

Department of Mathematics and Statistics
 Box 27, 124 Raymond Ave
 Vassar College, Poughkeepsie, NY 12604

Phone: (845) 518 6499
 Email: jihu@vassar.edu
 Web: <http://pages.vassar.edu/jihu/>

EDUCATION

Ph.D. in Statistical Science, *Duke University* 2015
 Advisor: Jerome P. Reiter
 Title: “Dirichlet Process mixture models for nested categorical data”

Certificate of College Teaching, *Duke Graduate School* 2015

M.S. in Statistical Science, *Duke University* 2013

B.S. in Computing Mathematics, *City University of Hong Kong* 2011
 First class honor, minor in Finance

POSITIONS

Associate Professor 2022 - present
Department of Mathematics and Statistics, Vassar College

Assistant Professor 2015 - 2022
Department of Mathematics and Statistics, Vassar College

Faculty Fellow 2020 - 2021
The National Center for Science and Engineering Statistics, National Science Foundation

Research Assistant 2012 - 2015
Department of Statistical Science, Duke University

Teaching Assistant 2011 - 2015
Department of Statistical Science, Duke University

Summer Intern Summer 2013
IIS Statistical Methodology Group, Novartis

GRANTS, FELLOWSHIPS, HONORS & AWARDS

SCHOLARLY RESEARCH

NSF National Center for Science and Engineering Statistics
 Project title: Advancing Communication, Collaboration, and Evidence-building through Synthetic Survey of Earned Doctorates (ACCESS).
 Role: Co-Principle Investigator.
 Period: Oct. 2023 - Oct. 2025. Amount: \$800,000.00

Alfred P. Sloan Foundation
 Project title: Statistical Privacy and Public Policy Workshop Series.

Role: Principle Investigator.

Period: 10/12/2022 - 12/31/2024. Amount: \$49,948.00

United States Department of Agriculture Economic Research Service

Project title: ERS Disclosure Review Assistance.

Role: Principle Investigator.

Period: 10/1/2022 - 9/30/2024. Amount: \$134,712.00

NSF National Center for Science and Engineering Statistics

Project title: Modern Disclosure Limitation and Privacy Protection Methods for Survey Data.

Role: Principle Investigator.

Period: 6/1/2020 - 12/31/2021. Amount: \$56,732.40

OpenDP Fellow, The Institute for Quantitative Social Science, Harvard University

Period: 6/7/2021 - 9/7/2021.

Research Committee Award, Vassar College

Role: Principal Investigator.

Period: 5/6/2019 - 12/31/2020. Amount: \$3,871.00

NSF ASA/NSF/BLS Research Fellowship

Project title: Synthetic Consumer Expenditure Survey Data at BLS.

Role: Principal Investigator.

Period: 2/1/2018 - 12/31/2019. Amount: \$88,531.00

Center for International Research on the Japanese Economy, University of Tokyo

Role: Visiting Scholar.

Period: 1/11/2018 - 1/25/2018. Amount: fully funded

Research Committee Award, Vassar College

Role: Principal Investigator.

Period: 5/5/2017 - 12/31/2018. Amount: \$2,940.45

ASA Student Paper Competition Award

2015

Joint Statistical Meetings

Institute for Employment Research (IAB), Germany

Role: Visiting Scholar.

Period: 7/1/2014 - 7/31/2014. Amount: fully funded

STATISTICS EDUCATION

Faculty Advisor to Honorable Mention Winner

2022

The Undergraduate Research Project Competition (USRESP), CAUSE & ASA

National Science Foundation

Project title: Advancing Bayesian Thinking in STEM.

Role: Principal Investigator (PIs: Dogucu (UC Irvine) and Herring (Duke University)).

Period: 12/15/2022 - 11/30/2024. Amount: \$300,000.00

Liberal Arts Collaborative for Digital Innovation

Project title: Introduction to Data Science.

Role: Co-Principal Investigator.

Period: 1/1/2019 - 9/30/2022. Amount: \$207,300.00

Liberal Arts Collaborative for Digital Innovation

Project title: Bayesian Statistics.

Role: Principal Investigator.

Period: 10/3/2018 - 6/30/2022. Amount: \$27,500.00

Grand Challenge Climate Change, Vassar College

Project title: CMPU/MATH 144 Foundations of Data Science.

Role: Co-Principal Investigator.

Period: 6/1/2019 - 6/30/2021. Amount: \$11,000.00

Faculty Advisor to Honorable Mention Winner

2020

*The Undergraduate Research Project Competition (USRESP), CAUSE & ASA***Liberal Arts Collaborative for Digital Innovation**

Project title: Data Confidentiality.

Role: Principal Investigator.

Period: 1/1/2020 - 6/30/2020. Amount: \$4,500.00

Liberal Arts Collaborative for Digital Innovation

Project title: Bayesian Inference with Python.

Role: Principal Investigator.

Period: 1/1/2020 - 6/30/2020. Amount: \$3,000.00

Faculty Advisor to Student 1st-place Winner

2019

*The Undergraduate Class Project Competition (USCLAP), intermediate statistics, CAUSE & ASA***Endowment for Strategic Faculty Support, Vassar College**

Role: Principal Investigator.

Period: 4/18/2017 - 12/31/2017. Amount: \$1,287.75

Faculty Advisor to Student 1st-place Winner

2016

*The Undergraduate Class Project Competition (USCLAP), intermediate statistics, CAUSE & ASA***OTHERS****Faculty Success Program, National Center for Faculty Development & Diversity**

Role: Participant.

Period: 1/1/2018 - 6/30/2018. Amount: fully funded

STATISTICAL CONSULTING**Statistical consultant (data confidentiality)***New York City Department of Health and Mental Hygiene*

2022 - present

Statistical consultant (missing data)*resonate*

2020

PUBLICATIONS

PEER-REVIEWED PAPERS AND CONFERENCE PROCEEDINGS

* indicates an undergraduate student co-author

SCHOLARLY RESEARCH

33. **Hu, J.** and Bowen, C. M. (forthcoming), “Advancing microdata privacy protection: a review of synthetic data”, to appear in *WIREs Computational Statistics*. [arXiv link](#)
32. **Hu, J.**, Williams, M. R., and Savitsky, T. D. (forthcoming), “Mechanisms for global differential privacy under Bayesian data synthesis”, to appear in *Data Privacy special issue at Statistica Sinica*. [arXiv link](#)
31. **Hu, J.** and Savitsky, T. D. (2023), “Bayesian data synthesis and disclosure risk quantification: an application to the Consumer Expenditure Surveys”, *Transactions on Data Privacy*, 16:2, 83-121. [Open Access](#)
30. Schneider, M. J, **Hu, J.**, Mankad, S., and Bale C. D. (2023), “Protecting the anonymity of online users through Bayesian data synthesis”, *Expert Systems With Applications*, 216, 119409.
29. Guo, S.* and **Hu, J.** (2023), “Data privacy protection through Bayesian data synthesis: a case study on Airbnb listings”, *The American Statistician*, 77(2), 192-200. [link to the published paper](#)
28. **Hu, J.**, Savitsky, T. D. and Williams, M. R. (2022), “Risk-efficient Bayesian data synthesis for privacy protection”, *Journal of Survey Statistics and Methodology*, 10(5), 1370-1399. [link to the published paper](#)
27. Cao, Y.* and **Hu, J.** (2022), “Privacy protection for youth risk behavior using Bayesian data synthesis: a case study to the YRBS”, *Privacy in Statistical Databases e-Proceedings*. [link to the published paper](#)
26. **Hu, J.**, Drechsler, J. and Kim, H. J. (2022), “Accuracy gains from privacy amplification through sampling for differential privacy”, *Journal of Survey Statistics and Methodology, Special Issue: Privacy, Confidentiality, and Disclosure Protection*, 10(3), 688-719. [link to the published paper](#)
25. **Hu, J.**, Savitsky, T. D. and Williams, M. R. (2022), “Private tabular survey data products through synthetic microdata generation”, *Journal of Survey Statistics and Methodology, Special Issue: Privacy, Confidentiality, and Disclosure Protection*, 10(3), 720-752. [link to the published paper](#)
24. Savitsky, T. D., Williams, M. R. and **Hu, J.** (2022), “Bayesian pseudo posterior mechanism under differential privacy”, *Journal of Machine Learning Research*, 23(55), 1-37. [Open Access](#)
23. **Hu, J.**, Akande, O., and Wang, Q. (2021), “Data imputation and data synthesis with the R package NPBayesImputeCat”, *The R Journal*, 13:2, 90-110. [Open Access](#)
22. Drechsler, J. and **Hu, J.** (2021), “Synthesizing geocodes to facilitate access to detailed geographical information in large-scale administrative data”, *Journal of Survey Statistics and Methodology*, 9(3), 523-548. [Open Access](#)
21. Hornby, R.* and **Hu, J.** (2021), “Identification risks evaluation of partially synthetic data with the IdentificationRiskCalculation R package”, *Transactions of Data Privacy*, 14:1, 37-52. [Open Access](#)

20. **Hu, J.**, Savitsky, T. D. and Williams, M. R. (2020), “Risk-weighted data synthesizers for microdata dissemination”, *Special Issue: A New Generation of Statisticians Tackles Data Privacy, CHANCE*, 33(4), 29-36. Open Access
19. Ros, K.* , Olsson, H.* and **Hu, J.** (2020), “Two-phase data synthesis for income: an application to the NHIS”, *Privacy in Statistical Databases e-Proceedings*. arxiv link
18. **Hu, J.** (2019), “Bayesian estimation of attribute and identification disclosure risks in synthetic data”, *Transactions on Data Privacy*, 12:1, 61-89. Open Access
17. **Hu, J.** and Hoshino, N. (2018), “The Quasi-Multinomial synthesizer for categorical data”, *Privacy in Statistical Databases (Lecture Notes in Computer Science 11126)*, ed. J. Domingo-Ferrer and F. Montes, Springer, 75-91.
16. Manrique-Vallier, D. and **Hu, J.** (2018), “Bayesian non-parametric generation of synthetic multivariate categorical data in the presence of structural zeros”, *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 181(3), 635-647.
15. **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(1), 183-200. Open Access
14. **Hu, J.** and Drechsler, J. (2015), “Generating synthetic geocoding information for public release”, In: *S. A. Europäische Kommission (Hrsg.), NTTS -Conferences on New Techniques and Technologies for Statistics*, 56-59.
13. **Hu, J.**, Reiter, J. P. and Wang, Q. (2014), “Disclosure risk evaluation for fully synthetic categorical data”, *Privacy in Statistical Databases (Lecture Notes in Computer Science 8744)*, ed. J. Domingo-Ferrer, Springer, 185-199.
12. **Hu, J.**, Mitra, R. and Reiter, J. P. (2013), “Are independent parameter draws necessary for multiple imputation?” *The American Statistician*, 67(3), 143-149.
11. **Hu, J.** and Reiter, J. P. (2013), “Non-parametric Bayesian model for generating synthetic household data”, *Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality 2013*, http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2013/Topic_2_Duke_University.pdf

STATISTICS EDUCATION

10. Kejzlar, V. and **Hu, J.** (forthcoming), “Introducing variational inference in statistics and data science curriculum”, *The American Statistician*. link to the published paper
9. **Hu, J.** and Dogucu, M. (2022), “Content and computing outline of two undergraduate Bayesian courses: tools, examples, and recommendations”, *Stat SDSS 2021 Special Issue*, 11(1), e452. Open Access
8. Dogucu, M. and **Hu, J.** (2022), “The current state of undergraduate Bayesian education and recommendations for the future”, *The American Statistician*, 76(4), 405-413. Open Access
7. Albert, J. and **Hu, J.** (2020), “Bayesian computing in the undergraduate statistics curriculum”, *Journal of Statistics Education*, 28(3), 236-247. Open Access
6. Johnson, A., Rundel, C., **Hu, J.**, Ross, K. and Rossman, A. (2020), “Teaching an undergraduate course in Bayesian statistics: a panel discussion”, *Journal of Statistics Education*, 28(3), 251-261. Open Access

5. **Hu, J.** (2020), “A Bayesian statistics course for undergraduates: Bayesian thinking, computing, and research”, *Journal of Statistics Education*, 28(3), 229-235. Open Access
4. Albert, J., Cetinkaya-Rundel, M., and **Hu, J.** (2020), “Online statistics teaching and learning”, *Teaching and Learning Mathematics Online*, ed. J. P. Howard and J. F. Beyers, Chapman and Hall/CRC, Boca Raton, Florida.
3. Garcia, S. R., **Hu, J.**, and Miller, S. J. (2020), “Upper level mathematics and statistics courses share across campuses”, *Teaching and Learning Mathematics Online*, ed. J. P. Howard and J. F. Beyers, Chapman and Hall/CRC, Boca Raton, Florida.
2. **Hu, J.** (2019), “Teaching upper level undergraduate statistics through a shared/hybrid model”, *CHANCE*, 32(2), 31-36.

BOOKS

1. Albert, J. and **Hu, J.** (2019), “Probability and Bayesian Modeling”, *Texts in Statistical Science*, CRC Press. Book Website

PROFESSIONAL SERVICE PUBLICATIONS

1. **Hu, J.** and Bowen, C. M. (2022), “Preserving privacy: human and computational resource limitations and what statisticians and data scientists can do.” *Amstat News*, September Issue.

SOFTWARE

* indicates an undergraduate student co-author

4. Hornby, R.* and **Hu, J.**, “AttributeRiskCalculation: Calculating Attribute Disclosure Risks in Synthetic Microdata”, GitHub link.
3. Hornby, R.* and **Hu, J.**, “IdentificationRiskCalculation: Calculating the Identification Risk in Partially Synthetic Microdata”, GitHub link.
2. Wang, Q., Akande, O., **Hu, J.**, Reiter, J. P. and Barrientos, A., “NestedCategBayesImpute: Modeling and Generating Synthetic Versions of Nested Categorical Data in the Presence of Impossible Combinations”, CRAN link.
1. Wang, Q., Manrique-Vallier, D., Reiter, J. P., and **Hu, J.**, “NPBayesImputeCat: Non-Parametric Bayesian Multiple Imputation for Categorical Data”, CRAN link.

TECHNICAL REPORTS

* indicates an undergraduate student co-author

2. Savitsky, T. D., **Hu, J.**, and Williams, M. R., “Re-weighting of vector-weighted mechanisms for utility maximization under differential privacy”. arXiv link
1. Hornby, R.* and **Hu, J.**, “Bayesian estimation of attribute disclosure risks in synthetic data with the AttributeRiskCalculation R package”. arXiv link

WORK IN PROGRESS

* indicates an undergraduate student co-author

2. **Hu, J.**, Slavkovic, A. and Charest, A. S., “An Introduction to Statistical Data Privacy: Synthetic Data and Differential Privacy”, book in progress (contract signed with CRC Press).
1. Albert, J. and **Hu, J.**, “Introduction to Bayesian Modeling”, book in progress (contract signed with CRC Press).

OTHER PROFESSIONAL ACTIVITY**EDITORIAL LEADERSHIP****Associate Editor**

Statistics and Public Policy 2023 - present

Survey Statistics, Journal of Statistics and Survey Methodology 2022 - present

Knowledge Management & Machine Learning, INFORMS Journal on Computing 2021 - present

Editorial Board

Mathematics Association of America Scatterplot 2021 - present

Transactions on Data Privacy 2019 - present

SOCIETY AND CONFERENCE LEADERSHIP**Program Chair**

American Statistical Association Social Statistics Section 2024

Steering Committee

American Statistical Association Privacy and Confidentiality Interest Group 2023 - present

Member

American Statistical Association Committee on Privacy and Confidentiality 2024 - present

Joint Statistical Meetings 2024 Program Committee 2023 - present

American Statistical Association Advisory Committee on Continuing Education 2023 - present

Chair

ISBA Section on Bayesian Education Research and Practice 2023 - present

Program Chair-Elect

American Statistical Association Social Statistics Section 2023

Co-organizer

NISS IOF on Advancing Demographic Equity with Privacy Preserving Methodologies 2022 - 2023

Co-chair

Privacy and Public Privacy Conference 2023 - present

USPROC + eUSR, CAUSE & ASA 2021 - present

Education Track, Symposium on Data Science and Statistics 2022 (SDSS) 2021 - 2022

Mentor

American Statistical Association SSDSE Mentoring Program 2021 - present

Breakout Room Lead and Facilitator

Prepare to Teach <https://preparingtoteach.org/> 2021

Treasurer

ISBA Section on Bayesian Education Research and Practice 2017 - 2019

INTER-COLLEGE ACTIVITY

Course Advisor

LACOL Summer Course on Applied Machine Learning 2022 - 2023

Advising Committee

LACOL Data Science in Liberal Arts 2018 - 2019

Breakout Session Lead and Facilitator

LACOL Workshop 2018

Vassar Representative

LACOL Quantitative Skills Working Group 2016 - 2018

VOLUNTEERING WORK

Journal Referee

Annals of Applied Statistics (4), Bayesian Analysis, CRC Press (5), Epidemiology, International Conference on Teaching Statistics, Journal of American Statistical Association, Journal of Computational and Graphical Statistics, Journal of Official Statistics (5), Journal of Royal Statistical Society - Series A, Journal of Royal Statistical Society - Series C, Journal of Statistics and Data Science Education (3), Journal of Statistics Education (5), Journal of Survey Statistics and Methodology (5), METRON, National Science Foundation (4), Pearson, Princeton University Press, Psychological Methods, Revista CEA, Science Advances, Software X, Springer Nature, Statistical Analysis and Data Mining, The American Statistician, The R Journal

Panelist

National Science Foundation (one panel)

Volunteer

Statistics Education Booth, Joint Statistical Meetings 2016 2016

SHORT COURSES AND WORKSHOPS

Co-Instructor, *International Society for Bayesian Analysis World Meeting* July 2024

- . Short Course - Bayesian Methods for Statistical Data Privacy
- Co-Instructor**, *ASA Traveling Course, Philadelphia and Southern Ontario Chapters* Fall 2023
 - . Short Course - Introducing Bayesian Statistical Analysis into Your Teaching
- Co-Instructor**, *Bayes BATS, bootcamp for STEM educators on Bayesian curriculum*
 - . Year 1, University of California, Irvine July 2023
- Instructor**, *Summer School at Universitat Politècnica de Catalunya* June 2023
 - . Short Course - Statistical Data Privacy
- Co-Instructor**, *ISI Online Course (virtual)* Jan. 2023
 - . Short Course - Bayesian Thinking: Fundamentals, Computation, and Hierarchical Modeling
- Co-Instructor**, *eCOTS 2022 (virtual)* May 2022
 - . Short Course - Introducing Bayesian Statistical Analysis into Your Teaching
- Co-Instructor**, *USCOTS 2021 (virtual)* June 2021
 - . Short Course - Introducing Bayesian Statistical Analysis into Your Teaching
- Co-Instructor**, *Joint Statistical Meetings (virtual)* Nov. 2020
 - . Short Course - Bayesian Thinking: Fundamentals, Computation, and Hierarchical Modeling
- Instructor**, *NCSES Lecture Series* June 2020
 - . Workshop - Data Privacy: Bayesian Data Synthesis and Differential Privacy
- Co-Instructor**, *Liberal Arts Collaborative for Digital Innovation* Mar. 2020
 - . Workshop - Liberal Arts Remote Teaching
- Instructor**, *Blended Learning in the Liberal Arts Conference* May 2019
 - . Workshop - Teaching a Shared/Hybrid/Online Course using Zoom
- Instructor**, *U.S. Bureau of Labor Statistics* Oct. 2018
 - . Short course - Introduction to Bayesian Inference in R
- Instructor**, *CIRJE, University of Tokyo* Jan. 2018
 - . Short course - The Dirichlet Process and DP Mixture Models

PRESENTATIONS

STATISTICAL MEETINGS

- “Recent Efforts in Statistical Privacy and Public Policy” Oct. 2023
FCSM Research & Policy Conference 2023, Hyattsville, MD, USA
- “Synthetic Data Generation for Survey Data - Discussion” Oct. 2023
FCSM Research & Policy Conference 2023, Hyattsville, MD, USA
- “Introducing Bayesian Methods in Statistics and Data Science Curriculum” Aug. 2023
Joint Mathematical Meetings (topic-contributed, organizer & discussant), Toronto, Canada

“Private Tabular Survey Data Products through Synthetic Microdata Generation” <i>CMStatistics 2022 (invited), hybrid</i>	Dec. 2022
“Private Tabular Survey Data Products through Synthetic Microdata Generation” <i>Fields Institute Workshop on Differential Privacy and Statistical Data Analysis, virtual</i>	July 2022
“Mechanisms for Global Differential Privacy under Bayesian Data Synthesis” <i>IMS Annual Meeting 2022, virtual</i>	June 2022
“Private Tabular Survey Data Products through Synthetic Microdata Generation” <i>Joint Mathematical Meetings (invited), virtual</i>	Apr. 2022
“Accuracy Gains from Privacy Amplification through Sampling for Differential Privacy” <i>Federal Committee on Statistical Methodology (FCSM) 2021, virtual</i>	Nov. 2021
“Data Privacy Protection and Utility Preservation through Bayesian Data Synthesis” <i>Conference in Honor of Fred Smith and Chris Skinner, Zoom</i>	July 2021
“Bayesian Pseudo Posterior Synthesis for Data Privacy Protection” <i>The Sixth International Conference on Establishment Statistics (invited), Zoom</i>	June 2021
“Examples from Two Undergraduate Bayesian Courses” <i>Symposium on Data Science & Statistics (peer-reviewed abstract), Zoom</i>	June 2021
“Two-phase Data Synthesis for Income: An Application to the NHIS” <i>Privacy in Statistical Databases 2020, Zoom</i>	Sept. 2020
“Thinking beyond the p-values: Advancing Bayesian Education for the Undergraduates” <i>Joint Statistical Meetings (invited, organizer & chair), Zoom</i>	Aug. 2020
“Risk-efficient Bayesian Data Synthesis for Privacy Protection” <i>Joint Statistical Meetings (invited), Zoom</i>	Aug. 2020
“Bayesian Pseudo Posterior Synthesis for Data Privacy Protection” <i>Joint Statistical Meetings, Denver, CO, USA</i>	Aug. 2019
“Teaching Upper-Level Statistics Courses Through a Shared/Hybrid Model” <i>Symposium on Data Science and Statistics, Bellevue, WA, USA</i>	May 2019
“The Quasi-Multinomial Synthesizer for Categorical Data” <i>Privacy in Statistical Databases 2018, Valencia, Spain</i>	Sept. 2018
“Synthetic Consumer Expenditure Survey Data at U. S. Bureau of Labor Statistics” <i>International Conference on Applied Mathematics (ICAM 2018), Hong Kong</i>	June 2018
“Teaching Upper-Level Statistics Courses Through a Shared/Hybrid Model” <i>Women in Statistics and Data Science, La Jolla, CA, USA</i>	Oct. 2017

- “Introducing Bayesian Statistics at Courses of Various Levels” Aug. 2017
Joint Statistical Meetings (organizer & discussant), Baltimore, MD, USA
- “Strategies to Facilitate Access to Geocoding Information Based on Synthetic Data” June 2017
The 1st International Conference on Econometrics and Statistics (invited), Hong Kong
- “Strategies to Facilitate Access to Geocoding Information Based on Synthetic Data” May 2017
3rd International Conference on Engineering and Computational Mathematics (invited), Hong Kong
- “Graduate Student Internships” Oct. 2016
Women in Statistics and Data Science (panelist), Charlotte, NC, USA
- “Making Sense of the Transition to Faculty Life” Oct. 2016
Women in Statistics and Data Science (panelist), Charlotte, NC, USA
- “Professional Opportunities at Smaller Colleges and Universities” Aug. 2016
Joint Statistical Meetings, Chicago, IL, USA
- “Dirichlet Process Mixture Models for Nested Unordered Categorical Data” June 2016
Frontiers of Statistics and Data Sciences (invited), Hong Kong
- “Dirichlet Process Mixture Models for Nested Unordered Categorical Data” Aug. 2015
Joint Statistical Meetings, Seattle, WA, USA
- “Nested Dirichlet Process Models for Household Data Synthesis” Oct. 2014
Advances in Interdisciplinary Statistics and Combinatorics, Greensboro, NC, USA
- “Disclosure Risk Evaluation for Fully Synthetic Categorical Data” Sept. 2014
Privacy in Statistical Databases 2014, Eivissa, Spain
- “Nested Dirichlet Process Models for Household Data Synthesis” Aug. 2014
Joint Statistical Meetings, Boston, MA, USA
- “Nested Dirichlet Process Models for Household Data Synthesis” Oct. 2013
UNECE/Eurostat Work Session on Statistical Data Confidentiality (invited), Ottawa, Canada
- “Nonparametric Bayesian Models for Generating Synthetic Household Data” Aug. 2013
Joint Statistical Meetings, Montreal, Canada

FEDEARL STATISTICAL AGENCIES

- “Private Tabular Survey Data Products through Synthetic Microdata Generation” Feb. 2021
National Center for Science and Engineering Statistics, National Science Foundation, Zoom
- “A Friendly Discussion on Differential Privacy” Sept. 2020
National Center for Health Statistics (invited, co-presenter), Zoom
- “Access to Restricted Data Files for Select Statistical Research Projects” July 2019

Consumer Expenditure Surveys Microdata Users' Workshop 2019 (panelist), Washington D.C., USA

“Using CE Microdata in Undergraduate Statistics Courses” July 2019
Consumer Expenditure Surveys Microdata Users' Workshop 2019, Washington D.C., USA

“Bayesian Pseudo Posterior Synthesis for Data Privacy Protection” Mar. 2019
USDA National Agricultural Statistical Service, Washington D.C., USA

“Bayesian Pseudo Posterior Synthesis for Data Privacy Protection” Mar. 2019
Center for Statistical Research and Methodology, U.S. Census Bureau, Washington D.C., USA

“Exploring Synthetic Consumer Expenditure Surveys Data” Dec. 2018
Disclosure Review Board Seminar (invited), U.S. Bureau of Labor Statistics, Washington D.C., USA

“The Synthetic Data Approach to Data Confidentiality” Dec. 2016
U.S. Bureau of Labor Statistics (invited), Washington DC, USA

UNIVERSITIES AND ORGANIZATIONS

“Private Tabular Survey Data Products through Synthetic Microdata Generation” June 2022
Statistics Seminar (invited), RAND, virtual

“Incorporating Disclosure Risk in Designing Data Synthesis Models” Jan. 2022
American Statistical Association Privacy Day Webinar (invited), virtual

“Data Privacy Protection through Bayesian Data Synthesis: A Case Study on Airbnb Listings”
Mathematics Seminar (invited), Skidmore College, Zoom Nov. 2021

“Data Privacy Protection through Bayesian Data Synthesis: A Case Study on Airbnb Listings”
Mathematics and Statistics Seminar (invited), Wellesley College, Zoom Nov. 2021

“Risk-efficient Bayesian Data Synthesis for Privacy Protection” Nov. 2020
Privacy Seminar (invited), Penn State University, Zoom

“Risk-efficient Bayesian Data Synthesis for Privacy Protection” Feb. 2020
Statistics Seminar (invited), University of Massachusetts Amherst, Amherst, MA, USA

“Generating Synthetic Family Income for the Consumer Expenditure Surveys” Feb. 2020
Statistics Seminar (invited), Smith College, Northampton, MA, USA

“Risk-efficient Bayesian Data Synthesis for Privacy Protection” Jan. 2020
Westat (invited), Rockville, MD, USA

“Bayesian Pseudo Posterior Synthesis for Data Privacy Protection” May 2019
Epidemiology and Biostatistics Seminar (invited), Drexel University, PA, USA

“Synthetic Consumer Expenditure Survey Data at U. S. Bureau of Labor Statistics” Oct. 2018
Statistics Seminar, Binghamton University, NY, USA

“Teaching Upper-Level Statistics Courses Through a Shared/Hybrid Model” <i>Statistics Seminar, Binghamton University, NY, USA</i>	Oct. 2018
“The Synthetic Data Approach to Data Confidentiality” <i>Workshop on Statistical Disclosure Control (invited), Kanazawa University, Japan</i>	Jan. 2018
“Dirichlet Process Mixture Models for Nested Unordered Categorical Data” <i>The Applied Statistics Workshop (invited), University of Tokyo, Japan</i>	Jan. 2018
“Teaching an Undergraduate Bayesian Statistics Course” <i>Statistical Science Department Seminar, Duke University, NC, USA</i>	Dec. 2017
“Dirichlet Process Mixture Models for Nested Unordered Categorical Data” <i>Department of Mathematics and Statistics (invited), Bowling Green State University, OH, USA</i>	Oct. 2017
“The Synthetic Data Approach to Data Confidentiality” <i>AALAC Workshop on Data Ethics, Pomona College, Claremont, CA, USA</i>	Feb. 2017
“Strategies to Facilitate Access to Geocoding Information Based on Synthetic Data” <i>NYU Center for Urban Science + Progress (invited), New York, NY, USA</i>	Oct. 2016
“Why and How to Make Fake Data” <i>Faculty Research Presentations, Vassar College, Poughkeepsie, NY, USA</i>	Aug. 2016
“Generating Synthetic Household Data for the Decennial Census” <i>Department of Mathematics (invited), City University of Hong Kong</i>	June 2015
“Missing Not at Random in SAS” <i>Norvatis Statistical Science VC, East Hanover, NJ, USA</i>	Aug. 2013

EDUCATION MEETINGS

“Bayesian Methods and the Statistics and Data Science Curriculum” <i>CAUSE and JSDSE webinar series (invited, penalist), Zoom</i>	Feb. 2021
“Summer Data Science Online: Building Bridges through Collaboration in the Liberal Arts” <i>OLC Innovate, Zoom</i>	June 2020
“LACOL 2020 Summer Data Science Panel” <i>LACOL 2020 Virtual Workshop, Zoom</i>	June 2020
“Summer Data Science (Online) by and for the Liberal Arts” <i>National Workshop on Data Science Education 2019 (lightning), UC Berkeley, CA, USA</i>	June 2019
“A Bayesian Course for Cross-Campus Share” <i>eCOTS 2018 (poster), virtual</i>	May 2018
“Teaching Upper-Level Statistics Courses Through a Shared/Hybrid Model” <i>Blended Learning in the Liberal Arts 2018, Bryn Mawr College, PA, USA</i>	May 2018

“Where’s the Remote? Exploring Upper Level Math/Stats Hybrid Course Sharing for the Liberal Arts”

EDUCAUSE Learning Initiative (ELI) Annual Meeting, New Orleans, LA, USA

Jan. 2018

COLLEGE ACTIVITY

TEACHING

Instructor, *Department of Mathematics and Statistics, Vassar College* 2015 - present

- . MATH 126 - Calculus IIA: Integration Theory
- . MATH 127 - Calculus IIB: Sequences and Series
- . MATH 141 - Introduction to Statistical Reasoning
- . CMPU/MATH 144 - Foundations of Data Science
- . MATH 220 - Multivariable Calculus
- . MATH 240 - Introduction to Statistics
- . MATH 241 - Probability (GitHub repo of course material)
- . MATH 242 - Applied Statistical Modeling
- . MATH 301 - Statistical Data Privacy (GitHub repo of course material)
- . MATH 341 - Statistical Inference
- . MATH 347 - Bayesian Statistics (GitHub repo of course material)

Supervisor of independent studies, *Vassar College*

- | | |
|---|-------------------------|
| . Attribute Disclosure Risk Evaluation for Synthetic Data | Spring 2021 |
| . Bayesian Estimation of Future Realized Volatility | Fall 2019 |
| . Bayesian Inferences with Python | Spring 2020 |
| . Bayesian Methods for Sparse Data | Spring 2020 |
| . Bayesian Network Analysis | Spring 2022 |
| . Bayesian Non-Parametric Models | Spring 2017 |
| . Bayesian Time Series | Fall 2017 & Spring 2019 |
| . Bayesian Variable and Model Selection | Fall 2019 |
| . Identification Risks of Partial Synthetic Data | Fall 2019 & Spring 2020 |
| . LACOL Course Developer | Spring 2020 |
| . Machine Learning Email Classification | Fall 2022 |
| . Pedagogical Partnership | Fall 2021 & Spring 2022 |
| . Python for Data Science | Spring 2020 |
| . Think Bayes | Spring 2021 |
| . Topics of Data Science | Fall 2019 |
| . Tree-Based Methods for Synthetic Data | Fall 2017 |

Instructor, *Liberal Arts Collaborative for Digital Innovation (LACOL)* 2017 - present

Bayesian Statistics, Course Sharing Project

Project info: <https://lacol.net/category/collaborations/course-sharing/>

Course material: YouTube playlists - Fall 2017, Spring 2019, Fall 2019, Spring 2022

Co-Instructor, *Liberal Arts Collaborative for Digital Innovation (LACOL)* 2018 - present

Introduction to Data Science, Course Sharing Project

Project info: <https://lacol.net/critical-ds/>

DEPARTMENTAL SERVICE**Committee Member**

<i>Colloquium series</i>	2021 - 2022
<i>Applied Mathematics faculty search</i>	2021 - 2022
<i>Administrative assistant search</i>	2019
<i>Statistics faculty search</i>	2019
<i>Applied Mathematics faculty search</i>	2016 - 2017

Major Academic Advisor

<i>Vassar College</i>	2016 - present
-----------------------	----------------

Correlate Academic Advisor

<i>Vassar College</i>	2016 - present
-----------------------	----------------

COLLEGE SERVICE**Committee Member**

<i>Steering Committee, Data Science and Society, Vassar College</i>	2022 - present
<i>Postdoc Search Committee, Data Science and Society, Vassar College</i>	2022
<i>Committee on Assessment, Vassar College</i>	2021 - 2022
<i>Rebalanced Curriculum Review Committee, Vassar College</i>	2021 - 2022
<i>Committee on Benefits, Vassar College</i>	2021 - 2022
<i>Science / Math Brainstorming Group (covid-19), Vassar College</i>	2020
<i>Fellowship Committee, Vassar College</i>	2019 - 2021
<i>Committee on Committees, Vassar College</i>	2016 - 2018
<i>Thompson-Bartlett Fellowship Selection Committee, Vassar College</i>	2017

Pre-major Academic Advisor

<i>Vassar College</i>	2016 - present
-----------------------	----------------

Co-organizer

<i>DataFest @ Vassar</i> http://pages.vassar.edu/datafest/	annually since 2016
<i>Women of Color in STEM Workshop, Women of Color Conference, CHAS</i>	2017
<i>Exploring Online Bridging Program, Vassar College</i>	2017

College Program Involvement

<i>The STEPP Program, Vassar College</i>	2021 - 2022
<i>HHMI Grand Challenges, Vassar College</i>	2019 - 2021

Faculty Host

<i>International Friendship Program, Vassar College</i>	2016 - present
---	----------------

Faculty Panelist/Speaker

<i>Gender Inclusivity in STEM dinner, Vassar College</i>	2019
<i>Journeys, Campus Life and Diversity Office, Vassar College</i>	2017
<i>Hidden Figures Strong House Screening, Vassar College</i>	2017
<i>Story Sharing Session, Chinese Student Community (CSC), Vassar College</i>	2017

<i>Discover America, Office of International Services, Vassar College</i>	2017
<i>Vassar Science Scholars Program, Vassar College</i>	2017
<i>Talking about Teaching session on international students, Vassar College</i>	2017
<i>Strong STEM Brunch Series (Computer Science and Math), Vassar College</i>	2015

ACTIVITY AND AWARDS BEFORE VASSAR

Instructor , <i>Department of Statistical Science, Duke University</i> . STA 101 - Data Analysis and Statistical Inference	Summer 2014
Guest Lecturer , <i>Department of Statistical Science, Duke University</i> . STA 111 - Probability and Statistical Inference (undergrad) . STA 130 - Probability and Statistics in Engineering (undergrad) . STA 723 - Statistics Case Studies (grad)	Spring 2015
Coursera Community Teaching Assistant , <i>Duke University</i> . Data Analysis and Statistical Inference	Fall 2014
Head Facilitator, Chinese Conversation Club <i>International House, Duke University</i>	2012 - 2015
Student Blogger <i>Student Affairs, Duke University</i>	2012 - 2013
Student Mentor for incoming PhD students <i>Department of Statistical Science, Duke University</i>	2012 - 2014
Statistical Analyst <i>Women in Science and Engineering Symposium, Duke University</i>	2012
Ph.D. First Year Fellowship <i>Department of Statistical Science, Duke University</i>	2011 - 2012
Scholarship for Mainland China Students <i>City University of Hong Kong</i>	2007 - 2011
Student Exchange Program Scholarship <i>College of Science and Engineering, City University of Hong Kong</i>	2009