# Lihua Lei

Sequoia Hall, 390 Serra Mall, Stanford University
Stanford, CA 94305-4020

Website: https://lihualei71.github.io
Email: lihualei@stanford.edu

## Academic Appointments

Assistant Professor of Economics — 7/2022 - present
Stanford Graduate School of Business

Postdoctoral Researcher, Stanford University — 9/2019 - 7/2022
Advisor: Emmanuel Candès

## Education

Ph.D. Statistics, University of California, Berkeley — 8/2014 - 8/2019
Advisors: Peter, J. Bickel, and Michael, I. Jordan

B.S. Statistics, Peking University — 9/2010 - 6/2014
Undergraduate Thesis Advisor: Song Xi, Chen

B.A. Economics (double major), Peking University — 9/2011 - 6/2014

## Awards

IMS (Institute of Mathematical Statistics) New Researcher Travel Award, 2022.

Rising Star in Data Science (University of Chicago), 2021.

Eric Lehmann Citation (Annual Dissertation Award in Theoretical Statistics), 2019.

Citadel Fellowship, 2017-2018.

Outstanding Graduate Student Instructor Award, 2016.

ICML (International Conference on Machine Learning) Travel Award, 2016.

ACIC (Atlantic Causal Inference Conference) Travel Award, 2018, 2022.

Scholarship from The Sally and Terry Speed Graduate Support Fund, 2015.

## Teaching Experiences

TA, Statistics 210A, Theoretical Statistics A, Fall 2015 (taught by Michael, I. Jordan).

TA, Statistics 210B, Theoretical Statistics B, Spring 2016 (taught by Martin, J. Wainright).

## Research Experiences

## Research Interests

Econometrics and Statistics; Machine Learning;
Causal Inference; Uncertainty Quantification;
Large-scale Inference; Networks; Optimization

## Journal Publications [1]

1. Horváth, S.*, **Lei, L.***, Richtárik, P., and Jordan, M. I. (2022) Adaptivity of Stochastic Gradient Methods for Nonconvex Optimization. *SIAM Journal on Mathematics of Data Science (SIMODS)*, 4(2), 634-648.

2. Fithian, W. and **Lei, L.** (2021) Conditional calibration for false discovery rate control under dependence. To appear in *Annals of Statistics (AoS)*.

3. Bates, S.*, Angelopoulos, A.*, **Lei, L.***, Malik, J., and Jordan, M. I. (2021) Distribution-Free, Risk-Controlling Prediction Sets. *Journal of the ACM (JACM)*, 68(6), 1–34.

4. Loper, J. H.*, **Lei, L.***, Fithian, W., and Tansey, W. (2021) Smoothed Nested Testing on Directed Acyclic Graphs. To appear in *Biometrika*.

5. **Lei, L.** and Candès, E. J. (2021) Conformal Inference of Counterfactuals and Individual Treatment Effects. *Journal of the Royal Statistical Society: Series B (JRSS-B)*. 83(5), 911-938.

6. **Lei, L.** and Ding, P. (2020) Regression Adjustment in Randomized Experiments With A Diverging Number of Covariates. *Biometrika*. 108(4), 815-828.

7. Li, T., **Lei, L.**, Bhattacharyya, S., Sarkar, P., Bickel, P. J., and Levina, E. (2020) Hierarchical community detection by recursive partitioning. To appear in *Journal of the American Statistical Association (JASA)*.

8. **Lei, L.** and Bickel, P. J. (2020) An Assumption-Free Exact Test For Fixed-Design Linear Models With Exchangeable Errors. *Biometrika*, 108(2), 397-412.

9. **Lei, L.**, Ramdas, A., and Fithian, W. (2020) A general interactive framework for FDR control under structural constraints. *Biometrika*, 108(2), 253-267.

10. **Lei, L.** and Jordan, M. I. (2020) On the Adaptivity of Stochastic Gradient-Based Optimization. *SIAM Journal on Optimization (SIOPT)*, 30(2), 1473-1500.

11. D'Amour, A., Ding, P., Feller, A., **Lei, L.**, and Sekhon, J. (2019) Overlap in High Dimensional Observational Studies. *Journal of Econometrics (JoE)*, 221(2), 644-654.

12. **Lei, L.**, and Fithian, W. (2018). AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *Journal of the Royal Statistical Society: Series B (JRSS-B)*, 80(4), 649-679.

13. **Lei, L.**, Bickel, P. J., and El Karoui, N. (2018). Asymptotics For High Dimensional Regression M-Estimates: Fixed Design Results. *Probability Theory and Related Fields (PTRF)*, 172(3-4), 983-1079.

[1] * = alphabetical ordering or equal contribution

14. Chen, S. X., **Lei, L.**, and Tu, Y. (2016). Functional Coefficient Moving Average Model with Applications to forecasting Chinese CPI. *Statistica Sinica*, 26, 1649-1672.

## Conference Publications

1. Elibol, M., **Lei, L.**, and Jordan, M. I. (2020). Variance Reduction with Sparse Gradients. To appear in *International Conference on Learning Representations (ICLR)*.

2. Ye, Y., **Lei, L.**, and Ju, C. (2018). HONES: A Fast and Tuning-free Homotopy Method For Online Newton Step. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*.

3. **Lei, L.**, Ju, C., Chen, J., and Jordan, M. I. (2017). Nonconvex Finite-Sum Optimization Via SCSG Methods. In *Proceedings of the 30th Advances in Neural Information Processing Systems (Neurips)*.

4. **Lei, L.** and Jordan, M. I. (2017). Less than a Single Pass: Stochastically Controlled Stochastic Gradient. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*.

5. **Lei, L.** and Fithian, W. (2016). Power of Ordered Hypothesis Testing. In *Proceedings of the 33th International Conference on Machine Learning (ICML)*.

## Under Revision

1. Arkhangelsky, D.*, Imbens, G. W.*, **Lei, L.***, and Luo, X.* (2021) Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *Arxiv e-prints, abs/2107.13737*. Under R&R at *Quantitative Economics (QE)*.

2. Bates, S.*, Candès, E. J.*, **Lei, L.***, Romano, Y.*, and Sesia, M.* (2021) Testing for Outliers with Conformal p-values. *ArXiv e-prints, abs/2104.08279*. Under Major Revision at *Annals of Statistics (AoS)*.

3. Candès, E. J.*, **Lei, L.***, Ren, Z.* (2021) Conformalized Survival Analysis. *ArXiv e-prints, abs/2103.09763*. Under Major Revision at *Journal of the Royal Statistical Society: Series B (JRSS-B)*.

## Preprints and Submissions

1. Angelopoulos, A.*, Bates, S.*, Candès, E. J.*, Jordan, M. I.*, and **Lei, L.*** (2021) Learn then Test: Calibrating Predictive Algorithms to Achieve Risk Control. *Arxiv e-prints, abs/2110.01052*.

2. Yang, C., **Lei, L.**, Ho, N., and Fithian, W. (2021) BONuS: Multiple multivariate testing with a data-adaptive test statistic. *ArXiv e-prints, abs/2106.15743*.

3. **Lei, L.***, Li, X.*, and Lou, X.* (2020) Consistency of Spectral Clustering on Hierarchical Stochastic Block Models. *ArXiv e-prints, abs/2004.14531*.

4. **Lei, L.** (2019) Unified $\ell_{2\to\infty}$ Eigenspace Perturbation Theory for Symmetric Random Matrices. *ArXiv e-prints, abs/1909.04798*.

### Software

1. `adaptMT`: `R` package on Adaptive P-value Thresholding (on CRAN);
2. `cfcausal`: `R` package on conformal inference of counterfactuals and individual treatment effects (on github);
3. `dbh`: `R` package on dependence-adjusted Benjamini-Hochberg and general step-up procedures (on github);
4. `mkn`: `R` package on multiple knockoffs based inference (on github).
5. `ovalue`: `R` package on distribution-free assessment of population overlap (O-values) for observational studies (on github).

## Professional Services

### Invited Talks (no interview/job talks)

1. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Joint Statistical Meetings*, Aug. 2022.
2. Adaptivity of Stochastic Gradient-based Methods. *International Conference on Continuous Optimization*, Jul. 2022.
3. Learn then Test: Calibrating Predictive Algorithms to Achieve Risk Control. *ICSA China Conference*, Jul. 2022.
4. Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *ICSA China Conference*, Jul. 2022.
5. Conformal Inference of Counterfactuals and Individual Treatment Effects. *First Workshop on Interactive Causal Learning*, Jun. 2022.
6. Conformal Inference of Counterfactuals and Individual Treatment Effects. *IMS Annual in Probability and Statistics*, Jun. 2022.
7. Cyclic Permutation Test for Linear Models. *Simons Institute for the Theory of Computing, University of California, Berkeley*, Jun. 2022.
8. Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *New England Statistics Symposium*, May. 2022.
9. Conformal Inference of Counterfactuals and Individual Treatment Effects. *New England Statistics Symposium*, May. 2022.
10. Learn then Test: Calibrating Predictive Algorithms to Achieve Risk Control. *Tsinghua IIIS Seminar on Foundations Of Data Science (FODS)*, Apr. 2022
11. Conformal Inference of Counterfactuals and Individual Treatment Effects. *CIRG Seminar, North Carolina University*, Apr. 2022.

12. Discussion of "Inference on weighted average value function in high-dimensional state space". *ASSA 2022 Annual Meeting*, Jan. 2022.

13. Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *One World YoungStatS (Webinar)*, Dec. 2021.

14. Testing for Outliers with Conformal p-values. *Royal Bank of Canada*, Dec. 2021.

15. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Microsoft Research*, Dec. 2021.

16. AdaPT: An interactive procedure for multiple testing with side information. *Journal of Royal Statistical Society Webinar*, Dec. 2021.

17. Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *Stanford University (Causal Inference Group)*, Oct. 2021.

18. Conformal Inference of Counterfactuals and Time-to-event Outcomes. *University of Washington (Statistics Seminar)*, Oct. 2021.

19. Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *University of California, Davis (Econometrics Seminar)*, Oct. 2021.

20. Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *Harvard University (Econometrics Workshop)*, Sep. 2021.

21. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Temple University (Statistics Seminar)*, Sep. 2021.

22. Conformal Inference of Counterfactuals and Individual Treatment Effects. *New York University (Data Science Lunch Seminar)*, Sep. 2021.

23. Conformal Inference of Counterfactuals and Individual Treatment Effects. *ICSA Applied Statistics Symposium*, Sep. 2021.

24. Testing for Outliers with Conformal p-values. *12th International Conference on Multiple Comparison Procedures (MCP)*, Aug. 2021.

25. Distribution-Free Assessment of Population Overlap in Observational Studies. *ICML "Neglected Assumptions of Causal Infernece" workshop*, Aug. 2021.

26. Conformalized Survival Analysis. *IFDS Summer Workshop*, Aug. 2021.

27. Conformal Inference of Counterfactuals and Individual Treatment Effects. *INFORMS Virtual Healthcare Conference*, Jul. 2021.

28. Conformal Inference of Counterfactuals and Individual Treatment Effects. *The Hong Kong University of Science & Technology*, Jun. 2021.

29. AdaPT: An interactive procedure for multiple testing with side information. *WNAR*, Jun. 2021.

30. Conformal Inference of Counterfactuals and Time-to-event Outcomes. *ETH Zürich (Young Data Science Seminar)*, Jun. 2021.

31. Hierarchical Community Detection for Heterogeneous and Multi-scaled Networks. *University of California, Berkeley (Neyman Seminar)*, May. 2021.

32. Conformal Inference of Counterfactuals and Time-to-event Outcomes. *Stanford University (Statistics Seminar)*, May. 2021.

33. Conformal Inference of Counterfactuals and Time-to-event Outcomes. *University of Chicago (Statistics Seminar)*, May. 2021.

34. Hierarchical Community Detection for Heterogeneous and Multi-scaled Networks. *Stanford University (RAIN Seminar)*, Apr. 2021.

35. Conformal Inference of Counterfactuals and Individual Treatment Effects. *U Chicago Crime Lab (Data Science Meeting)*, Apr. 2021.

36. Cyclic Permutation Test for Linear Models. *University of Chicago, Booth School of Business (Econometrics and Statistics Colloquium Workshop)*, Apr. 2021.

37. Conformal Inference of Counterfactuals and Individual Treatment Effects. *University of Maryland, Baltimore County (Statistics Seminar)*, Apr. 2021.

38. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Dataiku research webinar*, Apr. 2021.

39. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Northwestern University (Biostatistics & Statistics Joint Seminar Series)*, Apr. 2021.

40. Conformal Inference of Counterfactuals and Individual Treatment Effects. *University of Manchester (CfE Biostatistics Forum)*, Mar. 2021.

41. Conformal Inference of Counterfactuals and Individual Treatment Effects. *The London School of Hygiene & Tropical Medicine (LSHTM)*, Mar. 2021.

42. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Computer Vision Talks Webinar*, Mar. 2021.

43. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Facebook Core Data Science*, Mar. 2021.

44. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Univerity of California, Los Angeles (Big Data and ML Seminar)*, Feb. 2021.

45. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Univerity of Missouri (Biostatistics Group Seminar)*, Feb. 2021.

46. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Leiden Univerity (Stéphanie van der Pas and Aad van der Vaart's group meeting)*, Feb. 2021.

47. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Twitter*, Feb. 2021.

48. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Stitch Fix*, Feb. 2021.

49. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Rutgers University (Statistics Department Seminar)*, Feb. 2021.

50. Distribution-Free Assessment of Population Overlap in Observational Studies. *University of Pennsylvania (Statistics Department Seminar)*, Jan. 2021.

51. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Lyft*, Dec. 2020.

52. Conformal Inference of Counterfactuals and Individual Treatment Effects. *John Hopkins University (Causal Reading Group)*, Dec. 2020.

53. AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *Florida State University (Hongyuan Cao's Class, Guest Lecture)*, Nov. 2020.

54. Adaptivity of Stochastic Gradient-based Methods. *University of California, Berkeley (Jiantao Jiao's group)*, Oct. 2020.

55. Hierarchical Community Detection for Heterogeneous and Multi-scaled Networks. *University of North Carolina at Chapel Hill (Statistics Department Seminar)*, Oct. 2020.

56. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Peking University (Biostatistics Department Seminar)*, Oct. 2020.

57. Conformal Inference of Counterfactuals and Individual Treatment Effects. *The 2020 Pacific Causal Inference Conference*, Sep. 2020.

58. Conformal Inference of Counterfactuals and Individual Treatment Effects. *Online Causal Inference Seminar*, Jul. 2020.

59. AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *International Seminar on Selective Inference*, Jul. 2020.

60. AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *Southwestern University of Finance and Economics*, Jul. 2020.

61. Conformal Inference of Counterfactuals and Individual Treatment Effects. *2020 BAAI Conference*, Jun. 2020.

62. Fair and Efficient Allocation of Medical Resources During COVID-19. *Journal Club, Stanford University*, Jun. 2020.

63. Survival Analysis of COVID-19 Patients. *Data Studio, Stanford University*, May. 2020.

64. AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *University of California, Berkeley (Michael I. Jordan's group)*, May. 2020.

65. Fair and Efficient Allocation of Medical Resources During COVID-19. *University of California, Berkeley (Sandrine Dudoit's group)*, May. 2020.

66. AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *WNAR 2020 Annual Meeting*, Jun. 2020. (Cancelled due to COVID-19).

67. Unified $\ell_{2\to\infty}$ Eigenspace Perturbation Theory for Symmetric Random Matrices. *SIAM Conference on Mathematics of Data Science (MDS20)*, May. 2020. (Cancelled due to COVID-19).

68. BONuS: Adaptive Multiple Testing With Multivariate Test Statistics. *Banff Workshop*, Mar. 2020. (Cancelled due to COVID-19).

69. Model-Free Assessment of Overlap in Observational Studies. *University of California, Berkeley (Causal Reading Group)*, Jan. 2020.

70. Model-Free Assessment of Overlap in Observational Studies. *SAMSI Workshop At Duke University)*, Dec. 2019.

71. BONuS: Adaptive Multiple Testing With Multivariate Test Statistics. *Facebook Core Data Science*, Nov. 2019.

72. Doubly robust two-way fixed effects regression for panel data. *Third Annual Berkeley-Stanford Econometrics Jamboree)*, Oct. 2019.

73. Regression Adjustment in Randomized Experiments With A Diverging Number of Covariates. *Joint Statistical Meeting, Topic Contributed Session*, Aug. 2019.

74. Hierarchical community detection by recursive partitioning. *ICSA Applied Statistics Symposium*, Jun. 2019.

75. Model-Free Assessment of Overlap in Observational Studies. *Yale University (Quantitative Research Methods Workshop)*, Feb. 2019.

76. Hierarchical community detection by recursive partitioning. *University of California, Davis (Statistics Student Seminar)*, Jan. 2019.

77. Model-Free Assessment of Overlap in Observational Studies. *Stanford University (Stefan Wager's group)*, Jan. 2019.

78. Statistical Inference in Moderate Dimensions. *University of South California (Statistics Department Seminar)*, Dec. 2018.

79. Hierarchical community detection by recursive bi-partitioning. *University of Michigan (Liza Levina's group)*, Oct. 2018.

80. Statistical Inference in Moderate Dimensions. *University of Michigan (Student Seminar)*, Oct. 2018.

81. Regression Adjustment in Randomized Experiments With A Diverging Number of Covariates. *University of Michigan (Xuming He's group)*, Oct. 2018.

82. AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *International Conference on Econometrics and Statistics (EcoSta)*, Jun. 2018.

83. Regression Adjustment in Randomized Experiments With A Diverging Number of Covariates. *Atlantic Causal Inference Conference (ACIC)*, May. 2018.

84. Stochastically Controlled Stochastic Gradient (SCSG) Method. *University of California, Davis (Cho-Jui Hsieh's group)*, Mar. 2018.

85. AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *Stanford University (David Tse's group)*, Feb. 2018.

86. Asymptotics For High Dimensional Regression M-Estimates: Fixed Design Results. *Berkeley-Stanford Econometrics Jamboree*, Nov. 2017.

## Reviewing (reviews of revisions excluded)

**Statistics Journals (40):** (#papers in parentheses) Annals of Statistics (AoS, 5), Journal of the American Statistical Association (JASA, 6), Biometrika (6), Journal of the Royal Statistical Society-Series B (JRSS-B, 3), Journal of the Royal Statistical Society-Series A (JRSS-A, 1), Bernoulli (1), Electronic Journal of Statistics (EJS, 2), Biometrics (3), Journal of Causal Inference (JCL, 2), Stat (1), Springer Book (1), SpringerBriefs (1), International Journal of Biostatistics (IJB, 1), Journal of Computational and Graphical Statistics (JCGS, 2), Statistica Sinica (1), International Journal of Approximate Reasoning (IJA, 1), Statistics in Biosciences (SIBS, 1), Journal of Statistical Planning and Inference (JSPI, 1).

**Economics Journals (3):** (#papers in parentheses) Econometrica (1), Journal of Political Economy (JPE, 1), Journal of Applied Econometrics (JAE, 1).

**Other Journals (19):** SIAM Journal on Mathematics of Data Science (SIMODS, 1), Operations Research (1), Management Science (1), Journal of Machine Learning Research (JMLR, 7), Transactions on Pattern Analysis and Machine Intelligence (TPAMI, 3), Computational Optimization and Applications (COAP, 1), Optimization Methods and Software (GOMS, 1), Artificial Intelligence (1), Information and Inference: A Journal of the IMA (IMAIAI, 2), Vietnam Journal of Mathematics (VJOM, 1), IEEE Transactions on Information Theory (1).

**Conferences:** (year in parentheses) ICML (2019, 2020, 2021), NeurIPS (2019, 2020, 2021), COLT (2019, 2020, 2021), AISTATS (2019), UAI (2020, 2021)

## Other Academic Services

- Co-organizer of the International Seminar on Selective Inference (with Will Fithian, Rina Barber and Daniel Yekutieli).
- Co-organizer of the GRoup-Of-Why (GROW) seminar, a causal reading group at Stanford University (with Guillaume Basse and Dominik Rothenhäusler).
- Co-organizer of the Banff workshop "The Interface Between Selective Inference and Machine Learning" (with Will Fithian, Rina Barber and Daniel Yekutieli, cancelled due to the pandemic).
- Co-organizer of the Banff workshop "Causal Inference with Big Data" (with Peng Ding, Marloes Maathuis and Fabrizia Mealli, cancelled due to the pandemic).