

## CONTACT INFORMATION

RAND Corporation  
4570 Fifth Ave.  
Pittsburgh, PA 15213

jsnoke@rand.org  
www.joshuasnoke.com

## EDUCATION

**Pennsylvania State University**, University Park, PA  
*Ph.D.*, Statistics, August 2018  
Graduate Minor: Social Data Analytics  
Dissertation: *Statistical Data Privacy Methods for Increasing Research Opportunities*  
Advisor: Aleksandra B. Slavković

**Wheaton College**, Wheaton, IL  
*B.S.*, Mathematics and Economics, May 2013  
Honors: Magna Cum Laude

## PROFESSIONAL APPOINTMENTS

**RAND Corporation**, Pittsburgh, PA & Santa Monica, CA  
Associate Statistician August 2018 - Present  
Research Adjunct September 2016 - August 2017  
Summer Associate June 2016 - August 2016

**Pennsylvania State University**, University Park, PA  
Graduate Research Assistant  
*Department of Statistics* January 2018 - June 2018  
May 2017 - August 2017  
January 2014 - August 2014

Graduate Research Trainee  
*Big Data Social Science IGERT* September 2015 - May 2016  
September 2014 - May 2015

Graduate Teaching Assistant  
*Categorical Data Analysis* August 2017 - December 2017  
*Applied Data Mining and Statistical Learning* August 2017 - December 2017  
*Elementary Statistics* August 2013 - December 2013

**Administrative Data Research Centre - Scotland**, Edinburgh, UK  
Visiting Research Fellow May 2015 - August 2015

## PROFESSIONAL ACTIVITIES

**Journal Referee**  
*Journal of Survey Statistics and Methodology, IEEE Transactions on Information Forensics & Security, Statistical Science*

**Committee Member**  
*American Statistical Association (ASA)* October 2018 - Present  
Committee on Privacy and Confidentiality in Statistics

### Program Planning Committee

*Symposium on Data Science and Statistics 2020*

October 2019 - June 2020

### Subject Matter Expert

*National Institute of Standards and Technology*  
Differentially Private Synthetic Data Challenge

October 2018 - May 2019

## PEER REVIEWED PUBLICATIONS

**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. 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

## NON-PEER REVIEWED ARTICLES

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

## PAPERS IN PREP

Bowen, Claire M. and **Snoke, J.**. Comparative Study of Differentially Private Synthetic Data Algorithms and Evaluation Standards.

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>.

## FELLOWSHIPS & AWARDS

### Fellowships

Dissertation Fellowship, U.S. Census Bureau

August 2016 - July 2017

NSF Big Data Social Science IGERT Traineeship,  
Pennsylvania State University

September 2014 - August 2016

President's Award, Wheaton College

August 2009 - May 2013

### Research Awards

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

July 2016

Rao Prize Conference Student Poster Award May 2015  
*Estimation of Structural Equation Models for Vertically Partitioned Data*

## PRESENTED WORK

### Invited Talks

Simons Institute Program on Data Privacy: Foundations and Applications March 2019  
*Statistical Perspectives on Differentially Private Synthetic Data*

International Conference on Privacy in Statistical Databases September 2018  
*pMSE Mechanism: Differentially Private Synthetic Data with Maximal Distributional Similarity*

U.S. Census Bureau June 2018  
*Statistical Data Privacy Methods for Increasing Research Opportunities*

Joint Statistical Meetings July 2017  
*Discussant. Differential Privacy in Statistical Agencies: Present and Future*

International Statistical Institute World Statistics Congress July 2017  
*Perspective on Utility for Synthetic Data Sets*

International Conference on Privacy in Statistical Databases September 2016  
*Accurate Estimation of Structural Equation Models with Remote Partitioned Data*

Joint Statistical Meetings August 2016  
*General and Specific Utility Measures for Synthetic Data*

### Guest Lectures

Data Privacy Methods, Pennsylvania State University October 2014  
*Statistical Disclosure Control Methods*

### Contributed Talks

Joint Statistical Meetings August 2019  
*pMSE Mechanism: A Statistical Approach to Differentially Private Synthetic Data*

Joint Statistical Meetings August 2018  
*Using Neural Generative Models to Release Synthetic Twitter Corpora with Reduced Stylometric Identifiability of Users*

Joint Statistical Meetings July 2017  
*Secure Multiparty Maximum Likelihood Estimation with Partitioned Databases*

RAND Corporation August 2016  
*Social Media and Events Marketing for Army Recruiting*

FARR Institute International Conference August 2015  
*Utility of Synthetic Microdata for Health Research*

### Workshop Talks

Stochastic Modeling and Computational Statistics Seminar April 2017  
*Secure Multiparty Maximum Likelihood Estimation with Partitioned Databases*

Isaac Newton Institute Synthetic Data Workshop November 2016  
*Beyond Microdata: Approaches for Generating Synthetic Tweets*

### Poster Presentations

Department of Statistics 50th Anniversary Conference <i>Differentially Private Synthetic Data with Maximal Distributional Similarity</i>	May 2018
Big Data Social Science IGERT Poster Session <i>Statistical Disclosure Control for High Dimensional Data</i>	April 2016
Rao Prize Conference Big Data Social Science IGERT Poster Session <i>Estimation of Structural Equation Models for Vertically Partitioned Data</i>	May 2015 April 2015

## COMPUTING SKILLS

### Statistical Software

Proficient in R, SAS, STATA, Minitab

### Other Tools

Proficient in Git, Python, Excel  
Experience in Java, SQL

### Development

Contributor to R package *synthpop*

## ADDITIONAL PROFESSIONAL DEVELOPMENT

<b>Neighborhood Learning Alliance</b> , Pittsburgh, PA Summer Tutor	June 2013 - August 2013
<b>Boston University</b> , Boston, MA Summer Institute for Training in Biostatistics	June 2012 - July 2012
<b>Via Science</b> , Boston, MA Data Science Intern	August 2012 - November 2012
<b>Wheaton College</b> , Wheaton, IL Teaching Assistant <i>Intermediate Macroeconomics</i> <i>Calculus II</i>	August 2011 - May 2012 January - May 2011

## LAST UPDATED

November 19, 2019