Zijian Guo

Contact Information

Department of Statistics,
Rutgers University,
Piscataway, NJ 08854

(848)445-2690
zijguo@stat.rutgers.edu
http://statistics.rutgers.edu/home/zijguo/

Education

Ph.D. Statistics, University of Pennsylvania — 2017
Thesis advisor: T. Tony Cai
"Statistical Inference For High-Dimensional Linear Models"

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

Positions

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

Other Positions

Invited Research Fellow — Sep 2019
The Chan School of Public Health, Harvard University
Host: Tianxi Cai

Invited Research Fellow — Nov 2018
Forschungsinstitut für Mathematik, ETH, Zürich
Host: Peter Bühlmann

Invited Research Fellow — Aug 2017
Perelman School of Medicine, Upenn
Host: Hongzhe Li

Research Interests

High-dimensional statistical inference, causal inference, non-parametric statistics, and applications to genetics and health data.

- Confidence interval construction in high dimensions
- Causal inference with instrumental variables
- Semi-supervised inference
- Mediation analysis
- Additive models
- Heterogeneous data analysis

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.
- Dr. Chao Yong Chi-hsing Scholarship in Mathematics, CUHK — 2011
- Chung Chi Ivy League Exchange Scholarship, CUHK — 2010
- Dean's List, College of Arts and Science, University of Pennsylvania — 2010
- Dean's Honors List, Faculty of Science, CUHK — 2008, 2009
- Honors at Entrance to the Chinese University of Hong Kong (4 years) — 2008

PUBLICATIONS * indicates alphabetical ordering authorship; underline indicates the supervised students.

1. **Guo, Z.** (2021). Post-selection Problems for Causal Inference with Invalid Instruments: A Solution Using Searching and Sampling. *arXiv preprint arXiv:2104.06911.*
2. Hou, J., **Guo, Z.**, Cai, Tianxi (2021). Surrogate Assisted Semi-supervised Inference for High Dimensional Risk Prediction. *arXiv preprint arXiv:2105.01264.*
3. **Guo, Z.** (2020). Inference for High-dimensional Maximin Effects in Heterogeneous Regression Models Using a Sampling Approach. *arXiv preprint arXiv:2011.07568.*
4. **Guo, Z.**, Ćevid, D., & Bühlmann, P. (2020). Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding. *arXiv preprint arXiv:2004.03758.*
5. Li, S., & **Guo, Z.** (2020). Causal Inference for Nonlinear Outcome Models with Possibly Invalid Instrumental Variables. *arXiv preprint arXiv:2010.09922.*
6. Ma, R., **Guo, Z.**, Cai, T. T., and Li, H. (2020). Statistical Inference of Genetic Relatedness using High-Dimensional Logistic Regression.
7. *Cai, T. T., **Guo, Z.**, & Ma, R. (2020). Statistical Inference for High-Dimensional Generalized Linear Models with Binary Outcomes.
8. **Guo, Z.**, Rakshit, P., Herman, D., & Chen, J. (2019). Inference for Case Probability in High-dimensional Logistic Regression. *arXiv preprint arXiv:2012.07133.*
9. **Guo, Z.**, Renaux, C., Bühlmann, P., & Cai, T. T. (2019). Group Inference in High Dimensions with Applications to Hierarchical Testing. *arXiv preprint arXiv:1909.01503.*
10. **Guo, Z.** & Zhang, C. (2019). Local Inference in Additive Models with Decorrelated Local Linear Estimator. *arXiv preprint arXiv:1907.12732.*
11. **Guo, Z.** & Zhang, C. (2021+). Extreme Nonlinear Correlation for Multiple Random Variables and Stochastic Processes with Applications to Additive Models. *Stochastic Processes and their Applications*, to appear.
12. *Cai, Tianxi, Cai, T. T., & **Guo, Z.** (2021+). Optimal Statistical Inference for Individualized Treatment Effects in High-dimensional Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, to appear.
13. *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.
14. **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.
15. **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.
16. **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.
17. *Cai, T. T., & **Guo, Z.** (2018). Accuracy assessment for high-dimensional linear regression. *Annals of Statistics*, 46(4), 1807-1836.
18. **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.
19. 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.

20. *Cai, T. T., & **Guo, Z.** (2017). Confidence intervals for high-dimensional linear regression: Minimax rates and adaptivity. *Annals of Statistics*, 45(2), 615-646.
21. **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.
22. **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.
23. **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.

Grants

1. National Institute of Health
R01LM013614 "Semi-supervised Approaches to Denoising Electronic Health Records Data for Risk Prediction"
   - Role: Co-Principal Investigator (PI: Dr. Tianxi Cai)
   - Period: July 2021 to June 2025.
2. National Institute of Health
R01GM140463 "Predictive Modeling with High-Dimensional lncomplete Data."
   - Role: Principal Investigator
   - Period: July 2020 to June 2023.
3. National Science Foundation
DMS 1811857 "Inference in High-Dimensional Linear Models: Methods, Theory and Applications."
   - Role: Principal Investigator
   - Period: Aug 2018 to Aug 2021.
4. 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."
   - Role: Co-Principal Investigator (PI: Dr. Min-ge Xie)
   - Period: July 2020 to June 2023.
5. National Institute of Health
R56-HL-138306-01 "Statistics Methods for Analyzing Electronic Health Record Data."
   - Role: Co-Investigator (PI: Dr. Jinbo Chen)
   - Period: June 2018 to Aug 2018.
6. University of Pennsylvania Medical School
"Statistics Methods for Analyzing Electronic Health Record Data."
   - Role: Senior Investigator (PI: Dr. Jinbo Chen)
   - Period: June 2019.

Software

1. R package `maczic`: Mediation Analysis for Count and Zero-Inflated Count Data. Available at `https://cran.r-project.org/web/packages/maczic/index.html`.
2. R package `SIHR`: Statistical Inference in High-dimensional Regression. Available at `https://cran.r-project.org/web/packages/SIHR/index.html`
3. Control Function. Available at `https://github.com/zijguo/Control-function`
4. TSHT: Two Stage Hard Thresholding. Available at `https://github.com/hyunseungkang/invalidIV`
5. SpotIV: Semi-parametric outcome models with invalid IVs. Available at `https://github.com/saili0103/SpotIV`
6. SARM: Sampling Aggregation for Ridge-type Maximin effects. Available at `https://github.com/zijguo/Maximin-Inference`

R codes are available at `https://github.com/zijguo`.

Teaching Experience

Instructor

- Rutgers University

FSRM 588: Financial Data Mining	Fall 2021
STAT 384: Intermediate Statistical Analysis	Spring 2021
STAT 484: Basic Applied Statistics (Cross-listed with STAT 384)	
FSRM 588: Financial Data Mining	Spring 2021
FSRM 588: Financial Data Mining	Spring 2020
Instructor Rating: 4.50 out of 5.0	
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

Student Supervision

PhD Thesis Advisor: Prabrisha Rakshit (expected 2023)

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

PhD Thesis Committee: Sai Li (2018); Yisha Yao (2021); Xiaokang Luo (expected 2022).

Master and Undergraduate Research Advisor: Tai Yang (2018-2019); Saide Tang (2019-2020); Wei Yuan (2020-2021); Yunjiao Bai (2021); Shiyu Yang (2021); Zhenyu Wang (2021); Enyan Zhang (2021).

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); Lyujiangnan Ye (2021); Jung Hyun Kim (2021); Saide Tang (2021); Wei Yuan (2021).

Academic Service

1. Organizer of Causal Inference Reading Group (Joint with Nicole Pashley and Tirthankar Dasgupta), Department of Statistics, Rutgers
2. Department Retreat Chair (2019-2020), Department of Statistics, Rutgers
3. Department Seminar Chair (2018-2019), Department of Statistics, Rutgers
4. Other Rutgers Committee service:
   - Department website committee (2020-2021)
   - Department retreat committee (2017-2018)
   - FSRM/MSDS committee (2017-2018, 2018-2019, 2019-2020, 2020-2021)
   - Ph.D. Exam committee (2018-2019, 2019-2020)
   - Graduate Curriculum committee (2019-2020, 2020-2021)
5. Organizing Committee for 2019 Rutgers Statistics Symposium
6. Program Committee for ICSA 2019 11th International Conference
7. Local Organizing Committee for 2018 ICSA Applied Symposium.
8. Reviewer for the following journals: *Annals of Statistics*, *Journal of the American Statistical Association*, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *Biometrica*, *Journal of Machine Learning*, *Journal of Econometrics*, *Biometrics*,*Statistica Sinica*, *IEEE International Symposium on Information Theory*, *Journal of Applied Statistics*, *COLT*.

Invited Talks

1. Invited speaker (forthcoming), session on "High-dimensional statistical inference/learning", IMS Annual Meeting, London, UK, June 2022
2. Invited speaker (forthcoming), special session on "Statistical Learning" (organized by Peter Bühlmann, Giraud Christophe, Cun-Hui Zhang), AMS-EMS-SMF Joint International Meeting 2022, Grenoble, France, July 2022
3. Invited speaker (forthcoming), CMStatistics, Dec 2021
4. Invited speaker (virtual), JSM 2021, Seattle, USA, *"Inference for High-dimensional Maximin Effects in Heterogeneous Regression Models Using a Sampling Approach."* Aug 2021
5. Invited speaker (virtual), The First International Conference on Statistics and Related Fields, Luxembourg, *"Inference for High-dimensional Maximin Effects in Heterogeneous Regression Models Using a Sampling Approach."* , July 2021

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

6. Reading group meeting (virtual) led by Tianxi Cai, The Chan School of Public Health, Harvard University, *"Maximin Effect and Distributional Robustness: A Review and New Advances"*, June 2021
7. Department seminar (virtual), Department of Statistics, East China Normal University, Shanghai, China, *"Inference for High-dimensional Maximin Effects in Heterogeneous Regression Models Using a Sampling Approach."*, May 2021
8. Center for Causal Inference Seminar (virtual), University of Pennsylvania, USA, *"Post-selection Problems for Causal Inference with Invalid Instruments: A Solution Using Searching and Sampling."*, May 2021
9. Department Seminar (virtual), Department of Economics, Chinese University of Hong Kong, Hong Kong, China, *"Post-selection Problems for Causal Inference with Invalid Instruments: A Solution Using Searching and Sampling."*, April 2021
10. Department Seminar (virtual), Department of Statistics, Hong Kong University, Hong Kong, China, *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding."*, April 2021
11. Department Seminar (virtual), Medical School, University of Exeter, Exeter, UK, *"Inference for Non-linear Treatment Effects with Control Function Methods"*, Feb 2021
12. Invited talk (virtual), CMStatistics 2020, London, UK, *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding."*, Dec 2020
13. Department seminar (virtual), Department of Statistics, Cornell University, USA *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding."*, Oct 2020
14. Invited talk (virtual), JSM 2020, Philadelphia, USA, *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding and Measurement Errors."*, Aug 2020
15. Invited participant (virtual), Mathematical and Statistical Challenges in Uncertainty Quantification, Cambridge, UK, July 2020
16. Department seminar (virtual), Department of Statistics, UC Davis, USA, *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding and Measurement Errors."*, May 2020
17. Center for Causal Inference Seminar (virtual), University of Pennsylvania, USA, *"Doubly Debiased Lasso: High-Dimensional Inference under Hidden Confounding and Measurement Errors."*, May 2020
18. Biostatistics reading group (virtual, led by Jinbo Chen), University of Pennsylvania, USA, *"Group Inference in High Dimensions with Applications to Hierarchical Testing."*, May 2020
19. Department seminar, Department of Statistics, East China Normal University, Shanghai, China, *"Group Inference in High Dimensions with Applications to Hierarchical Testing"*, Dec 2019
20. Invited talk, 11th ICSA International Conference, Hangzhou, China, *"Group Inference in High Dimensions with Applications to Hierarchical Testing"*, Dec 2019
21. 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
22. 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
23. Department seminar, Department of Statistics, East China Normal University, Shanghai, China, *"Individualized Treatment Selection: A Hypothesis Testing Approach In High-dimensional Models"*, June. 2019

24. Invited talk, 2019 Hangzhou Data Science Conference, Hangzhou, China, *"Local Inference in High-dimensional Sparse Additive Modeling"*, May. 2019
25. 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
26. Department seminar, ISOM, HKUST, Hong Kong, China, *"Local Inference in High-dimensional Sparse Additive Modeling"*, May. 2019
27. Department seminar, Department of Statistics, University of Virginia, USA, *"Local Inference in High-dimensional Sparse Additive Modeling"*, March. 2019
28. Invited talk, 2019 ICSA Data Science Conference, Xishuangbanna, Yunnan, China. *"Individualized Treatment Selection: A Hypothesis Testing Approach In High-dimensional Models"*, Jan. 2019
29. 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
30. Department seminar, Seminar for Statistics, Department of Mathematics, ETH, Swiss, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Nov. 2018
31. Department seminar, Department of Mathematics, NJIT, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Nov. 2018
32. Department seminar, ORFE, Princeton, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Oct. 2018
33. Department seminar, ISOM, HKUST, Hong Kong, China, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, July. 2018
34. Department seminar, Department of Statistics, Naikai University, China, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, July. 2018
35. Invited talk, IMS Asia Pacific Rim Meeting, Singapore, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018
36. Invited talk, HongKong EcoStat Conference, Hong Kong, China, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018
37. Invited talk, ICSA Symposium 2018, New Brunswick, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018
38. Invited talk, Purdue Symposium on Statistics, USA, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, June. 2018
39. 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
40. Invited talk, Lorentz Center, Leiden University, Netherlands, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Apr. 2018

41. Department seminar, Department of Statistics, Columbia University, *"Semi-supervised Inference for Explained Variance in High-dimensional Linear Regression and Its Applications"*, Apr. 2018
42. Invited talk, Statistical Foundations of Uncertainty Quantification for Inverse Problem, Cambridge, *"Inference for Functionals in High-dimensional Linear Models"*, June. 2017
43. Seminar, Center for Statistical Methods in Big Data, University of Pennsylvania, *"Inference with High-dimensional Covariates and Possibly Invalid Instruments"*, Apr. 2017
44. Seminar, Institute of Data science, Fox Business School, Temple University, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Feb. 2017
45. Department seminar, Department of Biostatistics, UC Berkeley, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Feb. 2017
46. Department seminar, Department of Statistics, Rutgers, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Feb. 2017
47. Department seminar, Department of Statistics, University of Michigan, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017
48. Department seminar, Department of Statistics, University of Minnesota, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017
49. Department seminar, Department of Statistics, UIUC, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017
50. Department seminar, DPMMS, University of Cambridge, *"Inference for High Dimensional Linear Regression: Fundamental Limits and Algorithms"*, Jan. 2017
51. Department seminar, Department of Statistics, UC Santa Barbara, *"Inference for High Dimensional Linear Models: Fundamental Limits and Algorithms"*, Jan. 2017
52. Invited talk, Mathematical Meeting in Statistics, Fréjus, France, *"Optimal Estimation of Genetic Correlation in High-dimensional Linear Models"*, Dec. 2016
53. Econometrics Lunch, Department of Economics, University of Pennsylvania, *"Confidence Intervals for Treatment Effects in High-Dimensional Linear Models"*, Nov. 2016

Other Talks

1. Topic contributed talk, Joint Statistical Meetings, Baltimore, USA, *"Optimal Estimation of Co-Heritability in High-Dimensional Linear Models"*, Aug. 2017
2. Contributed talk, Joint Statistical Meetings, Chicago, USA, *"Accuracy Assessment for High-dimensional Linear Regression"*, Aug. 2016
3. Contributed talk, Eastern North American Region, Austin, USA, *"Confidence Intervals for High-Dimensional Linear Regression: Minimax Rates and Adaptivity"*, Mar. 2016
4. Poster presentation, John W. Tukey Conference, Princeton University, *"Confidence Intervals for High-Dimensional Linear Regression: Minimax Rates and Adaptivity"*, Sept. 2015
5. Contributed talk, Joint Statistical Meetings, Seattle, USA, *"Distance Matrix Estimation from Noisy Observation of Low Rank Position Matrix"*, Aug. 2015
6. Contributed talk, Joint Statistical Meetings, Boston, USA, *"Instrumental Variable Approach for Mediation Analysis of Count Model"*, Aug. 2014
7. Topic Contributed talk, Joint Statistical Meetings, Montreal, Canada, *" Instrumental Variable Approach for Mediation Analysis of Zero-Inflated Count Model"*, Aug. 2013
8. 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