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

Statistical Consultant 2022 - present
New York City Department of Health and Mental Hygiene

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

HONORS & AWARDS

Faculty Advisor to Honorable Mention Winner 2020 & 2022
The Undergraduate Research Project Competition (USRESP), CAUSE & ASA

Faculty Advisor to Student 1st-place Winner 2016 & 2019
The Undergraduate Class Project Competition (USCLAP), intermediate statistics, CAUSE & ASA

ASA Student Paper Competition Award 2015
Joint Statistical Meetings

GRANTS & FELLOWSHIPS

EXTERNAL ($543,368.40 in total as PI and co-PI)

Active

- 2023 - 2025, co-PI, "Advancing Communication, Collaboration, and Evidence-building through Synthetic Survey of Earned Doctorates (ACCESS)", **NSF National Center for Science and Engineering Statistics**, total award: **$800,000.00**; total to co-PI: **$71,032.00**.

- 2022 - 2026, PI, "ERS Disclosure Review Assistance", **United States Department of Agriculture Economic Research Service** , total award: **$148,212.00**; total to PI: **$148,212.00**.
- 2022 - 2025, PI, "Advancing Bayesian Thinking in STEM", **National Science Foundation**, total award: **$300,000.00**; total to PI: **$99,516.00**.

Completed

- 2022 - 2024, PI, "Statistical Privacy and Public Policy Workshop Series", **Alfred P. Sloan Foundation**, total award: **$49,948.00**; total to PI: **$22,500.00**.
- 2020 - 2021, PI, "Modern Disclosure Limitation and Privacy Protection Methods for Survey Data", **NSF National Center for Science and Engineering Statistics**, total award: **$56,732.40**; total to PI: **$56,732.40**.
- 2018 - 2019, PI, "Synthetic Consumer Expenditure Survey Data at BLS", **NSF ASA/NSF/BLS Research Fellowship**, total award: **$88,531.00**; total to PI: **$88,531.00**.
- 2019 - 2022, co-PI, "Introduction to Data Science", **Liberal Arts Collaborative for Digital Innovation**, total award: **$207,300.00**; total to co-PI: **$24,800.00**.
- 2018 - 2022, PI, "Bayesian Statistics", **Liberal Arts Collaborative for Digital Innovation**, total award: **$$27,500.00**; total to co-PI: **$24,500.00**.
- 2020, PI, "Data Confidentiality", **Liberal Arts Collaborative for Digital Innovation**, total award: **$$4,500.00**; total to co-PI: **$4,500.00**.
- 2020, PI, "Bayesian Inference with Python", **Liberal Arts Collaborative for Digital Innovation**, total award: **$$3,000.00**; total to co-PI: **$3,000.00**.

INTERNAL AT VASSAR COLLEGE ($19,099.20 in total as PI)

Completed

- 2019 - 2021, PI, "CMPU/MATH 144 Foundations of Data Science", **Grand Challenge Climate Change**, total award: **$11,000.00**; total to PI: **$11,000.00**.
- 2019 - 2020, PI, "Disclosure Risks Evaluation of Public Use Microdata", **Research Committee Award**, total award: **$3,871.00**; total to PI: **$3,871.00**.
- 2018 - 2019, PI, "Synthetic Consumer Expenditure Survey Data at BLS", **Research Committee Award**, total award: **$2,940.45**; total to PI: **$2,940.45**.
- 2017, PI, "Upper-level Mathematics and Statistics Course Sharing", **Endowment for Strategic Faculty Support**, , total award: **$1,287.75**; total to PI: **$1,287.75**.

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.**, Williams, M. R., and Savitsky, T. D. (2025), "Mechanisms for global differential privacy under Bayesian data synthesis", *Data Privacy special issue* at *Statistica Sinica*, 35, 1-22. Open Access

32. **Hu, J.** and Bowen, C. M. (2024), "Advancing microdata privacy protection: a review of synthetic data methods", *WIREs Interdisciplinary Reviews: Computational Statistics*, e1636. doi:10.1002/wics.1636.

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. Europaische 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.** (2024), "Introducing variational inference in statistics and data science curriculum", *The American Statistician*, 78(3), 359-367.

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

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

4. Savitsky, T. D., **Hu, J.**, and Williams, M. R., "Maximizing utility for vector-weighted pseudo posterior mechanisms under differential privacy", submitted. arXiv link

3. Immerwahr, S., **Hu, J.**, Deng, W. Q., Bholanath, T., Lundy De La Cruz, N. and He, F., "Disclosure risk evaluation and mitigation solutions for health survey microdata files: application to the New York City Community Health Survey", to be submitted.

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 - 2024

Editorial Board
Mathematics Association of America Scatterplot 2021 - present
Transactions on Data Privacy 2019 - present

SOCIETY AND CONFERENCE LEADERSHIP

Education Officer
American Statistical Association Survey Research Methods Section 2024 - present

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
American Statistical Association Advisory Committee on Continuing Education 2023 - present
Joint Statistical Meetings 2024 Program Committee 2023 - 2024

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-founder
Privacy and Public Policy Conference 2023 - present

Co-chair
Privacy and Public Policy Conference 2023 - 2024
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 2021

Treasurer
ISBA Section on Bayesian Education Research and Practice 2017 - 2019

VOLUNTEERING WORK

Journal Referee
Annals of Applied Statistics (4), Bayesian Analysis, Computers in Biology and Medicine, CRC Press (6), Epidemiology, International Conference on Teaching Statistics, Journal of American Statistical Association, Journal of Computational and Graphical Statistics, Journal of Official Statistics (6), 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), MAA Scatterplot, METRON, National Science Foundation (5), Patterns, 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 Advisory Panel (one panel)

National Science Foundation (two panels)

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
. Year 2, Vassar College July 2024

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

"New Explorations of Disclosure Avoidance Techniques and Applications - Organizer" Oct. 2024
FCSM Research & Policy Conference 2024, Hyattsville, MD, USA

"Best Practices and Challenges of Disclosure Reviews - Organizer & Chair" Oct. 2024
FCSM Research & Policy Conference 2024, Hyattsville, MD, USA

"New Bayesian Methods for Statistical Data Privacy - Organizer & Chair" Oct. 2024
FCSM Research & Policy Conference 2024, Hyattsville, MD, USA

"Evaluating Statistical Disclosure Control Techniques based on the Risk and Utility of Privacy-Protected Data" Aug. 2024
Joint Statistical Meetings (topic-contributed, chair), Portland, OR, USA

"Mechanisms for Global Differential Privacy under Bayesian Data Synthesis" June 2024
Statistical Society of Canada Annual Meeting 2024 (invited), St. John's, NL, Canada

"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 Statistical Meetings (topic-contributed, organizer & discussant), Toronto, Canada

"Private Tabular Survey Data Products through Synthetic Microdata Generation" Dec. 2022
CMStatistics 2022 (invited), hybrid

"Private Tabular Survey Data Products through Synthetic Microdata Generation" July 2022
Fields Institute Workshop on Differential Privacy and Statistical Data Analysis, virtual

"Mechanisms for Global Differential Privacy under Bayesian Data Synthesis" June 2022
IMS Annual Meeting 2022, virtual

"Private Tabular Survey Data Products through Synthetic Microdata Generation" Apr. 2022
Joint Mathematical Meetings (invited), virtual

"Accuracy Gains from Privacy Amplification through Sampling for Differential Privacy" Nov. 2021
Federal Committee on Statistical Methodology (FCSM) 2021, virtual

"Data Privacy Protection and Utility Preservation through Bayesian Data Synthesis" July 2021
Conference in Honor of Fred Smith and Chris Skinner, Zoom

"Bayesian Pseudo Posterior Synthesis for Data Privacy Protection" June 2021
The Sixth International Conference on Establishment Statistics (invited), Zoom

"Examples from Two Undergraduate Bayesian Courses" June 2021
Symposium on Data Science & Statistics (peer-reviewed abstract), Zoom

"Two-phase Data Synthesis for Income: An Application to the NHIS" Sept. 2020
Privacy in Statistical Databases 2020, Zoom

"Thinking beyond the p-values: Advancing Bayesian Education for the Undergraduates" Aug. 2020
Joint Statistical Meetings (invited, organizer & chair), Zoom

"Risk-efficient Bayesian Data Synthesis for Privacy Protection" Aug. 2020
Joint Statistical Meetings (invited), Zoom

"Bayesian Pseudo Posterior Synthesis for Data Privacy Protection" Aug. 2019
Joint Statistical Meetings, Denver, CO, USA

"Teaching Upper-Level Statistics Courses Through a Shared/Hybrid Model" May 2019
Symposium on Data Science and Statistics, Bellevue, WA, USA

"The Quasi-Multinomial Synthesizer for Categorical Data" Sept. 2018
Privacy in Statistical Databases 2018, Valencia, Spain

"Synthetic Consumer Expenditure Survey Data at U. S. Bureau of Labor Statistics" June 2018
International Conference on Applied Mathematics (ICAM 2018), Hong Kong

"Teaching Upper-Level Statistics Courses Through a Shared/Hybrid Model" Oct. 2017
Women in Statistics and Data Science, La Jolla, CA, USA

"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

FEDERAL STATISTICAL AND OTHER AGENCIES

"A Primer on Statistical Disclosure Avoidance Techniques" Feb. 2024
USDA Economic Research Service FoodAPS Disclosure Review Team, virtual

"Synthetic Data, Differential Privacy, and Disclosure Risk Assessment" Feb. 2024
USDA National Agricultural Statistical Service RDD Research Seminar (invited), Zoom

"Data Disclosure Risk Assessment and Mitigation Strategies" Nov. 2023
The Center for Population Health Data Science, New York City Department of Health and Mental Hygiene, virtual

"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

"Mechanisms for Global Differential Privacy under Bayesian Data Synthesis" Nov. 2024
StatSci Research Alumni Symposium (invited), Duke University, Durham, NC, USA

"Statistical Data Privacy: Synthetic Data, Differential Privacy, and Disclosure Risk Assessment"
Statistics Seminar (invited), UC Santa Cruz, virtual May 2024

"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" Oct. 2018
Statistics Seminar, Binghamton University, NY, USA

"The Synthetic Data Approach to Data Confidentiality" Jan. 2018
Workshop on Statistical Disclosure Control (invited), Kanazawa University, Japan

"Dirichlet Process Mixture Models for Nested Unordered Categorical Data" Jan. 2018
The Applied Statistics Workshop (invited), University of Tokyo, Japan

"Teaching an Undergraduate Bayesian Statistics Course" Dec. 2017
Statistical Science Department Seminar, Duke University, NC, USA

"Dirichlet Process Mixture Models for Nested Unordered Categorical Data" Oct. 2017
Department of Mathematics and Statistics (invited), Bowling Green State University, OH, USA

"The Synthetic Data Approach to Data Confidentiality" Feb. 2017
AALAC Workshop on Data Ethics, Pomona College, Claremont, CA, USA

"Strategies to Facilitate Access to Geocoding Information Based on Synthetic Data" Oct. 2016
NYU Center for Urban Science + Progress (invited), New York, NY, USA

"Why and How to Make Fake Data" Aug. 2016
Faculty Research Presentations, Vassar College, Poughkeepsie, NY, USA

"Generating Synthetic Household Data for the Decennial Census" June 2015

Department of Mathematics (invited), City University of Hong Kong

"Missing Not at Random in SAS" Aug. 2013
Norvatis Statistical Science VC, East Hanover, NJ, USA

EDUCATION MEETINGS

"Training the Trainers in Bayesian Methods" June 2024
Improving Undergraduate STEM Education Initiative Summit 2024, Washington D.C., USA

"Bayesian Methods and the Statistics and Data Science Curriculum" Feb. 2021
CAUSE and JSDSE webinar series (invited, penalist), Zoom

"Summer Data Science Online: Building Bridges through Collaboration in the Liberal Arts"
OLC Innovate, Zoom June 2020

"LACOL 2020 Summer Data Science Panel" June 2020
LACOL 2020 Virtual Workshop, Zoom

"Summer Data Science (Online) by and for the Liberal Arts" June 2019
National Workshop on Data Science Education 2019 (lightning), UC Berkeley, CA, USA

"A Bayesian Course for Cross-Campus Share" May 2018
eCOTS 2018 (poster), virtual

"Teaching Upper-Level Statistics Courses Through a Shared/Hybrid Model" May 2018
Blended Learning in the Liberal Arts 2018, Bryn Mawr College, PA, USA

"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

ACTIVITIES AT VASSAR COLLEGE

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
. MATH 301 - Topics in AI: Methods and Hidden Costs
. MATH 341 - Statistical Inference

. MATH 347 - Bayesian Statistics (GitHub repo of course material)
. MATH 388 - Statistical Data Privacy

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 - 2022
Bayesian Statistics, Course Sharing Project
Course material: YouTube playlists - Fall 2017, Spring 2019, Fall 2019, Spring 2022

Co-Instructor, *Liberal Arts Collaborative for Digital Innovation (LACOL)* 2018 - 2022
Introduction to Data Science, Course Sharing Project

DEPARTMENTAL SERVICE

Committee Member

Statistics faculty search	2024 - 2025
Statistics faculty search	2023
Colloquium series	2021 - 2022
Applied Mathematics faculty search	2021 - 2022
Administrative assistant search	2019
Statistics faculty search	2019
Applied Mathematics faculty search	2016 - 2017

Major Academic Advisor
Vassar College 2016 - present

Correlate Academic Advisor
Vassar College 2016 - present

COLLEGE SERVICE

Committee Member
Committee on Committees (Chair), Vassar College 2024 - present

Steering Committee, Data Science and Society, Vassar College 2022 - present
Postdoc Search Committee, Data Science and Society, Vassar College 2022
Committee on Assessment, Vassar College 2021 - 2022
Rebalanced Curriculum Review Committee, Vassar College 2021 - 2022
Committee on Benefits, Vassar College 2021 - 2022
Science / Math Brainstorming Group (covid-19), Vassar College 2020
Fellowship Committee, Vassar College 2019 - 2021
Committee on Committees, Vassar College 2016 - 2018
Thompson-Bartlett Fellowship Selection Committee, Vassar College 2017; 2024

Pre-major and Major Academic Advisor
Vassar College 2016 - present

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

College Program Involvement
Faculty Leadership Development Program, Vassar College 2024
The STEPP Program, Vassar College 2021 - 2022
HHMI Grand Challenges, Vassar College 2019 - 2021

Faculty Host
International Friendship Program, Vassar College 2016 - present

Faculty Panelist/Speaker
Gender Inclusivity in STEM dinner, Vassar College 2019
Journeys, Campus Life and Diversity Office, Vassar College 2017
Hidden Figures Strong House Screening, Vassar College 2017
Story Sharing Session, Chinese Student Community (CSC), Vassar College 2017
Discover America, Office of International Services, Vassar College 2017
Vassar Science Scholars Program, Vassar College 2017
Talking about Teaching session on international students, Vassar College 2017
Strong STEM Brunch Series (Computer Science and Math), Vassar College 2015

ACTIVITIES AND AWARDS BEFORE VASSAR COLLEGE

Instructor, *Department of Statistical Science, Duke University* Summer 2014
. STA 101 - Data Analysis and Statistical Inference

Guest Lecturer, *Department of Statistical Science, Duke University* Spring 2015
. STA 111 - Probability and Statistical Inference (undergrad)
. STA 130 - Probability and Statistics in Engineering (undergrad)
. STA 723 - Statistics Case Studies (grad)

Coursera Community Teaching Assistant, *Duke University* Fall 2014
. Data Analysis and Statistical Inference

Head Facilitator, Chinese Conversation Club 2012 - 2015

International House, Duke University

Student Blogger 2012 - 2013
Student Affairs, Duke University

Student Mentor for incoming PhD students 2012 - 2014
Department of Statistical Science, Duke University

Statistical Analyst 2012
Women in Science and Engineering Symposium, Duke University

Ph.D. First Year Fellowship 2011 - 2012
Department of Statistical Science, Duke University

Scholarship for Mainland China Students 2007 - 2011
City University of Hong Kong

Student Exchange Program Scholarship 2009
College of Science and Engineering, City University of Hong Kong