WSS NEWS

WASHINGTON STATISTICAL SOCIETY

December 2005

Annual Holiday Dinner!!!

Please come join your friends and colleagues for a celebration of the holiday season. The 2005 WSS Holiday Dinner will be held Wednesday, December 14, at the Gordon Biersch Brewery. Hope to see you there!

JUDGES FOR THE 2006 SCIENCE FAIRS

Volunteers are needed to represent the Washington Statistical Society next spring as judges in five regional science fairs in Northern Virginia, suburban Maryland, and the District of Columbia. Since 1986, WSS has provided special awards at these fairs to students whose projects demonstrate excellence in data analysis or the application of statistical methods. Those who have participated in this activity have very much enjoyed the opportunity to interact with the students and to observe the widely diverse projects which are presented. The fairs are held on Saturday mornings in mid-March to mid-April. The only time required is that one Saturday morning, plus one weekday lunchtime meeting to discuss judging strategy. If you would like to be a science fair judge next spring, please e-mail Robert Clickner at Robertclickner@westat.com by January 25, and include your e-mail address, work and home phone numbers, your fax number and your mailing address. If you judged last spring, there is no need to contact Bob unless your e-mail address or phone number has changed. If you have any questions, please call Bob at 301-294-2815.

WSS and Other Seminars
(All events are open to any interested persons)

December	
1 Thurs.	**Empirical Bayes Analysis of Bivariate Binary Data: An Application to Small Area Estimation**
8 Thurs.	**The Effects of Cell Collapsing in Poststratification**
12 Mon.	**Comparing Homeowner and Lender Estimates of Housing Wealth and Mortgage Terms**
15 Thurs.	**Data Collection and Statistical Issues in Surveying Cell and Landline Telephone Samples**
January	
17 Tues	**Statistical Methods for Alerting Algorithms in Biosurveillance**
19 Thurs.	**The Use of Contact History Data for Exploring Survey Nonresponse in Federal Demographic Surveys. (A Joint Seminar)**

Also available on the Web at the following URL: http://www.science.gmu.edu/~wss

Announcements

SIGSTAT Schedule for December 2005 - April 2006

December 14, 2005: SAS PROC MDC (http://www.sas.com)

The MDC (Multinomial Discrete Choice) procedure analyzes models where the choice set consists of multiple alternatives. The procedure supports conditional logit, mixed logit, heteroskedastic extreme value, nested logit, and multinomial probit models. Charlie Hallahan will be the speaker.

January 18, 2006: SAS PROC QUANTREG (http://www.sas.com)

The QUANTREG procedure models the effects of covariates on the conditional quantiles of a response variable by means of a quantile regression. Ordinary least-squares (OLS) regression models the relationship between one or more covariates X and the conditional mean of the response variable Y given X=x. Quantile regression extends the regression model to conditional quantiles of the response variable, such as the median or 90th percentile. Quantile regression is particularly useful when the rate of change in the conditional quantile, expressed by the regression coefficients, depends on the quantile. Charlie Hallahan will be the speaker.

February 8, 2006: SAS PROC ENTROPY (http://www.sas.com)

The ENTROPY procedure implements a parametric method of linear estimation based on Generalized Maximum entropy. The ENTROPY procedure is suitable when there are outliers in the data and robustness is required, or when the model is ill-posed or undetermined for the observed data, or for regressions involving small data sets. Charlie Hallahan will be the speaker.

March 15, 2006: SAS PROC GLIMMIX (http://www.sas.com)

The GLIMMIX procedure fits statistical models to data with correlations or nonconstant variability and where the response is not necessarily normally distributed. These models are known as generalized linear mixed models (GLMM). The GLMMs, like linear mixed models, assume normal (Gaussian) random effects. Conditional on these random effects, data can have any distribution in the exponential family. In the absence of random effects, the GLIMMIX procedure fits generalized linear models (fit by the GENMOD procedure). Charlie Hallahan will be the speaker.

April 19, 2006: Introduction to SAS Enterprise Guide 4.1 for Statistical Analysis (http://www.sas.com)

The demo begins with a quick tour through the layout of Enterprise Guide (EG) as a menu-based interface to SAS procedures. Emphasis in this workshop will be on the statistical capabilities of EG. In particular, a set of data (put together using enhanced features of the query builder task) will serve as a case study for a forecasting exercise. The demo will show how code generated by EG can be customized, stored, and rerun, and custom reports saved with new Report Controls Integration. Linda Atkinson and Charlie Hallahan will be the speakers.

SIGSTAT is the Special Interest Group in Statistics for the CPCUG, the Capital PC User Group, and WINFORMS, the Washington Institute for Operations Research Service and Management Science.

All meetings are in Room S3031, 1800 M St, NW from **12:00 to 1:00** (**note new time**). Enter the South Tower & take the elevator to the 3rd floor to check in at the guard's desk.

First-time attendees should contact Charlie Hallahan, 202-694-5051, hallahan@ers.usda.gov, and leave their name. Directions to the building & many links of statistical interest can be found at the

SIGSTAT website, **http://www.cpcug.org/user/sigstat/**.

Note from the WSS NEWS Editor

Items for publication in the February issue of the WSS NEWS should be submitted no later than December 27, 2005. E-mail items to Michael Feil at michael.feil@usda.gov.

Program Announcement

Title: **Empirical Bayes Analysis of Bivariate Binary Data: An Application to Small Area Estimation**

Chair: Van Parsons, National Center for Health Statistics

Speakers: Malay Ghosh, Department of Statistics, University of Florida

Date/Time: Thursday, December 1, 2005 / 2:00 - 3:30 p.m.

Location: National Center for Health Statistics, room 1407C,3311 Toledo Road, Hyattsville, (Metro: Green Line, Prince George's Plaza and then about a 10 minute walk) Note: please try to arrive 15-30 minutes early because of possible security screening delays.

Sponsors: WSS Methodology Section and NCHS/Office of Research and Methodology

Abstract: The paper provides an empirical Bayes (EB) analysis of bivariate binary data with application to small area estimation. Small area estimation is gaining increasing prominence in survey methodology. The need for such statistics is felt in both the public and private sectors. The reason behind its success is that the same survey data, originally targeted towards a higher level of geography (e.g. states) needs to be used also for producing estimates at a lower level of aggregation (e.g. counties, subcounties or census tracts). The direct estimates in such cases are unavailable (e.g. due to zero sample size) and almost always unreliable due to large standard errors and coefficients of variation arising from the paucity of samples in individual areas.

The motivating example in this study is to estimate jointly the proportion of newborns with low birthweight and infant mortality rate at low levels of geography such as districts within a state. The data from the infant mortality study was conducted by NCHS. The original survey was designed to obtain reliable estimates at the state level. The same data needs to be used to produce estimates at the district level. We have used an EB approach for the analysis of such data. We have found second order correct approximations of the mean squared errors (MSE's) of these estimators, and have derived estimators of these MSE's which are also correct up to the second order. The methodology is illustrated with some real data related to low birthweight and infant mortality.

Program Announcement

Title: **The Effects of Cell Collapsing in Poststratification**

Chair: Shail Butani, U.S. Bureau of Labor Statistics

Speakers: Jay J. Kim, National Center for Health Statistics, Centers for Disease Control
And
Richard Valliant, Survey Research Center, University of Michigan

Date/Time: Thursday, December 8, 2005 / 12:30 - 2 p.m.

Location: Bureau of Labor Statistics, Conference Center in G440. To be placed on the seminar list attendance list at the Bureau of Labor Statistics you need to e-mail your name, affiliation, and seminar name to wss_seminar@bls.gov (underscore after 'wss') by noon at least 2 days in advance of the seminar or call 202-691-7524 and leave a message. Bring a photo ID to the seminar. BLS is located at 2 Massachusetts Avenue, NE. Use the Red Line to Union Station.

Sponsor: Methodology Section, WSS

Abstract: Poststratification is a common method of estimation in household surveys. Cells are formed based on characteristics that are known for all sample respondents and for which external control counts are available from a census or another source. The inverses of the poststratification adjustments are usually referred to as coverage ratios. Coverage of some demographic groups may be substantially below 100 percent, and poststratifying serves to correct for biases due to poor coverage. A standard procedure in poststratification is to collapse or combine cells when the sample sizes fall below some minimum or the weight adjustments are above some maximum. Collapsing may decrease the variance of an estimate but may simultaneously increase its bias. We study the effects on bias and variance of this type of dynamic cell collapsing through simulation using a population based on the 2003 National Health Interview Survey.

Program Announcement

Title: **Comparing Homeowner and Lender Estimates of Housing Wealth and Mortgage Terms**

Speaker: Brian Bucks, Federal Reserve Board of Governors

Discussant: Scott Susin, U.S. Census Bureau

Chair: Linda Atkinson, Economic Research Service, USDA

Date/time: Monday, December 12, 2005 / 12:30 – 2:00 p.m.

Location: Bureau of Labor Statistics Conference Center, Room 10. To be placed on the seminar attendance list at the Bureau of Labor Statistics you need to e-mail your name, affiliation, and seminar name to wss_seminar@bls.gov (underscore after `wss') by noon at least 2 days in advance of the seminar or call 202-691-7524 and leave a message. Bring a photo ID to the seminar. BLS is located at 2 Massachusetts Avenue, NE. Take the Red Line to Union Station.

Sponsor: WSS Economics Section

Abstract: Much research on housing wealth relies on the assumption that households are able to report these data accurately. In this paper, we test the validity of this assumption by comparing homeowner-reported data on house values and mortgage terms from the Survey of Consumer Finances (SCF) to lender-reported data from the Office of Federal Housing Enterprise Oversight (OFHEO), the Loan Performance Corporation, and the Residential Finance Survey. We test the accuracy of the data in two ways. First, we compare the distributions of key variables in the homeowner- and lender-reported data. Second, we examine the internal edit codes in the SCF to assess respondent confidence in their answers.

We find that homeowners are able to report the broad features of their housing wealth rather well. An index of house value appreciation based on SCF data matches the aggregate OFHEO index fairly closely. This finding is consistent with other studies that suggest that owner assessments of house value are reasonably accurate. Homeowners are also able to report the maturity and type of their mortgage with a fair amount of accuracy. However, homeowners with adjustable-rate mortgages are less certain about many aspects of their mortgages.

These findings imply that homeowner-reported data are more useful for investigating some housing wealth questions than others. Studies of the effects of housing wealth on consumption, for example, can reasonably be based on homeowner-reported data. However, lender-reported data may be preferred for studies of the vulnerability of households to interest rate shocks.

Program Announcement

Title: **Data Collection and Statistical Issues in Surveying Cell and Landline Telephone Samples**

Chair: Brian J. Meekins, U.S. Bureau of Labor Statistics

Speaker: Michael Brick, Westat;

Date/Time: Thursday, December 15, 2005 / 12:30 - 2:00 p.m.

Location: Bureau of Labor Statistics, Conference Center in G440. To be placed on the seminar list attendance list at the Bureau of Labor Statistics you need to e-mail your name, affiliation, and seminar name to wss_seminar@bls.gov (underscore after 'wss') by noon at least 2 days in advance of the seminar or call 202-691-7524 and leave a message. Bring a photo ID to the seminar. BLS is located at 2 Massachusetts Avenue, NE. Use the Red Line to Union Station.

Sponsor: Methodology Section, WSS

Abstract: As an increasing proportion of the US population use cell phones for most or all of their personal telephone activities, research into conducting surveys that include cell phones is important. This talk reviews a dual frame survey of landline and cell phone numbers conducted in the summer of 2004 for the Joint Program in Survey Methodology. The goal of the survey was to evaluate the feasibility of including cellular phone numbers in a random digit dial telephone survey. As an introduction, a brief background on the status of coverage and usage by telephone service will given, followed by some of the key operational and statistical issues identified as a result of conducting the survey. Special attention is devoted to the statistical biases associated with the dual frame approach.

Program Announcement

Title: **Statistical Methods for Alerting Algorithms in Biosurveillance**

Chair: Myron Katzoff, National Center for Health Statistics

Speakers: Howard S. Burkom, National Security Technology Department, The Johns Hopkins University Applied Physics Laboratory

Date/Time: Tuesday, January 17, 2006 / 12:30 - 2:00 p.m.

Location: National Center for Health Statistics, room 1403A, 3311 Toledo Road, Hyattsville, MD (Metro: Green Line, Prince George's Plaza and then about a 10 minute walk). Note: please try to arrive 15-30 minutes early because of possible security screening delays.

Sponsors: WSS Section on National Defense and Homeland Security

Abstract: Syndromic surveillance involves the monitoring of available data sources for early warning of outbreaks of unspecified disease or of specified disease before the confirmation of identifying symptoms, with the objective to complement physician sentinel surveillance with false alarm rates acceptable to the public health infrastructure. Data sources include clinical data such as counts of syndrome-specific emergency department visits or physician office visits, and nonclinical data such as over-the-counter remedy sales and school/work absentee rates.

The terrorist attacks of 2001 added urgency to the development and activation of automation-aided biosurveillance, and system applications have extended to monitoring of natural public health threats such as the onset of influenza season as well as to recent new ones such as West Nile virus, the SARS epidemic, and a potential avian flu pandemic. Effective systems require a combination of expertise in medicine and epidemiology, in information technology, and in statistics and related fields of analysis.

A common approach among system developers has been to adapt chart-based methods from the field of statistical process control. Major obstacles to this approach are the evolving and often nonstationary input data streams, the uncertainty of the nature of the signal to be detected, and the presence of systematic or periodic behavior in the data background. Thus, robust detection performance, measured by timeliness and sensitivity at controlled alert rates, requires a combination of modeling and process control suitable to the characteristics of the monitored data.

The technical part of this presentation discusses several algorithmic approaches to the monitoring of syndromic time series, including adaptations of standard control charts, Riffenburgh's moving F statistic, and scan statistics. An interactive spreadsheet environment will be used to enable detailed examination of the positive and negative features of these methods on several data types. A generalized exponential smoothing approach to data modeling will be discussed, and a control chart derived from it will be used to illustrate a detection evaluation methodology.

Program Announcement

Title: **The Use of Contact History Data for Exploring Survey Nonresponse in Federal Demographic Surveys. (A Joint Seminar)**

Chair: John Dixon, U.S. Bureau of Labor Statistics

Speakers: Nancy Bates, U.S. Census Bureau
James M. Dahlhamer, National Center for Health Statistics/Centers for Disease Control and Prevention

Date/Time: Thursday, January 19, 2006 / 12:30 - 2 p.m.

Location: Bureau of Labor Statistics, Conference Center Room 9. To be placed on the seminar list attendance list at the Bureau of Labor Statistics you need to e-mail your name, affiliation, and seminar name to wss_seminar@bls.gov (underscore after 'wss') by noon at least 2 days in advance of the seminar or call 202-691-7524 and leave a message. Bring a photo ID to the seminar. BLS is located at 2 Massachusetts Avenue, NE. Use the Red Line to Union Station.

Sponsor: Methodology Section, WSS

Abstracts:

Reluctance to Participate in Federal Demographic Surveys: An Exploration of the National Health Interview Survey and Consumer Expenditure Survey using Survey Process Data

Nancy Bates and Andrea Piani
U.S. Census Bureau

In 2002-2003, the Census Bureau designed an automated contact history data collection system known as the Contact History Instrument or CHI. The CHI was developed to systematically record the number of contact attempts, mode, date and time of attempt and other details behind interim outcomes in personal visit surveys (e.g., reasons for refusals and strategies attempted).

Using CHI data from the 2005 National Health Interview Survey and the 2005 Consumer Expenditure Survey, we explore reasons why some households are reluctant to participate in the interview process. We investigate the extent of reluctance, what the most frequently cited reasons are, and whether these vary by characteristics such as survey topic, household composition, and other auxiliary variables such as region, urbanicity, or mode of contact. We also report how patterns of reluctance may change as the number of contacts increases. Finally we explore whether some reasons are more highly correlated with the decision to refuse the survey. In closing we offer recommendations how CHI data can be used as a feedback mechanism for improving field productivity and understanding the reasons people participate in federal surveys.

Developing Models of Initial Contact in the National Health Interview Survey (NHIS)

James M. Dahlhamer, Barbara J. Stussman, Catherine M. Simile and Beth Taylor
National Center for Health Statistics, Centers for Disease Control and Prevention

Response rates in government surveys have been declining over the past two decades raising concerns about the ability of survey estimates to accurately reflect the characteristics of the target population. One of the reasons for declining response rates is the reduced accessibility of households, arising, in part, from increased physical control of access to housing units and

household compositions in which no one is home for long periods of time. In an effort to achieve acceptable rates and quality of response, interviewers need to be as efficient as possible in contacting sample households so as to leave ample time for gaining respondent cooperation. The purpose of this study, therefore, is to identify factors that influence contactability.

The National Health Interview Survey (NHIS), an on-going population-based health survey conducted by the National Center for Health Statistics, Centers for Disease Control and Prevention, recently adopted the stand-alone, Blaise-based Contact History Instrument (CHI). Interviewers use CHI to record critical information on each contact attempt, including mode, date, and time of attempt, features of doorstep interactions, and strategies used for making contact and gaining cooperation. Using core survey and CHI data from the 2005 NHIS, models of initial contact with sample households are developed and tested. In addition to social-environmental (e.g., MSA status, region of residence) and household-level measures (e.g., the presence of children, household size, etc.) known to influence contactability, the role of interviewer strategies (e.g., time and mode of contact attempt, information-seeking behaviors) is assessed. By identifying attributes of difficult-to-contact households and the strategies for improving accessibility, survey procedures can be adjusted to improve the efficiency of field operations.

Employment

As a service to local statisticians, *WSS News* provides notification of employment opportunities and description of those seeking employment here in the Washington, DC, area. Readers are encouraged to take advantage of this feature of the newsletter. The deadline for inserting notices is five (5) weeks before the publication date. Those interested should email or call Anne Peterson, at apeterson@insightpolicyresearch.com or (703) 387-3032.

WESTAT

Westat is an employee-owned corporation headquartered in the suburbs of Washington, DC (Rockville, Maryland). We provide statistical consulting and survey research to the agencies of the U.S. Government and to a broad range of business and institutional clients. With a strong technical and managerial staff and a long record of quality research, Westat has become one of the leading survey research and statistical consulting organizations in the United States.

Our company was founded in 1961 by three statisticians. The current staff of more than 1,700 includes over 60 statisticians, as well as research, technical, and administrative staff. In addition, our professional staff is supported by data collection and processing personnel situated locally and in field sites around the country. The work atmosphere is open, progressive, and highly conducive to professional growth.

Our statistical efforts continue to expand in areas such as the environment, energy, health, education, and human resources. Westat statisticians are actively involved in teaching graduate-level courses in statistical methods and survey methodology in collaborative arrangements with area colleges and universities. We are currently recruiting for the following statistical position:

Survey Sampling Statistician (Job Code WSS/DRM/5001)

Three or more years of relevant experience in sample design and selection, frames development, weighting, imputation, and variance estimation. Must have a master's or doctoral degree in statistics and have excellent writing skills. Coursework in sample survey design highly desirable.

Westat offers excellent growth opportunities and an outstanding benefits package including life and health insurance, an Employee Stock Ownership Plan (ESOP), a 401(k) plan, flexible spending accounts, professional development, and tuition assistance. For immediate consideration, please send your cover letter, indicating the Westat Job Code, and resume by one of the following methods to:***[Job Code is REQUIRED to apply]*** Westat, Attn: Resume System, 1650 Research Boulevard, Rockville, MD 20850-3195; Email: resume@westat.com ; FAX: (888) 201-1452. We are an Equal Opportunity Employer.

Tenure Track Assistant Professor in Statistics or Mathematics

American University

The Department of Mathematics and Statistics in the College of Arts and Sciences at American University has an opening for a tenure track assistant professor in Statistics or Mathematics for Fall 2006.

Qualifications: earned doctorate in Mathematics or in Statistics by Fall 2006, as well as evidence of effective teaching and either a record of or the potential for continuing productive scholarship. Responsibilities: teaching undergraduate and graduate level mathematics or statistics courses; conducting research; advising and mentoring students, with particular sensitivity to women and

minority students; institutional service.

Submit letter of application and vitae to Search Committee, Department of Mathematics and Statistics, American University, 4400 Massachusetts Avenue NW, Washington, DC 20016-8050. Have official transcripts and three letters of reference sent directly to the department. At least one letter should specifically mention teaching experience.

All applicants are encouraged to review full application instructions, available at www.mathstat.american.edu/positions, or from the department at (202) 885-3124, or by email at mathstat@american.edu.

American University is an Equal Employment Opportunity / Affirmative Action employer, committed to a diverse faculty, staff, and student body.

Women and minority candidates are strongly encouraged to apply.

Financial Analyst/Manager

Our client, a well established, boutique management consulting firm, provides statistical, econometric and financial analysis; market research; forecasting; and strategic advice in a range of industries.

They are seeking a Financial Analyst/Manager to assist in a wide variety of analytical tasks including data analysis, cash flow, actuarial forecasting, econometrics, and a variety of research methods. The successful candidate will be responsible for rigorous quantitative analyses and will manage 1-3 people on consulting projects.

Qualifications include:

- 3+ years relevant experience
- Outstanding analytical skills
- Advanced Excel skills
- Strong communication skills
- Mathematics and/or statistics background is a plus but not required
- MBA preferred but not required

The company is located in Washington, DC. Travel less than 20%.

Salary $80,000-$110,000 + Generous benefits.

Please submit your qualifications, highlighting those relevant to above, to Tammy Tuan at tct@JustinBradley.com

Statisticians, Junior and Senior

Starting Date: Immediate

Location: Department of Transportation, Washington, DC

Education: MS or Ph.D. in Applied Statistics

Skills Required: SAS and Survey Sampling

Communication Skills: Good Oral and Written Communication

Salary: Open

Duties: Providing Statistical and Analytical support to the US Department of Transportation. Must

be able to do independent research.

Apply to: Mr. A. R. Char, Senior VP, URC Enterprises, Inc.; email: ar.char@verizon.net, Fax: 301-570-5622, Phone: 301-570-5622

NOTICE OF SENIOR-LEVEL VACANCY

The U.S. Government Accountability Office (GAO) has a senior-level, nonmanagerial opening for its Chief Statistician. The Chief Statistician serves as an expert advisor and chief consultant on the statistical aspects of GAO's work. The overall responsibilities are to provide leadership on GAO's application of statistical techniques and GAO's evaluation of federal statistical programs.

Please submit resumes to Dr. Nancy Kingsbury at KingsburyN@GAO.gov or call 202-512-2700. If you have any questions or wish to nominate someone, please contact Dr. Kingsbury.

P.O. Box 752
Suitland, MD 20752

NON-PROFIT ORG.
U.S. POSTAGE
PAID
FALLS CHURCH, VA
PERMIT NO. 186

Washington Statistical Society