Sizhu Lu

Email: sizhu_lu@berkeley.edu **Phone**: (857) 600-9526 **Address**: 397 Evans Hall, Berkeley, CA

Research interests

Causal inference, Semiparametric theory, Applied statistics

Education

University of California, Berkeley Berkeley, CA
Ph.D. candidate in Statistics
Advisor: Professor Peng Ding. *GPA: 4.00*

University of California, Berkeley Berkeley, CA
Ph.D. in Business Administration, Haas School of Business 08/2022
Advisor: Professor Amir Kermani

Peking University Beijing, China
B.S. in Mathematics and Applied Mathematics & B.A. in Finance 07/2017

Manuscripts

Design-based causal inference in bipartite experiments
Sizhu Lu*, Lei Shi*, Yue Fang, Wenxin Zhang, and Peng Ding, 2025+, *under review,* *arXiv preprint arXiv:2501.09844*

Estimating treatment effects with competing intercurrent events in randomized controlled trials
Sizhu Lu, Yanyao Yi, Yongming Qu, Karen Liu, Ting Ye, and Peng Ding, 2025+, *Major revision at Journal of the American Statistical Association,* *arXiv preprint arXiv:2503.03049*

Principal stratification with continuous post-treatment variables: nonparametric identification and semiparametric estimation
Sizhu Lu, Zhichao Jiang, and Peng Ding, 2025, *Journal of the Royal Statistical Society Series B: Statistical Methodology, in press*

Flexible sensitivity analysis for causal inference in observational studies subject to unmeasured confounding
Sizhu Lu and Peng Ding, 2023+, *under review,* *arXiv preprint arXiv:2305.17643*

The roles of estimands and assumptions in causal inference: Comment on "Chasing shadows: how implausible assumptions skew our understanding of causal estimands"
Sizhu Lu and Peng Ding, 2025, *Statistics in Biopharmaceutical Research, in press*

Selected research in progress

Estimating within-cluster and between-cluster spillover effects in randomized saturation designs
joint work with Lei Shi and Peng Ding

Randomized saturation designs are two-stage experiments: they first randomly assign treatment probabilities over the clusters and then randomly assign the treatment to the units within the clusters. The existing literature on randomized saturation designs focuses on estimating within-cluster spillover effects by assuming away between-cluster spillover effects. However, the units may interact across clusters in practice. A leading example is that some units are geographically close to each other, so spillover effects arise across clusters. Based on the potential outcomes framework, we formulate the causal inference problem of estimating within-cluster and between-cluster spillover effects in randomized saturation designs. We clarify the causal estimands and establish rigorous statistics theory for inference. We also apply our theory to analyze a recent randomized saturation design of cash transfer on household expenditure in Kenya.

Two-phase sampling for the local average treatment effect: efficient estimation and optimal design
joint work with Peng Ding
Instrumental variable methods are widely used for causal inference with unmeasured confounding, yet practical applications often involve partial data collection due to cost or design constraints. Motivated by modern large-scale randomized experiments with noncompliance and outcomes that are costly to measure, we study the identification and estimation of the local average treatment effect under two-phase sampling. We first show that the canonical two-stage least squares estimator must be weighted properly, and that the corresponding standard error must be modified according to the two-phase sampling design. We then propose a semiparametrically efficient and multiply robust estimator based on the efficient influence function. More importantly, we argue that two-phase sampling can utilize post-treatment variables that are predictive of the outcome, and discuss the corresponding estimation and inference problems. Our theory not only provides the basis for the optimal two-phase sampling but also covers the setting of missing outcomes under the local average treatment effect framework.

Simple yet efficient weighting estimation for the marginal quantile treatment effect (QTE) in randomized experiments
joint work with Peng Ding

Causal effect estimation with competing intercurrent events: treatment policy and hypothetical strategies
joint work with Yanyao Yi, Yongming Qu, Ting Ye, and Peng Ding

Honors and scholarships

Graduate Division Conference Travel Grant, UC Berkeley	2024
San Francisco ASA JSM Travel Award	2024
Society for Political Methodology Polmeth Meeting Travel Award	2024

Outstanding Graduate Student Instructor	2023
Outstanding Graduate of Beijing (Top 1%)	2017
Merit Student of Peking University (Top 2%)	2014 – 2016

Teaching experience

Graduate Student Instructor, Department of Statistics, UC Berkeley

STAT 230A: Linear Models	Spring 2021, Spring 2023
STAT 156/256: Causal Inference	Fall 2022, Fall 2023
STAT 215B: Statistical Models: Theory and Application	Spring 2024
STAT 135: Concepts of Statistics	Spring 2025

Graduate Student Instructor, Haas School of Business, UC Berkeley

MBA 283: Real Estate Finance and Securitization	Fall 2019
MFE 230I: Fixed Income Markets	Summer 2020
MFE 230M: Asset-Backed Security Markets	Fall 2020 – 2022

Instructor, Haas School of Business, UC Berkeley

Finance Net Present Value, Business Academy for Youth	Summer 2019

Talks and Posters

Principal stratification with continuous post-treatment variables

Talk at School of Management and Economics, CUHK	05/2024
Poster at 2024 American Causal Inference Conference	05/2024
Poster at Winter Workshop: Causal inference and its applications, University of Florida	01/2024
Poster at 2023 CLIMB retreat, UC Berkeley	11/2023

Flexible sensitivity analysis for causal inference in observational studies subject to unmeasured confounding

Talk at 2024 INFORMS, Invited paper session	10/2024
Poster at 2024 PolMeth	07/2024
Talk at Center for Statistical Science, Tsinghua University	06/2023

Estimating treatment effects with competing intercurrent events in randomized controlled trials

Talk at 2024 JSM Topic-Contributed Paper Session	08/2024
Talk at Eli Lilly and Company (G4S Seminar)	05/2024
Talk at 2024 Berkeley Statistics Annual Research Symposium	04/2024

Design-based causal inference in bipartite experiments

Poster at 2025 Berkeley Statistics Annual Research Symposium	09/2025
Poster at Experimental Designs in the Era of Artificial Intelligence Workshop	03/2025
Poster at 2024 Stanford Berkeley Joint Colloquium	10/2024

Estimating within-cluster and between-cluster spillover effects in randomized saturation designs

Talk at the Emory Causal Network Analysis Workshop	08/2025

Reviewing service	Journal of the American Statistical Association, Annals of Statistics, Biometrika, Journal of the Royal Statistical Society, Series B, Journal of Causal Inference, Statistica Sinica, Journal of Educational and Behavioral Statistics, Biometrical Journal, Biostatistics, Journal of Computational and Graphical Statistics, Sociological Methods and Research.

Industry experience	**Netflix**	Los Gatos, CA
	Experimentation and Causal Inference Intern	05/2023 – 08/2023
	Eli Lilly and Company	Berkeley, CA
	Statistician Co-op Intern	01/2024 – 05/2024
	Academic Contractor: Consulting Statistician	09/2024 – Present

Programming skills	Proficient in: R, Python, MATLAB, and Stata.