Zijian Guo

Contact Information	Hill Center 110 Frelinghuysen Road Piscataway, NJ 08854	(848)445-2690 zijguo@stat.rutgers.edu http://statistics.rutgers.edu/home/zijguo/

Research Interests High-dimensional inference, causal inference, econometrics, nonparametric inference and applications to health and genetics.

- Inference in high-dimensional regression
- Causal inference with instrumental variables
- Additive models
- Mediation analysis
- Efficient data imputation
- Applications to genetics and health record data

Positions

Assistant Professor — Sep 2017-present
Department of Statistics
Rutgers, the State University of New Jersey

Academic Visits

Visiting Scholar — Sep 2019
TH Chan School of Public Health, Harvard University
Host: Tianxi Cai

Visiting Scholar — Nov 2018
Forschungsinstitut für Mathematik, ETH, Zürich
Host: Peter Bühlmann

Visiting Scholar — Aug 2017
Perelman School of Medicine, Upenn
Host: Hongzhe Li

Education

Ph.D. Statistics, University of Pennsylvania — 2017
Thesis advisor: T. Tony Cai

B.S. Mathematics, The Chinese University of Hong Kong — 2012

Publication Published/Accepted Papers (* indicates alphabetical ordering authorship)

1. *Cai, T. T., & **Guo, Z.** (2020). Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(2), 391-419.
2. **Guo, Z.**, Wang, W., Cai, T. T., & Li, H. (2019). Optimal estimation of genetic relatedness in high-dimensional linear models. *Journal of the American Statistical Association*, 114(525), 358-369.
3. **Guo, Z.**, Kang, H., Cai, T. T., & Small, D. S. (2018). Testing Endogeneity with High Dimensional Covariates. *The Journal of Econometrics*, 207(1), 175-187.
4. **Guo, Z.**, Kang, H., Cai, T. T., & Small, D. S. (2018). Confidence Interval for Causal Effects with Invalid Instruments using Two-Stage Hard Thresholding. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(4), 793-815.
5. *Cai, T. T., & **Guo, Z.** (2018). Accuracy assessment for high-dimensional linear regression. *Annals of Statistics*, 46(4), 1807-1836.
6. **Guo, Z.**, Small, D. S., Gansky, S. A., & Cheng, J. (2018). Mediation analysis for count and zero-inflated count data without sequential ignorability and its application in dental studies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(2), 371-394.

7. Cheng, J., Cheng N. F., **Guo, Z.**, Gregorich, S., Amid I. I., & Gansky, S. A. (2018). Mediation analysis for count and zero-inflated count data. *Statistical Methods in Medical Research*, 27(9), 2756-2774.
8. *Cai, T. T., & **Guo, Z.** (2017). Confidence intervals for high-dimensional linear regression: Minimax rates and adaptivity. *Annals of Statistics*, 45(2), 615-646.
9. **Guo, Z.**, & Small, D. S. (2016). Control function instrumental variable estimation of nonlinear causal effect models. *Journal of Machine Learning Research*, 17(100), 1-35.
10. **Guo, Z.**, Cheng, J., Lorch, S. A., & Small, D. S. (2014). Using an instrumental variable to test for unmeasured confounding. *Statistics in Medicine*, 33(20), 3528-3546.
11. **Guo, Z.**, Kogan, R., Qiu, H., & Strichartz, R. S. (2014). Boundary value problems for a family of domains in the Sierpinski gasket. *Illinois Journal of Mathematics*, 58(2), 497-519.

Technical Reports (* indicates alphabetical ordering authorship)

12. *Cai, Tianxi, Cai, T. T., & **Guo, Z.** (2019). Individualized Treatment Selection: An Optimal Hypothesis Testing Approach In High-dimensional Models. Under major revision at *Journal of the Royal Statistical Society: Series B*.
13. **Guo, Z.** & Zhang, C. (2019). Extreme Nonlinear Correlation for Multiple Random Variables and Stochastic Processes with Applications to Additive Models. Under major revision at *Stochastic Processes and their Applications*.
14. **Guo, Z.**, Renaux, C., Bühlmann, P., & Cai, T. T. (2019). Group Inference in High Dimensions with Applications to Hierarchical Testing. *Technical Report*.
15. **Guo, Z.** & Zhang, C. (2019). Local Inference in Additive Models with Decorrelated Local Linear Estimator. *Technical Report*.
16. **Guo, Z.**, Rakshit, P., Herman, D., & Chen, J. (2019). Inference for Case Probability in High-dimensional Logistic Regression. *Technical Report*.
17. **Guo, Z.**, Ćevid, D., & Bühlmann, P. (2020). Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding and Measurement Errors. *Technical Report*.
18. *Cai, T. T., **Guo, Z.**, & Ma, R. (2020). Statistical Inference for High-Dimensional Generalized Linear Models with Binary Outcomes. *Technical Report*.
19. Li, S., & **Guo, Z.** (2020). Causal Inference for Nonlinear Outcome Models with Possibly Invalid Instrumental Variables. *Technical Report*.

Grants

1. National Institute of Health #1R01GM140463-01 "Predictive Modeling with High-Dimensional lncomplete Data."
   - Period: July 2020 to June 2023.
   - Role: Principal Investigator
2. National Science Foundation #DMS 1811857 "Inference in High-Dimensional Linear Models: Methods, Theory and Applications."
   - Period: Aug 2018 to July 2021.
   - Role: Principal Investigator
3. National Science Foundation #DMS 2015373 "Repro Sampling Method: A Transformative Artificial-Sample-Based Inferential Framework with Applications to Discrete Parameter, High-Dimensional Data, and Rare Events Inferences"
   - Period: July 2020 to June 2023.
   - Role: Co-Principal Investigator
4. National Institute of Health #R56-HL-138306-01 "Statistics Methods for Analyzing Electronic Health Record Data."

- Period: June 2018 to Aug 2018.
- Role: Co-Investigator

5. Upenn Medical School "Statistics Methods for Analyzing Electronic Health Record Data" Subcontract.
   - Period: June 2019.
   - Role: Senior Investigator

SOFTWARE

R codes are available at `https://statistics.rutgers.edu/home/zijguo/Software.html`

1. **Two Stage Hard Thresholding (TSHT)**
2. **Functional Inference in High-dimensional Regression (FIHR)**
3. **Semi-parametric outcome models with invalid IVs (SpotIV)**

TEACHING EXPERIENCE

Instructor

- Rutgers University

FSRM 588: Financial Data Mining	Spring 2020
FSRM 588: Financial Data Mining	Fall 2019
Instructor Rating: 4.75 out of 5.0	
STAT 594: Advanced Modern Statistical Inference II	Spring 2019
Instructor Rating: 4.82 out of 5.0	
FSRM 588: Financial Data Mining	Fall 2018
Instructor Rating: 4.71 out of 5.0	
FSRM 588: Financial Data Mining	Fall 2017
Instructor Rating: 4.82 out of 5.0	

- The Wharton School, University of Pennsylvania

STAT 111 : Introductory Statistics	Summer 2016
Instructor Rating: 3.6 out of 4.0	

Recitation Instructor Fall 2014

The Wharton School, University of Pennsylvania

STAT 111: Introductory Statistics

Teaching Assistant

The Wharton School, University of Pennsylvania

STAT 102: Business Statistics	Spring 2017
STAT 970: Mathematical Statistics	Fall 2016
STAT 622: Statistical Modeling	Spring 2016
STAT 550: Mathematical Statistics	Fall 2015

Students Supervision

PhD Thesis Advisor: Prabrisha Rakshit (expected 2023)

PhD Students and Post Doc Mentor[1]: Domagoj Cevid (ETH); Jue Hou (Harvard); Sai Li (Upenn); Molei Liu (Harvard); Rong Ma (Upenn); Claude Renaux (ETH); Ye Tian (Rutgers); Lu Wang (Upenn); Zheshi Zheng (Rutgers).

PhD Thesis Committee: Sai Li (2018); Yisha Yao (expected 2020)

Master Thesis Advisor: Yankun Xu (2018); Yangdi Li (2018); Guanyu Huang (2018); Haoze Tang (2018); Wenzhe Zhang (2018); Yaran Su (2018); Xinyi Zhang (2018); Yuan Liang (2019); Hequan Zhang (2019); Qiaochu Chen (2019); Jiamin Deng (2019); Zeen Huo (2019); Jianyu Li (2020); Junjie Chen (2020).

Honors and Awards

- ICSA New Researcher Award, ICSA 2019 — Dec. 2019
- IMS travel Award, JSM — Aug. 2017
- President Gutmann Leadership Award, University of Pennsylvania — Apr. 2017
- J. Parker Bursk Prize — Sept. 2016
  Awarded by the Statistics Department at the Wharton School for excellence in research.
- Statistics in Epidemiology Young Investigator Award, JSM — Aug. 2013
  Awarded by the American Statistical Association section on Statistics in Epidemiology for the paper "Using an instrumental variable to test for unmeasured confounding."
- Chung Chi College Departmental Prize, CUHK — 2011
- Dr. Chao Yong Chi-hsing Scholarship in Mathematics, CUHK — 2011
- Chung Chi Traveling Award in Mathematics, CUHK — 2011
- Chung Chi Ivy League Exchange Scholarship, CUHK — 2010
- Caring Alumni Student Exchange Scholarship, CUHK — 2010
- Dean's List, College of Arts and Science, UPenn — 2010
- Dean's Honors List, Faculty of Science, CUHK — 2008, 2009
- Chung Chi College Scholarship, CUHK — 2009
- Honors at Entrance to the Chinese University of Hong Kong (4 years) — 2008

Academic Service

- Organizer of Causal Inference Reading Group (Joint with Nicole Pashley and Tirthankar Dasgupta), Department of Statistics, Rutgers
- Department Retreat Chair (2019-2020), Department of Statistics, Rutgers
- Department Seminar Chair (2018-2019), Department of Statistics, Rutgers
- Other Rutgers Committee service: Department retreat (2017-2018); FSRM (2017-2018, 2018-2019, 2019-2020); Ph.D. Exam (2018-2019, 2019-2020); Graduate Curriculum (2019-2020);
- Organizing Committee for 2019 Rutgers Statistics Symposium
- Program Committee for ICSA 2019 11th International Conference
- Local Organizing Committee for 2018 ICSA Applied Symposium.
- Session organizer for 2019 Hangzhou Data Science Conference
- Session organizer for 2018 ICSA Applied Symposium.
- Reviewer for the following journals: *Annals of Statistics*, *JASA*, *JRSSB*, *Biometrica*, *Statistica Sinica*, *IEEE International Symposium on Information Theory*, *Journal of Applied Statistics*, *Biometrics*, *Journal of Machine Learning*, *COLT*.

[1] I am mentoring these PhD students or Post Doc for one or multiple projects but not their advisor

TALKS

• Department seminar (virtual), Department of Statistics, Cornell University, USA *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding."*, Oct 2020

• Invited talk (virtual), JSM 2020, Philadelphia, USA, *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding and Measurement Errors."*, Aug 2020

• Department seminar (virtual), Department of Statistics, UC Davis, USA *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding and Measurement Errors."*, May 2020

• Upenn causal reading group (virtual), Upenn, USA, *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding and Measurement Errors."*, May 2020

• Department seminar, Department of Statistics, East China Normal University, Shanghai, China, *"Group Inference in High Dimensions with Applications to Hierarchical Testing"*, Dec 2019

• Invited talk, 11th ICSA International Conference, Hangzhou, China, *"Group Inference in High Dimensions with Applications to Hierarchical Testing"*, Dec 2019

• Invited talk, International Statistical Conference in Memory of Professor Sik-Yum Lee, Hong Kong, China, *"Group Inference in High Dimensions with Applications to Hierarchical Testing"*, Dec 2019

• Causal reading group (led by James Robins), School of Public Health, Harvard University *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Sep. 2019

• Department seminar, Department of Statistics, East China Normal University, Shanghai, China, *"Individualized Treatment Selection: A Hypothesis Testing Approach In High-dimensional Models"*, June. 2019

• Invited talk, 2019 Hangzhou Data Science Conference, Hangzhou, China, *"Local Inference in High-dimensional Sparse Additive Modeling"*, May. 2019

• Department seminar, School of Data Science, City University of Hong Kong, Hong Kong, China, *"Individualized Treatment Selection: A Hypothesis Testing Approach In High-dimensional Models"*, May. 2019

• Department seminar, ISOM, HKUST, Hong Kong, China, *"Local Inference in High-dimensional Sparse Additive Modeling"*, May. 2019

• Department seminar, Department of Statistics, University of Virginia, USA, *"Local Inference in High-dimensional Sparse Additive Modeling"*, March. 2019

• Invited Speaker, 2019 ICSA Data Science Conference, Xishuangbanna, Yunnan, China. *"Individualized Treatment Selection: A Hypothesis Testing Approach In High-dimensional Models"*, Jan. 2019

• Young Research Session, Memorial Workshop for Lawrence D. Brown, University of Pennsylvania, USA. *"Individualized Treatment Selection: A Hypothesis Testing Approach In High-dimensional Models"*, Nov. 2018

• Seminar for Statistics, Department of Mathematics, ETH, Swiss, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Nov. 2018

• Department seminar, Department of Mathematics, NJIT, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Nov. 2018

• Department seminar, ORFE, Princeton, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Oct. 2018

• Department seminar, ISOM, HKUST, Hong Kong, China, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, July. 2018

• Department seminar, Department of Statistics, Naikai University, China, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*,

July. 2018

• Invited talk, IMS Asia Pacific Rim Meeting, Singapore, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018

• Invited talk, HongKong EcoStat Conference, Hong Kong, China, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018

• Invited talk, ICSA Symposium 2018, New Brunswick, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018

• Invited talk, Purdue Symposium on Statistics, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018

• Invited talk, 2018 Hangzhou Data Science Conference, Hangzhou, China, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, May. 2018

• Invited talk, Lorentz Center, Leiden University, Netherlands, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Apr. 2018

• Department seminar, Department of Statistics, Columbia University, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Apr. 2018

• Topic contributed talk, Joint Statistical Meetings, Baltimore, USA, *"Optimal Estimation of Co-Heritability in High-Dimensional Linear Models"*, Aug. 2017

• Invited talk, Statistical Foundations of Uncertainty Quantification for Inverse Problem, Cambridge, *"Inference for Functionals in High-dimensional Linear Models"*, June. 2017

• Seminar, Center for Statistical Methods in Big Data, University of Pennsylvania, *"Inference with High-dimensional Covariates and Possibly Invalid Instruments"*, Apr. 2017

• Seminar, Institute of Data science, Fox Business School, Temple University, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Feb. 2017

• Department seminar, Department of Biostatistics, UC Berkeley, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Feb. 2017

• Department seminar, Department of Statistics, Rutgers, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Feb. 2017

• Department seminar, Department of Statistics, University of Michigan, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017

• Department seminar, Department of Statistics, University of Minnesota, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017

• Department seminar, Department of Statistics, UIUC, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017

• Department seminar, DPMMS, University of Cambridge, *"Inference for High Dimensional Linear Regression: Fundamental Limits and Algorithms"*, Jan. 2017

• Department seminar, Department of Statistics, UC Santa Barbara, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017

• Invited talk, Mathematical Meeting in Statistics, Fréjus, France, *"Optimal Estimation of Genetic Correlation in High-dimensional Linear Models"*, Dec. 2016

• Econometrics Lunch, Department of Economics, University of Pennsylvania, *"Confidence Intervals for Treatment Effects in High-Dimensional Linear Models"*, Nov. 2016

• Contributed talk, Joint Statistical Meetings, Chicago, USA, *"Accuracy Assessment for High-dimensional Linear Regression"*, Aug. 2016

• Contributed talk, Eastern North American Region, Austin, USA, *"Confidence Intervals for High-Dimensional Linear Regression: Minimax Rates and Adaptivity"*, Mar. 2016

• Poster presentation, John W. Tukey Conference, Princeton University, *"Confidence Intervals*

for High-Dimensional Linear Regression: Minimax Rates and Adaptivity", Sept. 2015

• Contributed talk, Joint Statistical Meetings, Seattle, USA, *"Distance Matrix Estimation from Noisy Observation of Low Rank Position Matrix"*, Aug. 2015

• Contributed talk, Joint Statistical Meetings, Boston, USA, *"Instrumental Variable Approach for Mediation Analysis of Count Model"*, Aug. 2014

• Topic Contributed talk, Joint Statistical Meetings, Montreal, Canada, *" Instrumental Variable Approach for Mediation Analysis of Zero-Inflated Count Model"*, Aug. 2013

• Poster presentation, Atlantic Causal Inference Conference, Harvard University, *"Control Function Instrumental Variable Estimation of Nonlinear Causal Effect Models"*,May. 2013

MEMBERSHIPS

• American Statistical Association

• Institute of Mathematical Statistics

• International Chinese Statistical Association

• The Econometric Society