GSS/SSS SEPTEMBER 2011 NEWSLETTER
Government Statistics and Social Statistics Sections
American Statistical Association

GSS web site: http://www.amstat.org/sections/sgovt/
SSS web site: http://www.amstat.org/sections/ssoc/

GSS Publications Officer: Sonya L. Vartivarian
(VartivarianS@gao.gov)

SSS Publications Officer: Joanna M. Turner
(turn0053@umn.edu)

Newsletter Editor: Natalya Verbitsky-Savitz
(NVSavitz@Mathematica-MPR.com)

IN THIS ISSUE:

- **CONGRATULATIONS TO GSS-SPONSORED AWARD WINNERS FOR 2011!** **NEW!**
- **THE COUNCIL OF PROFESSIONAL ASSOCIATIONS ON FEDERAL STATISTICS: JUNE 2011 MEETING** **NEW!**
- **FORECASTING CITRUS PRODUCTION IN FLORIDA** **NEW!**
- **ANNOUNCEMENT: JPSM-CENSUS DISTINGUISHED LECTURE BY J.N.K. RAO** **NEW!**

~~~~~~

CONGRATULATIONS TO GSS-SPONSORED AWARD WINNERS FOR 2011! NEW!

Submitted by Steven Paben, 2011 GSS Chair

The JSM Meetings in Miami Beach served as a great opportunity to recognize all of the Government Statistics Section (GSS) sponsored and co-sponsored award winners for 2011. The Roger Herriot Award was presented at a special session to Michael Messner of the EPA. This award, which includes a $1000 honorarium and a citation, is to recognize individuals who develop unique and innovative approaches to the solution of statistical problems in federal programs. In particular, Michael was recognized for his work on Statipedia.

The GSS poster competition winners were Emily O'Malley Olsen and Sherry Everett Jones from the Centers for Disease Control and Prevention. Their award winning poster was entitled "Method Comparison for Assessing Trends Over Time of Age of First Cigarette Use Among High-School Students in the U.S. – Youth Risk Behavior Survey, 1991-2009."

At a special session co-sponsored with the Social Statistics Section and the Survey Research Methods Section, the five student paper competition winners presented their papers. The overall quality of the presentations was quite impressive.

At the GSS Business Meeting, the Wray Jackson Smith Scholarship Award was presented to Rebecca Medway, a doctoral student in the Joint Program in Survey Methodology at the University of Maryland. She plans to use the $1000 award to attend the 2012 JSM and present findings related to her research on the potential tradeoff between nonresponse bias and measurement error when incentives are used. We also re-recognized Jenise Swall of the EPA as

the winner of the Jeanne Griffith Mentoring Award. Finally, we announced this year's winner of the Pat Doyle Award for services to the GSS, our illustrious newsletter editor Natalya Verbitsky-Savitz!

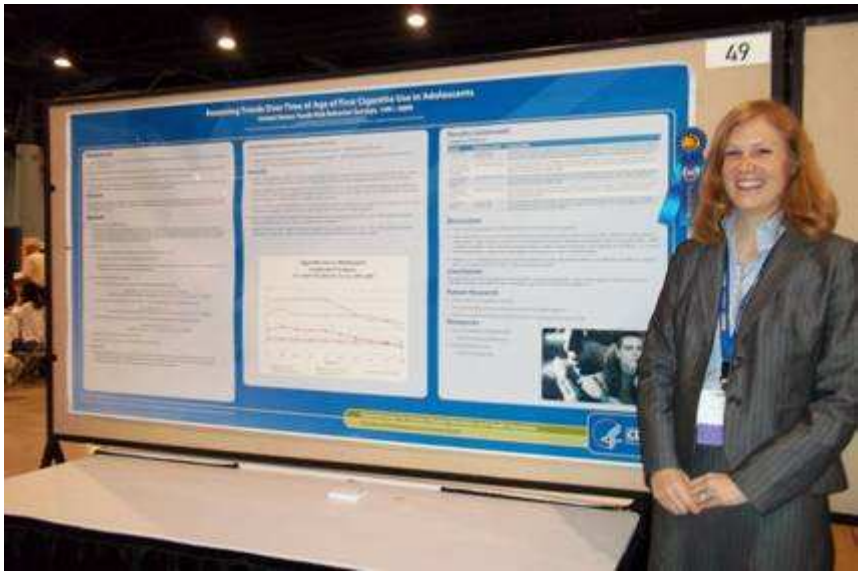


Photo: Emily O'Malley Olsen (CDCP) after winning the GSS Poster Competition at JSM 2011.

~~~~~~

THE COUNCIL OF PROFESSIONAL ASSOCIATIONS ON FEDERAL STATISTICS: JUNE 2011 MEETING NEW!

Submitted by Robert Lussier, COPAFS representative, Government Statistics Section

Note: The Council of Professional Associations on Federal Statistics (COPAFS) acts as the advocate for the development and dissemination of high-quality federal statistics. Member organizations include professional associations, businesses, research institutes, and others interested in Federal statistics. Through COPAFS, members have an opportunity to review and have an impact on issues including timeliness, quality, confidentiality, and the relevance of data. COPAFS holds quarterly meetings, the last one being on June 3 2011. Detailed minutes, together with copies of the overheads used by the presenters can be found on COPAFS website ***www.copafs.org***

As part of his Executive Director's Report, Ed Spar first noted that while the FY2011 budgets have finally passed, it is not yet clear exactly how funding will be distributed. Amid all the question marks, indications are that the Census Bureau will receive funding for annual updates to the Master Address File and for some expansion of the American Community Survey (ACS) sample. There is concern about what happens if such funding is not provided for 2012. Late word is that the Stat Abstract may not be eliminated; instead, it may take the form of links to data on the web, rather than a finished publication.

Spar observed that the situation at National Center for Health Statistics (NCHS) has improved somewhat, as they have some supplemental funding from the Affordable Care Act. They should be able to maintain and possibly even upgrade the National Health Interview Survey and the National Health and Nutrition Examination Survey.
~~~~~~

The Office of Government Ethics has proposed a change to a rule that limits the ability of federal employees to serve in official capacities in nonprofit organizations. This includes the permission to serve as board members. Spar commented that it could be helpful to organizations like COPAFS.

Senator Tom Coburn has issued a report in which he charges the National Science Foundation (NSF) with waste and mismanagement. Coburn denigrates the role of social sciences, and calls for elimination of NSF's Social, Behavioral and Economics Directorate.

Spar then introduced Adrienne Pilot of the Council of Economic Advisors, who described an initiative that would permit additional data sharing between the Census Bureau, Bureau of Labor Statistics (BLS) and Bureau of Economic Analysis (BEA). As Pilot described it, the Census Bureau's business register is comingled with Internal Revenue Service (IRS) files, and therefore falls under IRS restrictions. The Census and BLS business lists could be brought together, and the differences reconciled to produce an improved database. The next step is to identify legislation to which this measure can be attached.

A Review of the Upcoming 2012 Census of Agriculture

Cynthia Clark, National Agricultural Statistics Service (NASS), explained that the agency has two major funding line items – the agricultural estimates program and the census of agriculture (taken every five years). The 2011 budget came in with less funding than in 2010: the census is moving ahead with the basic plan, but the estimates program took a significant hit.

In planning the 2012 census, input has been sought from numerous users. Internal users have called for information on animal grazing, rented land by utilization, and expanded information on renewable energy. Recommendations from external users include expanded information on agroforestry, counts of farms that sell to intermediary outlets, a count of on-farm packing facilities, and farm use of broadband Internet.

Data collection tests are taking place. Clark described a number of tests related to increasing response. With the results of these tests in mind, 2012 data collection will include autodial or postcard pre-census notification, on-line web reporting, initial mail-out of questionnaires, postcard or autodial reminder, a second questionnaire mailing, and the use of certified mailing to significant operations.

Milestones include online web reporting available November 2012, initial questionnaire mailing in December 2012, data collection through May 2013, and the release of results in early 2014.

Population Considerations for the Nuclear Field

Danny Smith, consultant in the nuclear field, described events that took place at three nuclear facilities: the Shoreham nuclear plant on Long Island Sound (about 60 miles from Manhattan), the Fukushima Daiichi plant on the east coast of Japan, and the Chernobyl plant in Ukraine.

Rethinking Urban-Rural and the Barriers between Statistical and Programmatic Uses

Michael Ratcliffe, U.S. Census Bureau, explained that the Census Bureau has not been taking programmatic uses into account when developing statistical geographic area concepts or when delineating areas. The purpose was solely the tabulation and presentation of data. However, the Census Bureau is increasingly aware of the need to consider how areas are used.

Reviewing recent urban/rural classifications, Ratcliffe noted that they tend to be dichotomous, with rural and nonmetropolitan defined as residual categories. Looking to the future, Ratcliffe said that thought is being given to the development of an urban-rural continuum that would better reflect the variety of urban, suburban, exurban and rural landscapes, as well as the functional relationships between urban areas. The meaningful subdivision of larger agglomerations is another topic for consideration.

John Cromartie, Economic Research Service, called the idea of a single definition needlessly limiting, and argued that research on rural issues requires different perspectives. In fact, dozens of definitions exist, including those established by the Census Bureau, Office of Management and Budget (OMB), U.S. Department of Agriculture (USDA), U.S. Department of Housing and Urban Development (HUD), Department of Health and Human Services (HHS), and others. The differences between these definitions boil down to two questions. First, for any entity, where is the boundary between urban and rural, and second, what is the minimum population size for an entity to be considered urban?

On the question of boundary, Cromartie explained that urban boundaries can be based on administrative areas (such as cities), land use (the view from an airplane), or economic factors (such as commuting or labor force). On the question of minimum population size, Cromartie noted that any rural definition includes some towns and villages below a chosen population threshold. The Census Bureau has long used 2,500 as the urban population threshold. Cromartie argued for the need for higher urban-size thresholds. He also noted that USDA has adjusted its threshold upward over the years, but not the Census Bureau.

Cromartie concluded by suggesting that we could improve the efficiency of federal programs by adopting multiple urban/rural definitions, targeted to multiple purposes, and that the choice of definition should be driven by the objectives of the program or application.

Overview of New Interactive Bureau of Economic Analysis (BEA) Tables

Thomas Dail from BEA, described a soon to be released web-based data access tool that provides enhanced access to BEA data that previously have been available only through five separate applications. There was a soft launch in April, and the full launch is scheduled for June 10. The system can be accessed at www.bea.gov/itable.

Benefits of the new system include data access with a single tool that enables more efficient table creation, downloads to more formats, customized tables and charts, saving and exporting charts, and even the forwarding of tables via social media.

Dail demonstrated the system and wrapped up by announcing that a Version 2.0 is already being developed for 2012, and will provide enhanced features, such as the ability to perform calculations.

This concluded the June meeting. The next COPAFS meeting will take place September 9, 2011.

~~~~~~

FORECASTING CITRUS PRODUCTION IN FLORIDA

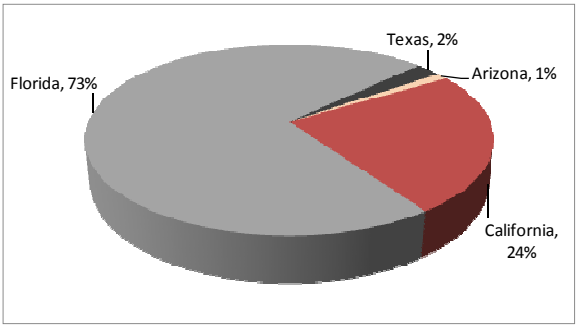
Submitted by Jeff Geuder, Director, USDA/NASS – Florida Field Office

The National Agricultural Statistics Service (NASS) is the statistical data collection and reporting arm of the USDA. Headquartered in Washington DC, the agency maintains 46 Field Offices around the country to administer the survey program and provide official estimates of agricultural production. The Florida Field Office works cooperatively with the Florida Department of Agriculture and Consumer Services (FDACS) to provide statistics on agricultural production of fruit, vegetables, floriculture crops, cattle (beef and dairy), as well as statistics pertaining to chemical usage, labor, and prices.

FORECASTING FLORIDA'S CITRUS PRODUCTION

During the last decade, nearly three quarters of all U.S. citrus was grown in Florida (Figure 1). The relative size of the Florida citrus crop demands that statistically accurate forecasts be made starting in October of each year and continuing throughout the marketing year. In order to provide these statistically accurate forecasts, NASS has developed a survey program based on objective counts and measurements (rather than subjective reporting from individual producers).

Figure 1. U.S. Citrus Production: Percent of Total (2001-2010)

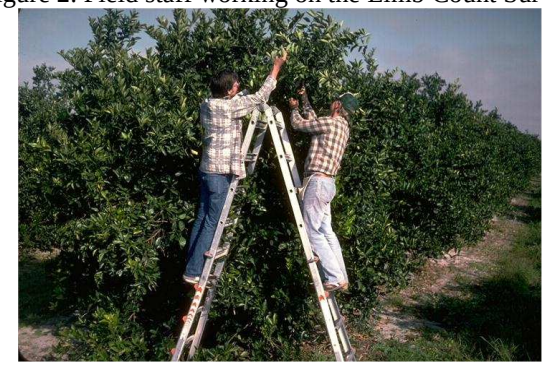


There are four basic components of the citrus production forecast: (1) number of bearing trees; (2) average number of fruit per tree; (3) average fruit size; and (4) fruit loss from droppage. NASS conducts a series of surveys to estimate each of these components.

Commercial Tree Inventory: The commercial tree inventory is conducted every year to provide an estimate of the number bearing trees. It also provides a sampling frame for the objective measurement surveys (described below). Digitized grove boundaries are used to generate maps, which field crews use to check for changes from the previous tree inventory, typically old trees being removed or new trees being planted. The resulting database consists of every block of citrus in the State, with information on size, location, variety of fruit, and age of the trees.
~~~~~~

Survey To Estimate Fruit Per Tree: Another survey is conducted every year in August and September to provide an estimate of the average number of fruit per tree. A stratified sample of over 3,200 groves is selected from the tree inventory database. Survey crews visit each selected grove and use a random selection process to select two trees within the grove. On each selected tree, a sample limb representing approximately 10 percent of the bearing surface of each tree is randomly selected. Fruit is then counted on this sample limb and the fruit counts are expanded by the reciprocal of the probability of selection to provide an estimate of the total number of fruit on the tree (Figure 2).

Figure 2. Field staff working on the Limb Count Survey



Survey To Estimate Fruit Size: Citrus fruit size varies from season to season, depending on weather and cultural practices. A fruit size survey is conducted monthly, starting in September and continuing throughout the season, using a stratified sample of approximately 1,800 groves selected from the tree inventory database. Field crews randomly select two trees in each of the sample groves and randomly select a limb to measure the size of each piece of fruit on the limbs. NASS uses a 10-year regression model to project each month's survey measurements to final size at harvest.

Survey To Estimate Fruit Loss: A certain amount of fruit drops off citrus trees during a growing season. Since the initial estimate of fruit per tree is made in late summer, NASS conducts a monthly survey to track the drop rate throughout the year. Using the same sample of 1,800 groves, field crews randomly select trees and branches for this survey. Beginning in September and then at monthly intervals, fruit on the sample branches are counted. NASS uses a 10-year regression model to project each month's survey measurements to final drop at harvest.

Figure 3. Accuracy of October and January Forecasts, Oranges (1996-2010)

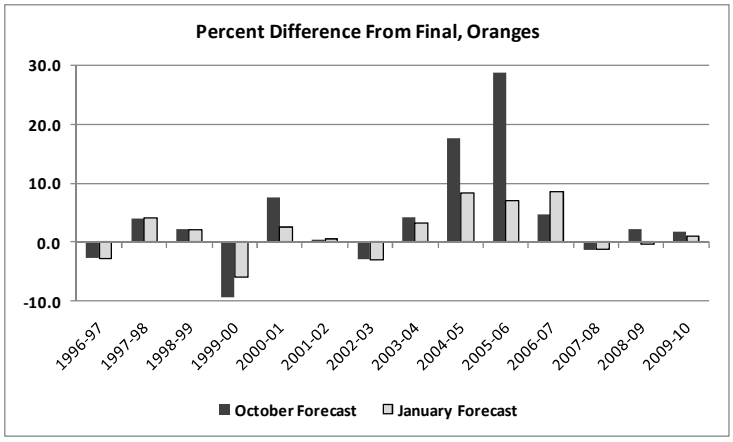
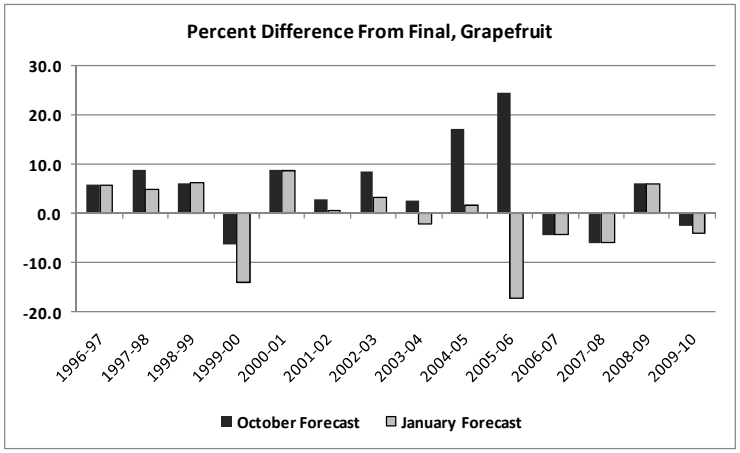


Figure 4. Accuracy of October and January Forecasts, Grapefruit (1996-2010)



ACCURACY OF MONTHLY FORECASTS

Several factors influence the accuracy of the monthly production forecasts. Weather events (such as hurricanes and freezes) have a significant effect on the final production of citrus each year. Figure 3 shows the how the October and January forecasts for oranges compare to the final production over the past 10 seasons. Figure 4 shows the same data for grapefruit. In most seasons, the October forecast is within 5-10 percent of the final production. The notable exceptions are two seasons affected by hurricanes (2004-05 and 2005-06) and a season in which there was a record low loss due to fruit drop (1999-2000).

~~~~~~
~~~~~~

ANNOUNCEMENT

JPSM-Census Distinguished Lecture

Dr. J.N.K. Rao is a Distinguished Research Professor at Carleton University, Ottawa, Canada, a Consultant to Statistics Canada, and a Member of Statistics Canada's Advisory Committee on Methodology. Among the awards and honors, Professor Rao has received the Waksberg Award for Survey Methodology, the Gold Medal of the Statistical Society of Canada, election to the Royal Society of Canada, and Honorary Doctorate of the University of Waterloo. He has made fundamental contributions to the design-based classical theory of sampling, to the foundations of sampling during the debates of the 1960s and 70s, to a variety of aspects of variance estimation, to the analysis of complex survey data, and to small area estimation.

Date: Friday, September 30, 2011
Time: 3:00 PM - 5:00 PM (Reception: 5:00 PM – 6:00 PM)
Location: 1524 Van Munching Hall, University of Maryland, College Park
Title: Estimation of Complex Small Area Parameters with Application to Poverty Indicators
Presenter: Dr. J.N.K. Rao

Chair: Dr. Graham Kalton, Senior Vice President at the Westat

Discussants:

1. Dr. Peter Lanjouw, Manager, Poverty and Inequality Group,
   Development Economics Research Group (DECRG), the World Bank
2. Dr. Partha Lahiri, Professor, Joint Program in Survey Methodology,
   University of Maryland, College Park

Abstract: Model-based small area estimation has largely focused on means or totals, using either area level models or unit level models. Empirical best linear unbiased prediction (EBLUP), empirical Bayes or empirical best (EB) and hierarchical Bayes (HB) methods have been extensively used for point estimation and for measuring the variability of the estimators. Primary purpose of this presentation is to study the estimation of complex non-linear small area parameters by using EB and HB methods. Our methodology is generally applicable, but we focus on measures of poverty indicators, in particular on the class of poverty measures called FGT poverty measures (Foster, Greer and Thorbecke, 1984). The World Bank has been releasing small area estimates of the FGT measures for several countries, using the methodology of Elbers, Lanjouw and Lanjouw (2003). The ELL methodology assumes a unit level nested error linear regression model that combines both census and survey data and produces simulated censuses of the variables of interest using the bootstrap. Estimates for any desired small areas are produced from the simulated censuses. The average of the resulting estimates is taken as the area estimate and the variance of the estimates is taken as a measure of variability of the area estimate. We present EB estimation of FGT poverty measures for small areas using best prediction methodology based on the joint predictive density of the non-observed values given the observed

data, assuming normality for a suitably transformed value of the variable of interest, for example log of the welfare variable. A nested error linear regression unit-level model with random small area effects is assumed on the transformed variable. We show that values from the joint predictive density under the unit level model can be obtained by generating only univariate normal variables. For comparison with the ELL method, we assume the same model with small area random effects for ELL, although ELL did not include small area effects in their models. We use a parametric bootstrap method for estimating the mean squared error (MSE) of the EB estimators. We develop a census EB method that can be used when the sample data cannot be linked to census auxiliary data. We also study HB estimation under normality, assuming a diffuse prior on the model parameters. We show that the posterior mean and the posterior variance of small area parameters can be obtained using a grid method that avoids the use of Monte Carlo Markov chain methods for generating values from the posterior density of the parameter of interest. If the distribution of random effects and/or unit errors in the unit level model deviate significantly from normality, then the normality-based EB or HB estimators can be biased under significant skewness. We extend the EB method to skew normal random effects and/or unit errors. We present the results of a model-based simulation study on the relative performance of EB, ELL and HB estimators.

Contact Person: Dr. Yang Cheng, Branch Chief, Governments Division, US Census Bureau
yang.cheng@census.gov

For direction, parking, and other information:
http://www.jpsm.umd.edu/jpsm/?events/specialevents/distinguished_lecture_2011_09_30/index.htm

~~~~~~
~~~~~~

GOVERNMENT STATISTICS SECTION

The mission of the Government Statistics Section (GSS) is to promote the use of sound statistical theory and methods in the production of data at all levels of government--be it Federal, State, local, or international; assist in the broad dissemination of those data; and encourage good statistical practice by all users. Areas of interest for the Section include all that involve the production, dissemination, and application of governmental statistics, including concern with statistical policy issues, quality and usefulness of governmental data products, special problems of State and local data, comparability of data among different countries, and the role of professional statisticians in the public sector.

OFFICERS AND CONTACTS

Steven P. Paben
Chair, 2011
Paben.Steven@BLS.gov

Brian A. Harris-Kojetin
Chair-elect, 2011
Brian_A._Harris-Kojetin@omb.eop.gov

John S. Dixon
Past Chair, 2011
Dixon.John@BLS.gov

Iris M. Shimizu
Program Chair, 2011
IShimizu@cdc.gov

Daniell S. Toth
Program Chair-elect, 2011
Toth.Daniell@bls.gov

Kevin Cecco
Secretary/Treasurer, 2011-2012
kevin.cecco@irs.gov

Rick Peterson
Staff Liaison
Rick@amstat.org

Kennon Copeland
Council of Sections Representative, 2010-2012
Copeland-Kennon@norc.org

Robert Lussier
COPAFS Representative, 2011-2012
RobertMLussier@videotron.ca

Carol C. House
COPAFS Representative, 2010-2011
HousCa@gmail.com

Sonya Vartivarian
Publications Officer, 2007-2012
VartivarianS@gao.gov

Natalya Verbitsky-Savitz
Newsletter Editor, 2009-2011
NVSavitz@Mathematica-MPR.com

Bill Wong
Assistant Editor, *Amstat Online*
bjwg@yahoo.com

Tai Phan
GSSLIST Coordinator
Tai.Phan@ed.gov

2010 GSS CORPORATE SPONSORS

Bureau of Economic Analysis (BEA), U.S. Department of Commerce
Bureau of Labor Statistics (BLS), U.S. Department of Labor
Division of Science Resources Statistics (SRS), National Science Foundation (NSF)
National Agricultural Statistics Service (NASS), U.S. Department of Agriculture (USDA)
National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC)
Statistics of Income Division, Internal Revenue Service (IRS)
U.S. Census Bureau
U.S. Social Security Administration

2011 GSS CORPORATE SPONSORS

Bureau of Economic Analysis (BEA), U.S. Department of Commerce
Bureau of Justice Statistics (BJS), U.S. Department of Justice
Bureau of Labor Statistics (BLS), U.S. Department of Labor
Division of Science Resources Statistics (SRS), National Science Foundation (NSF)
National Agricultural Statistics Service (NASS), U.S. Department of Agriculture (USDA)
National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC)

SOCIAL STATISTICS SECTION

The Social Statistics Section seeks to advance research in social statistics, both in areas which involve the use of methods of statistical inquiry, and in those which involve the use of statistical data and the development of statistical measurement. This section will also plan for active participation in the affairs of the American Statistical Association by those interested in these matters and for representation of activities in this major field in the program of the ASA.

OFFICERS AND CONTACTS

John Thompson
Chair, 2011
Thompson-John@norc.org

Kirsten K. West
Secretary/Treasurer, 2010-2011
Kirsten.K.West@census.gov

Joseph Salvo
Chair-elect, 2011
JSalvo@planning.nyc.gov

Robert Santos
Council of Sections Representative, 2010-2012
RSantos@urban.org

Sharon M. Stern
Past Chair, 2011
Sharon.M.Stern@census.gov

Joanna M. Turner
Publications Officer, 2011-2012
turn0053@umn.edu

Nancy Clusen
Program Chair, 2011
NClusen@Mathematica-MPR.com

Rick Peterson
Staff Liaison
Rick@amstat.org

Elizabeth Stuart
Program Chair-elect, 2011
estuart@jhsph.edu