

CONTACT INFORMATION

RAND Corporation
4570 Fifth Ave.
Pittsburgh, PA 15213

jsnoke@rand.org
www.joshuasnoke.com

EDUCATION

Pennsylvania State University

Ph.D., Statistics

08.2018

Graduate Minor: Social Data Analytics

Dissertation: *Statistical Data Privacy Methods for Increasing Research Opportunities*

Advisor: Aleksandra B. Slavković

Wheaton College

B.S., Mathematics and Economics

05.2013

Honors: Magna Cum Laude

RESEARCH POSITIONS

RAND Corporation

Associate Statistician

08.2018 - Present

Research Adjunct

09.2016 - 08.2017

Summer Associate

06.2016 - 08.2016

Pennsylvania State University

Graduate Research Assistant, Department of Statistics

01.2018 - 06.2018

05.2017 - 08.2017

01.2014 - 08.2014

Graduate Research Trainee, Big Data Social Science IGERT

09.2014 - 05.2016

Administrative Data Research Centre - Scotland

Research Fellow

05.2015 - 08.2015

PEER REVIEWED PUBLICATIONS

Bowen, C. M. and **Snoke, J.** (2021). Comparative Study of Differentially Private Synthetic Data Algorithms from the NIST PSCR Differential Privacy Synthetic Data Challenge. *Journal of Privacy and Confidentiality* 11 (1). <https://doi.org/10.29012/jpc.748>.

Snoke, J., Raab, G. M., Nowok, B., Dibben, C. and Slavković, A. (2018). General and specific utility measures for synthetic data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(3), pp.663-688. doi:10.1111/rssa.12358

Snoke, J., Brick, T. R., Slavković, A. and Hunter, M. D. (2018). Providing accurate models across private partitioned data: Secure maximum likelihood estimation. *The Annals of Applied Statistics*, 12(2), pp.877-914. doi:10.1214/18-AOAS1171.

Snoke, J. and Slavković, A. (2018). *pMSE* Mechanism: Differentially Private Synthetic Data with Maximal Distributional Similarity. In *Privacy in Statistical Databases*. pp.138-159. Springer. doi:10.1007/978-3-319-99771-1

Snoke, J., Brick, T. R. and Slavković, A. (2016). Accurate Estimation of Structural Equation Models with Remote Partitioned Data. In *Privacy in Statistical Databases*. pp.190-209. Springer. doi:10.1007/978-3-319-45381-1

ADDITIONAL REPORTS & ARTICLES

Snoke, J. and Bowen, C. M. (2020). How Statisticians Should Grapple with Privacy in a Changing Data Landscape. *CHANCE*, 33(4), 6-13. <https://doi.org/10.1080/09332480.2020.1847947>

Kaufman, J. H., Diliberti, M., Hunter, G. P., Grant, D. M., Hamilton, L. S., Schwartz, H. L., Setodji, C. M., **Snoke, J.**, and Young, C. J. (2020). COVID-19 and the state of K-12 schools: results and technical documentation from the fall 2020 American Educator Panels COVID-19 surveys. *RAND Corporation*.

Ayer, L., Schultz, D. J., Abbott, M., Barnes-Proby, D., Chan, W. Y., Dunbar, M. S., Hoch, E., Liu, H. H., Martineau, M., Ohana, E., Siconolfi, D., **Snoke, J.**, Stevens, C., and Towe, V. L. (2020). Mental health task-shifting in community-based organizations: implementation, impact, and cost-evaluation of the Connections to Care Program. *RAND Corporation*.

Snoke, J. and Bowen, C. M. (2019). Differential Privacy: What Is It? *Amstat News*. 26-28. Issue #501. <https://magazine.amstat.org/blog/2019/03/01/differentialprivacy/>

Ororbia II, A. G., Linder, F. and **Snoke, J.**. Using Neural Generative Models to Release Synthetic Twitter Corpora with Reduced Stylometric Identifiability of Users. Preprint available at <https://arxiv.org/abs/1606.01151>

PROFESSIONAL ACTIVITIES

Professional Service

Human Subjects Protection Committee 06.2019 - Present
RAND Corporation

Committee on Privacy and Confidentiality 10.2018 - Present
American Statistical Association (ASA)

Subject Matter Expert for Differentially Private Synthetic Data Challenges 10.2018 - 05.2019
National Institute of Standards and Technology (NIST) 10.2020 - Present

Academic Journal Reviewer

Annals of Applied Statistics
Biostatistics
Harvard Data Science Review
IEEE Transactions on Information Forensics & Security
Journal of Machine Learning
Journal of Privacy and Confidentiality
Journal of Research on Educational Effectiveness
Journal of the Royal Statistical Society: Series A
Journal of Survey Statistics and Methodology
Statistical Science
Statistics and Computing
Science Advances

Conference Program Committees

Symposium on Data Science and Statistics (SDSS) 06.2020

ACM-IMS Foundations of Data Science (FODS) 10.2020

FELLOWSHIPS & AWARDS

Spotlight Award 01.2021
For Leadership within the RAND Statistics Group
RAND Corporation

Dissertation Fellowship 08.2016 - 07.2017
U.S. Census Bureau

NSF Big Data Social Science IGERT Traineeship 09.2014 - 08.2016
Pennsylvania State University

Joint Statistical Meetings ASA Student Paper Award 07.2016
General and Specific Utility Measures for Synthetic Data
Social Statistics, Government Statistics, and Survey Research Methods Section

Rao Prize Conference Student Poster Award 05.2015
Estimation of Structural Equation Models for Vertically Partitioned Data
Pennsylvania State University

President's Award 08.2009 - 05.2013
Wheaton College

GRANTS & CONTRACTS

Development and Deployment of a Validation Server for Administrative Tax Data
Funder: Urban Institute, Prime: Sloan Foundation - \$100,000 11.2020 - 10.2022

Planning and Development of a Prototype Privacy-Preserving Validation Server for Administrative Tax Data
Funder: Urban Institute, Prime: National Science Foundation - \$15,000 10.2020 - 09.2021

Comparative Study of Differentially Private Synthetic Data Algorithms and Evaluation Standards
Funder: National Institute of Standards and Technology - \$100,000 03.2020 - 11.2020

SELECTED PRESENTATIONS

Invited Talks

Society for Industrial and Applied Mathematics: Mathematics of Data Science 05.2020
Comparative Study of Differentially Private Synthetic Data Algorithms and Evaluation Standards

Simons Institute Program on Data Privacy: Foundations and Applications 03.2019
Statistical Perspectives on Differentially Private Synthetic Data

International Conference on Privacy in Statistical Databases 09.2018
pMSE Mechanism: Differentially Private Synthetic Data with Maximal Distributional Similarity

U.S. Census Bureau 06.2018
Statistical Data Privacy Methods for Increasing Research Opportunities

Joint Statistical Meetings 07.2017
Discussant. Differential Privacy in Statistical Agencies: Present and Future

International Statistical Institute World Statistics Congress <i>Perspective on Utility for Synthetic Data Sets</i>	07.2017
International Conference on Privacy in Statistical Databases <i>Accurate Estimation of Structural Equation Models with Remote Partitioned Data</i>	09.2016
Joint Statistical Meetings <i>General and Specific Utility Measures for Synthetic Data</i>	08.2016
Workshop Talks	
Stochastic Modeling and Computational Statistics Seminar <i>Secure Multiparty Maximum Likelihood Estimation with Partitioned Databases</i>	04.2017
Isaac Newton Institute Synthetic Data Workshop <i>Beyond Microdata: Approaches for Generating Synthetic Tweets</i>	11.2016

TEACHING APPOINTMENTS

Pennsylvania State University

Graduate Teaching Assistant

Categorical Data Analysis

08.2017 - 12.2017

Applied Data Mining and Statistical Learning

08.2017 - 12.2017

Elementary Statistics

08.2013 - 12.2013

Wheaton College

Teaching Assistant

Intermediate Macroeconomics

08.2011 - 05.2012

Calculus II

01.2011 - 05.2011

COMPUTING SKILLS

Software

Proficient in R, Python, Git, Excel

Familiarity with SAS, STATA, Minitab, Java, SQL

Development

Contributor to R package *synthpop*