tuesday, December 10, 1985; 12:00-1:30 p.m.
LOCATION: Martin Luther King Memorial Library, Room 315, 901 G Street, N.W.
ABSTRACT: The Survey of Income and Program Participation and its predecessor, the 1979 Income Survey Development Program Research Panel (ISDP), are the first large scale nationally-representative surveys to permit measuring intra-year changes in composition and economic status of households. This paper report on research to define households over time using data from the 1979 ISDP. The research examined alternative definitions of longitudinal households and measures of annual income and poverty status under each definition, in order to shed empirical light on problems of presenting annual statistics for part-year households and related problems of longitudinal measurement.

****

TOPIC: Data Linkage at IRS
SPEAKERS: Peter Sailer, Thomas B. Petska, and Marvin Schwartz, IRS
CHAIR: Warren Buckler, Social Security Administration
DISCUSSANT: Doug Sater, Bureau of the Census
DATE AND TIME: Tuesday, December 10, 1985; 12:00-1:30 p.m.
LOCATION: IRS Auditorium, 7th Floor, 1111 Constitution Avenue, N.W.
ABSTRACT: The Workshop on Exact Matching Methodologies, held last May, brought together a great number of people interested in matching techniques. This session focuses on several different areas of application currently underway at IRS. Pete Sailer will describe some panel studies based on individual income tax returns; Tom Petska will talk about some linkage efforts using business tax returns; and Marvin Schwartz will discuss some studies underway to improve wealth estimates from estate tax returns.

****

(CONTINUED)

TOPIC: Identifying Measurement Error in a Consumer Unit's Report of Expenditures
SPEAKER: Clyde Tucker, Bureau of Labor Statistics
CHAIR: Elisabeth Martin, Bureau of Social Science Research
DISCUSSANT: Kent Marquis, Bureau of the Census
DATE AND TIME: Wednesday, December 11, 1985; 1:00-2:30 p.m.
LOCATION: GAO Auditorium, 441 G Street, N.W. (Call 523-1760 to arrange building entry.)
ABSTRACT: The estimation of sampling error is relatively straightforward even in the case of a complex sample survey design. Unfortunately, the population estimates from sample surveys are usually subject to other, nonsampling errors which are not so easily estimated. This paper attempts to identify and quantify the nonsampling error arising from the measurement of expenditures for food in the 1980-81 Consumer Expenditure Diary Survey conducted by the Bureau of the Census for the Bureau of Labor Statistics. This approach relies on the assumption that there is a relationship between the errors in consumers' expenditures reports and the ways in which the consumers report information. Using this assumption, measurement error is modeled as a micro latent variable based on an analysis of the latent structure among various indicators of the pattern of information reported by individual respondents.

****

ANNOUNCEMENTS

Conference on State Statistics

The National Governor's Association Center for Policy Research in cooperation with the ASA's Committees on Government Statistics and Subnational Statistics, the Council of Professional Assocations on Federal Statistics and other State associations is sponsoring a conference entitled "State Statistics--III." It will be held on Wednesday, January 8 through Friday, January 10, 1986 at the Omni Shoreham Hotel. The conference will focus on activities currently conducted within States to coordinate and manage information resources and with Federal-State issues of concern to the States. The audience will include a mix of state and local officals from the legislative and executive branches, private sector data users and producers, federal agency personnel, congressional representatives and others interested in our national statistical system and the state role in that system. The cost of the conference is $125 on or before December 23 and $175 after December 23. For further information on the Conference program, call Lorriane Amico at National Governor's Assocation (202) 624-5346. Questions concerning logistics may be directed to Jan Kary Dunlavey at 202-624-5347.

October Board Meeting Notes

A special feature of the October 29th was a report by Charles Mann as Chairperson of the WSS Strategic Planning Committee. The Committee will be formulating a long-range plan for WSS. Of particular importance will be the Society's future emphasis on short courses and other special events.

ANNOUNCEMENTS (Continued)

A Note From The President---Mary Grace Kovar

The WSS would like to become more active in promoting statistical awareness and excellence among high school students through local science fairs. Since we have not done this before, we should probably start small and add activities in later years. For example, when the slide presentation that is being developed by the Cincinnati Chapter is ready next year, we might develop a speakers' bureau and WSS members could use it to make presentation at individual high schools.

Susan Ellenberg suggested that one way of beginning would be to present prizes at the science fairs held in the local school districts. She suggested that a book, such as Statistics: A Guide to the Unknown would make a good prize. The Board voted to award prizes at district or regional science fairs in the Washington area this school year. The Washington area is defined by the District of Columbia and those school districts adjoining it.

WSS members will have to make it possible. If you are willing to be a judge, contact Susan Ellenberg with your name, address, telephone number, and the school district where you would like to judge. It won't take much of your time, probably an afternoon, and it might be a lot of fun. Susan can be reached at 496-4836 (daytime) or 299-9039 (evenings).

Second Annual Research Conference Scheduled for March 23-26, 1986

Planning for the Census Bureau's Second Annual Research Conference (ARC II) at the Sheraton International Conference Center, Reston, Virginia is well underway. The theme for ARC II will be nonsampling error. Invited paper sessions include discussions on causes, measurement, and reducing the effects of nonsampling error; quality issues; nonsampling error considerations for longitudinal data; and public attitudes and survey response issues. Research in specific program areas will be discussed in sessions devoted to the 1990 Decennial Census, international economic input/output analysis, new methods of estimating household and population characteristics, economic census and survey evaluation studies, and research on the Survey of Income and Program Participation and the services sector of our economy. The program will also contain a track of four sessions on undercoverage which will include discussions of census undercount research and research into problems caused by undocumented immigration.

This year's Shirley Kallek Memorial Lecture will feature Professor James Durbin from the London School of Economics and Political Science, and dinner presentations by John G. Keane, Director of the Bureau of the Census, and Mr. Vincent P. Barabba, Executive Director of Market Research and Planning at General Motors (and former Director of the Bureau of the Census).

For additional information, write the Center for Survey Methods Research, Bureau of the Census, Washington, D.C. 20233. A limited supply of the Proceedings from ARC I are available on a first come/first served basis.

EMPLOYMENT COLUMN

Deadline for inserting notices is five (5) weeks before the publication date

Send notices and request to:
Evelyn R. Kay
520 22nd Street, N.W.
Washington, D.C. 20037 202/331-1153

JOB OPENINGS

Research Biostatistician (GS-12): M.S. or Ph.D with SAS programming skills. Unique opportunity to collaborate in analysis of extensive database on neurological and neurobehavioral sequelae of focal penetrating brain wounds, including anatomic (CT scan) brain structure-function relationships. Send resume/SF-171/info to Dr. A. Salazar, Vietnam Head Injury Study, HSHL-CI, Walter Reed Army Medical Center, Washington, D.C. 20307-5001, (202) 576-1348.

Mathematical Statistician GS-1529-12/13: M.S. or Ph.D in statistics with extensive knowledge and experience in the application of probability theory, statistical theory, and theory of experimental design. Requires knowledge of advanced univariate and multivariate statistical methods, nonparametric methods, FORTRAN programming, and use of statistical analysis software packages. Send SF-171 to: Ms. Patricia A. Reissig, Administrative Support Division, U.S. Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, Maryland 20814-2797, or call (202) 295-1630.

Supervisor Statistician (Health), GM-15: The Interview and Examination Statistics Program, of the National Center for Health Statistics is seeking persons to apply for the position of Supervisor Statistician (Health), (Director, Division of Health Interview Statistics). Applicants must have experience in planning, coordinating and directing large-scale population-based health interview surveys, and experience in providing leadership to professional staff. In addition, it is desirable for applicants to have administrative management ability, ability to conduct methodological research, ability to consult with professionals in a variety of scientific fields and settings and, ability to analyze statistical data in published reports.

This position is at the GM-15 grade level with the salary range from $52,262 to $67,940. Interested persons should send a Standard Form 171 (Resumes will not be accepted) to: OASH Personnel Office, Room 17A-08, 5600 Fishers Lane, Rockville, MD 20857. When applying cite vacancy announcement number 85H117 and position title. For additional information contact 301-443-1986. Technical inquiries should be directed to Robert Fuchsberg or Owen Thornberry 301-436-7085. Applications will not accepted after December 31, 1985.

Mathematical and Survey Statisticians (GS-5 through 12): The Energy Information Administration (EIA) has several positions and anticipates future positions for mathematical statisticians and survey statisticians at the GS-5 through GS-12 level. EIA's Office of Statistical Standards is particularly interested in a GS-7/9/11 survey statistician with both a statistical and computer programming background, and a GS-11/12 Ph.D level mathematical statistician, also with computer programming experience. Send completed SF-171 to Nancy Kirkendall, Senior Mathematical Statistician, Office Statistical Standards (EI-70), 1000 Independence Avenue, S.W., Washington, D.C. 20585.

Senior Consulting Statistician: Statistical computing, strong background in mathematical statistics and probability theory, operating knowledge of experimental design, analysis of variance, regression, multivariate, categorical data, and time series analysis. Apply to UC Davis, Employment Office, Davis, CA. 95616 for Job #1437 by 1/10/86. Up to three letters of reference.

JOB APPLICANT

Listed below is a brief description of the qualifications of an applicant seeking employment. Employers interested in interviewing applicant should notify Mrs. Kay of their interest by CODE NUMBER. The request should be by mail and should include the employer's name, organization, and telephone number. The applicant will be notified of the employer's interest and initiation of any further contact will be left to the applicant. All contacts will be confidential.

Code Number:	86-05
Position wanted:	Survey statistician/data analyst in D.C. or Maryland suburbs.
Experience:	29 years in most aspects of survey research and data analysis for Federal Government as Survey Statistician, Mathematical Statistician and Analytical Statistician, including Contract and Project Officer. Extensive experience in survey data editing and database development.
Education:	BBA, City College of New York, 1950. Graduate courses - 45 credits in Statistics, 25 credits in Economics.

WASHINGTON STATISTICAL SOCIETY
P.O. Box 23502
Washington, D.C. 20026