CAUSAL Lab@SUFE Onboarding Guide (V1.1)

Zhiheng Zhang

July 10, 2025

Welcome to our lab!

This is the on-boarding guide for interdisciplinary research on causality. **It is mandatory reading for all interns, master's students, and PhD students** in CAUSAL lab@SUFE. We first provide instructions on preparing for causality research, and then directly present frontier topics in causality research to promote further comprehension and collaboration.

Disclaimer: This document reflects the primary research directions of our lab. It does **not** represent the only or the optimal pathway to learning causal inference.

1 Getting Started with Causal Inference: A Guide for Students in CS, Statistics, and Economics

We follow a fundamental causal learning progression: from introductory to proficient, and ultimately to advanced mastery. I will outline tailored learning paths for two groups: CS(ML) and Stats(Econ).

Requirement: Each Student with either a CS(ML) or Stats(Econ) background are required to select and thoroughly read at least one resource from Section 1.1 and Section 1.2, respectively.

1.1 Introductory part

- [PM18] The Book of Why: The New Science of Cause and Effect (CS & Stats).
- [HR22] What If: The Counterfactuals of Modern Epidemiology (CS & Stats).

1.2 Formal part

- [Din24] The First Course in Causal Inference (Stats).
- [WB24] Causal Inference: A Statistical Learning Approach (Stats).
- https://web.stanford.edu/~swager/stats361.pdf (CS & Stats).
- Online courses of ML and Causality:
  - Susan Athey: https://www.gsb.stanford.edu/faculty-research/labs-initiatives/sil/research/methods/ai-machine-learning/short-course?utm_source=chatgpt.com(CS & Stats).
  - Brady neal: https://www.bilibili.com/video/BV1nZ4y1K78i/?vd_source=ace5e5a325c2dff9ac92e50fa21 & Stats).
  - Chernozhukov et al: http://chapters.causalml-book.org/CausalML_book_2022.pdf.

1.3 Advanced part

- Online causality inference seminar(YouTube): https://www.youtube.com/channel/UCiiOj5GSES6uw21kfXnxj3A (CS & Stats).

- Literature: Students could read from the following literature to follow up on frontiers:
Journal: JASA, AOS, JRSSB, Biometrika, PNAS;
Conference: ICML, NeurIPS, ICLR, UAI, AISTATS, COLT

Additional comment:

- If you come from a mathematical major, you should also take the class on (high-dimensional) statistics and probability; otherwise, you are required to run the code demo in one of the above course links.

- You should develop an intuitive understanding of the statistical ideas behind causality and be able to critically compare the strengths and limitations of Judea Pearl's causal graphical framework and Donald Rubin's potential outcomes framework (refer to [RR13]).

2 Frontier Research Topics under Active Development

In the study of causality, on the one hand, aiming at the real-world setting, there are two kinds of methods: **observational studies** and **experimental design**. On the other hand, aiming at the objectives, there are three central pillars: Identification, Estimation, and Learning. Each corresponds to a classical type of causal inference task[1].

However, we do not wish to define our causal research merely through the conventional paradigm of scenario-based or goal-specific classifications. The long-term vision of our lab is to build a unified and simple causal framework, combined with disciplinary researchers, that widely empowers both theory and real-world applications:

- On the theoretical side: to write killer papers under the simplest and most elegant assumptions;

- On the applied side: to write healer papers upon real-world pain points by tackling the most pressing and complex challenges in industry/society.

Inspired by it, we separate the taste of research into the progressive parts: **rethink, innovate, embrace, and apply**, focusing on the procedure **input, structure, and output**. I have listed the key papers that are recommended to be read.

2.1 Rethink Assumptions

Challenge conventional wisdom and commonly held assumptions (investigate the definitions and implications of these assumptions on your own). See the video [2].

- **Violate the OVERLAP assumption** [KSU24], [CKM+25], [JRYW22].
- **Violate the SUTVA assumption** [Leu22], [LWZ24], [MT21], [Viv25] (also other series of papers of Davide Viviano).
- **Violate the UNCONFOUNDEDNESS assumption** [CKM+25], [Zha], [KMU21], [MGTT18].

2.2 Innovate with Statistical/ML Methods

Modern causal frameworks are driven by new statistical and ML innovations.

- **Conformal prediction** [LC21].
- **Prediction-powered inference** [WSF+24].
- **(Prediction and then) Optimization** [BM25], [WHBG24], [EG22].

[1] Identification This addresses the fundamental question: "Can the causal quantity of interest be expressed as a function of the observed data distribution under certain assumptions?" It focuses on deriving conditions under which causal effects are identifiable, often using tools like causal graphs, do-calculus, and potential outcomes. Estimation Once a causal estimand is identified, estimation concerns "How can we construct statistical estimators to recover it from data?" This step bridges theory and practice, involving methods such as inverse probability weighting, g-computation, doubly robust estimators, and semiparametric theory. Learning Learning moves beyond point estimation to ask "How can we make optimal or robust decisions based on causal knowledge?" This includes topics like policy learning, individualized treatment rules, off-policy evaluation, and online/active experimentation, often leveraging machine learning and reinforcement learning frameworks.

[2] (https://www.youtube.com/watch?v=LJIoFKBxCKE)

• **Optimal transport** [JLS23], [LGBG25].
• **Representation learning (including negative part)** [MFF23], [MFSF25].
• **Mini-max optimality** [JS24].
• **RL/Online learning/Sequential inference (trade-off)** [SLW23], [ZW24], [NFBR25] (also other series of papers of Elias Bareinboim).
• **Double machine learning** [CCD+17], [DBKR24].

2.3 Embrace Constraints

Focus on real-world constraints and make reasonable solutions.
• **Noisy scenario** [GYWJ22], [ZDC+23].
• **Active sampling (Limited samples)** [KOKI24], [ZWLL], [WCG+25], [HSSZ24].
• **Privacy** [OA25], [SHF25].
• **Functional data** [THZY25].
• **Time delay** [ZBMC+25].
• **Heterogeneity and fusion** [DY25], [LD24], [XKP+23], [LRVP25].

2.4 Apply to Impact

Apply to the real world.
• **Foundation model(LLM)** [JYG+25], [GJR24], [PLOZ24].
• **Multi agent and game** [LWL+24], [LWW+], [SNH24], [MWX21], [ATM21].
• **Recommendation system** [ZHHJ24], [BPT18], [YLN+23], [LRW+25].
• **Ride-sharing** [DWZ+24].

References

[ATM21] Eric Auerbach and Max Tabord-Meehan. The local approach to causal inference under network interference. *arXiv preprint arXiv:2105.03810*, 2021.

[BM25] Eli Ben-Michael. Partial identification via conditional linear programs: estimation and policy learning. *arXiv preprint arXiv:2506.12215*, 2025.

[BPT18] Omer Ben-Porat and Moshe Tennenholtz. A game-theoretic approach to recommendation systems with strategic content providers. *Advances in Neural Information Processing Systems*, 31, 2018.

[CCD+17] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and causal parameters. Technical report, 2017.

[CKM+25] Yang Cai, Alkis Kalavasis, Katerina Mamali, Anay Mehrotra, and Manolis Zampetakis. What makes treatment effects identifiable? characterizations and estimators beyond unconfoundedness. *arXiv preprint arXiv:2506.04194*, 2025.

[DBKR24] Abhinandan Dalal, Patrick Blöbaum, Shiva Kasiviswanathan, and Aaditya Ramdas. Anytime-valid inference for double/debiased machine learning of causal parameters. *arXiv preprint arXiv:2408.09598*, 2024.

[Din24] Peng Ding. *The First Course in Causal Inference*. CRC Press, Boca Raton, 2024.

[DWZ+24] Runpeng Dai, Jianing Wang, Fan Zhou, Shikai Luo, Zhiwei Qin, Chengchun Shi, and Hongtu Zhu. Causal deepsets for off-policy evaluation under spatial or spatio-temporal interferences. *arXiv preprint arXiv:2407.17910*, 2024.

[DY25] Edgar Dobriban and Mengxin Yu. Symmpi: predictive inference for data with group symmetries. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, page qkaf022, 2025.

[EG22] Adam N Elmachtoub and Paul Grigas. Smart “predict, then optimize”. *Management Science*, 68(1):9–26, 2022.

[GJR24] Yu Gui, Ying Jin, and Zhimei Ren. Conformal alignment: Knowing when to trust foundation models with guarantees. *arXiv preprint arXiv:2405.10301*, 2024.

[GYWJ22] Wenshuo Guo, Mingzhang Yin, Yixin Wang, and Michael Jordan. Partial identification with noisy covariates: A robust optimization approach. In *Conference on causal learning and reasoning*, pages 318–335. PMLR, 2022.

[HR22] Miguel A. Hernán and James M. Robins. *What If: The Counterfactuals of Modern Epidemiology*. CRC Press, Boca Raton, 2022.

[HSSZ24] Christopher Harshaw, Fredrik Sävje, Daniel A Spielman, and Peng Zhang. Balancing covariates in randomized experiments with the gram–schmidt walk design. *Journal of the American Statistical Association*, 119(548):2934–2946, 2024.

[JLS23] Wenlong Ji, Lihua Lei, and Asher Spector. Model-agnostic covariate-assisted inference on partially identified causal effects. *arXiv preprint arXiv:2310.08115*, 2023.

[JRYW22] Ying Jin, Zhimei Ren, Zhuoran Yang, and Zhaoran Wang. Policy learning” without” overlap: Pessimism and generalized empirical bernstein’s inequality. *arXiv preprint arXiv:2212.09900*, 2022.

[JS24] Jikai Jin and Vasilis Syrgkanis. Structure-agnostic optimality of doubly robust learning for treatment effect estimation. *arXiv preprint arXiv:2402.14264*, 2024.

[JYG+25] Wenlong Ji, Weizhe Yuan, Emily Getzen, Kyunghyun Cho, Michael I Jordan, Song Mei, Jason E Weston, Weijie J Su, Jing Xu, and Linjun Zhang. An overview of large language models for statisticians. *arXiv preprint arXiv:2502.17814*, 2025.

[KMU21] Nathan Kallus, Xiaojie Mao, and Masatoshi Uehara. Causal inference under unmeasured confounding with negative controls: A minimax learning approach. *arXiv preprint arXiv:2103.14029*, 2021.

[KOKI24] Masahiro Kato, Akihiro Oga, Wataru Komatsubara, and Ryo Inokuchi. Active adaptive experimental design for treatment effect estimation with covariate choices. *arXiv preprint arXiv:2403.03589*, 2024.

[KSU24] Samir Khan, Martin Saveski, and Johan Ugander. Off-policy evaluation beyond overlap: Sharp partial identification under smoothness. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp, editors, *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 23734–23757. PMLR, 21–27 Jul 2024.

[LC21] Lihua Lei and Emmanuel J Candès. Conformal inference of counterfactuals and individual treatment effects. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 83(5):911–938, 2021.

[LD24] Sijia Li and Rui Duan. Efficient collaborative learning of the average treatment effect under data sharing constraints. *arXiv preprint arXiv:2410.02941*, 2024.

[Leu22] Michael P Leung. Causal inference under approximate neighborhood interference. *Econometrica*, 90(1):267–293, 2022.

[LGBG25] Sirui Lin, Zijun Gao, Jose Blanchet, and Peter Glynn. Estimation of optimal causal bounds via covariate-assisted optimal transport. *arXiv preprint arXiv:2506.00257*, 2025.

[LRVP25] Quinn Lanners, Cynthia Rudin, Alexander Volfovsky, and Harsh Parikh. Data fusion for partial identification of causal effects. *arXiv preprint arXiv:2505.24296*, 2025.

[LRW+25] Xiang Li, Feng Ruan, Huiyuan Wang, Qi Long, and Weijie J Su. A statistical framework of watermarks for large language models: Pivot, detection efficiency and optimal rules. *The Annals of Statistics*, 53(1):322–351, 2025.

[LWL+24] Anjie Liu, Jianhong Wang, Haoxuan Li, Xu Chen, Jun Wang, Samuel Kaski, and Mengyue Yang. Attaining humans desirable outcomes in human-ai interaction via structural causal games. *arXiv preprint arXiv:2405.16588*, 2024.

[LWW+] Anjie Liu, Jianhong Wang, Zhuohan Wang, Xinrui Yang, Haoxuan Li, Xu Chen, Samuel Kaski, Jun Wang, and Mengyue Yang. Attaining human's desirable outcomes in indirect human-ai interaction via multi-agent influence diagrams.

[LWZ24] Xin Lu, Yuhao Wang, and Zhiheng Zhang. Adjusting auxiliary variables under approximate neighborhood interference. *arXiv preprint arXiv:2411.19789*, 2024.

[MFF23] Valentyn Melnychuk, Dennis Frauen, and Stefan Feuerriegel. Bounds on representation-induced confounding bias for treatment effect estimation. *arXiv preprint arXiv:2311.11321*, 2023.

[MFSF25] Valentyn Melnychuk, Dennis Frauen, Jonas Schweisthal, and Stefan Feuerriegel. Orthogonal representation learning for estimating causal quantities. *arXiv preprint arXiv:2502.04274*, 2025.

[MGTT18] Wang Miao, Zhi Geng, and Eric J Tchetgen Tchetgen. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*, 105(4):987–993, 2018.

[MT21] Yunpu Ma and Volker Tresp. Causal inference under networked interference and intervention policy enhancement. In *International Conference on Artificial Intelligence and Statistics*, pages 3700–3708. PMLR, 2021.

[MWX21] Evan Munro, Stefan Wager, and Kuang Xu. Treatment effects in market equilibrium. *arXiv preprint arXiv:2109.11647*, 5, 2021.

[NFBR25] Georgy Noarov, Riccardo Fogliato, Martin Bertran, and Aaron Roth. Stronger neyman regret guarantees for adaptive experimental design. *arXiv preprint arXiv:2502.17427*, 2025.

[OA25] Yuki Ohnishi and Jordan Awan. Locally private causal inference for randomized experiments. *Journal of Machine Learning Research*, 26(14):1–40, 2025.

[PLOZ24] Chanwoo Park, Xiangyu Liu, Asuman Ozdaglar, and Kaiqing Zhang. Do llm agents have regret? a case study in online learning and games. *arXiv preprint arXiv:2403.16843*, 2024.

[PM18] Judea Pearl and Dana Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York, 2018.

[RR13] Thomas S. Richardson and James M. Robins. Single world intervention graphs: A primer. *arXiv preprint arXiv:1301.4700*, 2013.

[SHF25] Maresa Schröder, Justin Hartenstein, and Stefan Feuerriegel. Private: Differentially private confidence intervals for average treatment effects. *arXiv preprint arXiv:2505.21641*, 2025.

[SLW23] David Simchi-Levi and Chonghuan Wang. Multi-armed bandit experimental design: Online decision-making and adaptive inference. In *International Conference on Artificial Intelligence and Statistics*, pages 3086–3097. PMLR, 2023.

[SNH24] Ana-Andreea Stoica, Vivian Y Nastl, and Moritz Hardt. Causal inference from competing treatments. *arXiv preprint arXiv:2406.03422*, 2024.

[THZY25] Ruoxu Tan, Wei Huang, Zheng Zhang, and Guosheng Yin. Causal effect of functional treatment. *Journal of Machine Learning Research*, 26(91):1–39, 2025.

[Viv25] Davide Viviano. Policy targeting under network interference. *Review of Economic Studies*, 92(2):1257–1292, 2025.

[WB24] Stefan Wager and Guillaume Basse. *Causal Inference: A Statistical Learning Approach*. Cambridge University Press, 2024. Forthcoming.

[WCG+25] Hechuan Wen, Tong Chen, Mingming Gong, Li Kheng Chai, Shazia Sadiq, and Hongzhi Yin. Enhancing treatment effect estimation via active learning: A counterfactual covering perspective. *arXiv preprint arXiv:2505.05242*, 2025.

[WHBG24] Zhenyu Wang, Yifan Hu, Peter Bühlmann, and Zijian Guo. Causal invariance learning via efficient optimization of a nonconvex objective. *arXiv preprint arXiv:2412.11850*, 2024.

[WSF+24] Yuxin Wang, Maresa Schröder, Dennis Frauen, Jonas Schweisthal, Konstantin Hess, and Stefan Feuerriegel. Constructing confidence intervals for average treatment effects from multiple datasets. *arXiv preprint arXiv:2412.11511*, 2024.

[XKP+23] Ruoxuan Xiong, Allison Koenecke, Michael Powell, Zhu Shen, Joshua T Vogelstein, and Susan Athey. Federated causal inference in heterogeneous observational data. *Statistics in Medicine*, 42(24):4418–4439, 2023.

[YLN+23] Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. How bad is top-k recommendation under competing content creators? In *International Conference on Machine Learning*, pages 39674–39701. PMLR, 2023.

[ZBMC+25] Kelly W Zhang, Thomas Baldwin-McDonald, Kamil Ciosek, Lucas Maystre, and Daniel Russo. Impatient bandits: Optimizing for the long-term without delay. *arXiv preprint arXiv:2501.07761*, 2025.

[ZDC+23] Zhiheng Zhang, Quanyu Dai, Xu Chen, Zhenhua Dong, and Ruiming Tang. Robust causal inference for recommender system to overcome noisy confounders. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2349–2353, 2023.

[Zha] Zhiheng Zhang. Tight partial identification of causal effects with marginal distribution of unmeasured confounders. In *Forty-first International Conference on Machine Learning*.

[ZHHJ24] Ruohan Zhan, Shichao Han, Yuchen Hu, and Zhenling Jiang. Estimating treatment effects under recommender interference: A structured neural networks approach. *arXiv preprint arXiv:2406.14380*, 2024.

[ZW24] Zhiheng Zhang and Zichen Wang. Online experimental design with estimation-regret trade-off under network interference. *arXiv preprint arXiv:2412.03727*, 2024.

[ZWLL] Zhiheng Zhang, Haoxiang Wang, Haoxuan Li, and Zhouchen Lin. Active treatment effect estimation via limited samples.