NSF-Census Research Network
Newsletter
Vol. 2, Issue 3

Research Node Focus: Carnegie Mellon University

Stephen (Steve) Fienberg, Maurice Falk University Professor of Statistics and Social Science at Carnegie Mellon University (CMU) and **Bill Eddy**, John C. Warner Professor of Statistics (Emeritus) at CMU, have a long-standing interest in everything to do with census taking. They have spent several decades working on statistical issues for the U.S. Census. Fienberg served two terms, and Eddy served two terms, as Chair of the Committee on National Statistics (CNSTAT) for the National Academy of Science's National Research Council.

In fact, in 1999 Fienberg and Margo Anderson wrote a book, *Who Counts?*, that talks about the 1990 decennial census and the efforts made to use sampling to adjust census results for the differential undercount.

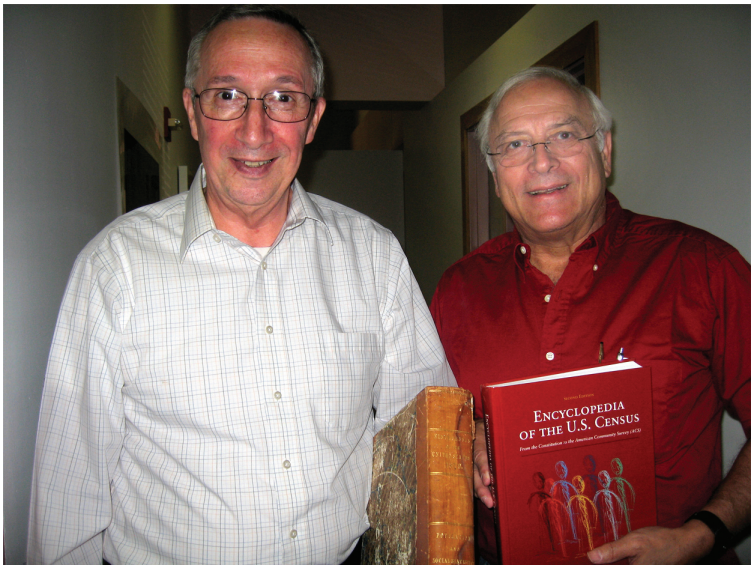
Fienberg's research has long focused on how to conduct a census, and why a census is actually a statistical activity. As such, what is the formal way to think about it? Fienberg is a renowned veteran in census research. He has been studying census surveys since the 1970s. He looked at what methodologies one should use in interpreting censuses. He has been studying census surveys since the 1970's. He looked at what methodologies one should use in interpreting censuses.

As chair of CNSTAT in the '80s, Fienberg convened a panel that was the trigger to census evaluation. "The first really big census panel on new methodology for census taking which came out in 1985 was done under my auspices." Almost two decades later, before Eddy was leading CNSTAT, he was a member of another panel evaluating the results of the 2000 census; the panel's work included elaborate field visits during census taking as well as an assessment of the methodology and how well it worked.

CMU's node is focusing on three main issues, including the cost of census taking, the diminishing response rates and privacy issues.

"We are looking at things such as record linkage, and thinking about how to benchmark the census and how to decide if something is accurate or not accurate. These are all themes that show up from the work that Bill and I had done over the years," said Fienberg.

Another important aspect that CMU's node is working on is education and training. Eddy is making projects related to census work available to summer undergraduates. "Rebecca Nugent, who is one on the NCRN team of researchers, oversees the whole statistics undergraduate program



Bill Eddy and Steve Fienberg. Bill is holding a report from the 9th Census in 1870, and Steve is holding the a Encyclopedia of the Census in which he has an article.

which has grown from a rather small number to about 250 majors. They all get exposed in one way or another about issues in methodology in survey and census taking," explained Fienberg, "We want to get the next generation of statisticians and informed citizens engaged in the importance of these issues."

One of the summer projects this year that the undergraduate students worked on was to recreate the website that the U.S. Census will use in 2020 to conduct online surveying. "Using the Internet is a horrible break from tradition that causes the U.S. Census staff immense mental pain because if somebody says they live in a particular place to give their information, they have no way to know if that is true,"

Continued on page 2

In This Issue

Research Node - CMU (Continued)

From page 1

said Eddy. Another term for this phenomena is "non-ID" because there is no way to know where the person is really located.

The students wanted to see if there was a better response rate if it was known up front that the questions being asked were for the government organization or not. In about five weeks they built a replica of the U.S. Census' website for the 2020 Census. Frank McPhillips from the U.S. Census told Eddy that he was flabbergasted at the amount of work the students were able to accomplish in just five weeks.

Fienberg said, "Our graduate students are now moving on in their careers. Mauricio Sadinle just took a position at Duke and NISS as a postdoctoral fellow. Sam Ventura has just joined our department faculty on a multi-year appointment, and Rebecca Steorts who was a postdoc on our project is now on the faculty at Duke." Eddy and Fienberg hope that everyone involved in their NCRN project will continue to explore census and survey research issues in their new roles.

"There is a giant gulf between the way the agencies go about their business and the ways in which we in the department think about statistical methodology more broadly," noted Fienberg.

Another area that CMU is focusing on is record linkage, which is taking information from two files that are overlapping and combining the information. The U.S. Census takes a section of the census block, it does the census again for this block and then puts the data together to figure out how accurate the census data is. But there are a couple of problems with this methodology. One is that you pair up records for different people, or you fail to match records that actually belong to the same person. And, once the match has been made, most statisticians and practitioners have treated the matching process as if it had been without error. CMU's team is looking at record linkage from very fresh perspectives, and is especially interested in the propagation of matching uncertainty into subsequent analyses.

Another area that the team is researching is how to handle the large amount of data that is produced by the various censuses and surveys, especially when it comes to record linkage. Take, for example, a file that has 100,000 records in it and another file that has 200,000 records in it and now you are trying to link the two together. You have to match every one of the 100,000 records to every one of the 200,000 records. That takes a lot of computing power. But, the U.S. population is over 300 million, so scaling this up to a national level is really hard. "A lot of the activity we have been doing is focusing on how to make record linkage methods scale," said Fienberg. They have been using data other than the U.S. Census data to experiment with their methods because working with the Census data is sometimes tricky. One of the data sets they have used is the number of people who have been killed in Syria. There are seven of these lists and the team has focused heavily on combining four of the seven lists and to do it for the multiple files all at once.

Zack Kurtz in his PhD thesis work as part of the project worked on multiple recapture estimation motivated by looking at three or more lists. Once you've linked them, how do you estimate how many people you didn't see? This is also known as the problem of multiple systems estimation and we have studied it in different ways over several decades at CMU. Zack's work has extended the known approaches in interesting ways of relevance to census work.

The other major area of research CMU is doing involves privacy and research. There are two kinds of intrusions. One is someone trying to hack into the files. The second kind of intrusion is the data that are made available and by doing record linkage, an intruder can gain information about people by using the information that is in a supposedly protected data set.

Alessandro Acquisti, one of the co-principal investigators, who is professor of Information Technology, and his former Ph.D. student and postdoc, Laura Brandimarte, have been focused on how people respond to the issues and concerns about privacy and confidentiality when they actually interact with a governmental agency. They have done some fascinating experiments about this, particularly about people sharing geography online or in some kind of electronic form. They have been deeply involved with the discussion about online forms. They are also looking at online geographic information and the way in which it could be used to help identify the location of individuals responding to an online Census form. Laura is joining the faculty at the University of Arizona this fall, but will remain as a collaborator to the CMU team.

There is another project that the CMU team would like to accomplish where they would take the Pennsylvania drivers license records and match them with the American Community Survey (ACS) and the 2010 U.S. Census data so they can actually scale up some of the theories they have written, but moving the data set from Pennsylvania to the U.S. Census has been held up for three years, so it is uncertain as to whether the team will be able to work on this project or not.

The NCRN team at CMU can be proud that the work they have accomplished over four decades or more will continue to be carried on by the next generation of statisticians thanks to the great mentoring by this veteran team.

Communicating Uncertainty in Official Statistics

On June 18, **Charles Manski**, Board of Trustees Professor in Economics at Northwestern University, made a plenary presentation at the 2015 Policy Summit of the Federal Reserve Bank of Cleveland. Manski described aspects of his NCRN research on how uncertainty from measurement affects public policy. Erica Groshen, commissioner of Labor Statistics, Bureau of Labor Statistics, and Jeffrey Kling, associate director for Economic Analysis, Congressional Budget Office, also participated in the session. Manski's presentation was based on his forthcoming article that will be published in September in the *Journal of Economic Literature*, entitled, "Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern."

Over fifty years ago, Oskar Morgenstern, a German-born economist who is most famously known for helping to develop game theory and the foundations of decision theory with John von Neumann, wrote a book in which he urgently argued that government statisticians measure error in official economic statistics and report these measurements to the public (*On the Accuracy of Economic Observations, 1963)*. Manski's article reviews the state of affairs through various federal statistical agencies on how they are reporting potential errors today.

Federal statistical agencies typically report official economic statistics as point estimates and don't provide accompanying measures of error. Users of the statistics may incorrectly view them as error free, or may not understand the magnitude of error. Manski's article discusses issues such as survey nonresponse which can greatly skew the results in a survey. Many federal agencies' news releases never mention when there is a high nonresponse rate. Manski also examines issues such as seasonal adjustment and data revisions, as other areas in which agencies should measure and report potential errors.

Kristen Olson Named as New PI of the Nebraska Node

Kristen Olson is the new PI of the Nebraska node. Olson is an Associate Professor of Sociology at the University of Nebraska-Lincoln. Her research focuses on interviewer effects, survey nonresponse, mixed mode surveys, and causes of measurement errors. Olson is the Membership and Chapter Relations Chair of the American Association for Public Opinion Research, current Past-President of the Midwest Association for Public Opinion Research, Associate Editor of Research Syntheses for Public Opinion Quarterly, and Associate Editor, Journal of Official Statistics.

Kristen Olson is the new PI for the Nebraska research node.

Save the Date!

The next NCRN meeting will be held December 14-15, 2015 at the Bureau of Labor Statistics in Washington DC. More information about the meeting will be shared soon.

Node News

David Folch is the new Principal Investigator of the Spatial Sciences Census Research Node (SSCRN) at University of Colorado at Boulder and University of Tennessee. **Seth Spielman**, who is taking a leave of absence, remains connected as Co-Principal Investigator.

Scott Holan from the Missouri node was promoted to Professor from Associate Professor. His new position begins September 1.

Matthew Simpson PhD – Iowa State University (Co-major Economics and Statistics) joined the Missouri Node as a postdoc in August.

The Nebraska Node welcomes **Adam Eck** as a postdoctoral research associate for the 2015-2016 academic year. Eck is a computer scientist trained at University of Nebraska-Lincoln under NCRN co-PI LeenKiat Soh.

Jolene D. Smyth and **Kristen Olson** (Nebraska node) were selected by the European Survey Research Association as having the best paper published in the journal *Survey Research Methods* in 2013-2014.

Duke/NISS supported students Tyler Ransom (Economics) and Monika Hu (Statistics) received PhD degrees from Duke, and Olanrewaju Akande and Sohae Oh received Masters degrees from Duke.

Noel Cressie who is also from the Missouri node, is a Member of the Scientific Committee, 2016 Spatial Econometrics Association Conference, to be held in Rome.

Hang Kim who was a postdoctoral student for the Duke/NISS node, has accepted a position as assistant professor at the University of Cincinnati.

The Census - Research Data Center (RDC) will be opening at University of Missouri this Fall. Researchers from Missouri University and around the Midwest will be able to access millions of files of census bureau data for research projects ranging from public health issues to economics.

Fienberg Receives the Jerome Sacks Award for Cross-Disciplinary Research

Stephen E. Fienberg received the 2015 Jerome Sacks Award for Cross-Disciplinary Research. The announcement was made by Nell Sedransk, Director of the National Institute of Statistical Sciences (NISS), during a reception at the Joint Statistical Meetings in Seattle, Washington. Sedransk said Fienberg is receiving this award, "For a remarkable career devoted to the development and application of statistical methodology to solve problems for the benefit of society, including aspects of human rights, privacy and confidentiality, forensics, survey and census-taking, and more; and for exceptional leadership in a variety of professional and governmental organizations, including in the founding of NISS."

NISS Chair Mary Batcher gives Steve Fienberg the Jerome Sacks Award for Cross-Disciplinary Research

Fienberg is Maurice Falk University Professor of Statistics and Social Science in the Department of Statistics, the Machine Learning Department, and Cylab, and co-Director of the Living Analytics Research Center at Carnegie Mellon University. He has authored more than 20 books and 500 papers and related publications. Fienberg's interest in confidentiality and disclosure limitation ties both to surveys and censuses and also to categorical data analysis. It also led him to work on a multiple year project with NISS entitled "NISS Digital Government Project."

Fienberg is originally from Toronto, Canada. He received his bachelor's degree in mathematics and statistics from the University of Toronto, then earned his master's degree and his doctorate in statistics from Harvard University. He joined the CMU faculty in 1980. Amongst his many accolades, he has served as the co-Chair of the National Academy of Sciences Report Review Committee from 2008-2016. He is the co-Founder of the Journal of Privacy and Confidentiality, and the Editor-in-Chief of the Annals of Applied Statistics and Founding Editor of the Annual Review of Statistics and its Application. He is a fellow of the International Society for Bayesian Analysis and was President from 1996-1997; a fellow of the Institute of Mathematical Statistics and President from 1998-1999; and he is also a Fellow of the Royal Society of Canada, the American Association for the Advancement of Science,

Continued on page 10

Publications

The following are the most recent additions to publications produced by the research nodes within NCRN. A comprehensive list can be found here. (http://www.ncrn.info/documents/bibliographies)

Abowd, John M., and Ian M. Schmutte. Economic Analysis and Statistical Disclosure Limitation. Cornell University Preprint 1813:40581, 2015, available at http://hdl.handle.net/1813/40581.

Abowd, John M., Kevin L. McKinney, and Ian M. Schmutte. Modeling Endogenous Mobility in Wage Determination. Cornell University Preprint 1813:40306, 2015, available at http://hdl.handle.net/1813/40306.

Akande, O. A Comparison of Multiple Imputation Methods for Categorical Data (Master's Thesis), *Statistical Science*. Duke University Masters, 2015.

Arab, A., M.B. Hooten, and C.K. Wikle. "Hierarchical Spatial Models." In *Encyclopedia of Geographical Information Science*. Springer, 2015.

Bradley, J.R., N. Cressie, and T. Shi. "Comparing and selecting spatial predictors using local criteria." *Test* 24, no. 1 (2015): 1-28. DOI: 10.1007/s11749-014-0415-1.

Burden, S., N. Cressie, and D.G. Steel. "The SAR model for very large datasets: A reduced-rank approach." *Econometrics* 3, no. 2 (2015): 317-338. DOI: 10.3390/econometrics3020317.

Bradley, J. R., S. H. Holan, and C.K. Wikle. "Multivariate Spatio-Temporal Models for High-Dimensional Areal Data with Application to Longitudinal Employer-Household Dynamics." *ArXiv* (2015).

Cressie, N., and R. L. Chambers. "Comment: Spatial sampling designs depend as much on "how much?" and "why?" as on "where?"." *Bayesian Analysis* (2015).

Cressie, N., and R. L. Chambers. "Comment on Article by Ferreira and Gamerman." *Bayesian Analysis* 10, no. 3 (2015): 741-748. DOI: doi:10.1214/15-BA944B.

Cressie, N., and E.L. Kang. "Hot enough for you? A spatial exploratory and inferential analysis of North American climate-change projections." *Mathematical Geosciences* (2015). DOI: 10.1007/s11004-015-9607-9.

Cressie, N., and S. Burden. "Evaluation of diagnostics for hierarchical spatial statistical models." In *Geometry Driven Statistics*, 241-256. Chinchester: Wiley, 2015.

Cressie, N., and S. Burden. "Figures of merit for simultaneous inference and comparisons in simulation experiments." *Stat* 4, no. 1 (2015): 196-211. DOI: 10.1002/sta4.88.

Fienberg, S. E.. "Moving Toward the New World of Censuses and Large-Scale Sample Surveys: Methodological Developments and Practical Implementations." *Journal of Official Statistics* (2015).

Fosdick, B. K., M. De Yoreo, and J. P. Reiter. "Categorical data fusion using auxiliary information." *ArXiv* (2015).

Hahn, P. R., J. S. Murray, and I. Manolopoulou. "Flexible prior specification for partially identified nonlinear regression with binary responses." *Journal of the American Statistical Association* (forthcoming).

H. Shaefer, Luke. Introduction to The Survey of Income and Program Participation (SIPP). University of Michigan Preprint 1813:40169, 2015, available at http://hdl.handle.net/1813/40169.

Hu, J., J.P. Reiter, and Q. Wang. "Dirichlet Process Mixture Models for Nested Categorical Data." *ArXiv* (2015).

Hu, J.. Dirichlet Process Mixture Models for Nested Categorical Data (Ph.D. Thesis), *Statistical Science*. Duke University Ph.D., 2015, available at http://dukespace.lib.duke.edu/dspace/handle/10161/9933.

Publications Continued

Murray, J. S., and J. P. Reiter. "Multiple Imputation of Missing Categorical and Continuous Values via Bayesian Mixture Models with Local Dependence." *arXiv* (2015).

Olson, K., and J.D. Smyth. "The Effect of CATI Questionnaire Design Features on Response Timing." *Journal of Survey Statistics and Methodology* 3, no. 3 (2015): 361-396. DOI: 10.1093/jssam/smv021

Porter, A.T., S.H. Holan, and C.K. Wikle. "Bayesian Semiparametric Hierarchical Empirical Likelihood Spatial Models." *Journal of Statistical Planning and Inference* 165 (2015): 78-90. DOI: 10.1016/j.jspi.2015.04.002.

Porter, A.T., S.H. Holan, and C.K. Wikle. "Multivariate Spatial Hierarchical Bayesian Empirical Likelihood Methods for Small Area Estimation." *STAT* 4, no. 1 (2015): 108-116. DOI: 10.1002/sta4.81.

Quick, Harrison, Scott H. Holan, Christopher K. Wikle, and Jerome P. Reiter. "Bayesian Marked Point Process Modeling for Generating Fully Synthetic Public Use Data with Point-Referenced Geography." *Spatial Statistics* (2015). DOI: 10.1016/j.spasta.2015.07.008.

Ransom, T. Dynamic Models of Human Capital Accumulation (Ph.D. Thesis), *Economics*. Duke University Ph.D., 2015, available at http://dukespace.lib.duke.edu/dspace/handle/10161/9929.

Ryan, M., J.R. Bradley, T. Oswald, C.K. Wikle, and S.H. Holan. An Analysis of Bullying and Suicide in the United States using a Non-Gaussian Multivariate Spatial Model, *The National Conference On Undergraduate Research (NCUR)*. Eastern Washington University, 2015, available at https://www.cur.org/ncur_2015/account/?search=true.

Schifeling, T. S., and J. P. Reiter. "Incorporating marginal prior information into latent class models." *Bayesian Analysis* (2015). DOI: doi:10.1214/15-BA959.

Seeskin, Zachary H., and Bruce D. Spencer. Effects of Census Accuracy on Apportionment of Congress and Allocations of Federal Funds, *IPR Working Paper Series*. Northwestern University, Institute for Policy Research Working Paper WP-15-05, 2015, available at http://www.ipr.northwestern.edu/publications/papers/2015/ipr-wp-15-05.html.

Siddique, J., J. P. Reiter, A. Brincks, R. Gibbons, C. Crespi, and C. H. Brown. "Multiple imputation for harmonizing longitudinal non-commensurate measures in individual participant data meta-analysis." *Statistics in Medicine* (2015). DOI: 10.1002/sim.6562.

Spencer, Bruce D., Julian May, Steven Kenyon, and Zachary H. Seeskin. Cost-Benefit Analysis for a Quinquennial Census: The 2016 Population Census of South Africa., *IPR Working Paper Series*. Northwestern University, Institute for Policy Research Working Paper WP-15-06, 2015, available at http://www.ipr.northwestern.edu/publications/papers/2015/ipr-wp-15-06.html.

Wikle, C.K. "Hierarchical models for uncertainty quantification." In Handbook of Uncertainty Quantification." Springer, 2015.

Wikle, C.K. "Modern Perspectives on Statistics for Spatio-Temporal Data." *WIRES Computational Statistics* 7, no. 1 (2015): 86-98. DOI: 10.1002/wics.1341.

Wildhaber, M.L., C.K. Wikle, E.H. Moran, C.J. Anderson, K.J. Franz, and R. Dey. "Hierarchical, stochastic modeling across spatiotemporal scales of large river ecosystems and somatic growth in fish populations under various climate models: Missouri River sturgeon example." *Geological Society* (2015).

Publications Continued

Wildhaber, M.L., R. Dey, C.K. Wikle, C.J. Anderson, E.H. Moran, and K.J. Franz. "A stochastic bioenergetics model based approach to translating large river flow and temperature in to fish population responses: the pallid sturgeon example." *Geological Society* 408 (2015). DOI: 10.1144/SP408.10.

Wilson, C. R., and D. G. Brown. "Change in Visible Impervious Surface Area in Southeastern Michigan Before and After the "Great Recession:" Spatial Differentiation in Remotely Sensed Land-Cover Dynamics." *Population and Environment* 36, no. 3 (2015): 331-355. DOI: 10.1007/s11111-014-0219-y.

Wu, G., S.H. Holan, C.H. Nilon, and C.K. Wikle. "Bayesian Binomial Mixture Models for Estimating Abundance in Ecological Monitoring Studies." *Annals of Applied Statistics* 9 (2015): 1-26. DOI: 10.1214/14-AOAS801.

Yang, W. H., C.K. Wikle, S.H. Holan, K. Sudduth, and D.B. Meyers. "Bayesian Analysis of Spatially-Dependent Functional Responses with Spatially-Dependent Multi-Dimensional Functional Predictors." *Statistica Sinica* 25 (2015). DOI: 10.5705/ss.2013.245w .

Presentations

Scott Holan, Missouri, gave a talk on "Bayesian Lattice Filters for Time-Varying Autoregression and Time-Frequency Analysis," at the University of Illinois, Urbana-Champaign IL on April 2.

Jerry Reiter delivered the seminar, "Protecting confidentiality in an era with no privacy," in the BSSR Lecture Series at the National Institutes of Health on April 9, 2015.

On April 27, **Chris Wikle**, Missouri, gave a talk on "Predictive modeling of environmental and agricultural characteristics utilizing spatio-temporal models" at the 27th annual Conference on Applied Statistics in Agriculture, in Manhattan, KS.

He also gave a talk " Using quadratic nonlinear statistical emulators to facilitate ocean biogeochemical data assimilation." at the Midwest Mathematics and Climate Conference, University of Kansas, Lawrence, KS.

The Michigan node sponsored a one-day SIPP training workshop at Census Bureau Headquarters on May 15.

The Nebraska node team collectively presented 13 papers at the 70th Annual Meeting of the American Association for Public Opinion Research in Hollywood, FL, May 14-17. These presentations include:

- **Arunhachalam, H., G. Atkin, A. Eck, D. Wettlaufer, L.-K. Soh,** and **R. F. Belli**. "I Know What You Did Next: Predicting Respondent's Next Activity Using Machine Learning"
- **Atkin, G., H. Arunachalam, A. Eck, D.Wettlaufer, L.-K. Soh,** and **R. F. Belli**. "Using Machine Learning Techniques to Predict Respondent Type from A Priori Demographic Information"
- **Belli, R. F., L. D. Miller, L.-K. Soh**, and **T. Al Baghal**. "Using Data Mining to Examine Interviewer-Respondent Interactions in Calendar Interviews"
- **Cordova Cazar, A. L.,** and **Belli, R. F.** "The Use of Paradata to Evaluate Interview Complexity and Data Quality (in Calendar and Time Diary Surveys)."
- **Deal, C. E., Kirchner, A., Cordova Cazar, A. L., Ellyne, L.,** and **Belli, R. F.** "Changing 'Who' or 'Where': Implications for Data Quality in the American Time Use Survey"
- **Kirchner, Antje** and **Kristen Olson**. "Effects of interviewer and respondent behavior on data quality: An investigation of question types and interviewer learning."
- **McCutcheon, Allan L.** "Survey Informatics: The Future of Survey Methodology and Survey Statistics Training in the Academy?"
- **McCutcheon, Allan L.** "The Role of Device Type and Respondent Characteristics in Internet Panel Survey Breakoff."

Presentations (Continued)

- **Olson, K.** and **Smyth, J. D.** "Why Do Interviewers Speed Up? An Examination of Changes in Interviewer Behaviors over the Course of the Survey Field Period."
- **Smyth, J. D.** and **Olson, K.** "Recording What the Respondent Says: Does Question Format Matter?"
- **Soh, L.-K., A. Eck,** and **A. L. McCutcheon**. "Predicting Breakoff Using Sequential Machine Learning Methods"
- **Wang, Mengyang, Allan L. McCutcheon,** and **Laura Allen**. "Grids and Online Panels: A Comparison of Device Type from a Survey Quality Perspective."
- **Wettlaufer, D., H. Arunachalam, G. Atkin, A. Eck, L.-K. Soh,** and **R. F. Belli**. "Determining Potential for Breakoff in Time Diary Survey Using Paradata."

On May 26, **Scott Holan**, Missouri, presented "Time-Frequency Estimation and Time-Frequency Functional Data Models," at the Indo-US Workshop on Time Series Analysis, in Pune, India.

Chris Wikle and **Noel Cressie,** Missouri node, were involved with the plenary presentation on Environmental Informatics at the Southern Regional Council on Statistics Summer Research Conference, Carolina Beach on June 7.

Scott Holan, Missouri, "Bayesian Analysis of Spatially-Dependent Functional Responses with Spatially-Dependent Multi-Dimensional Functional Predictors," International Chinese Statistical Association (ICSA) Applied Symposium/ Graybill Conference, Fort Collins, CO on June 15.

Dr. Antje Kirchner and **Dr. Kristen Olson**, University of Nebraska, gave a presentation "Effects of interviewer and respondent behavior on data quality: An investigation of question types and interviewer learning" at the European Survey Research Association (ESRA)Annual Meeting July 14, 2015. **Kirchner** also gave a talk entitled "Using machine learning to correct for survey nonresponse bias" at the same conference on July 16, 2015.

Margaret Levenstein of the University of Michigan presented "Harnessing Data In the Wild" reporting on the Michigan node's use of social media to track labor markets at the ESRA n Reykjavik, July 17, 2015.

A session on "Synthetic establishment microdata around the world" was organized at the International Statistical Institute's 60th World Statistics Congress in Rio de Janeiro, by Lars Vilhuber (Cornell). Amongst the presenters were: **John M. Abowd** (NCRN-Cornell) and Kevin L. McKinney (U.S. Census Bureau), "Noise Infusion as a Confidentiality Protection Measure for Graph-based Statistics" ; **Lars Vilhuber** (NCRN-Cornell) and Javier Miranda (U.S. Census Bureau), "Using partially synthetic data to replace suppression in the Business Dynamics Statistics"; **Satkartar Kinney** (NCRN-NISS), **Jerry Reiter** (NCRN-Duke), and Javier Miranda (U.S. Census Bureau), "Improving the Synthetic Longitudinal Business Database: Synthesizing Firms". The meeting was held July 26-31.

Noel Cressie (Missouri Node) also spoke at the World Statistical Congress. His presentation was entitled "Spatio-Temporal Data Fusion for Big Data and its Application to Satellite Remote Sensing". **Mauricio Sadinle**, formerly with CMU, now at Duke/NISS, presented a poster "Bayesian Estimation of Bipartite Matchings for Record Linkage" at the World Statistical Congress as well. And Larry Cox (Duke/NISS) taught a short course on "Data Anonymization Balancing Data Confidentiality and Data Quality."

Presentations (Continued)

JSM 2015

JSM 2015 was held in Seattle, Washington August 9 - 12 and several of the research nodes made presentations. See box for full details of the talks.

Sunday, August 9:

Advancing Research in Total Survey Error: Establishing New Links Between Multiple Sources of Survey Error **— Topic Contributed Papers**

Chair: **Brady T. West, University of Michigan Institute for Social Research;** Organizer: **Brady T. West, University of Michigan Institute for Social Research**

Speakers:

- "Do Interviewers with High Cooperation Rates Behave Differently? Interviewer Cooperation Rates and Interview Behaviors", **Kristen Olson, University of Nebraska - Lincoln; Jolene D. Smyth, University of Nebraska - Lincoln; Antje D. Kirchner, University of Nebraska - Lincoln**
- "Decomposing Mobile Versus PC Web Mode Effects in a Probability Web Panel", Christopher Antoun
- "The Effects of Nonresponse Error and Measurement Error on Estimates of Regression Coefficients", Antje Kirchner; Barbara Felderer, University of Mannheim
- "To Allow or Disallow Smartphone Participation in Web Surveys: Choosing Between the Potential for Coverage and Measurement Error", Gregg Peterson, Institute for Social Research; Jamie Griffin, Institute for Social Research; John LaFrance, Market Strategies International; JiaoJiao Li, Market Strategies International

Discussant: Jill Dever, RTI International

Monday, August 10:

Bayesian Approaches to Record Linkage — Invited Papers

Chair: **Stephen Fienberg, Carnegie Mellon University;** Organizer: **Mauricio Sadinle, Carnegie Mellon University**

Speakers:

- "Methods for Quantifying Conflict Casualties in Syria", **Rebecca Steorts, Carnegie Mellon University; Samuel Ventura, Carnegie Mellon University; Mauricio Sadinle, Carnegie Mellon University; Stephen Fienberg, Carnegie Mellon University**;
- "Adjusting for Errors in Blocking Variables in Record Linkage", **Nicole Dalzell, Duke University; Jerry Reiter, Duke University**;
- "Joint modeling for record linkage and statistical analysis", Michael Larsen, George Washington University;
- "Simultaneously Propagating Uncertainty and Providing Inference on Linked Data via Bayesian Graphical Record Linkage", Andrea Tancredi, Sapienza University of Rome; Rebecca Steorts, Carnegie Mellon University; Brunero Liseo, Sapienza University of Rome

Discussant: William E Winkler, US Census Bureau

Has Informed Consent Outrun Its Usefulness? — Invited Panel

Committee on Privacy and Confidentiality, Section on Statistical Consulting, Scientific and Public Affairs Advisory Committee

Chair and Organizer: **Alan F. Karr, RTI International**

Has Informed Consent Outrun Its Usefulness?

- Franke Kreuter, Joint Program in Survey Methdology

Presentations (Continued)

- Helen Nissenbaum, New York University
- Julia Lane, American Institutes for Research
- Marjory Blumenthal, Office of Science and Technology Policy

Tuesday, August 11:

Protecting Privacy While Maximizing the Utility of Government Data: Evaluating Current Approaches and Exploring New Alternatives — Topic Contributed Papers

Government Statistics Section, Section on Statistical Consulting, Committee on Applied Statisticians, Scientific and Public Affairs Advisory Committee
Organizer(s): Michael B. Hawes, U.S. Department of Education
Chair(s): Michael B. Hawes, U.S. Department of Education

- Standards and Guidelines for Data Disclosure Control: A Review — Thomas Krenzke, Westat ; Shep Roey, Westat ; Lin Li, Westat ; Jane Li, Westat
- Quality Assessment of the National Survey on Drug Use and Health (NSDUH) Public Use Files — Neeraja Sathe, RTI International ; Feng Yu, RTI International ; Lanting Dai, RTI International ; Jonaki Bose, Substance Abuse and Mental Health Services Administration ; Art Hughes, Substance Abuse and Mental Health Services Administration
- Assessing the Utility of Top-Coded Consumer Expenditure Survey Data — Daniel Yang, Bureau of Labor Statistics ; Daniell Toth, Bureau of Labor Statistics
- Developing and Testing the Microdata Analysis System — Michael Freiman, U.S. Census Bureau ; Amy Lauger, U.S. Census Bureau ; Marlow Lemons, U.S. Census Bureau ; Bryan Schar, U.S. Census Bureau
- Synthetic Longitudinal Business Databases for International Comparisons — Joerg Drechsler, Institute for Employment Research ; **Lars Vilhuber, Cornell University**

Sacks Award (Continued)

from page 4

the American Statistical Association, the American Academy of Political and Social Science, and a member of the U.S. National Academy of Sciences.

The Jerome Sacks Award for Cross Disciplinary Research was created in 2001 in honor of Jerome (Jerry) Sacks, the founding director of NISS. As the Sacks award recipient, Fienberg receives $1,000, and his name is added to a plaque at NISS that lists all recipients of the award.

The NCRN Newsletter is published quarterly by the NCRN Coordinating Office. You may reach us at:
275 Ives Hall,
Cornell University, Ithaca, NY 14853
(607) 255-2744
info@ncrn.info
www.ncrn.info
Jamie Nunnelly - Managing Editor