Effects of a Government-Academic Partnership:

Has the NSF-Census Bureau Research Network Helped

Improve the U.S. Statistical System?

by

Daniel H. Weinberg, DHW Consulting and U.S. Census Bureau (retired)
John M. Abowd, Cornell University and U.S. Census Bureau
Robert F. Belli, University of Nebraska
Noel Cressie, University of Wollongong and University of Missouri
David C. Folch, Florida State University
Scott H. Holan, University of Missouri and U.S. Census Bureau
Margaret C. Levenstein, University of Michigan
Kristen M. Olson, University of Nebraska and U.S. Census Bureau
Jerome P. Reiter, Duke University and U.S. Census Bureau
Matthew D. Shapiro, University of Michigan
Jolene Smyth, University of Nebraska
Leen-Kiat Soh, University of Nebraska
Bruce D. Spencer, Northwestern University
Seth E. Spielman, University of Colorado
Lars Vilhuber, Cornell University and U.S. Census Bureau
Christopher K. Wikle, University of Missouri

11 September 2018

ABSTRACT

The National Science Foundation-Census Bureau Research Network (NCRN) was established in 2011 to create interdisciplinary research nodes on methodological questions of interest and significance to the broader research community and to the Federal Statistical System (FSS), particularly to the Census Bureau. The activities to date have covered both fundamental and applied statistical research and have focused at least in part on the training of current and future generations of researchers in skills of relevance to surveys and alternative measurement of economic units, households, and persons. This paper focuses on some of the key research findings of the eight nodes, organized into six topics: (1) Improving census and survey data-quality and data collection methods; (2) Using alternative sources of data; (3) Protecting privacy and confidentiality by improving disclosure avoidance; (4) Using spatial and spatio-temporal statistical modeling to improve estimates; (5) Assessing data cost and data-quality tradeoffs; and (6) Combining information from multiple sources. The paper concludes with an evaluation of the ability of the FSS to apply the NCRN's research outcomes and suggests some next steps, as well as the implications of this research-network model for future federal government research initiatives.

ACKNOWLEDGEMENTS

This paper began as a May 8, 2015 presentation to the National Academies of Science's Committee on National Statistics by two of the principal investigators of the National Science Foundation-Census Bureau Research Network (NCRN) – John Abowd and the late Stephen Fienberg (Carnegie Mellon University). The authors acknowledge the contributions of the principal and co-principal investigators of the NCRN who are not co-authors of the paper (William Block, William Eddy, Alan Karr, Charles Manski, Nicholas Nagle, and Rebecca Nugent), the comments of Patrick Cantwell, Constance Citro, Adam Eck, Brian Harris-Kojetin, Eloise Parker, the editor (Ting Yan), and two anonymous referees. We also thank Handan Xu, who organized the references. We note with sorrow the deaths of Stephen Fienberg and Allan McCutcheon, two of the original NCRN principal investigators. The conclusions reached in this paper are not the responsibility of the National Science Foundation (NSF), the Census Bureau, or any of the institutions to which the authors belong. The NCRN was supported by NSF grants to the participating institutions: 1129475 to Northwestern University; 1130706 to Carnegie Mellon University; 1131500 to University of Michigan-Ann Arbor; 1131848 to Cornell University; 1131897 to Duke University and National Institute of Statistical Sciences (NISS); 1132008 to University of Colorado-Boulder; 1132015 to University of Nebraska-Lincoln; 1132031 to University of Missouri; and 1507241 for the Coordinating Office (Cornell, Duke, and NISS). The principal investigators also wish to acknowledge Cheryl Eavey's sterling grant administration on behalf of the NSF, and Daniel Weinberg and Nancy Bates (the latter assisted by Krista Park and Renee Ellis) in their role as Census Bureau liaison for the NCRN program.

Effects of a Government-Academic Partnership: Has the NSF-Census Bureau Research Network Helped Improve the U.S. Statistical System?

1. INTRODUCTION

A key problem that statistics agencies around the world face is the decline in participation in household and business surveys over the past 25 years (Groves 2017; Tourangeau and Plewes 2013; Williams and Brick 2018), which lowers the quality and increases the cost of official statistics. Meanwhile, large-scale data and computationally intensive methods, popularly known as "big data," are laying the foundation for a paradigm shift in the way statistical information is conceptualized, produced, and used. The U.S. Census Bureau and its partner, the U.S. National Science Foundation (NSF), recognized a need for the U.S. Federal Statistical System (FSS) to adapt and evolve. The development and reporting of official statistics by government agencies relies heavily on the foundation provided by academic (and self-generated) basic research. Therefore, in 2011, these partners established the NSF-Census Bureau Research Network (NCRN), a novel program of grants to academic institutions that married basic research activities to the applied needs of governmental statistical agencies.

With funding largely from the Census Bureau, NSF disseminated a call for proposals in September 2010 to create research nodes, each of which was to be staffed by teams of researchers conducting interdisciplinary research and educational activities on methodological questions of interest and significance to the broader research community and to the FSS, particularly the Census Bureau. To encourage fresh and innovative approaches of broad applicability, the solicitation posed a wide range of federal statistical problems without specifying the approaches (see the list in online Appendix A). After peer review of the proposals, the NSF made grant awards to six "medium" and two "small" nodes: Carnegie Mellon University, University of Colorado-Boulder joint with the University of Tennessee (a small node), Cornell University, Duke University joint with the National Institute of Statistical Science (NISS), University of Michigan-Ann Arbor, University of Missouri, University of Nebraska-Lincoln, and Northwestern University (small). A second solicitation, to establish a Coordinating Office for the NCRN led to a separate award to Cornell and Duke/NISS (see http://www.ncrn.info). Initial awards were made in October 2011 for a 5-year period. Supplemental awards and no-cost extensions allowed parts of the network to be funded through September 2018. Aggregate funding for the network was approximately $25.7 million.

The network includes several investigators with decades of direct collaboration with the FSS. But it also includes many more scholars, from the agencies and from academia, who only recently have invested in understanding the uses of the statistical products as well as the methods used to produce them. This focus has produced innovative applications and new methodologies that are immediately applicable to current systems. It also advanced the NCRN goal of engaging new researchers – both experienced and at the start of their careers – in research relevant to the future of the FSS.

The activities to date have covered both fundamental and applied statistical research and have focused at least in part on the training of current and future generations of researchers in skills of relevance to surveys and alternative measurements of economic units, households, and persons. The results of "basic" research covered by this grant program are described in the more than 400 papers sponsored by the NCRN program and published as preprints or in academic journals (see https://archives.vrdc.cornell.edu/ncrn.info/documents/bibliographies/ for a complete list as of April 2018). Many of these research products have "applied" implications important to FSS agencies.

The remainder of this paper will be in four parts. The next section discusses in brief some of the key research findings of the eight nodes, organized into six topics: (1) Improving census and survey data-quality and data collection methods; (2) Using alternative sources of data; (3) Protecting privacy and confidentiality by improving disclosure avoidance; (4) Using spatial and spatio-temporal statistical modeling to improve estimates; (5) Assessing data-cost and data-quality tradeoffs; and (6) Combining information from multiple sources. Section 3 explores collaborations across nodes and with federal agencies. The paper concludes with an evaluation of the ability of the FSS to apply the NCRN's research outcomes and suggests some next steps, as well as the implications of this research-network model for future federal government-academia collaborations. Online Appendix B discusses education activities and outcomes, and new software developed.

2. SELECTED RESEARCH OF THE NCRN NODES

We focus on the network's contributions in six main areas, acknowledging that there is some overlap among them.

A. Improving Census and Survey Data-Quality and Data Collection Methods

Given the importance of the Census Bureau's core mission, it is perhaps not surprising that a good deal of NCRN research focused on improving its data collection methods. It is clear to both academic researchers and Census Bureau professionals that one important path to a less expensive decennial census in 2020 is through the use of more up-to-date technology. The traditional Census Bureau approach is being rethought, especially since there will be widespread use of online census forms. Such broad census design issues have been the focus of the Carnegie Mellon node in its interaction with Census Bureau researchers. NCRN research on the effects of different types of census errors on the resulting allocations of funds and representation, which has taken place at the Northwestern node and is described further in subsection 2.E below, provides guidance on where to focus error-reducing resources. Improving the Census was a touchstone of the late Stephen Fienberg's career; his vision for the future of the Census is summarized in his 2013 Morris Hansen Lecture (Fienberg 2014).

By studying survey data, paradata, and audio recordings, Nebraska-node researchers have consistently found that the design of the questions plays a greater role in predicting survey data quality indicators (e.g., item non-response, response timing) and interviewer and respondent behaviors during a survey (e.g., exact question reading, provision of adequate answers) than characteristics of interviewers or respondents (Olson et al. 2018a; Olson and Smyth 2015; Smyth and Olson 2018; Timbrook et al. 2016). For example, Olson and Smyth (2015) found that 53% of the variance in response time in a telephone survey was due to the questions compared to only 3% due to interviewers and 7% due to respondents, and that this "question" variance can be largely explained by question features such as complexity (complex questions take longer) and sensitivity (sensitive questions are quicker). Similarly, Olson et al. (2018b) found that between 23% and 76% of the variance in respondent answering behaviors can be attributed to the questions compared to almost zero due to interviewers and 6% to 19% due to respondents themselves. In addition, they found that *interviewer* behaviors and communication processes are affected by those of respondents (Timbrook et al. forthcoming), and that *respondent* communication and cognitive processes are affected by respondent-interviewer interactions (Belli et al. 2013; Belli and Baghal 2016; Charoenruk and Olson forthcoming; Kirchner et al.

2017; Kirchner and Olson 2017; Olson et al. 2016, 2018b; Timbrook et al. forthcoming).

Nebraska node analysis of paradata has helped to better understand other aspects of interviewer/respondent interactions, including respondent retrieval patterns and prompts, which are especially relevant for questionnaire design in calendar and time diary interviewing (Atkin et al. 2014; Baghal et al. 2014; Belli and Baghal 2016; Olson and Parkhurst 2013). These findings have direct application in the Survey of Income and Program Participation (SIPP) and the American Time Use Survey (ATUS). Specifically, in a validation study of calendar interviewing, Belli and colleagues (Belli et al. 2013, 2016) found that whereas the use of parallel and sequential retrieval probes and strategies (which associate past contemporaneous and temporally-ordered events used by interviewers and respondents respectively) are associated with better data quality, interviewer parallel probes are unexpectedly associated with poorer data quality when each is soon followed by a respondent parallel retrieval strategy. With the ATUS, results from Kirchner et al (forthcoming) indicate that the resolution of initially missing reports of activities during a day is associated with respondents' engagement to report changes in who was present and where activities took place.

Such work with paradata is also relevant for designing and building computer-assisted telephone instruments that make recommendations to the interviewer (Arunhachalam et al. 2015; Atkin et al. 2015). For instance, Nebraska-node researchers used paradata to develop an intelligent agent that monitors interview progress and makes recommendations to the interviewer to help streamline data entry, improve the effectiveness and efficiency of interviewer-software interactions, and predict respondent breakoffs in web surveys (Eck et al. 2015; Eck and Soh 2017) In particular, Eck et al. (2015) used sequential machine learning with Markov chains to learn conditional probabilities of sequences leading to survey outcomes such as breakoff in paradata, and they used recurrent neural networks to learn the likelihood of breakoff using 23 instances of the Gallup Web Panel from 2012-2014. Between 56% and 75% of breakoff cases were identified with high precision (above 80%) using the Markov chain model, and 77% to 89%

of breakoff cases were identified with even better precision (above 92%) using the recurrent neural network model. One interesting corollary of an increased use of paradata in adaptive surveys is the necessity of organizing the storage, retrieval, and increased complexity in analytic tools needed for use of such data for analysis of large surveys (Olson and Parkhurst 2013), such as the multimodal American Community Survey (ACS). Editing the data for consistency to eliminate obvious errors (e.g., children older than their parents, pregnant males) is important.

The Duke-NISS node has been working on methods that improve how FSS agencies handle missing and faulty values. Murray and Reiter (2016) develop a flexible engine for multiple imputation or missing multivariate continuous and categorical variables, which they apply to impute missing items in data from the Survey of Income and Program Participation. Their model blends mixtures of multinomial distributions with mixtures of multivariate normal regression models in one joint model. In this way, the model adapts to the distributional features of the observed data, allowing it to automatically capture nonlinearities and interaction effects across entire multivariate distributions. Using simulations, they show that their model produces multiple-imputation confidence intervals and distribution estimates with better properties (e.g., smaller mean squared errors and closer to nominal coverage rates) than intervals based on general location models or chained equations, which are the default standards in multiple imputation of mixed data.

As another example of improved imputation methods, White et al. (2018) adapt regression trees as engines for imputation of missing items in the Census of Manufactures. They demonstrate improvements over existing imputation routines for this central data product, which historically have been based on mean and ratio imputations. Other relevant works include methods for handling non-ignorable nonresponse (Sadinle and Reiter 2017, 2018) for imputation of missing items in household data (Hu et al. 2018), and for imputation-based approaches for deciding whether or not to stop data collection (Paiva and Reiter 2017).

For decades, FSS agencies have based their statistical editing practices on the principles elucidated by Fellegi and Holt (1976). Reiter and his colleagues have developed methods to improve on these time-honored methods by using Bayesian approaches to allow stochastic editing to create multiply imputed, plausible datasets, building on ideas in Ghosh-Dastidar and Schafer (2003). The approaches are based on hierarchical models that include (a) flexible multivariate models for the true data values, with support restricted to feasible values, (b) models for errors given the latent true values, and (c) models for the reported values when errors are made. Traditional single-error localization and imputation procedures lead researchers to underestimate uncertainty. By assuming stochastic models for measurement errors, this alternative approach generates many plausible “corrected” datasets, thereby propagating uncertainty about error localization, and fully leverages information in the observed data to inform the edits and imputations. These developments include the use of such methods for both numerical-valued economic data (Kim et al. 2015) and categorical-valued demographic data (Manrique-Vallier and Reiter 2018). Using empirical examples and simulations with data from the Economic Census and the ACS, they further demonstrate that the stochastic edit imputation routines can result in secondary data files with smaller mean squared errors and closer to nominal coverage rates than methods based on Fellegi-Holt systems. The Census Bureau has begun a project to incorporate these methods into its 2017 Economic Census by using integrated edit, imputation, and confidentiality protection based on synthetic data models developed by Kim et al. (2015). The methods will permit publication of North American Product Classification System estimates and their margins of error without pre-specifying the table layout, as is currently done for the North American Industrial Classification System tabulations, and this project illuminates how more accurate modern methods can substitute for less accurate but convenient historical ones.

The Cornell node collaborated with the Census Bureau's Longitudinal Employer-Household Dynamics Program. This program publishes quarterly statistics using administrative records from state unemployment insurance record systems integrated with censuses, surveys and administrative records from the Census Bureau's household and business production systems. McKinney et al. (2017) produced the first total error analysis of the publications from these data.

B. Using Alternative Sources of Data

Censuses and surveys are not the only ways to collect information about the population and the economy. Independent sources can potentially provide useful data, such as from administrative records collected by governments for their own purposes (e.g. property assessments to levy real estate taxes or program applications to obtain benefits) and information provided by individuals in the course of their everyday activities (ranging from Twitter and Facebook posts to traffic-monitoring stations).

Making use of such information (particularly administrative records) in a statistical-agency environment typically requires record linkage, though there are cases where such information can be used without linkage (such as the Census Bureau's use of income tax records from the Internal Revenue Service for small businesses to avoid burdensome interviews). Record linkage is a critical component of the efforts to reduce census costs and, potentially, to improve accuracy. Of course, administrative records data have their limitations. As Groves and Harris-Kojetin (2017, p. 3–12) point out: "Administrative data can have many limitations including: (1)

lack of quality control, (2) missing items or records (i.e., incompleteness), (3) differences in concepts between the program and what the statistical agency needs, (4) lack of timeliness (e.g., there may be long lags in receiving some or all of the data), and (5) processing costs (e.g., staff time and computer systems may be needed to clean and complete the data)."

Record linkage (or matching) occurs at virtually every stage of operational and experimental census designs. Specifically, when the household address frame is the primary control system, record linkage occurs every time this frame is updated, primarily in the operation known as deduplication. The Census Bureau obtains a semiannual list of every address to which the U.S. Postal Service delivers (or plans to deliver) mail and, after removal of commercial and governmental addresses, this list is used to update the Master Address File, which is used both to carry out a population and housing census and as a sampling frame for ongoing household surveys. Also, when the first decennial census contact is not from a traditional mail-in mail-back form, record linkage occurs when the responses are integrated as they are received, especially if they are received without a decennial census identification (ID) code.

Traditionally, the address on the mail-back form links directly to the Master Address File, linking the geography for the household to the accuracy of the Master Address File. When the first contact is via an online form (IP address) or cell phone (cellular location services), this information must be linked to the Master Address File. In the 2020 Census, Internet response can take one of two forms, called ID and non-ID. In the ID form, the respondent enters the encrypted Master Address File identifier supplied on the invitation to take the census. In the non-ID form, the respondent supplies a residential address directly. Processing the non-ID cases uses this alternative address information. Record linkage is expected to play a critical role in the non-ID processing. It will also likely play a critical role in the non-response follow-up stage via the use of information from multiple administrative record lists to complete the form in the absence of directly collected data (or supplementary to an incomplete report). Additionally, record linkage is one of an intruder's possible methods for attempting to break the confidentiality of released data, and thus one must assess the risk of confidentiality breaches from published tables and public-use microdata samples.

All of these (and other) record linkage applications can be quantitatively improved using new tools that simultaneously link more than two lists, while deduplicating each of the lists. The solutions provide conceptual generalizations of the familiar Fellegi-Sunter (1969) method for two lists (or deduplication of a single list) that are computationally feasible for application at the scale of the decennial census (Sadinle 2017; Sadinle and Fienberg 2013; Steorts et al. 2016). Further, the new methods acknowledge and propagate the uncertainty from the matching process into subsequent analyses. Improved record linkage can also improve the data needed to handle nonresponse to the census and to surveys, often by providing data for a particular address from administrative records, but also by providing data for modeling non-respondents.

Particularly relevant for the Census Bureau is combining these issues into useful statistical models and methods. Fienberg (2015) presents a discussion of the value of addressing (1) record linkage methods for three or more files, (2) combining duplicate detection and record linkage, (3) propagating duplicate detection and record linkage error into subsequent calculations, and (4) measuring both erroneous enumerations and omissions.

Record linkage is also important for business data. In collaboration with the University of Michigan's Sloan Foundation-funded Census-enhanced Health and Retirement Study, the Michigan node developed and tested methods for probabilistic linkage of the employers of Health and Retirement Study respondents to the Census Business Register. This work addresses the complexity and benefits of linking household and business data to better understand employment of older Americans. The record linkage research confronts the difficulty of how individuals report their place of employment and how it is represented in administrative data. The approach taken highlights the importance of accounting for errors in matching records and of using probabilistic techniques to reflect these errors in subsequent analyses (Abowd and Schmutte 2016). This research also produced new software for standardizing business names, a necessary step in linking organizational data (Wasi and Flaaen 2015).

The second alternative source of data for statistical agencies is “non-designed data,” also sometimes termed “organic data,” “third-party data,” “naturally occurring data,” or “data in the wild,” such as from social media like Twitter or transaction data that are digital traces of people’s and businesses’ daily activities (bank and credit card transactions, shopping, turning on lights, etc.). The key issue is not yet whether those data can replace data that FSS agencies use to report key social, economic, housing, and demographic indicators, but whether those data can provide useful indicators and checks on traditional time series, or produce measures at lower cost, greater frequency, more geographic detail, or in conjunction with survey data to reduce respondent burden. Note however that their use in official statistics could easily be jeopardized by changes in methodology by the independent provider, or even its discontinuation, as well as the proprietary nature of its collection and dissemination.

Account data. Data on consumers’ transactions and balances can provide high-frequency and high-quality measures of spending, income, and assets that are difficult to measure accurately using surveys, which rely on infrequent self-reports from relatively small samples of individuals. In collaboration with a Sloan Foundation-funded database development project, the Michigan node pioneered the use of comprehensive account data from linked checking and credit card accounts to confront the difficulties of using such naturally occurring account data to produce economically meaningful measurements and to study economic behavior and outcomes. Gelman et al. (2014) show that account data drawn from a large sample of users of a financial services application can be broadly representative of the U.S. population. They use this newly developed data infrastructure to shed light on the excess sensitivity of spending to predictable income, show how households accommodate short-run drops in liquidity (Gelman et al. forthcoming) and how spending responds to a permanent change in gasoline prices (Gelman et al. 2016).

The use of transaction and balance data have great promise to improve spending and income measures published by the FSS. Spending reports are either based on very aggregate store-level data (the Census Bureau Advance Monthly Retail Trade Report) or surveys of consumers (the U.S. Bureau of Labor Statistics Consumer Expenditure Survey). Both these surveys suffer from declining response rates and other data quality problems. Income reports when benchmarked to Internal Revenue Service tax data (such as the U.S. Bureau of Economic Analysis National Income and Product Accounts and its monthly Personal Income and Outlays) show survey underreporting. Tax data are inherently annual and available to the FSS only with a considerable lag and substantial disclosure limitations. On the other hand, transaction data are available daily, with high precision, for large samples of individuals, with great detail on location and type of spending, and with almost no lag.

Social media data. Official statisticians understand the framework in which a time series indicator like new unemployment insurance claims can be used to measure change. The population at risk is all statutory employees covered by state unemployment insurance systems. When the indicator goes down, fewer such employees filed new claims for unemployment insurance. What does an increase in Tweets about “job loss” mean? The Michigan node developed a predictive model to assess this question. Job-loss Tweets do forecast the changes in official new claims for unemployment insurance, particularly upward spikes, allowing one to capture turning points in economic activity that are often missed or captured only with a long lag using traditional approaches (Antenucci et al. 2013, 2014). The project developed a real-time predictor of unemployment insurance claims and maintains a website giving weekly updates (see econprediction.eecs.umich.edu).

An ongoing challenge to the use of social media data, in particular for measurement over time, is that while there is an enormous amount of this type of cross-section data, no particular social media platform has existed long enough to capture an entire business cycle, let alone multiple such transitions. Thus, the development of measures from social media data, requires the systematic use of prior knowledge about the structure of the economy, such as how job flows change over the business cycle, akin to the use of seasonality adjustments. Without a benchmark reference, how can the predictive model detect a change in the weights it attaches to its inputs? The Michigan node is now addressing this issue with the development of an interactive model that allows those with domain expertise to provide benchmark datasets and economic concepts for measurement to a large archive of unstructured, web-based (social media and imaging) data in order to generate and archive new time series measures.

Researchers are also investigating natural-language processing of social media, transaction, and accounting data to help better understand economic measurement. This research is of interest to the Bureau of Labor Statistics, the Census Bureau, the Bureau of Economic Analysis, and the U.S. Federal Reserve Board.

The impact of non-designed data on economic statistics and policy analysis. The FSS largely relies on its ongoing data collections for official time series because of the need for consistency over long time periods. This consistency is especially important to policymakers (Yellen 2017). Nonetheless, official statistics are making increased used of non-designed economic data for price and value measurement (U.S. Bureau of Economic Analysis Advisory Committee 2017; U.S. Federal Economic Statistics Advisory Committee 2015). The Census Bureau, Bureau of Labor Statistics, and Bureau of Economic Analysis are currently making substantial use of commercial data in their programs. The work of the Michigan node has addressed questions of representativeness, timeliness, and coverage that are essential for using these data more systematically in official statistics (Gelman et al. 2014).

NCRN work on economic indicators has focused on the question of whether novel economic indicators have incremental information that could be of use to policymakers. Antenucci et al. (2014) show that the social media index constructed from tweets has supplemental explanatory power for nowcasting new claims for unemployment insurance beyond the consensus forecast of experts. Hence, even with the short time series of data available from social media, there is evidence that social media data can be used by policymakers or market participants to extract information about the state of economy. That paper also shows preliminary evidence of the shift in the relationship between vacancies and unemployment known as the Beveridge curve that is an ingredient to understanding the recovery from the Great Recession.

Non-designed data can also be used to provide policymakers with information not readily available in official statistics because they are insufficiently granular. Former Federal Reserve Chair Yellen (Yellen 2017) cites two examples of such research that were relevant for Federal Reserve monitoring of the economy: the analysis of by Federal Reserve staff of the effects of Hurricane Matthew (Aladangady et al. 2016) and the analysis by Michigan node researchers of the effects of the 2014 gasoline price decline on consumer spending (Gelman et al. 2016).

Other non-designed data. Another use of auxiliary data comes from combining area-level covariates measured over space and/or time with tabulated survey estimates within a hierarchical model-based framework. For example, (Porter et al. 2015a) models the ACS 5-year period estimates of mean per capita income in Missouri counties and utilizes percentage of unemployed individuals in each county as auxiliary information, also obtained from the ACS. Another salient example comes from the Missouri node's use of social media (functional time series) data from Google trends (Porter et al. 2014). The approach extends the traditional Fay-Herriot model to the spatial setting using functional and/or image covariates. A natural use for this methodology could be to incorporate remote sensing data as image covariates to augment information obtained from federal surveys or to assist with in-office address canvassing.

Work by Michigan and Cornell researchers contributes to our understanding of multiply-sourced data. Michigan-node work compared survey (SIPP) and administrative (Longitudinal Employer-Household Dynamics--LEHD) measures of the causes of job loss and studied the implications for estimates of the response of earnings to job loss (Flaaen et al. forthcoming), developed and studied a measure of firm quality based on the ability of firms in the LEHD to attract and retain workers (Sorkin 2018), and developed explanations of the divergence of survey (Health and Retirement Survey) and administrative (Social Security) measures of earnings (Hudomiet 2015). Cornell-node work investigated the coherence of ACS and administrative reports of workplace location (Green et al. 2017).

The Missouri node has proposed improvements to the statistics created from the LEHD database (Bradley et al. 2015a, 2017) that make use of multivariate spatio-temporal statistical modeling. The Census Bureau and the Missouri and Cornell nodes are collaborating to enhance the precision of the disseminated estimates.

C. Protecting Privacy and Confidentiality by Improving Disclosure Avoidance

Privacy is about what information a respondent is willing to share while confidentiality is about the ethical and statutory requirements to keep personal data from unauthorized disclosure to a third party. Three different approaches to confidentiality protection span the ongoing work of the nodes in this area: data swapping (historically the Census Bureau method of choice to date for both the decennial census and the ACS), multiple imputation (involving the preparation of synthetic datasets), and the more recently developed method of differential privacy that emanates from cryptography and computer science and offers the strongest possible privacy guarantees. However, differential privacy has not yet been proven to work for all kinds of data releases that the Census Bureau is accustomed to producing (Abowd and Schmutte 2016). The *Journal of Privacy and Confidentiality* devoted an entire issue (2015-2016, volume 7, issue 2) to differential privacy; see also Murray (2015).

Both the Carnegie Mellon and Cornell nodes have contributed to the "the economics of privacy." Acquisti et al. (2015, 2016) highlight how the economic analysis of privacy evolved over time, as advancements in information technology raised increasingly nuanced and complex issues. They highlight three themes: (1) Characterizing a single unifying economic theory of privacy is hard, because privacy issues of economic relevance arise in widely diverse contexts; (2) There are theoretical and empirical situations where the protection of privacy can both enhance and detract from individual and societal welfare; and (3) Consumers' ability to make informed decisions about their privacy is severely hindered because they are often in a position of imperfect or asymmetric information regarding when their data is collected, for what purposes, and with what consequences.

But a much larger social issue also concerns researchers in the network. What are the appropriate tradeoffs between data confidentiality and data accuracy? As Abowd and Schmutte (2017) show, public statistics will be under-provided by private suppliers, and welfare losses from the under-provision can be substantial (see also Abowd and Schmutte forthcoming). But a key contribution of theirs is that the question cannot be answered from the technology of statistical disclosure limitation or privacy-preserving data mining. It requires understanding how the citizen consumers of an agency's statistics value data accuracy when they must pay with some loss of privacy. All the players in this arena, public and private, understand the risks associated with direct privacy breaches far better than they understand how to measure a society's preferences for public data that can only be produced with some privacy loss. Changes to the current paradigm may require new legislation.

Among the network's new contributions in this area is a focus on quantifying the disclosure risks associated with large-scale record linkage, such as that proposed for the 2020 Census, and on producing accurate statistics that control that risk in a quantifiable way. Much of the NCRN research on disclosure avoidance addresses how to combine statistical disclosure limitation with correct analysis of the published data, including understanding the uncertainty introduced through probabilistic data linkage or model-based data imputation (Kim et al. 2016).

Several of the network's researchers have worked on extending prior work on the use of synthetic data as a disclosure avoidance technique (Kinney et al. 2011). The Cornell and Duke-NISS nodes have continued supporting the Census Bureau in learning from and extending the use of synthetic data (Kinney et al. 2014; Miranda and Vilhuber 2016; Vilhuber et al. 2016). Researchers from the Duke-NISS node are collaborating with the Census Bureau on creating a synthetic-data version of the 2017 Census of Manufactures. The Duke-NISS node has also developed and extended techniques for securely and privately providing users with feedback on the quality of their inferences from the synthetic data (Chen et al. 2016). These include differentially private statistical significance tests, Receiver Operating Characteristic curves (Park et al. 2004), and plots of residuals versus predicted values for linear and logistic regression; an R software package is under development. Synthetic data techniques have caught the attention of the popular press (Callier 2015). Finally, the Missouri and Duke-NISS nodes have collaborated to propose disclosure avoidance methods for spatially correlated data (Quick et al. 2015b, 2018).

How can the transparency of research using an agency's confidential data be increased, for instance to ensure reproducibility? Scientific integrity requires curation of the provenance of the data used in such research. In turn, reproducibility of the use of confidential data ultimately improves its quality. But confidentiality concerns have often proven an impediment to achieving these goals. Expansive codebooks or detailed metadata is subject to the same confidentiality constraints as the actual data. For instance, Internal Revenue Service regulations prevent the naming of certain variables (columns) in the data, and yet codebooks need to be complete and accurate. Similarly, standard summary statistics include ranges (maxima and minima) and percentiles, which are subject to disclosure avoidance measures. These constraints are not handled well (or at all) by traditional tools for data documentation, and are hard to verify in user-generated documents. Researchers at the Cornell node have proposed enhancing various standards for curating metadata in a way that respects these confidentiality constraints imposed on the curators (Abowd et al. 2012; Lagoze et al. 2013a, 2014). A software system to implement the enhancement was developed, the Cornell Comprehensive Extensible Data Documentation and Access Repository, and is used to disseminate various codebooks (SIPP "Synthetic Beta File" and the Census Bureau's Synthetic Longitudinal Business Database). Additional work aims to further expand the standard to embed provenance information, allowing researchers to tie diverse public-use and synthetic data products to common confidential source files (Lagoze et al. 2013b).

D. Using Spatial and Spatio-Temporal Statistical Modeling to Improve Estimates

The ACS design explicitly combines spatial and temporal information to produce annual and 5-year estimates for many subpopulations. These estimates are released with associated margins of error (MOEs that define 90-percent confidence intervals). Working with current ACS data, researchers at the Missouri node and the Colorado-Tennessee node have each developed new spatial techniques for aggregating and disaggregating the basic ACS estimates geographically. In particular, the Missouri node introduced methodology that uses a Bayesian spatio-temporal model that can create estimates over customized (user-defined) geographies and/or times, with associated measures of uncertainty (Bradley et al. 2015b). Public-use software for implementing their approach was presented at the 2017 Joint Statistical Meetings (Raim et al. 2017).

A key challenge in working with ACS estimates is that the ACS reporting uses geography (census block groups and tracts) previously used only for decennial census long form estimates; yet small geographies have large margins of error (Folch et al. 2016; Spielman et al. 2014). Interviews conducted with urban planners (frequent users of small area ACS data) show that while they are often aware of this problem they ignore it (Jurjevich et al. 2018). For example, a survey respondent (planner at a regional planning agency), noting that margins of error (MOEs) from the ACS were sometimes larger than the estimates themselves, said: “I should not use the data or provide a range from 0-200, but often I don't have the time to look in detail at the MOEs for as many geographies and years of data that we have to provide data for. It gets overlooked much too often but it's hard to have a good solution when there isn't better data available." The Colorado-Tennessee node also conducted usability studies of ACS data through an experiment that monitored keystrokes, mouse movement, and eye movement. They found that when confronted with uncertain data on a familiar city, subjects tended to substitute their local knowledge of the community for the data when making decisions; but when they did not know the city, uncertainty in the data created variability in outcomes of the assigned task (Griffin et al. 2014). Combined, these results indicate that there is both a need and a demand for tools to help end users communicate ACS data uncertainty and to make the estimates more usable for analysis.

That node took two approaches to that task. First, the node developed software that groups demographically similar and spatially adjacent census tracts (or any census geography) into "regions" (Folch and Spielman 2014; Spielman and Folch 2015). As tracts are grouped together, variances of the estimates typically decreases. Since ACS uncertainty varies from attribute to attribute, the user can select the particular attributes relevant for their research question to generate the maximum number of regions, where each region's attributes meet a data-quality threshold. (Data and interactive visualizations are available for four data scenarios on all U.S. metropolitan statistical areas at www.reducinguncertainty.org.) The second approach uses multivariate statistical clustering to group demographically similar census tracts into latent classes. This approach was used to make a broad hierarchical classification of all U.S. census tracts (Spielman and Singleton 2015). This data product is published and distributed by Carto, a New York based mapping startup (available at carto.com/data-observatory).

In an effort to design "optimal" statistical geographies, the Colorado-Tennessee node has examined the spatial structure of the American population by measuring the sensitivity of census estimates to gerrymandering. That is, it assesses the effect of altering the boundaries of census tracts. The answer, while preliminary, seems to be “quite a lot in some places.” For example, over 10% percent of census tracts saw changes of 10% or greater in a measure of segregation (entropy) as the result of changing the tract boundary while keeping the total population constant (Fowler et al. 2018).

Taking a different approach, the Missouri node has developed a statistical framework for regionalization of multiscale spatial processes (Bradley et al. 2016a). The proposed method directly addresses the important modifiable areal unit problem (MAUP) and the ecological fallacy problems associated with multiscale spatial data and introduces a criterion for assessing spatial aggregation error. This criterion, called CAGE (Criterion for spatial AGgregation Error), is then minimized to produce an *optimal* statistical regionalization. The impact of such methodology has significant implications for various FSS stakeholders. For example, various ACS data-users wishing to aggregate tabulations across geographies (using the methods discussed in (Bradley, Wikle, and Holan 2015b) can evaluate to what extent valid inferences can be made; R software packages for CAGE and spatio-temporal change-of-support are currently under development (e.g., see Raim et al. 2017).

Results can be directly referenced to identifiable inputs in the statistical system and reproduced reliably from those inputs. Advances in the curation of the metadata help ensure that the agency’s use of these methods can be audited and its published results can be reproduced. Reproducibility is not always possible for data analysis based on commercial data such as Google Trends, but the Michigan node’s research using Twitter feeds can be reproduced because they post their underlying data.

The Missouri node has also been actively engaged in developing hierarchical statistical models that leverage different sources of dependence (e.g., multivariate, spatial, and spatio-temporal) to improve the precision of estimates from various data products. Broadly speaking, many of the proposed techniques can be viewed as natural generalizations of the methods currently used for small-area estimation by most statistical agencies. That is, they are generalizations of the Fay-Herriot (1979) model (Bradley et al. forthcoming, 2015a; b, 2016a; b; Cressie and Zammit-Mangion 2016; Porter et al. 2014, 2015a; b; c; Sengupta and Cressie 2013a; b); for additional details see online Appendix C. The Missouri node has developed the hierarchical-statistical-modeling approach in ways that will give federal statistical agencies a distinct advantage for their data products over commercial value-added re-sellers of the same data. This advantage stems directly from the agency's access to and use of the complete set of geographic identifiers and original data values in doing the calculations and then applying statistical disclosure limitation to the outputs (Quick et al. 2015b, 2018). The methodologies developed at the Missouri node typically use the Census Bureau geography definitions, but they provide the flexibility to depart from this restriction. In other words, the proposed methods retain the ability to operate from customized geographies and/or temporal supports through the use of a change-of-support approach (Bradley et al. 2015b, 2016b); see Appendix C. Furthermore, the small area estimates come with a measure of their uncertainty that allows prediction intervals to be constructed.

There are numerous examples of multiple surveys disseminating related demographic variables that are measured over space and/or time. The Missouri node's methodology combines the disseminated estimates from these surveys to produce estimates with higher precision. Additionally, in cases where estimates are disseminated with incomplete spatial and/or temporal coverage, the Missouri node's approach leverages various sources of dependence to produce estimates at every spatial location and at every time point. The approach for combining the multiple surveys is developed as a fully Bayesian model. The proposed methodology is demonstrated by jointly analyzing period estimates from the Census Bureau's ACS and concomitant estimates obtained from the Bureau of Labor Statistics Local Area Unemployment Statistics program (Bradley et al. 2016a).

More generally, the Missouri node uses spatial, spatio-temporal, and/or multivariate dependence structures to generate point-in-time estimates of subpopulation quantities and to provide an associated measure of uncertainty. (Traditional Fay-Herriot small-area estimates are a special case.) Flexible models have been introduced that allow estimation for both Gaussian and non-Gaussian settings (Bradley et al. forthcoming, 2015a; b, 2016a; b, 2017, 2018; Porter et al. 2014, 2015c; a; b; Sengupta and Cressie 2013a). Extensions of the method can be used to incorporate other variables from the frame, or related frames. For example, Bradley et al. (2016a) introduce a multivariate mixed-effect spatio-temporal model that combines estimates from the Bureau of Labor Statistics' Local Area Unemployment Statistics with estimates from the ACS, to produce estimates that have significantly improved precision over using either survey individually. Cressie and Zammit-Mangion (2016) take a conditional approach to multivariate modeling in the Gaussian setting.

Visualization constitutes another important component in the analysis of spatial and spatio-temporal data. Using the ACS, Lucchesi and Wikle (2017) develop and present methods for simultaneously visualizing areal (spatial) data and its uncertainty using bivariate choropleth maps, map pixelation, glyph rotation, as well as animations. (See online Appendix D for further discussion and examples.) Spatial data can also be used to provide timely information about changing economic conditions. In work by the Michigan node that combines the themes of non-designed data and geospatial analysis, Wilson and Brown (2015) use satellite imagery to show how the "Great Recession" affected southern Michigan by measuring changes in visible impervious surface area (VISA). The paper shows that VISA (e.g., structures and paved roads and parking lots) declined from before the Great Recession (2001-2005) to after (2006-2011). This novel application of satellite imagery provides a new tool for measuring changes in economic activity.

E. *Assessing Data Cost and Data Quality Tradeoffs*

Fundamental problems for the U.S. federal statistical system (and for government statistical agencies around the world) include how to understand the value of the statistics they produce, how to compare value to cost in order to guide rational setting of statistical priorities, how to increase value for given cost, and how to better communicate the value of their data programs to those who set their budgets. The market does not provide a measure of value because government statistical data are public goods, so to understand their value it is necessary to understand how the statistics are used, and what would occur if the statistics were available with different data quality characteristics. The Northwestern node extended and applied statistical decision theory, including cost-benefit analysis, to attack such basic questions.

Spencer et al. (2017) develop a cost-benefit analysis for the 2016 quinquennial census of South Africa to an alternative of no census. They measured benefits arising from more accurate allocations of funds due to improved population numbers. Improved fund allocation was also a consideration for similar analyses in the United Kingdom and New Zealand, which assumed that the fund allocation formulas optimized social welfare when applied to error-free statistics. In contrast, Spencer et al. explicitly allowed for willingness to pay for improved accuracy in allocations.

The 2020 U.S. Census is highly cost-constrained relative to previous censuses, and there is uncertainty about the quality of the census attainable for the allowed cost. Seeskin and Spencer (2015) considered alternative specifications of census quality and modeled the effects on (1) the funding allocation of more than $5 trillion over the decade of the 2020s, and (2) the distribution of seats in the U.S. House of Representatives in 2022. They allowed for vectors of errors in census state population sizes to have arbitrary means, standard deviations, and correlations, and to be either multivariate normally distributed or multivariate-*t* on four degrees of freedom. For a given cost-quality relationship, their analysis permits estimation of the distortions in distributions of funds and seats that arise for a given cost, in order to reveal the tradeoffs. For example, when the average standard deviation of a state's population is 2% of its actual population, the expected number of seats going to the wrong state is about 6.5, and the expected amount of misallocated federal funds over the 10-year intercensal period is $40 billion. The expected absolute deviations in apportionments and in allocations both increased approximately linearly with the average relative standard deviation of state population numbers. Seeskin and Spencer (2018) extend the analysis of changes in apportionment caused by census error, using short term projections of state populations based on the Census Bureau's postcensal population estimates for 2017, and assuming that patterns of error in 2020 state populations are similar to those measured for the 2010 census, except that the magnitudes may be larger. They found that when 3 House seats are shifted, the losing states are Texas (2 seats) and Florida (1 seat).

In other work at the Northwestern node, Manski (2015) distinguishes transitory statistical uncertainty, permanent statistical uncertainty, and conceptual uncertainty. He illustrated how each arises as the Bureau of Economic Analysis periodically revises Gross Domestic Product estimates, the Census Bureau generates household income statistics from surveys with nonresponse, and the U.S. Bureau of Labor Statistics seasonally adjusts employment statistics. He anchors his discussion of communication of uncertainty in the contribution of Morgenstern (1963), who argued forcefully for agency publication of error estimates for official economic statistics (as is done by the Census Bureau for monthly and quarterly economic indicators releases). In a related technical article, Manski (2016) elaborates on the theme of communicating uncertainty in official statistics, focusing on the permanent statistical uncertainty created by survey nonresponse. In current work, Manski is focusing on the crucial survey design question regarding how much data to collect and how much effort to expend to enhance the quality of the collected data when faced with a fixed budget. Dominitz and Manski (2017) use decision theory with a minimax regret principle for choosing between a high-cost high-accuracy survey and a low-cost low-accuracy one, where low accuracy is considered in two ways – imprecise survey responses and unit non-response.

F. *Combining Information from Multiple Sources*

Distinguished from record linkage, which attempts to combine data sources in a way that matches information from multiple sources, better estimates can be made by combining information from multiple sources by modeling. One particular extant example is the Census Bureau's Small Area Income and Poverty Estimates program (http://www.census.gov/did/www/saipe). The Missouri node has expanded this research field by developing a hierarchical Bayesian approach using geography and/or time to enhance model estimation and prediction (Bradley et al. 2015b), in effect creating powerful spatio-temporal mixed effects models that include Fay-Herriot (1979) models as a special case. Given the available surveys, the conditional distributions of the latent processes of interest are used for statistical inference. To demonstrate the proposed methodology, researchers from the Missouri node have jointly analyzed period estimates from multiple surveys (Bradley et al. 2016a). For example, the proposed model combines data from the ACS and the Local Area Unemployment Statistics program to provide improved estimates of unemployment rates.

Other ways to improve socio-economic estimates from the ACS involve models and data internal to the Census Bureau. For example, should modeling using external data sources be used to improve upon the direct survey estimates available from a household survey, and should survey-based (direct) estimates and model-based estimates, and indeed mixed (weighted) estimates, all be produced, or would confidentiality suggest limiting the types of data (and variables) that are modeled? The experience of the Census Bureau with its Small Area Income and Poverty and Health Insurance Estimates programs to address this question is relevant, as it attempts to expand the modeling to unemployment rates (noted above) and to the estimation of jurisdictions required to offer multi-lingual ballots under Section 203 of the 1965 Voting Rights Act. Modeling can be used to generate new ACS estimates, other than those published for fixed geographies and fixed time periods (currently 1 year and 5 years), say a 4-year period estimate for a particular combination of census tracts representing a neighborhood (Bradley et al. 2015b).

3. THE IMPORTANCE OF COLLABORATION

As the NCRN matured, the opportunities and desirability of direct collaboration across the nodes and with the FSS agencies (particularly the Census Bureau) became more apparent. We focus first on inter-nodal collaborations, some of which resulted from movement of students between nodes (e.g., from being post-doctoral fellow at one node to then being a faculty member at another node). It is likely that inter-nodal collaborations took place only because these universities were linked through the NCRN, especially through the biennial meetings convened by the NCRN Coordinating Office (mostly at the Census Bureau), since the topics chosen by the nodes did not overlap very much (a conscious decision by the NCRN program sponsors).

Examples of inter-nodal collaborations include: (1) Duke-NISS and Missouri on generating synthetic geographies; (2) Duke-NISS and Carnegie Mellon on improvements to Fellegi-Sunter (1969) matching models; (3) Duke-NISS and Cornell on continued development of synthetic establishment data; (4) Missouri and most of the other nodes at the 2016 "Workshop on Spatial and Spatio-Temporal Design and Analysis for Official Statistics"; (5) Michigan, Carnegie Mellon, Cornell, and Duke-NISS on evaluating methods for probabilistic linkage; (6) Michigan and Cornell on implementing model-based probabilistic linkage for economic units, enhancing surveys with measures from administrative data, and evaluating quality of survey measures using administrative data; (7) Michigan and Duke-NISS on SIPP training; (8) Nebraska and Carnegie Mellon regarding the development of an automated calendar for survey use; and (9) Missouri and Cornell on spatio-temporal models for the LEHD program.

One of the most active collaborations between Census Bureau and nodal researchers was the Summer Working Group for Employer List Linking (SWELL). The purpose of this group, which included researchers from the Michigan, Carnegie Mellon, and Cornell nodes and Census Bureau staff, was to develop tools for linking person-level survey responses to employer information in administrative records files using probabilistic record linkage. Once the linkage was accomplished, there were four areas of potential payoff: (1) production of a research-ready crosswalk between survey responses and administrative employer records including quality metrics to help users assess the probability that a particular link is correct; (2) comparison of self-reporting to administrative measures (e.g., location, earnings, firm size, industry, layoffs) enabling the enhancement of data quality by improving edits and imputations; (3) creation of improved or new measures available to users without increasing respondent burden; and (4) investigation of new research questions that could not be answered by either dataset alone (e.g., through creation of new variables and longitudinal outcomes or histories). The group has produced software (in SAS and STATA) for standardizing business names to allow improved linkages between survey reports of business names and administrative data from those employers (for the STATA version, see Wasi and Flaaen 2015). The research also helps to improve the Census Bureau's ability to design employer surveys that sample firms based on the composition of their employees, so that there can be better and more representative estimates of the characteristics of the employers of American workers. This successful collaboration was only possible because of the existence of a Federal Statistics Research Data Center (FSRDC) at each location, allowing the sharing of data and research in real time. Despite the seasonality implied by its name, it is an ongoing collaboration.

Other examples of direct collaborations of node researchers with Census Bureau staff include the following: (1) development of a model to predict 2020 Census quality, as measured by the accuracy of the state population totals (Northwestern); (2) assessment of respondent comfort with geolocation of their home (Carnegie Mellon); (3) improvements in multiple file matching methods to aid the 2020 Census (Carnegie Mellon); (4) research to better understand residential mobility (Colorado-Tennessee); (5) imputations for missing business and demographic estimates (Duke-NISS); (6) development of methods for creation of synthetic business data (Duke-NISS); (7) creation of a synthetic data version of the 2017 Economic Censuses (Duke-NISS); (8) improvements in confidentiality protection of demographic data (Cornell); (9) participation in the Census Bureau's ACS Data Products Design working group (Colorado-Tennessee); (10) provision of advice on plans for 2020 Census operations, specifically on geographic targeting for the communications campaign, non-response follow-up, and coverage measurement (Colorado-Tennessee); (11) development of an imputation methodology for the Monthly Advance Retail Trade Survey, development of model-based statistical methodology for in-office address canvasing, and implementation of space-time methodology using ACS estimates (Missouri); (12) provision of advice to Census Bureau staff on revising the American Time Use Survey user interface where SIPP Event History Calendar navigation patterns are shown to be associated with data quality, which have potential implications for interviewer training (Nebraska); and (13) working with the Census Bureau's Center for Survey Measurement to assist with detecting measurement error through paradata (Nebraska).

There are still challenges for the transfer of the new technologies and for approaches to practical implementation. Most likely to produce technology transfers is *direct* collaboration between Census Bureau staff and node researchers. Because of the challenges in implementing many of the collaborative innovations, they produce fewer scientific publications but ensure that the research bears direct fruit within the FSS agencies. Many of the NCRN researchers now collaborate in solving ongoing implementation issues because NCRN greatly expanded FSS access to academic collaborators. Online Appendix E lists both active NCRN-FSS collaborations and collaborations that have led to changes in FSS production processes.

The SWELL does demonstrate the value of collaboration between academics and FSS staff when there are common scientific goals, especially where these intersect with operational requirements of the FSS. On the geography front, researchers affiliated with the Colorado-Tennessee node are collaborating with the U.S. Geological Survey (Wood et al. 2015), Oak Ridge National Laboratory, and the U.S. Forest Service to improve their use of small-area data. Researchers from the Missouri node collaborated with the U.S. Centers for Disease Control and

Prevention on methodology for disclosure avoidance (Quick et al. 2015a).

One possible amelioration of this lack of direct collaboration would be through co-location. Several individuals have attempted to take the results of their basic research and assist the Census Bureau in implementing their results by working on-site at the Census Bureau. One common approach has been for these individuals to become temporary federal employees, either through the Intergovernmental Personnel Act, as “Schedule A” employees, or through summer student employment or fellowships (such as dissertation fellowships) or the “Summer at Census” program. Still others have become off-site collaborators, working on such projects as improving the American Time Use Survey time diaries collected by the Census Bureau for the Bureau of Labor Statistics, improving the SIPP Event History Calendar for the Census Bureau, and revising the Census of Manufactures edit and imputation and data-dissemination strategies. Other topics that these “partially resident” researchers are working on include capture-recapture methodology (relevant for the estimation of census error), small-area estimation for the ACS and other surveys, improving editing and imputation for missing data, improving record-linkage practices allowing for uncertainty, implementing better storage paradigms for paradata, determining how to use paradata to identify problems, and improving the LEHD database. Other collaborations include matching the SIPP to the LEHD database (including development of a new-firm quality measure), improving the measurement of pension buyouts, SWELL, linking import-export data to the Longitudinal Business Database and to non-Census Bureau data on multinationals, to allow new types of research (but available only to Census Bureau and FSRDC researchers).

4. LESSONS LEARNED

The NCRN has been recognized with the 2017 *Statistical Partnerships among Academe,* *Industry, and Government* (SPAIG) award from the American Statistical Association "for addressing methodological questions of interest to the federal statistical system and training future generations to design, conduct, analyze, and report official statistics." The network nodes have individually been productive, both in the basic and the applied research domains, with many publications, including many in high-impact journals. Cross-node and government-university collaborations have occurred that probably would not have happened in the absence of a network, encouraged by the semi-annual open NCRN meetings (mainly at the Census Bureau).

Yet, improvements are desirable and possible. We believe that there are four valuable lessons that have been learned about government-academic research partnerships.

First, better coordination between the agency and academic partners leads to more useful research outcomes. One suggestion is that "ways be found to facilitate not only the ability of academic scholars to spend time working within … government agencies, but also that key agency career researchers be encouraged and detailed to spend significant periods of time at the university-based research nodes where they can actively participate in the development of methodologies and basic science advances being pioneered there" [from *NCRN Reverse Site Visit Report*, February 2015]. As noted above, the Census Bureau has already implemented part-time employment relationships, allowing the agency to bring the university-based researchers onto their agency teams directly. Moreover, better dissemination and communication across FSS agencies, perhaps through the Interagency Committee on Statistical Policy (chaired by the U.S. Chief Statistician), would facilitate greater utilization of other relevant research as well. Should a similar government-academic partnership be pursued in the future, we encourage the government agencies to think about likely collaborations in advance. We note that increased participation of different FSS agencies in the FSRDC network will also support the dissemination of research relevant to the entire FSS.

Second, the cross-fertilization that will result from academics working in close collaboration with government researchers will further enhance technology transfer. It is not enough for academics to invent new and useful methods if it is difficult for the relevant agencies to adopt those new methods. Adoption of several new techniques emanating from the NCRN nodes is well underway at the Census Bureau, due in large part to those same researchers assisting the Census Bureau with the adoption.

Third, it is important to think through the issues of academic access to confidential data in advance. While participating in the FSRDC program (then the Census Bureau RDC program) was not a requirement for a grant, all but one of the nodes without an RDC eventually joined that program and their research benefitted from access to restricted data. The FSRDCs could also provide a convenient way for Census Bureau staff to work in an academic setting for extended periods without losing touch with ongoing agency activities that might require access to confidential data. Furthermore, the FSRDC program can be used to link together collaborators from many locales, whether at the host academic institution or not.

Fourth, the ability of FSS agencies to hire students trained as statisticians, whether through government-academic partnerships or otherwise, needs to be improved. Such students have skills most other potential hires do not have, and hiring them can enhance the integration of research results into FSS practices. The main impediments are threefold: the hiring process is complex, the federal wage structure is often not competitive with the industry or academic labor markets, and many students are foreign nationals and therefore not eligible under current rules. One mechanism to consider is a periodic virtual hiring seminar for math- and data-oriented students, perhaps run jointly by FSS agencies under the auspices of the Chief Statistician.

Fifth, big data is ubiquitous in the lives of households and businesses. NCRN research is helping statistical agencies implement the use of non-designed data in official statistics and helping them to be better prepared for ongoing changes that are inevitable as agencies rely less on surveys and more on naturally occurring data.

In closing, we note the (inevitable) challenges of managing a network comprising researchers from many disciplines spread across both academia and government. Breaking the disciplinary silos, to engage in true cross-disciplinary research, is a challenge under any circumstances, and previous NSF-funded networks have certainly encountered the same challenges. Add to that the difficulty of bridging the gap between theory and practice, and the various gaps in expectations between academic researchers and government practitioners, and it is clear that any such project can take a while to produce results. Moreover, the path from preliminary results to applied research is sometimes hard to execute, even if it is a clear goal of the academic researcher. A key insight is to keep the network participants talking with one another and the sponsoring agencies; the NCRN's semi-annual meetings were more frequent than those of many other networks, and hence they may have led to a faster convergence of ideas and language.

Overcoming the challenges to cross-disciplinary collaboration created a unique research situation. NSF often recognizes the long-term aspect of creating effective collaborations when creating centers of excellence, but these are not typically initiated in collaboration with a non-grant-making agency like the Census Bureau, and the budgetary intricacies of an NSF-agency collaboration are challenging. Nevertheless, any future attempt at creating a network similar in scale and breadth to the NCRN should consider addressing the budgetary issues for at least a 10-year horizon.

REFERENCES

Abowd, J. M., and Schmutte, I. M. (forthcoming), "An Economic Analysis of Privacy Protection and Statistical Accuracy as Social Choices," *American Economic Review*. https://doi.org/10.1257/aer.20170627.

Abowd, J. M., and Schmutte, I. M. (2016), "Economic Analysis and Statistical Disclosure Limitation," *Brookings Papers on Economic Activity (Spring)*, 221–293.

Abowd, J. M., and Schmutte, I. M. (2017), "Revisiting the Economics of Privacy: Population Statistics and Confidentiality Protection as Public Goods.," Labor Dynamics Institute Document 37.

Abowd, J. M., Vilhuber, L., and Block, W. C. (2012), "A Proposed Solution to the Archiving and Curation of Confidential Scientific Inputs," *Privacy in Statistical Databases*, 216–225.

Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015), "Privacy and Human Behavior in the Age of Information," *Science*, 347, 509–514.

Acquisti, A., Taylor, C., and Wagman, L. (2016), "The Economics of Privacy," *Journal of Economic Literature*, 54, 442–492.

Aladangady, A., Shifrah, A.-D., Dunn, W., Feiveson, L., Lengermann, P., and Sahm, C. (2016), "The Effect of Hurricane Matthew on Consumer Spending," Board of Governors of the Federal Reserve System.

Antenucci, D., Cafarella, M., Levenstein, M., Ré, C., and Shapiro, M. D. (2013), *Ringtail: Feature Selection for Easier Nowcasting*, WebDB, https://www.cs.stanford.edu/people/chrismre/papers/webdb_ringtail.pdf.

Antenucci, D., Cafarella, M., Levenstein, M., Ré, C., and Shapiro, M. D. (2014), *Using Social Media to Measure Labor Market Flows*, Working Paper, NBER, http://www.nber.org/papers/w20010.

Arunhachalam, H., Atkin, G., Eck, A., Wettlaufer, D., Soh, L.-K., and Belli, R. F. (2015), "I Know What You Did Next: Predicting Respondent's Next Activity Using Machine Learning.," presented at the *70th Annual Conference of the American Association for Public Opinion Research (AAPOR)*.

Atkin, G., Arunachalam, H., Eck, A., Soh, L.-K., and Belli, R. F. (2014), "Designing an Intelligent Time Diary Instrument: Visualization, Dynamic Feedback, and Error Prevention and Mitigation.," presented at the *69th Annual Conference of the American Association for Public Opinion Research (AAPOR)*, Anaheim CA.

Atkin, G., Arunachalam, H., Eck, A., Wettlaufer, D., Soh, L.-K., and Belli, R. F. (2015), "Using Machine Learning Techniques to Predict Respondent Type from A Priori Demographic Information.," presented at the *70th Annual Conference of the American Association for Public Opinion Research (AAPOR)*, Hollywood FL.

Baghal, A., Tarek, R. F. B., Phillips, A. L., and Ruther, N. (2014), "What are You Doing Now? Activity Level Responses and Errors in the American Time Use Survey.," *Journal of Survey Statistics and Methodology*, 2, 519–537.

Belli, R. F., and Baghal, T. A. (2016), "Parallel Associations and the Structure of Autobiographical Knowledge.," *Journal of Applied Research in Memory and Cognition*, 5, 150–157.

Belli, R. F., Bilgen, I., and Baghal, T. A. (2013), "Memory, Communication, and Data Quality in Calendar Interviews.," *Public Opinion Quarterly*, 77, 194–219.

Belli, R. F., Miller, L. D., Baghal, T. A., and Soh, L.-K. (2016), "Using Data Mining to Predict the Occurrence of Respondent Retrieval Strategies in Calendar Interviewing: The Quality of Retrospective Reports," *Journal of Official Statistics*, 32, 579–600.

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

Annals of Applied Statistics, 9, 1761–1791.
Bradley, J. R., Holan, S. H., and Wikle, C. K. (2016a), "Multivariate Spatio-Temporal Survey Fusion with Application to the American Community Survey and Local Area Unemployment Statistics.," *STAT*, 5, 224–233.
Bradley, J. R., Holan, S. H., and Wikle, C. K. (2017), *Bayesian Hierarchical Models with Conjugate Full-Conditional Distributions for Dependent Data from the Natural Exponential Family.*, arXiv, , https://arxiv.org/abs/1701.07506.
Bradley, J. R., Holan, S. H., and Wikle, C. K. (2018), "Computationally Efficient Multivariate Spatio-Temporal Models for High-Dimensional Count-Valued Data (with Discussion)," *Bayesian Analysis*, 13, 253--310. https://doi.org/10.1214/17-BA1069.
Bradley, J. R., Wikle, C. K., and Holan, S. H. (forthcoming), "Hierarchical Models for Spatial Data with Errors that are Correlated with the Latent Process," *Statistica Sinica*. https://doi.org/10.5705/ss.202016.0230.
Bradley, J. R., Wikle, C. K., and Holan, S. H. (2015b), "Spatio-Temporal Change of Support with Application to American Community Survey Multi-Year Period Estimates," *Stat*, 4, 255–270. https://doi.org/10.1002/sta4.94.
Bradley, J. R., Wikle, C. K., and Holan, S. H. (2016b), "Bayesian Spatial Change of Support for Count-Valued Survey Data with Application to the American Community Survey.," *Journal of the American Statistical Association*, 111, 472–487.
Callier, V. (2015), "How Fake Data Could Protect Real People's Privacy," *The Atlantic*.
Charoenruk, N., and Olson, K. (forthcoming), "Do Listeners Perceive Interviewers' Attributes from their Voices and Do Perceptions Differ by Question Type?," *Field Methods*.
Chen, Y., Machanavajjhala, A., Reiter, J. P., and Barrientos, A. F. (2016), "Differentially Private Regression Diagnostics.," *IEEE 16th International Conference on Data Mining (ICDM)*, 81–90. https://doi.org/10.1109/ICDM.2016.0019.
Cressie, N., and Zammit-Mangion (2016), "Multivariate Spatial Covariance Models: A Conditional Approach," *Biometrika*, 103, 915–935.
Dominitz, J., and Manski, C. F. (2017), "More Data or Better Data? A Statistical Decision Problem," *Review of Economic Studies*, 84, 1583–1605. https://doi.org/10.1093/restud/rdx005.
Eck, A., and Soh, L.-K. (2017), "Sequential Prediction of Respondent Behaviors Leading to Error in Web-based Surveys," presented at the *Annual meeting of the American Association for Public Opinion Research*, New Orleans LA.
Eck, A., Soh, L.-K., McCutcheon, A. L., and Belli, R. F. (2015), "Predicting Breakoff Using Sequential Machine Learning Methods," presented at the *Annual meeting of the American Association for Public Opinion Research*, Hollywood FL.
Fay, I., Robert E., and Herriot, R. A. (1979), "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data," *Journal of the American Statistical Association*, 74, 269–277.
Fellegi, I. P., and Holt, D. (1976), "A Systematic Approach to Automated Edit and Imputation," *Journal of the American Statistical Association*, 71, 17–35.
Fellegi, I. P., and Sunter, A. B. (1969), "A Theory for Record Linkage," *Journal of the American Statistical Association*, 40, 1163–1210.
Fienberg, S. (2014), "Envisioning the 2030 Census," *23rd Annual Morris Hansen Lecture*.
Fienberg, S. E. (2015), "Discussion [of Special Issue on Coverage Problems in Administrative Sources].," *Journal of Official Statistics*, 31, 527–535.
Flaaen, A., Shapiro, M. D., and Sorkin, I. (forthcoming), "Reconsidering the Consequences of Worker Displacements: Firm versus Worker Perspective," *American Economic Journal: Macroeconomics*.

https://doi.org/10.3386/w24077.
Folch, D. C., Arribas-Bel, D., Koschinsky, J., and Spielman, S. E. (2016), “Spatial Variation in the Quality of American Community Survey Estimates.,” *Demography*, 53, 1535–1554.
Folch, D. C., and Spielman, S. E. (2014), “Identifying Regions Based on Flexible User Defined Constraints.,” *International Journal of Geographical Information Science*, 28, 164–184. https://doi.org/doi: 10.1080/13658816.2013.848986.
Fowler, C., Spielman, S. E., Folch, D. C., and Nagle, N. N. (2018), *Who Are the People in My Neighborhood? The ‘Contextual Fallacy’ of Measuring Individual Context with Census Geographies*, Working Papers, U.S. Census Bureau, Center for Economic Studies, https://ideas.repec.org/p/cen/wpaper/18-11.html.
Gelman, M., Gorodnichenko, Y., Kariv, S., Koustas, D., Shapiro, M., Silverman, D., and Tadelis, S. (2016), *The Response of Consumer Spending to Changes in Gasoline Prices*, Working Paper, Cambridge, MA: NBER, http://www.nber.org/papers/w22969.pdf. https://doi.org/10.3386/w22969.
Gelman, M., Kariv, S., Shapiro, M. D., Silverman, D., and Tadelis, S. (forthcoming), “How Individuals Smooth Spending: Evidence from the 2013 Government Shutdown Using Account Data.,” *Journal of Public Economics*.
Gelman, M., Kariv, S., Shapiro, M. D., Silverman, D., and Tadelis, S. (2014), “Harnessing Naturally Occurring Data to Measure the Response of Spending to Income.,” *Science*, 345, 212–215.
Ghosh-Dastidar, B., and Schafer, J. L. (2003), “Multiple Edit/Multiple Imputation for Multivariate Continuous Data,” *Journal of the American Statistical Association*, 98, 807–817.
Green, A. S., Kutzbach, M. J., and Vilhuber, L. (2017), *Two Perspectives on Commuting: A Comparison of Home to Work Flows Across Job-Linked Survey and Administrative Files.*, Papers, U.S. Census Bureau Center for Economic Studies Discussion, https://ideas.repec.org/p/cen/wpaper/17-34.html.
Griffin, A. L., Spielman, S. E., Jurjevich, J., Merrick, M., Nagle, N. N., and Folch, D. C. (2014), “Supporting Planners’ Work with Uncertain Demographic Data,” presented at the *GIScience 2014 Uncertainty Workshop*, Vienna, Austria, http://cognitivegiscience.psu.edu/uncertainty2014/papers/griffin_demographic.pdf.
Groves, R. M. an. B. A. H.-K. (ed.) (2017), *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy*, Washington DC: The National Academies Press.
Groves, R. M., and Harris-Kojetin, B. A. (eds.) (2017), *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy*, Washington DC: The National Academies Press.
Hu, J., Reiter, J. P., and Wang, Q. (2018), “Dirichlet Process Mixture Models for Modeling and Generating Synthetic Versions of Nested Categorical Data,” *Bayesian Analysis*, 13, 183–200. https://doi.org/10.1214/16-BA1047.
Hudomiet, P. (2015), “Four Essays in Unemployment, Wage Dynamics and Subjective Expectations.,” PhD Thesis, University of Michigan, http://hdl.handle.net/2027.42/113598.
Jurjevich, J., Griffin, A. L., Spielman, S. E., Folch, D. C., and Merrick, M. (2018), “Navigating Statistical Uncertainty: How Urban and Regional Planners Understand and Work with American Community Survey (ACS) Data for Guiding Policy,” *Journal of the American Planning Association*.
Kim, H. J., Cox, L. H., Karr, A. F., Reiter, J. P., and Wang, Q. (2015), “Simultaneous Editing and Imputation for Continuous Data.,” *Journal of the American Statistical Association*, 110, 987–999.
Kim, H. J., Reiter, J. P., and Karr, A. F. (2016), “Simultaneous Edit-Imputation and Disclosure Limitation for Business Establishment Data.,” *Journal of Applied Statistics*, 45. https://doi.org/10.1080/02664763.2016.1267123.
Kinney, S. K., Reiter, J. P., and Miranda, J. (2014), “SynLBD 2.0: Improving the Synthetic Longitudinal Business Database.,” *Statistical Journal of the International Association for Official Statistics*, 30,

129–135.
Kinney, S. K., Reiter, J. P., Reznek, A. P., Miranda, J., Jarmin, R. S., and Abowd, J. M. (2011), "Towards Unrestricted Public Use Business Microdata: The Synthetic Longitudinal Business Database," *International statistical review*, 79, 362–384. https://doi.org/10.1111/j.1751-5823.2011.00153.x.
Kirchner, A., Belli, R. F., Córdova-Cazar, A. L., and Deal, C. E. (forthcoming), "Memory Gaps in the American Time Use Survey: Investigating the Role of Retrieval Cues and Respondents' Level Of Effort," *Survey Research Methods.*
Kirchner, A., and Olson, K. (2017), "Experience or Cooperation? Examining Changes of Interview Length over the Course of the Field Period.," *Journal of Survey Statistics and Methodology*, 5, 84–108.
Kirchner, A., Olson, K., and Smyth, J. (2017), "Do Interviewer Post-Survey Evaluations of Respondents Measure Who Respondents Are or What They Do? A Behavior Coding Study.," *Public Opinion Quarterly forthcoming*, 81, 817–846.
Lagoze, C., Block, W. C., Williams, J., Abowd, J., and Vilhuber, and L. (2013a), "Data Management of Confidential Data," *International Journal of Digital Curation*, 8, 265–278.
Lagoze, C., Vilhuber, L., Williams, J., Perry, B., and Block, W. C. (2014), "CED2AR: The Comprehensive Extensible Data Documentation and Access Repository.," London, United Kingdom. https://doi.org/10.1109/JCDL.2014.6970178.
Lagoze, C., Willliams, J., and Vilhuber, L. (2013b), "Encoding Provenance Metadata for Social Science Datasets.," Communications in Computer and Information Science, eds. J. Greenberg and E. Garoufallou, Springer International Publishing.
Lucchesi, L. R., and Wikle, C. K. (2017), "Visualizing Uncertainty in Areal Data with Bivariate Choropleth Maps, Map Pixelation, and Glyph Rotation," *Stat*, 6, 292–302. https://doi.org/10.1002/sta4.150.
Manrique-Vallier, D., and Reiter, J. P. (2018), "Bayesian Simultaneous Edit and Imputation for Multivariate Categorical Data.," *Journal of the American Statistical Association*, 112, 1708–1719.
Manski, C. F. (2015), "Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern.," *Journal of Economic Literature*, 53, 631–53.
Manski, C. F. (2016), "Credible Interval Estimates for Official Statistics with Survey Nonresponse.," *Journal of Econometrics*, 191, 293–301.
McKinney, K., Green, A., Vilhuber, L., and Abowd, J. (2017), *Total Error and Variability Measures with Integrated Disclosure Limitation for Quarterly Workforce Indicators and LEHD Origin Destination Employment Statistics in OnThe Map*, Document, Labor Dynamics Institute, https://digitalcommons.ilr.cornell.edu/ldi/46.
Miranda, J., and Vilhuber, L. (2016), "Using Partially Synthetic Microdata to Protect Sensitive Cells in Business Statistics.," *Statistical Journal of the International Association for Official Statistics*, 32, 69–80.
Morgenstern, O. (1963), *On the Accuracy of Economic Observations*, Princeton: Princeton University Press.
Murray, J. S. (2015), "Probabilistic Record Linkage and Deduplication after Indexing, Blocking, and Filtering.," *Journal of Privacy and Confidentiality*, 7, 3–24.
Murray, J. S., and Reiter, J. P. (2016), "Multiple Imputation of Missing Categorical and Continuous Values via Bayesian Mixture Models with Local Dependence.," *Journal of the American Statistical Association*, 111, 1466–1479.
Olson, K., Cochran, B., and Smyth, J. D. (2018a), "Item Location, the Interviewer-Respondent Interaction, and Responses to Battery Questions in Telephone Surveys," *Sociological Methodology*, Advance online publication. https://doi.org/10.1177/0081175018778299.
Olson, K., Ganshert, A., and Smyth, J. D. (2018b), "The Effects of Respondent and Question

Characteristics on Respondent Answering Behaviors in Telephone Interviews," *Journal of Survey Statistics and Methodology*, Advance online publication. https://doi.org/10.1093/jssam/smy006.

Olson, K., Kirchner, A., and Smyth, J. D. (2016), "Do Interviewers with High Cooperation Rates Behave Differently? Interviewer Cooperation Rates and Interview Behaviors.," *Survey Practice*, 9, 1–11.

Olson, K., and Parkhurst, B. (2013), "Collecting Paradata for Measurement Error Evaluation," in *Improving Surveys with Paradata: Analytic Uses of Process Information*, ed. F. Kreuter, New York: John Wiley and Sons.

Olson, K., and Smyth, J. D. (2015), "The Effect of CATI Questionnaire Design Features on Response Timing.," *Journal of Survey Statistics and Methodology*, 3, 361–396.

Paiva, T., and Reiter, J. P. (2017), "Stop or Continue Data Collection: A Nonignorable Missing Data Approach to Continuous Data.," *Journal of Official Statistics*, 33, 579–599.

Park, S. H., Goo, J. M., and Jo, C.-H. (2004), "Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists," *Korean Journal of Radiology*, 5, 11. https://doi.org/10.3348/kjr.2004.5.1.11.

Porter, A. T., Holan, S. H., and Wikle, C. K. (2015a), "Multivariate Spatial Hierarchical Bayesian Empirical Likelihood Methods for Small Area Estimation.," *STAT*, 4, 108–116.

Porter, A. T., Holan, S. H., and Wikle, C. K. (2015b), "Bayesian Semiparametric Hierarchical Empirical Likelihood Spatial Models.," *Journal of Statistical Planning and Inference*, 165, 78–90.

Porter, A. T., Holan, S. H., Wikle, C. K., and Cressie, N. (2014), "Spatial Fay–Herriot Models for Small Area Estimation with Functional Covariates.," *Spatial Statistics*, 10, 27–42.

Porter, A. T., Wikle, C. K., and Holan, S. H. (2015c), "Small Area Estimation via Multivariate Fay-Herriot Models with Latent Spatial Dependence.," *Australian and New Zealand Journal of Statistics*, 57, 15–29.

Quick, H., Holan, S. H., and Wikle, C. K. (2015a), "Zeros and Ones: A Case for Suppressing Zeros in Sensitive Count Data with an Application to Stroke Mortality," *STAT*, 4, 255–270.

Quick, H., Holan, S. H., and Wikle, C. K. (2018), "Generating Partially Synthetic Geocoded Public Use Data with Decreased Disclosure Risk Using Differential Smoothing.," *Journal of the Royal Statistical Society - Series A*. https://doi.org/10.1111/rssa.12360.

Quick, H., Holan, S. H., Wikle, C. K., and Reiter, J. P. (2015b), "Bayesian Marked Point Process Modeling for Generating Fully Synthetic Public Use Data with Point-Referenced Geography.," *Spatial Statistics*, 14, 439–451.

Raim, A. M., Holan, S. H., Bradley, J. R., and Wikle, C. K. (2017), "A Model Selection Study for Spatio-Temporal Change of Support," ASA Proceedings of the Joint Statistical Meetings, American Statistical Association.

Sadinle, M. (2017), "Bayesian Estimation of Bipartite Matchings for Record Linkage.," *Journal of the American Statistical Association*, 112, 600–612.

Sadinle, M., and Fienberg, S. E. (2013), "A Generalized Fellegi-Sunter Framework for Multiple Record Linkage with Application to Homicide Record Systems," *Journal of the American Statistical Association*, 108, 385–397.

Sadinle, M., and Reiter, J. P. (2017), "Itemwise Conditionally Independent Nonresponse Modeling for Multivariate Categorical Data.," *Biometrika*, 104, 207–220.

Sadinle, M., and Reiter, J. P. (2018), "Sequential Identification of Nonignorable Missing Data," *Statistica Sinica*.

Seeskin, Z. H., and Spencer, B. D. (2015), *Effects of Census Accuracy on Apportionment of Congress and Allocations of Federal Funds.*, Working Paper, Northwestern University, Institute for Policy Research Working.

Seeskin, Z. H., and Spencer, B. D. (2018), *Balancing 2020 Census Cost and Accuracy: Consequences for*

Congressional Apportionment and Fund Allocations, Working Paper, Northwestern University, Institute for Policy Research Working.

Sengupta, A., and Cressie, N. (2013a), "Hierarchical Statistical Modeling of Big Spatial Datasets Using the Exponential Family of Distributions.," *Spatial Statistics*, 4, 14–44. https://doi.org/10.1016/j.spasta.2013.02.002.

Sengupta, A., and Cressie, N. (2013b), "Empirical Hierarchical Modelling for Count Data using the Spatial Random Effects Model," *Spatial Economic Analysis*, 8, 389–418. https://doi.org/10.1080/17421772.2012.760135.

Smyth, J., and Olson, K. (2018), "The Effects of Mismatches between Survey Question Stems and Response Options on Data Quality and Responses," *Journal of Survey Statistics and Methodology*. https://doi.org/10.1093/jssam/smy005.

Sorkin, I. (2018), "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*. https://doi.org/10.1093/qje/qjy001.

Spencer, B. D., May, J., Kenyon, S., and Seeskin, Z. (2017), "Cost-Benefit Analysis for a Quinquennial Census," *Journal of Official Statistics*, 33.

Spielman, S. E., and Folch, D. C. (2015), "Reducing Uncertainty in the American Community Survey through Data-Driven Regionalization," *PLOS ONE*, (A. R. Hernandez Montoya, ed.), 10, e0115626. https://doi.org/10.1371/journal.pone.0115626.

Spielman, S. E., Folch, D., and Nagle, N. (2014), "Patterns and Causes of Uncertainty in the American Community Survey.," *Applied Geography*, 46, 147–157.

Spielman, S. E., and Singleton, A. (2015), "Studying Neighborhoods Using Uncertain Data from the American Community Survey: A Contextual Approach.," *The Annals of the Association of American Geographers*, 102, 1003–1025.

Steorts, R. C., Hall, R., and Fienberg, S. E. (2016), "A Bayesian Approach to Graphical Record Linkage and Deduplication.," *Journal of the American Statistical Association*, 111, 1660–1672.

Timbrook, J., Smyth, J. D., and Olson, K. (2016), "Are Self-Description Scales Better than Agree/Disagree Scales in Mail and Telephone Surveys?," presented at the *Annual meeting of the Midwest Association for Public Opinion Research*, Chicago IL.

Timbrook, J., Smyth, J., and Olson, K. (forthcoming), "Why do Mobile Interviews Take Longer? A Behavior Coding Perspective.," *Public Opinion Quarterly*.

Tourangeau, R., and Plewes, T. J. (eds.) (2013), *Nonresponse in Social Science Surveys: A Research Agenda*, Washington, DC: The National Academies Press.

U.S. Bureau of Economic Analysis Advisory Committee (2017), "Measuring Quality Adjusted Prices in the 21st Century."

U.S. Federal Economic Statistics Advisory Committee. (2015), "Commercial Big Data and Official Economic Statistics."

Vilhuber, L., Abowd, J. M., and Reiter, J. P. (2016), "Synthetic Establishment Microdata Around the World.," *Statistical Journal of the International Association of Official Statistics*, 32, 65–68.

Wasi, N., and Flaaen, A. (2015), "Record Linkage Using Stata: Preprocessing, Linking, and Reviewing Utilities.," *Stata Journal*, 15, 672–697.

White, T. K., Reiter, J. P., and Petrin, A. (2018), "Imputation in U.S. Manufacturing Data and Its Implications for Productivity Dispersion," *Review of Economics and Statistics*, online. https://doi.org/10.1162/REST_a_00678.

Williams, D., and Brick, J. M. (2018), "Trends in U.S. Face-To-Face Household Survey Nonresponse and Level of Effort," *Journal of Survey Statistics and Methodology*, 6, 186–211. https://doi.org/10.1093/jssam/smx019.

Wilson, C. R., and Brown, D. G. (2015), "Change in Visible Impervious Surface Area in Southeastern

Michigan Before and After the ‘Great Recession’.,” *Population and Environment*, 36, 331–355.
Wood, N. J., Jones, J., Spielman, S., and Schmidtlein, M. C. (2015), “Community Clusters of Tsunami Vulnerability in the US Pacific Northwest.”
Yellen, J. R. (2017), *The Economic Outlook and the Conduct of Monetary Policy. Remarks at the Stanford Institute for Economic Policy. Text and video transcript.*