

## 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

Research Adjunct

Summer Associate

August 2018 - Present

September 2016 - August 2017

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

**Committee Member**

*American Statistical Association (ASA)*

Committee on Privacy and Confidentiality in Statistics

October 2018 - Present

**Journal Referee**

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

## 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

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  
*Estimation of Structural Equation Models for Vertically Partitioned Data*

May 2015

## PRESENTED WORK

### Invited Talks

Simons Institute Program on Data Privacy: Foundations and Applications <i>Statistical Perspectives on Differentially Private Synthetic Data</i>	March 2019
International Conference on Privacy in Statistical Databases <i>pMSE Mechanism: Differentially Private Synthetic Data with Maximal Distributional Similarity</i>	September 2018
U.S. Census Bureau <i>Statistical Data Privacy Methods for Increasing Research Opportunities</i>	June 2018
Joint Statistical Meetings <i>Discussant. Differential Privacy in Statistical Agencies: Present and Future</i>	July 2017
International Statistical Institute World Statistics Congress <i>Perspective on Utility for Synthetic Data Sets</i>	July 2017
International Conference on Privacy in Statistical Databases <i>Accurate Estimation of Structural Equation Models with Remote Partitioned Data</i>	September 2016
Joint Statistical Meetings <i>General and Specific Utility Measures for Synthetic Data</i>	August 2016
<b>Guest Lectures</b>	
Data Privacy Methods, Pennsylvania State University <i>Statistical Disclosure Control Methods</i>	October 2014
<b>Contributed Talks</b>	
Joint Statistical Meetings <i>Using Neural Generative Models to Release Synthetic Twitter Corpora with Reduced Stylometric Identifiability of Users</i>	August 2018
Joint Statistical Meetings <i>Secure Multiparty Maximum Likelihood Estimation with Partitioned Databases</i>	July 2017
RAND Corporation <i>Social Media and Events Marketing for Army Recruiting</i>	August 2016
FARR Institute International Conference <i>Utility of Synthetic Microdata for Health Research</i>	August 2015
<b>Workshop Talks</b>	
Stochastic Modeling and Computational Statistics Seminar <i>Secure Multiparty Maximum Likelihood Estimation with Partitioned Databases</i>	April 2017
Isaac Newton Institute Synthetic Data Workshop <i>Beyond Microdata: Approaches for Generating Synthetic Tweets</i>	November 2016
<b>Poster Presentations</b>	
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	May 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

**Neighborhood Learning Alliance**, Pittsburgh, PA

Summer Tutor

June 2013 - August 2013

**Boston University**, Boston, MA

Summer Institute for Training in Biostatistics

June 2012 - July 2012

**Via Science**, Boston, MA

Data Science Intern

August 2012 - November 2012

**Wheaton College**, Wheaton, IL

Teaching Assistant

*Intermediate Macroeconomics*

*Calculus II*

August 2011 - May 2012

January - May 2011

## LAST UPDATED

March 28, 2019