COURSE READINGS

Adaptive Confidence Intervals

1. Wager, Stefan, and Guenther Walther. "Uniform Convergence of Random Forests via Adaptive Concentration." *Working Paper* (2015). http://arxiv.org/pdf/1503.06388.pdf

Applications

1. Winkler, W. E. (2014), “Matching and record linkage.” *WIREs Comp Stat*, 6: 313–325. doi: 10.1002/wics.1317 https://www.census.gov/srd/papers/pdf/rr93-8.pdf
2. Gilchrist DS, Sands EG. “Something to Talk About: Social Spillovers in Movie Consumption.” *Journal of Political Economy*. Forthcoming. http://scholar.harvard.edu/files/dgilchrist/files/socialspillovers.pdf

Average Treatment Effects in Low Dimensions

1. Imbens G. “Matching Methods in Practice.” *Journal of Human Resources*, 2015;50(2):373-419. http://jhr.uwpress.org/cgi/reprint/50/2/373
2. Imbens G. “Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review.” *Review of Economics and Statistics*, 2004;86(1):4-29. https://doi.org/10.1162/003465304323023651

Bayesian Methods

1. Hugh A. Chipman, Edward I. George and Robert E. McCulloch. “BART: BAYESIAN ADDITIVE REGRESSION TREES” *The Annals of Applied Statistics* Vol. 4, No. 1 (March 2010), pp. 266-298
2. Hill, J. L. (2011). “Bayesian nonparametric modeling for causal inference.” *Journal of Computational and Graphical Statistics, 20*(1), 217. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1198/jcgs.2010.08162
3. Brown, L., Greenshtein, E. “Nonparametric Empirical Bayes and Compound Decision Approaches to Estimation of a High-Dimensional Vector of Normal Means” *The Annals of Statistics* Vol. 37, No. 4 (Aug., 2009), pp. 1685-1704. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/30243684
4. Robert, C. “Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction.” *Chance*, Volume 25, Issue 3, 2012, pp 59-61. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1080/09332480.2012.726568

Boosting

1. Zhang, Tong, and Bin Yu. “Boosting with Early Stopping: Convergence and Consistency”. *The Annals of Statistics* 33.4 (2005): 1538–1579. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/3448617

Cross-Validation

1. Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. Statistics Surveys, 4, 40–79. doi:10.1214/09-SS054 http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1214/09-SS054

2. Banko, M., & Brill, E. (2001). Scaling to very very large corpora for natural language disambiguation. Proceedings of the 39th Annual Meeting on Association for Computational Linguistics. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.3115/1073012.1073017
3. Cawley, G., & Talbot, N. (2007). "Preventing over-fitting during model selection via Bayesian regularisation of the hyper-parameters." *The Journal of Machine Learning Research*, 8, 841–861. http://dl.acm.org/citation.cfm?id=1248690
4. Syed, Ali R., "A Review of Cross Validation and Adaptive Model Selection." Thesis, *Georgia State University*, 2011. http://scholarworks.gsu.edu/math_theses/99
5. Szafranski, M., Grandvalet, Y., & Morizet-Mahoudeaux, P. (2008). "Hierarchical penalization." 1–8. https://papers.nips.cc/paper/3338-hierarchical-penalization.pdf
6. Zheng, A., & Bilenko, M. (2013). Lazy paired hyper-parameter tuning. *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*. http://dl.acm.org/citation.cfm?id=2540404

Heterogeneous Treatment Effects: Adaptive Experiments and Contextual Bandits

1. Li, Lihong, et al. "A contextual-bandit approach to personalized news article recommendation." Proceedings of the 19th international conference on World wide web. ACM, 2010. https://arxiv.org/abs/1003.0146
2. Bastani, Hamsa, and Mohsen Bayati. "Online Decision-Making with High-Dimensional Covariates." (2015)

Heterogeneous Treatment Effects: Multiple Testing

1. Wager, Stefan, and Guenther Walther. "Uniform Convergence of Random Forests via Adaptive Concentration." *arXiv preprint* (2015). http://arxiv.org/abs/1503.06388#
2. List, John A., Azeem M. Shaikh, and Yang Xu. "Multiple Hypothesis Testing in Experimental Economics". No. w21875. *National Bureau of Economic Research*, 2016. http://www.nber.org/papers/w21875
3. Deloukas, Panos, et al. "Large-scale association analysis identifies new risk loci for coronary artery disease." *Nature genetics* 45.1 (2013): 25-33. http://www.nature.com/ng/journal/v45/n1/abs/ng.2480.html
4. Heckman, J., Pinto, R., Shaikh, A., and Yavitz, A. (2011). Inference with imperfect randomization: The case of the Perry preschool program. Technical report, National Bureau of Economic Research. http://www.nber.org/papers/w16935
5. Lee, S. and Shaikh, A. M. (2014). Multiple testing and heterogeneous treatment effects: Re-evaluating the effect of progresa on school enrollment. Journal of Applied Econometrics, 29(4):612–626 http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=95682761&site=ehost-live&scope=site
6. Romano, J. andWolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. Journal of the American Statistical Association, 100(469):94–108 http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/27590521
7. Bonferroni, C. E. ``Il calcolo delle assicurazioni su gruppi di teste.'' In *Studi in Onore del Professore Salvatore Ortu Carboni*. Rome: Italy, pp. 13-60, 1935.
8. Holm, Sture. "A Simple Sequentially Rejective Multiple Test Procedure". Scandinavian Journal of Statistics 6.2 (1979): 65–70. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/4615733

Heterogeneous Treatment Effects: Optimal Policy Estimation

1. Athey, Susan, and Stefan Wager. "Efficient Policy Learning." *arXiv preprint arXiv:1702.02896* (2017). https://arxiv.org/abs/1702.02896
2. Dudík, Miroslav, John Langford, and Lihong Li. "Doubly robust policy evaluation and learning." *arXiv preprint arXiv:1103.4601* (2011). http://arxiv.org/pdf/1103.4601.pdf
3. Beygelzimer, Alina, and John Langford. "The offset tree for learning with partial labels." *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009. http://arxiv.org/pdf/0812.4044.pdf
4. Bhattacharya, Debopam, and Pascaline Dupas. "Inferring welfare maximizing treatment assignment under budget constraints." *Journal of Econometrics*167.1 (2012): 168-196. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1016/j.jeconom.2011.11.007
5. Dehejiia, Rajeev. "Program evaluation as a decision problem." *Journal of Econometrics*, Volume 125, Issues 1–2, March–April 2005, Pages 141–173. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1016/j.jeconom.2004.04.006
6. Manski, C. F. (2004). "STATISTICAL TREATMENT RULES FOR HETEROGENEOUS POPULATIONS." *Econometrica, 72*(4), 1221-1246. http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/203889282?accountid=14026
7. Keisuke Hirano and Jack R. Porter. "Asymptotics for Statistical Treatment Rules." *Econometrica*, Vol. 77, No. 5 (Sep., 2009), pp. 1683-1701. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/25621374

Heterogeneous Treatment Effects: Treatment Effect Estimation

1. Hartford, Jason, et al. "Counterfactual Prediction with Deep Instrumental Variables Networks." *arXiv preprint arXiv:1612.09596* (2016). https://arxiv.org/abs/1612.09596
2. Wager, Stefan, and Susan Athey. "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests." *arXiv preprint arXiv:1510.04342* (2015). http://arxiv.org/abs/1510.04342
3. Athey, Susan, and Guido W. Imbens. "Recursive Partitioning for Heterogeneous Causal Effects." *stat* 1050 (2015): 30. http://arxiv.org/abs/1504.01132
4. Athey, Susan, Julie Tibshirani, and Stefan Wager. "Solving Heterogeneous Estimating Equations with Gradient Forests." *arXiv preprint arXiv:1610.01271* (2016). https://www.gsb.stanford.edu/gsb-cmis/gsb-cmis-download-auth/425771
5. Cook, David I., Val J. Gebski, and Anthony C. Keech. "Subgroup analysis in clinical trials." *Medical Journal of Australia* 180.6 (2004): 289 http://obgyn.queensu.ca/assets/pdf/Subgroup_analysis.pdf
6. WILLKE, RJ; et al. From concepts, theory, and evidence of heterogeneity of treatment effects to methodological approaches: a primer. *BMC Medical Research Methodology*. 12, 1, 185-196, Jan. 2012. http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=85621819&site=ehost-live&scope=site
7. Lee, Myoung-Jae. "Non-parametric Tests for Distributional Treatment Effect for Randomly Censored Responses". Journal of the Royal Statistical Society. Series B (Statistical Methodology) 71.1 (2009): 243–264. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/20203886
8. Crump, Richard K. et al.. "Nonparametric Tests for Treatment Effect Heterogeneity". *The Review of Economics and Statistics* 90.3 (2008): 389–405. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/40043155
9. Zeileis, Achim, Torsten Hothorn, and Kurt Hornik. "Model-based Recursive Partitioning". *Journal of Computational and Graphical Statistics* 17.2 (2008): 492–514. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/27594318

10. Taddy, Matt, et al. "A nonparametric Bayesian analysis of heterogeneous treatment effects in digital experimentation." *arXiv preprint arXiv:1412.8563* (2014). http://arxiv.org/pdf/1412.8563.pdf
11. Weisberg, H. I., & Pontes, V. P. (2015). Post hoc subgroups in clinical trials: Anathema or analytics? *Clinical Trials, 12*(4), 357-364. http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/1698505816?accountid=14026
12. Lu Tian, Ash A. Alizadeh, Andrew J. Gentles & Robert Tibshirani. "A Simple Method for Estimating Interactions Between a Treatment and a Large Number of Covariates." *Journal of the American Statistical Association*, Volume 109, Issue 508, 2014, 1517-1532. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1080/01621459.2014.951443
13. Assmann, S. F., Pocock, S. J., Enos, L. E., & Kasten, L. E. (2000). Subgroup analysis and other (mis)uses of baseline data in clinical trials. The Lancet, 355(9209), 1064-9. http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/198999014?accountid=14026
14. Foster, J. C., Taylor, J. M.G. and Ruberg, S. J. (2011), "Subgroup identification from randomized clinical trial data." *Statist. Med*., 30: 2867–2880. doi:10.1002/sim.4322. http://onlinelibrary.wiley.com/doi/10.1002/sim.4322/abstract
15. Green, D. P., & Kern, H. L. (2012). Modeling heterogeneous treatment effects in survey experiments with bayesian additive regression trees. *Public Opinion Quarterly, 76*(3), 491. http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ufh&AN=82329247&site=ehost-live&scope=site
16. Signorovitch, J. "Identifying Informative Biological Markers in High-dimensional Genomic Data and Clinical Trials" *Harvard University*, 2007. ISBN 9780549036524
17. Su, Xiaogang, et al. "Subgroup analysis via recursive partitioning." The Journal of Machine Learning Research 10 (2009): 141-158. http://www.jmlr.org/papers/volume10/su09a/su09a.pdf
18. Imai, Kosuke; Ratkovic, Marc. Estimating treatment effect heterogeneity in randomized program evaluation. Ann. Appl. Stat. 7 (2013), no. 1, 443--470. doi:10.1214/12-AOAS593. http://projecteuclid.org/euclid.aoas/1365527206
19. Rosenblum, M., & van der Laan, ,M.J. (2011). Optimizing randomized trial designs to distinguish which subpopulations benefit from treatment. Biometrika, 98(4), 845. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1093/biomet/asr055
20. Benjamini, Yoav, and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". Journal of the Royal Statistical Society. Series B (Methodological) 57.1 (1995): 289–300. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/2346101
21. Jason Hartford, Greg Lewis, Kevin Leyton-Brown, Matt Taddy, "Counterfactual Prediction with Deep Instrumental Variables Networks" https://arxiv.org/abs/1612.09596

High Dimensional Methods Causal

1. Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. "LASSO methods for Gaussian instrumental variables models." (2011). http://arxiv.org/pdf/1012.1297.pdf
2. Caner, Mehmet, and Q. Fan. *The adaptive lasso method for instrumental variable selection*. Working Paper, North Carolina State University, 2010. http://apps.olin.wustl.edu/MEGConference/Files/pdf/2010/1.pdf
3. Gautier, E., & Tsybakov, A. (2014). *High-dimensional instrumental variables regression and confidence sets*. St. Louis: Federal Reserve Bank of St Louis. http://arxiv.org/abs/1105.2454
4. Bai, Jushan, and Serena Ng. "Selecting instrumental variables in a data rich environment." *Journal of Time Series Econometrics* 1.1 (2008). http://www.columbia.edu/~sn2294/papers/ivboost.pdf

5. Patorno, Elisabetta, et al. "Studies with many covariates and few outcomes: selecting covariates and implementing propensity-score–based confounding adjustments." *Epidemiology* 25.2 (2014): 268-278. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1097/EDE.0000000000000069
6. Rassen, Jeremy A., et al. "Covariate selection in high-dimensional propensity score analyses of treatment effects in small samples." *American journal of epidemiology* 173.12 (2011): 1404-1413. http://aje.oxfordjournals.org/content/173/12/1404
7. Hansen, C., & Kozbur, D. (2014). Instrumental variables estimation with many weak instruments using regularized JIVE. *Journal of Econometrics, 182*(2), 290. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/doi:10.1016/j.jeconom.2014.04.022
8. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology (Cambridge, Mass)*. 2009;20(4):512-522. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3077219/
9. Toh, S., García Rodríguez, L. A. and Hernán, M. A. (2011), Confounding adjustment via a semi-automated high-dimensional propensity score algorithm: an application to electronic medical records. Pharmacoepidem. Drug Safe, 20: 849–857. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1002/pds.2152
10. Westreich, D., Lessler, J., & Funk, M. J. (2010). Propensity score estimation: Neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *Journal of Clinical Epidemiology, 63*(8), 826-33. doi: http://dx.doi.org/10.1016/j.jclinepi.2009.11.020 http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/1033201130?accountid=14026
11. Ellis, Alan R., et al. "Confounding control in a nonexperimental study of STAR* D data: logistic regression balanced covariates better than boosted CART." *Annals of epidemiology* 23.4 (2013): 204-209. http://onlinelibrary.wiley.com/doi/10.1002/pds.2152/abstract
12. Wyss, Richard, et al. "The role of prediction modeling in propensity score estimation: an evaluation of logistic regression, bcart, and the covariate-balancing propensity score." *American journal of epidemiology* 180.6 (2014): 645-655. http://aje.oxfordjournals.org/content/180/6/645
13. Pirracchio, Romain, Maya L. Petersen, and Mark van der Laan. "Improving propensity score estimators' robustness to model misspecification using super learner." *American journal of epidemiology* 181.2 (2015): 108-119. http://aje.oxfordjournals.org/content/181/2/108
14. Chernozhukov, Victor, Christian Hansen, and Martin Spindler. "Valid Post-Selection and Post-Regularization Inference: An Elementary, General Approach." (2014). http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1146/annurev-economics-012315-015826
15. Chan, K. C. G., Yam, S. C. P. and Zhang, Z. (2015), Globally efficient non-parametric inference of average treatment effects by empirical balancing calibration weighting. Journal of the Royal Statistical Society: Series B (Statistical Methodology). doi: 10.1111/rssb.12129 http://onlinelibrary.wiley.com/doi/10.1111/rssb.12129/abstract
16. Bloniarz, Adam, et al. "Lasso adjustments of treatment effect estimates in randomized experiments." *arXiv preprint arXiv:1507.03652* (2015). http://arxiv.org/pdf/1507.03652.pdf
17. Farrell, Max H. "Robust inference on average treatment effects with possibly more covariates than observations." *Journal of Econometrics* 189.1 (2015): 1-23. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1016/j.jeconom.2015.06.017
18. IMAI, K; RATKOVIC, M. Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 76, 1, 243-263, Jan. 2014. ISSN: 13697412. http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=93468695&site=ehost-live&scope=site
19. Javanmard, Adel, and Andrea Montanari. "Confidence intervals and hypothesis testing for high-dimensional regression." *The Journal of Machine Learning Research* 15.1 (2014): 2869-2909. http://arxiv.org/pdf/1306.3171.pdf

20. Setoguchi, S., Schneeweiss, S., Brookhart, M. A., Glynn, R. J. and Cook, E. F. (2008), Evaluating uses of data mining techniques in propensity score estimation: a simulation study. Pharmacoepidem. Drug Safe., 17: 546–555. doi: 10.1002/pds.1555. http://onlinelibrary.wiley.com/doi/10.1002/pds.1555/abstract
21. Westreich, Daniel, Justin Lessler, and Michele Jonsson Funk. "Propensity score estimation: machine learning and classification methods as alternatives to logistic regression." Journal of clinical epidemiology 63.8 (2010): 826. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2907172/
22. Cannas, Massimo, Bruno Arpino, and Francesco Billari. "Machine learning techniques for Propensity score matching with clustered data. A simulation study." *46TH SCIENTIFIC MEETING OF THE ITALIAN STATISTICAL SOCIETY*. 2012. http://www.academia.edu/download/41744099/2234-3870-1-PB.pdf
23. Garbe, E., et al. "High-dimensional versus conventional propensity scores in a comparative effectiveness study of coxibs and reduced upper gastrointestinal complications." *European journal of clinical pharmacology* 69.3 (2013): 549-557. http://link.springer.com/article/10.1007/s00228-012-1334-2
24. Kreif, Noémi, et al. "Evaluating treatment effectiveness under model misspecification: a comparison of targeted maximum likelihood estimation with bias-corrected matching." *Statistical methods in medical research* (2014): 0962280214521341. http://smm.sagepub.com/content/early/2014/04/21/0962280214521341
25. WATKINS, S; et al. An Empirical Comparison of Tree-Based Methods for Propensity Score Estimation. *Health Services Research*. 48, 5, 1798-1817, Oct. 2013. ISSN: 00179124. http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=90469069&site=ehost-live&scope=site
26. Angelosante, D., & Giannakis, G. (2009). RLS-weighted lasso for adaptive estimation of sparse signals", 34th *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009)* *http://tinyurl.com/mlmofdq*
27. Zhao, Qingyuan. "Covariate Balancing Propensity Score by Tailored Loss Functions." *arXiv preprint arXiv:1601.05890* (2016). http://arxiv.org/pdf/1601.05890.pdf
28. Zubizarreta, José R. "Stable weights that balance covariates for estimation with incomplete outcome data." *Journal of the American Statistical Association* 110.511 (2015): 910-922. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1080/01621459.2015.1023805
29. Kang, Joseph D. Y., and Joseph L. Schafer. "Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data". *Statistical Science* 22.4 (2007): 523–539. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/27645858
30. S. Athey, G. Imbens, S. Wager: "Efficient Inference of Average Treatment Effects in High Dimensions via Approximate Residual Balancing," No. 3408. 2016
31. Y. Ning, S. Peng, K. Imai. "High Dimensional Propensity Score Estimation via Covariate Balancing" Working Paper, May 9, 2017
32. J. Hartford, G. Lewis, K. Leyton-Brown, M. Taddy. "Counterfactual Prediction with Deep Instrumental Variables Networks" Working Paper, Dec 30, 2016.
33. V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2017, December) "Double/Debiased Machine Learning for Treatment and Causal Parameters." https://arxiv.org/abs/1608.00060

Introduction and Overviews

1. Hal Varian, "Big Data: New Tricks for Econometrics," The Journal of Economic Perspectives, 28 (2), Spring 2014, 3-27 http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/1522171656?accountid=14026
2. Top 10 algorithms in data mining http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/204300634?accountid=14026
3. Donoho, David. "50 Years of Data Science," (2015) http://www.ccs.neu.edu/course/cs7280sp16/CS7280-Spring16_files/50YearsOfDataScience.pdf

4. Breiman, Leo. “Statistical Modeling: The Two Cultures.” Statist. Sci. 16 (2001), no. 3, 199--231. doi:10.1214/ss/1009213726. http://projecteuclid.org/download/pdf_1/euclid.ss/1009213726
5. Einav, Liran, Levin, Jonathan. “Economics in the age of big data”, Science 07 Nov 2014: Vol 346, Issue 6210 http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1126/science.1243089
6. Mitchell, T. (2006). “The discipline of machine learning.”, Working Paper, CMU-ML-01-108 http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf
7. Athey, Susan. "Beyond prediction: Using big data for policy problems." *Science* 355.6324 (2017): 483-485.

Matrix Completion and Factor Models

1. S. Athey, M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi (2017) ``Matrix Completion Methods for Causal Panel Data Models." https://arxiv.org/abs/1710.10251
2. J. Bai (2009), ``Panel data models with interactive fixed effects." Econometrica, 77(4): 1229--1279.
3. E. Candes and B. Recht (2009) ``Exact matrix completion via convex optimization." Foundations of Computational mathematics, 9(6):717-730.

Modeling Text and Language

1. Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." Journal of machine learning research 3.Jan (2003): 993-1022. http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf
2. Blei, David M. "Probabilistic topic models." Communications of the ACM 55.4 (2012): 77-84. https://cacm.acm.org/magazines/2012/4/147361-probabilistic-topic-models/fulltext

Network Experiments

1. Ugander, Johan, et al. "Graph cluster randomization: Network exposure to multiple universes." *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 329-227, 2013. http://arxiv.org/abs/1305.6979 (Durable: http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1145/2487575.2487695)
2. Athey, Susan, Dean Eckles, and Guido W. Imbens. “Exact P-values for Network Interference.” No. w21313. *National Bureau of Economic Research*, 2015. http://arxiv.org/abs/1506.02084 (http://www.nber.org/papers/w21313)
3. Bakshy, Eytan, Dean Eckles, and Michael S. Bernstein. "Designing and deploying online field experiments." *Proceedings of the 23rd international conference on World wide web*. ACM, 283-292, 2014. http://arxiv.org/abs/1409.3174 (Durable: http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1145/2566486.2567967)
4. Eckles, Dean, Brian Karrer, and Johan Ugander. "Design and analysis of experiments in networks: Reducing bias from interference." *arXiv preprint arXiv:1404.7530* (2014). http://arxiv.org/abs/1404.7530
5. Bakshy, Eytan, and Dean Eckles. "Uncertainty in online experiments with dependent data: An evaluation of bootstrap methods." *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1303-1311, 2013. http://arxiv.org/abs/1304.7406 (Durable: http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1145/2487575.2488218)
6. Bakshy, Eytan, et al. "Social influence in social advertising: evidence from field experiments." *Proceedings of the 13th ACM Conference on Electronic Commerce*. ACM, 146-161, 2012. http://arxiv.org/abs/1206.4327 (Durable: http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1145/2229012.2229027)
7. Taylor, Sean J., Eytan Bakshy, and Sinan Aral. "Selection effects in online sharing: Consequences for peer adoption." *Proceedings of the fourteenth ACM conference on Electronic commerce*. ACM, 821-836, 2013.

http://arxiv.org/abs/1311.2878 (Durable: http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1145/2482540.2482604) *DOI not working

8. Bakshy, Eytan, et al. "The role of social networks in information diffusion." *Proceedings of the 21st international conference on World Wide Web*. ACM, 519-528, 2012. http://arxiv.org/abs/1201.4145 (durable: http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1145/2187836.2187907)

Nonparametric Methods-Kernels and KNN

1. Bierens, Herman J. "Kernel estimators of regression functions." *Advances in econometrics: Fifth world congress*. Vol. 1. 1987. http://www.personal.psu.edu/hxb11/KERNEL.PDF
2. Mack, Y. P., and Murray Rosenblatt. "Multivariate k-nearest neighbor density estimates." *Journal of Multivariate Analysis* 9.1 (1979): 1-15. http://www.sciencedirect.com/science/article/pii/0047259X79900654
3. Biau, Gérard, Frédéric Cérou, and Arnaud Guyader. "On the rate of convergence of the bagged nearest neighbor estimate." *The Journal of Machine Learning Research* 11 (2010): 687-712. https://hal.archives-ouvertes.fr/file/index/docid/363875/filename/RR.pdf
4. Samworth, Richard J.. "OPTIMAL WEIGHTED NEAREST NEIGHBOUR CLASSIFIERS". *The Annals of Statistics* 40.5 (2012): 2733–2763. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/41806553
5. Scornet, Erwan, Gérard Biau, and Jean-Philippe Vert. "Consistency of random forests." *The Annals of Statistics* 43.4 (2015): 1716-1741. http://projecteuclid.org/euclid.aos/1434546220

Optimization Methods for Causal Inference

1. S. Athey, G. Imbens, S. Wager: "Efficient Inference of Average Treatment Effects in High Dimensions via Approximate Residual Balancing," No. 3408. 2016
2. Zubizarreta, José R. "Using mixed integer programming for matching in an observational study of kidney failure after surgery." *Journal of the American Statistical Association* 107.500 (2012): 1360-1371.

Prediction Policy

1. Kleinberg, Jon, et al. "Prediction policy problems." *The American Economic Review* 105.5 (2015): 491-495. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1257/aer.p20151023
2. Glaeser, Edward L., et al. *Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life*. No. w21778. National Bureau of Economic Research, 2015. https://research.hks.harvard.edu/publications/getFile.aspx?Id=1285
3. Glaeser, Edward, Andrew Hillis, Scott Duke Kominers, and Michael Luca. "Crowdsourcing City Government: Using Tournaments to Improve Inspection Accuracy." *American Economic Review: Papers and Proceedings* (forthcoming). http://www.nber.org/papers/w22124.pdf
4. Chalfin, Aaron, et al. "Productivity and Selection of Human Capital with Machine Learning." *American Economic Review: Papers and Proceedings*. Vol. 94. No. 16-069. Harvard Business School Working Paper, 2013. http://www.people.hbs.edu/mluca/PredictiveHiring.pdf
5. Goel, Sharad, Justin M. Rao, and Ravi Shroff. "Personalized Risk Assessments in the Criminal Justice System." https://5harad.com/papers/risky.pdf
6. Munoz, C., Smith, M., Patil, DJ. 2016. "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights". *The White House*

Regression Discontinuity

1. Guido Imbens, Thomas Lemieux, (2007) "Regression Discontinuity Designs: A Guide to Practice," http://www.nber.org/papers/w13039

Regularized Regression- Causal

1. Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. "Inference for high-dimensional sparse econometric models." (2011). http://arxiv.org/abs/1201.0220v1
2. Belloni, Alexandre, Victor Chernozhukov and Christian Hansen. 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives*, 28(2):29-50. https://www.aeaweb.org/articles?id=10.1257/jep.28.2.29
3. Belloni A, Chernozhukov V, Hansen C. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *Review of Economic Studies*. April 2014;81(2):608-650 http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eoh&AN=1475036&site=ehost-live&scope=site

Regularized Regression- Prediction

1. Zou, Hui. "The Adaptive Lasso and Its Oracle Properties". *Journal of the American Statistical Association* 101.476 (2006): 1418–1429 http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/27639762
2. Breiman, L., and D. Freedman. "How Many Variables Should Be Entered in a Regression Equation?" *Journal of the American Statistical Association* 78.381 (1983): 131–136. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/2287119
3. Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. "Inference for high-dimensional sparse econometric models." (2011). http://arxiv.org/abs/1201.0220v1
4. Tibshirani, Robert. "Regression Shrinkage and Selection via the Lasso: A Retrospective". *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 73.3 (2011): 273–282. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/41262671
5. Belloni, Alexandre, and Victor Chernozhukov. "High dimensional sparse econometric models: An introduction." *Springer Berlin Heidelberg*, 2011. http://arxiv.org/abs/1106.5242
6. Bühlmann, Peter. Statistical significance in high-dimensional linear models. Bernoulli 19 (2013), no. 4, 1212--1242. doi:10.3150/12-BEJSP11. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.3150/12-BEJSP11
7. Fan, Jianqing, and Runze Li. "Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties". *Journal of the American Statistical Association* 96.456 (2001): 1348–1360. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/3085904
8. Fan, J., & Lv, J. (2010). A Selective Overview of Variable Selection in High Dimensional Feature Space. Statistica Sinica, 1–44. http://arxiv.org/abs/0910.1122
9. Fan, J., Lv, J., & Qi, L. (2011). Sparse High Dimensional Models in Economics. Annual Review of Economics, 3, 291–317. doi:10.1146/annurev-economics-061109-080451 http://www-bcf.usc.edu/~jinchilv/publications/ARE-FLQ11.pdf
10. Park, T., & Casella, G. (2008). The Bayesian Lasso. Journal of the American Statistical Association, 103(482), 681–686. doi:10.1198/016214508000000337 http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/27640090
11. Zhao, P., & Yu, B. (2006). "On model selection consistency of Lasso." *The Journal of Machine Learning Research*. http://www.jmlr.org/papers/v7/zhao06a.html

Regularized Regression- Statistical Theory

1. Berk, Richard, et al. "Valid post-selection inference." *The Annals of Statistics* 41.2 (2013): 802-837. http://projecteuclid.org/euclid.aos/1369836961

2. Chernozhukov, V., Hansen, C., & Spindler, M. (2015). Valid post-selection and post-regularization inference: An elementary, general approach. *Annual Review of Economics, 7*, 649ff. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1146/annurev-economics-012315-015826
3. Taylor, J., & Tibshirani, R. J. (2015). Statistical learning and selective inference. *Proceedings of the National Academy of Sciences of the United States of America, 112*(25), 7629. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1073/pnas.1507583112

Robustness

1. Athey, Susan and Guido Imbens. 2015. "A Measure of Robustness to Misspecification." American Economic Review, 105(5):476-80 http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1257/aer.p20151020
2. Leamer, Edward E. 1982. "Sets of Posterior Means with Bounded Variance Priors." Econometrica 50 (2): 725-736. http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/214661074?accountid=14026
3. Leamer, Edward E. 1983. "Let's Take the Con out of Econometrics." American Economic Review 73 (1): 31-43. http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eoh&AN=0136421&site=ehost-live&scope=site
4. White, Halbert. 2000. "A Reality Check for Data Snooping." Econometrica 68 (5): 1097-1126. http://gsbproxy.stanford.edu/login?url=http://search.proquest.com/docview/203869138?accountid=14026

Supervised Learning-Neural Nets

1. Y. LeCun, Y. Bengio and G. Hinton, (2015) ``Deep learning" *Nature.* Vol. 521(7553): 436-444. https://www.nature.com/articles/nature14539
2. I. Goodfellow, Y. Bengio, and A. Courville (2016) ``Deep Learning." MIT Press.

Trees and Forests: Forests

1. Breiman, Leo. "Random forests." *Machine learning* 45.1 (2001): 5-32. http://link.springer.com/article/10.1023/A%3A1010933404324
2. Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods, 14*(4), 323-348. http://gsbproxy.stanford.edu/login?url=http://dx.doi.org/10.1037/a0016973

Trees and Forests: Statistical Theory

1. Biau, Gérard. "Analysis of a random forests model." *The Journal of Machine Learning Research* 13.1 (2012): 1063-1095. http://www.jmlr.org/papers/volume13/biau12a/biau12a.pdf
2. Kim, H., Wei-Yin, L., Yu-Shan, S., & Chaudhuri, P. (2007). Visualizable and interpretable regression models with good prediction power. *IIE Transactions, 39*(6), 565. http://gsbproxy.stanford.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=24471426&site=ehost-live&scope=site
3. Wager, Stefan, Trevor Hastie, and Bradley Efron. "Confidence intervals for random forests: The jackknife and the infinitesimal jackknife." *The Journal of Machine Learning Research* 15.1 (2014): 1625-1651. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4286302/

4. Wager, Stefan, and Susan Athey. "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests." *arXiv preprint arXiv:1510.04342* (2015). http://arxiv.org/abs/1510.04342
5. Athey, Susan, and Guido W. Imbens. "Recursive Partitioning for Heterogeneous Causal Effects." *stat* 1050 (2015): 30. http://arxiv.org/abs/1504.01132
6. Wager, Stefan, and Guenther Walther. "Uniform Convergence of Random Forests via Adaptive Concentration." *arXiv preprint arXiv:1503.06388* (2015). http://arxiv.org/pdf/1503.06388.pdf
7. Denil, Misha, David Matheson, and Nando De Freitas. "Narrowing the gap: Random forests in theory and in practice." *arXiv preprint arXiv:1310.1415*(2013). http://arxiv.org/pdf/1310.1415.pdf

Trees and Forests: Trees

1. Breiman, Leo, et al. "Classification and Regression Trees". *Chapman and Hall/CRC*; 1 edition (1984)
2. Segal, Mark Robert. "Regression Trees for Censored Data". *Biometrics* 44.1 (1988): 35–47. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/2531894

Unsupervised Learning: Community Detection

1. Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America, 103*(23), 8577. http://gsbproxy.stanford.edu/login?url=http://www.jstor.org/stable/30050231