



NSF-Census Research Network Newsletter

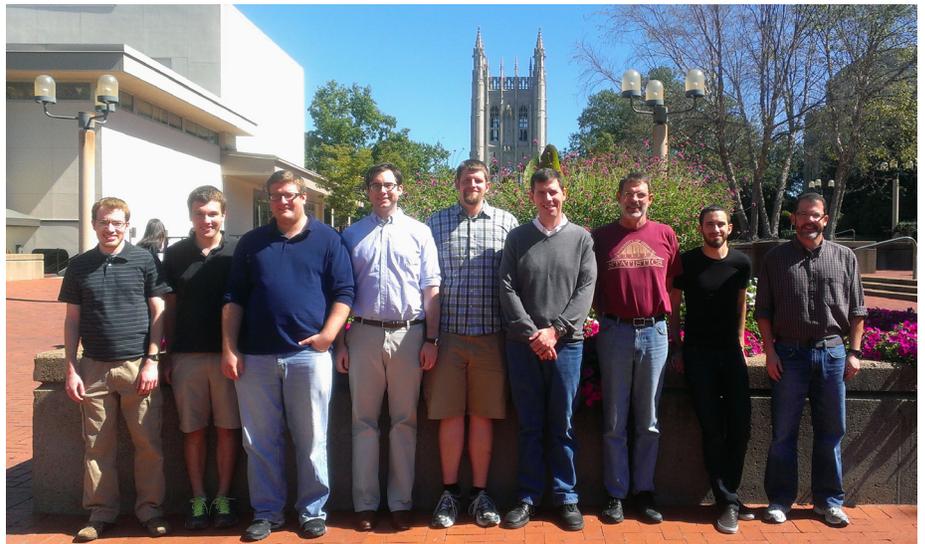
Vol. 1, Issue 1

NCRN Focus on University of Missouri Node

Scott Holan, Chris Wikle, and Noel Cressie lead the University of Missouri research node for NCRN; Holan is the Principal Investigator. A central theme of the node has been research on spatial statistical methodology for use with the American Community Survey (ACS), an ongoing survey that replaced the Decennial Census long form. The ACS collects information on many different aspects of individual and household demographics, including race, gender, age, family, income and benefits, education, veteran status, disabilities, commuting and much more. The ACS is used by many other federal, state and local agencies to determine funding levels for programs and establish priorities based on communities' needs. It is also used by non-government organizations to generate reports, determine whether a community is a good location to establish a new business, and identify issues that are important to a particular community.

The ACS produces one-year, three-year and five-year period estimates, depending on the resolution of the geography. One-year estimates are generated for areas with over 65,000 people, whereas three-year estimates are generated for areas with populations between 20,000 and 65,000. Finally, five-year estimates are generated for all geographies.

The Missouri node is developing spatial and spatio-temporal models for use in federal surveys. One component of this research consists of developing spatio-temporal models that allow one to compare trends across different geographies that may be measured on different scales. "For example, when looking at estimation of an undersampled, or non-sampled small area, why not look at the neighboring areas and/or times? In other words, by looking at the neighboring areas, one can potentially obtain information that may predict what's happening in the area of interest. Similarly one might consider last month's estimate to predict what might happen in this month for something like unemployment levels," said Scott Holan. This is a method called "borrowing strength." That is, areas that are nearby



Left-Right: Christopher Hassett, Patrick McDermott, Jonathan Bradley, Aaron Porter, Trevor Oswald, Scott Holan, Noel Cressie, Harrison Quick, and Christopher Wikle - (not shown Guohui Wu).

one another in space and/or time may be more alike than areas that are far apart. So, one utilizes this dependence structure to improve the precision in estimation of quantities of interest.

Holan noted that their node's focus on the ACS was due to its importance to the U.S. and the potential benefits that could be achieved through the use of spatial and spatio-temporal methodology. The Census Bureau has a rich history of research in time series methodology. However, there has been less focus on spatial and spatio-temporal methodology. "The ACS has a natural spatial structure," noted Holan, "and we believe that anything one can do to improve it is important." The models that node members are developing are of independent interest and will translate over to other federal statistical survey work.

Another area of study that the node is looking at is whether using social media and crowd sourcing with aggregate data may also help to reduce uncertainty in the estimates. This is an exciting area and will hopefully help federal statistical agencies to use big data in a new way.

In addition to the investigators, the group now has three postdoctoral fellows, along with three graduate research assistants working on NCRN-related projects. To date,

(Continued on page 2)

NCRN Calendar of Events

Feb 28 TCRN/SIPP workshop

The Triangle Census Research Network (TCRN) and the Duke Initiative in Survey Methodology at the Social Science Research Institute at Duke University, in collaboration with the Institute for Social Research at the University of Michigan, is holding a three-day workshop February 28 - March 2, 2014, in Durham NC. Funding for the workshop is provided via grant No. SES 1131897 from the National Science Foundation and the U.S. Bureau of the Census to the TCRN.

The workshop will introduce participants to the use of micro-data from the Survey of Income and Program Participation (SIPP) and provide the necessary foundations for participants to conduct their own SIPP-based research project. The SIPP collects longitudinal subannual data on respondents' income, labor force activity, household composition, health, migration, and eligibility for and participation in programs (e.g. TANF, WIC, Medicare, Medicaid, and numerous others). As such, it provides unique opportunities to examine the social and economic well-being of U.S. residents, and changes in residents' experiences over time.

H. Luke Shaefer, Assistant Professor of Social Work at the University of Michigan and National Poverty Center (NPC) Research Affiliate, will lead the workshop in collaboration with researchers at Duke University, the U.S. Census Bureau and other nationally recognized SIPP scholars. The Triangle Census Research Network will pay travel, lodging, and meal costs for a limited number of participants.

May 22-23, 2014

NCRN Spring Meeting, at U.S. Census Bureau Headquarters in Washington DC
See related article on [page 4](#) for details about the meeting.

September 11-12, 2014

NCRN Fall Meeting in New York City
Planning has started for the Fall 2014 meeting in NYC.

[Click here](#) to go to the calendar of events.

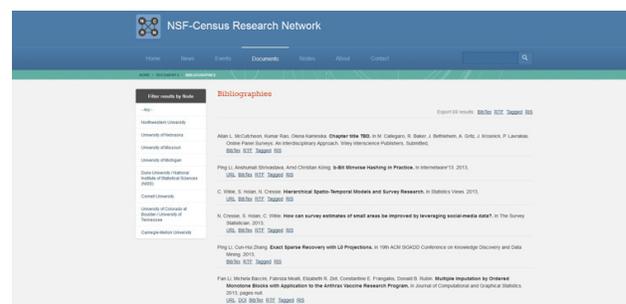
Website Features NCRN-Related Publications

The NCRN Coordinating Office (NCRN-CO) added a new feature to the NCRN-CO website. A comprehensive bibliography of all the publications that were developed as part of the NCRN grant is now available on the website at ncrn.info/documents/bibliographies.

One can view the comprehensive list, or filter the list to look at each research node. The list can be downloaded in BibTex format, along with other formats including RTF, Tagged and RIS. When available, a URL link points to the online version of the article.

Listed separately are working papers and preprints hosted on the NCRN-CO eCommons, at ncrn.info/documents/working-papers.

The lists will be updated on a regular basis. Please do submit new items as soon as you learn about them to nunnelly@niss.org. For publications, bibliographic database extracts are preferred (BibTex is much preferred). For preprints and working papers, the actual paper should be sent to nunnelly@niss.org, or uploaded directly to eCommons.



Missouri Node-Cont.

(from page 1)

the node has had four undergraduate research assistants, three of whom are now pursuing graduate studies in statistics. Education is a large component of the node and, through education and training activities, the node expects to equip some of the next generation of statisticians with the skills to work on spatial and spatio-temporal problems in federal statistics.

Past NCRN-Related Events

Summer at Census Talk

Matthew Shapiro and **Margaret Levenstein** from University of Michigan spoke on July 25 at the Summer at Census. Their talk was “CenHRS: Integrating HRS with Census Employer Data to Enhance Understanding of the Employment Prospects of Older Americans.”

This project is aimed to advance our understanding of the employment prospects and conditions of older Americans. The project will use Health and Retirement Study (HRS) data in combination with Census’s extensive data on employers and their workforces to develop new methods and data products. The primary goal of CenHRS is to identify employer and workforce characteristics important to the health and retirement behavior of older Americans. The project is supported by a grant from the Sloan Foundation.

HRS surveys more than 26,000 Americans over the age of 50 every two years. Sponsored by the National Institute of Aging and the Social Security Administration, it is a large-scale longitudinal project that studies the labor force participation and health transitions that individuals undergo toward the end of their work lives and in the years that follow. Since its launch in 1992, the HRS has collected extensive survey information on income and wealth, physical and mental health, employment and other activities, and linked these to individual-level administrative data from Social Security and Medicare. Over 1,000 scholarly articles have been published using HRS data; HRS has more than 10,000 registered users. Surveys modeled on HRS are currently underway in 25 countries on five continents.

This project will use the Business Register and Longitudinal Business Database, the Economic Census and related firm and establishment surveys, and the Longitudinal Employer Household Database to create contextual variables relating to a broad range of characteristics of firms at which HRS respondents worked during their lifetime. These measures will incorporate information on employee turnover and tenure; earnings distributions, age and other demographic characteristics of the firm’s workforce; business characteristics such as firm and establishment size, productivity, and export activity; and measures based on geography such as distance to work. One outcome will be a new dataset, CenHRS, which will include aggregate measures of the characteristics of the current and prior employers and co-workers of HRS respondents. Prototypes of the CenHRS dataset will be analyzed to determine which of these characteristics are of particular salience and how best to measure them. These measures will expand the research questions that can be addressed with HRS and provide the basis for new Census statistics on the characteristics of older workers and their employers. The project will also analyze differences in measures of administrative and respondent data for variables such as earnings, employer

size, and labor force attachment, and compare those differences to those for Census surveys linked to administrative data such as SIPP and ACS.

The talk described and reported the results from joint research with Kristin McCue (U.S. Census Bureau), **John Abowd** (Cornell University), Nada Wasi (University of Michigan), and David Weir (University of Michigan).

Summer at Census Talk II

Ashwin Machanavajjhala from Duke University also presented a talk for the Summer at Census series. His talk was “Tuning Privacy-Utility Tradeoffs in Statistical Databases Using Policies.”

Privacy definitions provide ways for trading off the privacy of individuals in a statistical database for the utility of downstream analysis of the data. One notion that has become the gold-standard is differential privacy, and it exposes one knob ϵ for tuning this trade-off. While ϵ is intuitive, recent work has shown it does not sufficiently capture the diversity in the privacy utility trade-off space - it provides insufficient utility in some applications and insufficient privacy when data are correlated.

Machanavajjhala presented Blowfish, a class of privacy definitions that provides a richer interface for trading-off privacy for utility. In particular, we allow data publishers to extend differential privacy using a policy, which specifies which information must be kept secret, and what constraints may be known about the data. While the former allows increased utility by not protecting certain properties about individuals, the latter provides added protection against adversary who know correlations in the data (arising from constraints). He formalized privacy policies and presented novel algorithms that explore new points in the privacy-utility trade-off space.

2013 Federal Committee on Statistical Methodology (FCSM) Research Conference

This three-day conference, hosted by The Council of Professional Associations on Federal Statistics (COPAFS) had several NCRN-related presentations including:

- “Boosting Algorithms for Edit and Imputation of Multiple-response Variables”, **Ping Li** (Cornell University), **John Abowd** (Cornell University)
- “Managing Confidentiality and Provenance Across Mixed Private and Publicly-Accessed Data and Metadata” **Carl Lagoze** (Michigan), **William Block** (Cornell Institute of Social and Economic Research), **Jeremy Williams** (Cornell Institute of Social and Economic Research), **John Abowd** (Cornell), **Lars Villhuber** (Cornell)

(Continued on page 4)

Past NCRN-Related Events (Continued)

- “Statistical Disclosure Limitation and Edit-Imputation” **Alan Karr** (NISS), **Hang Joon Kim** (Duke and NISS), **Jerome Reiter** (Duke), **Larry Cox** (NISS). This session was organized by Karr and Reiter.
- “Simultaneous Edit-Imputation for Continuous Microdata” **Hang Joon Kim** (Duke and NISS), **Jerome Reiter** (Duke), **Alan Karr** (NISS), **Larry Cox** (NISS)
- “Simultaneous Edit-Imputation for Categorical Microdata” **Daniel Manrique-Vallier** (Indiana), **Jerome Reiter** (Duke)
- “State Based Intraclass Correlation Values for Planning Group-Randomized Trials in Education: Within and Between District Estimates” **Eric Hedberg** (NORC), **Larry Hedges** (Northwestern)

Other Presentations

Bruce Spencer and **Zach Seeskin** from Northwestern University delivered a talk: “Overview of the NCRN Project - Research on Cost-Benefit Analysis for the 2020 Census.”

John Abowd from Cornell University spoke on September 24 at the Harvard Law School. He was part of the lighting talks of a session on Integrating Approaches to Privacy Across Research Lifecycle.

Steve Fienberg, Carnegie Mellon University, presented “Developing Confidentiality and Privacy Protection Methods That Scale” at the Statistics2013 Capstone event in London, England on November 12.

Carl Lagoze, University of Michigan and a member of the Cornell node, gave a talk on “Encoding Provenance Metadata for Social Science Datasets” at the 7th Metadata and Semantics Research Conference in Greece on November 19.

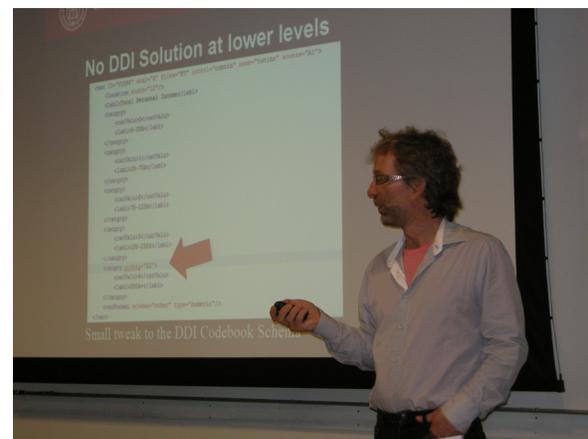
NCRN Spring Meeting at US Census Bureau

When the U.S. Census Bureau closed during the 2013 government shutdown, the [technical mini-symposium](#) jointly planned by the NCRN and the Census Bureau had to be cancelled at short notice.

NCRN is now planning on holding a technical mini symposium at the U.S. Census Bureau headquarters in Washington DC May 22-23. However, the program has not been finalized. Current plans are for two one-half day sessions focused on topics of broad interest to the nodes, the Census Bureau and the federal statistical system. Presentations will highlight research performed by the NCRN nodes, with opportunity for questions and comments from agency researchers.

The meeting will be free and open to the public. The meeting is an excellent opportunity for interested researchers and practitioners from the federal statistical system in Washington DC to learn about the NCRN.

Photos from May 2013



Lars Vilhuber (Cornell).



Hang Kim (NISS and Duke).

NCRN-Related Publications

- McCutcheon, Allan L., Kumar Rao, and Olena Kaminska. "Chapter title TBD." In *Online Panel Surveys: An Interdisciplinary Approach*, edited by M. Callegaro, R. Baker, J. Bethlehem, A. Gritz, J. Krosnick and P. Lavrakas. Wiley Interscience Publishers, Submitted.
- Hu, J., R. Mitra, and J.P. Reiter. "Are independent parameter draws necessary for multiple imputation?" *The American Statistician* 67 (2013): 143-149.
- Banerjee, A., J. Murray, and D. B. Dunson. "Bayesian learning of joint distributions of objects." In *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics (AISTATS) 2013*, to appear., 2013.
- Li, Ping, Anshumali Shrivastava, and Arnd Christian König. "b-Bit Minwise Hashing in Practice." In *Internetware'13.*, 2013.
- Shrivastava, Anshumali, and Ping Li. "Beyond Pairwise: Provably Fast Algorithms for Approximate K-Way Similarity Search." In *Neural Information Processing Systems (NIPS).*, 2013.
- Spielman, S.E., and Patrick Harrison. "The Co-Evolution of Residential Segregation and the Built Environment at the Turn of the 20th Century: A Schelling Model." In *Transactions in GIS.*, 2013.
- Lagoze, Carl, William C. Block, Jeremy Williams, John M. Abowd, and Lars Vilhuber. "Data Management of Confidential Data." *International Journal of Digital Curation* 8 (2013): 265-278.
- H. Shaefer, Luke, Xiaoqing Song, and Trina R. William Shanks. "Do single mothers in the United States use the Earned Income Tax Credit to reduce unsecured debt?" *Review of Economics of the Household* (2013): 659-680.
- , S.H. Holan, and M. Wildhaber. "Ecological Prediction with Nonlinear Multivariate Time-Frequency Functional Data Models." *Journal of Agricultural, Biological, and Environmental Statistics* 18 (2013): 450-474.
- Lagoze, Carl, Jeremy Williams, and Lars Vilhuber. "Encoding Provenance Metadata for Social Science Datasets." In *Metadata and Semantics Research*, edited by Emmanouel Garoufallou and Jane Greenberg, 123-134. Vol. 390. *Communications in Computer and Information Science* 390. Springer International Publishing, 2013.
- Lagoze, Carl, William C. Block, Jeremy Williams, and Lars Vilhuber. "Encoding Provenance of Social Science Data: Integrating PROV with DDI." In *5th Annual European DDI User Conference.*, 2013.
- Steorts, Rebecca C., and Malay Ghosh. "On estimation of mean squared errors of benchmarked and empirical bayes estimators." *Statistica Sinica* 23 (2013): 749-767.
- Li, Ping, and Cun-Hui Zhang. "Exact Sparse Recovery with L0 Projections." In *19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.*, 2013.
- Wu, G., S.H. Holan, and C.K. Wikle. "Hierarchical Bayesian Spatio-Temporal Conway-Maxwell Poisson Models with Dynamic Dispersion." *Journal of Agricultural, Biological, and Environmental Statistics* 18 (2013): 335-356.
- Wikle, C., S. Holan, and N. Cressie. Hierarchical Spatio-Temporal Models and Survey Research In *Statistics Views.*, 2013.
- Sengupta, A., and N. Cressie. "Hierarchical Statistical Modeling of Big Spatial Datasets Using the Exponential Family of Distributions." *Spatial Statistics* 4 (2013): 14-44.
- Cressie, N., S. Holan, and C. Wikle. How can survey estimates of small areas be improved by leveraging social-media data? In *The Survey Statistician.*, 2013.
- Spielman, S.E., and J. Logan. "Identifying Neighborhoods Using High Resolution Population Data." *Annals of the Association of American Geographers* 103 (2013): 67-84.
- Brandimarte, Laura, Alessandro Acquisti, and George Loewenstein. "Misplaced confidences: Privacy and the control paradox." *Social Psychological and Personality Science* 4 (2013): 340-347.
- Li, Fan, Michela Baccini, Fabrizia Mealli, Elizabeth R. Zell, Constantine E. Frangakis, and Donald B. Rubin. "Multiple Imputation by Ordered Monotone Blocks with Application to the Anthrax Vaccine Research Program." *Journal of Computational and Graphical Statistics* (2013): null.
- Si, Y., and J.P. Reiter. "Nonparametric Bayesian multiple imputation for incomplete categorical variables in large-scale assessment surveys." *Journal of Educational and Behavioral Statistics* 38 (2013): 499-521.
- Antenucci, Dolan, Erdong Li, Shaobo Liu, Bochun Zhang, Michael J. Cafarella, and Christopher Ré. "Ringtail: a generalized nowcasting system." *WebDB* 6 (2013): 1358-1361.

Antenucci, Dolan, Michael J. Cafarella, Margaret C. Levenstein, Christopher Ré, and Matthew Shapiro. "Ringtail: Feature Selection for Easier Nowcasting." *WebDB* (2013): 49-54.

Holan, S., and C. Wikle. "Spatio-temporal Design: Advances in Efficient Data Acquisition." In *Spatio-temporal Design: Advances in Efficient Data Acquisition*, edited by Jorge Mateu and Werner Muller, 269-284. Wiley, 2013.

Wilson, C. R. Using Satellite Imagery to Evaluate and Analyze Socioeconomic Changes Observed with Census Data., 2013.

Nardi, Yuval, Robert Hall, and Stephen E. Fienberg. "Achieving both valid and secure logistic regression analysis on aggregated data from different private sources." *Journal of Privacy and Confidentiality* 4 (2012): 189.

Holan, S., W. Yang, D. Matteson, and C.K. Wikle. "An Approach for Identifying and Predicting Economic Recessions in Real-Time Using Time-Frequency Functional Models." *Applied Stochastic Models in Business and Industry* 28 (2012): 485-499.

McElroy, T., and S. Holan. Asymptotic Theory of Cepstral Random Fields. University of Missouri, 2012.

Wang, J., and S. Holan. "Bayesian Multi-Regime Smooth Transition Regression with Ordered Categorical Variables." *Computational Statistics and Data Analysis* 56 (2012): 4165-4179.

Charest, Anne-Sophie. Confidentiality and Privacy Protection in a Non-US Census Context. Carnegie Mellon University, 2012.

Fienberg, Stephen E. "Counting the people." In Nathan and Beatrice Keyfitz Lecture in Mathematics and the Social Sciences. Toronto, Canada: Fields Institute, 2012.

Charest, Anne-Sophie. Creation and Analysis of Differentially-Private Synthesis Datasets. Carnegie Mellon University, 2012.

Xiaolin, Yang, Stephen E. Fienberg, and Alessandro Rinaldo. "Differential Privacy for Protecting Multi-dimensional Contingency Table Data: Extensions and Applications." *Journal of Privacy and Confidentiality* 4 (2012): 101-125.

Charest, Anne-Sophie. "Differential Privacy for Synthetic Datasets." In *Proceedings of the Survey Research Section of the SSC*. Guelph, Ontario, 2012.

Ventura, Samuel, Rebecca Nugent, and Erich R. H. Fuchs. "Disambiguating USPTO Inventors with Classification Models Trained on Comparisons of Labeled Inventor Records." In Conference Presentation Classification Society Annual Meeting, Carnegie Mellon University., 2012.

Brandimarte, Laura, and Alessandro Acquisti. "The Economics of Privacy." In *The Oxford Handbook of the Digital Economy*, edited by Martin Peitz and Joel Waldfogel, 547-570. Oxford University Press, 2012.

Charest, Anne-Sophie. "Empirical Evaluation of Statistical Inference from Differentially-Private Contingency Tables." In *Privacy in Statistical Databases*, edited by Josep Domingo-Ferrer and Ilenia Tinnirello, 257-272. Vol. 7556. Springer, 2012.

Li, Ping, and Cun-Hui Zhang. "Entropy Estimations Using Correlated Symmetric Stable Random Projections." In *Advances in Neural Information Processing Systems* 25, edited by P. Bartlett, F.C.N. Pereira, C.J.C. Burges, L. Bottou and K.Q. Weinberger, 3185-3193., 2012.

Manrique-Vallier, D., and J.P. Reiter. "Estimating identification disclosure risk using mixed membership models." *Journal of the American Statistical Association* 107 (2012): 1385-1394.

Steorts, Rebecca C., and Malay Ghosh. "On Estimation of Mean Squared Errors of Benchmarked and Empirical Bayes Estimators." In *2012 Joint Statistical Meetings*. San Diego, CA, 2012.

H. Shaefer, Luke, and Kathryn Edin. Extreme Poverty in the United States, 1996 to 2011. University of Michigan, 2012.

Sun, Xu, Anshumali Shrivastava, and Ping Li. "Fast Multi-task Learning for Query Spelling Correction." In *The 21st ACM International Conference on Information and Knowledge Management (CIKM 2012)*, 285-294., 2012.

Shrivastava, Anshumali, and Ping Li. "Fast Near Neighbor Search in High-Dimensional Binary Data." In *The European Conference on Machine Learning (ECML 2012)*., 2012.

Sadinle, Mauricio. "A Generalized Fellegi-Sunter Framework for Multiple Record Linkage with Application to Homicide Records Systems." *Journal of the American Statistical Association* 108 (2012): 385-397.

Li, Ping, Anshumali Shrivastava, and Arnd Christian König. "GPU-based minwise hashing: GPU-based minwise hashing." In *Proceedings of the 21st World Wide Web Conference (WWW 2012) (Companion Volume)*, 565-566., 2012.

Sengupta, A., and N. Cressie. Hierarchical Statistical Modeling of Big Spatial Datasets Using the Exponential Family of Distributions. The Ohio State University, 2012.

Reiter, J.P., and S.K. Kinney. "Inferentially valid partially synthetic data: Generating from posterior predictive distributions not necessary." *Journal of Official Statistics* 28 (2012): 583-590.

- Sadinle, Mauricio. "Logit-Based Confidence Intervals for Single Capture-Recapture Estimation." In *American Statistical Association Pittsburgh Chapter Banquet*. Pittsburgh, PA, 2012.
- Shalizi, Cosma, and Rebecca Nugent. "Maintaining Quality in the Face of Rapid Program Expansion." In *2012 Joint Statistical Meetings*. San Diego, CA, 2012.
- Ventura, Samuel, Rebecca Nugent, and Erich R. H. Fuchs. "Methods Matter: Revamping Inventor Disambiguation Algorithms with Classification Models and Labeled Inventor Records." In *Conference Presentation Academy of Management Annual Meeting*. Boston, MA, 2012.
- Sadinle, Mauricio. "MulFiles Record Linkage Using a Generalized Fellegi-Sunter Framework." In *Conference Presentation Classification Society Annual Meeting*, Carnegie Mellon University, 2012.
- Li, Ping, Art Owen, and Cun-Hui Zhang. "One Permutation Hashing." In *Advances in Neural Information Processing Systems 25*, edited by P. Bartlett, F.C.N. Pereira, C.J.C. Burges, L. Bottou and K.Q. Weinberger, 3122-3130., 2012.
- Fienberg, Stephen E. "Privacy in a world of electronic data: Whom should you trust?" *Notices of the AMS* 59 (2012): 479.
- Fienberg, Stephen E., and Jiashun Jin. "Privacy-preserving data sharing in high dimensional regression and classification settings." *Journal of Privacy and Confidentiality* 4 (2012): 221.
- Abowd, John M., Lars Vilhuber, and William Block. "A Proposed Solution to the Archiving and Curation of Confidential Scientific Inputs." In *Privacy in Statistical Databases*, edited by Josep Domingo-Ferrer and Ilenia Tinnirello, 216-225. Vol. 7556. Lecture Notes in Computer Science 7556. Springer Berlin Heidelberg, 2012.
- Sun, Xu, Anshumali Shrivastava, and Ping Li. "Query spelling correction using multi-task learning." In *Proceedings of the 21st World Wide Web Conference (WWW 2012)(Companion Volume)*, 613-614., 2012.
- Holan, S., W. Yang, D. Matteson, and C. Wikle. "Rejoinder: An approach for identifying and predicting economic recessions in real time using time frequency functional models." *Applied Stochastic Models in Business and Industry* 28 (2012): 504-505.
- Holan, S., and C.K. Wikle. "Semiparametric Dynamic Design of Monitoring Networks for Non-Gaussian Spatio-Temporal Data." In *Spatio-temporal Design: Advances in Efficient Data Acquisition*, edited by Jorge Mateu and Werner Muller, 269-284. Wiley, 2012.
- Adjerid, Idris, Alessandro Acquisti, and Laura Brandimarte. "Sleight of Privacy." In *Conference on Web Privacy Measurement.*, 2012.
- Kurtz, Zachary. Smooth Post-Stratification in Multiple Capture Recapture. Carnegie Mellon University, 2012.
- Fienberg, Stephen E. "Statistics in Service to the Nation." In *Presentation Samuel S. Wilks Lecture*. Princeton, NJ, 2012.
- Fienberg, Stephen E. "Teaching about Big Data: Curricular Issues." In *2012 Joint Statistical Meetings*. San Diego, CA, 2012.
- Srivastava, Radheshushka, Ping Li, and Debasis Sengupta. "Testing for Membership to the IFRA and the NBU Classes of Distributions." *Journal of Machine Learning Research - Proceedings Track for the Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS 2012)* 22 (2012): 1099-1107.
- Spielman, Seth, David Folch, John Logan, and Nicholas Nagle. "Thinking inside the box: Mapping the microstructure of urban eEnvironment (and why it matters)." In *AutoCarto 2012*. Columbus, Ohio, 2012.
- Hall, Robert, and Stephen E. Fienberg. "Valid Statistical Inference on Automatically Matched Files." In *Privacy in Statistical Databases*, edited by Josep Domingo-Ferrer and Ilenia Tinnirello, 131-142. Springer, 2012.
- H. Shaefer, Luke, and Marci Ybarra. "The welfare reforms of the 1990s and the stratification of material well-being among low-income households with children." *Children and Youth Services Review* 34 (2012): 1810-1817.
- Fienberg, Stephen E. "Comment on Gates: Toward a Reconceptualization of Confidentiality Protection in the Context of Linkages with Administrative Records." *Journal of Privacy and Confidentiality* 3 (2011): 65.
- Hall, Robert, Stephen E. Fienberg, and Yuval Nardi. "Secure multiparty linear regression based on homomorphic encryption." *Journal of Official Statistics* 27 (2011): 669.
- Manrique-Vallier, D., and J.P. Reiter. "Bayesian estimation of discrete multivariate latent structure models with structural zeros." *Journal of Computational and Graphical Statistics*: to appear.