Applied Data Analytics

Data analysis — Interpretation challenges

Observing, intervening, and counterfactuals

Hans-Martin von Gaudecker and Aapo Stenhammar

So what can we learn from data alone?

"Observing"

- Conditional mean

- How does Y change on average when X changes in the population we study?

- Does **not** imply anything about effects of **intervening** in X

What can we do?

1. Study interventions in X with a known allocation mechanism for X

    → (Randomised) controlled trials

2. Exploit variation that comes close to this setting, then do econometrics

    → Encouragements that people choose different X

    → Sudden policy changes

    → Cutoffs in eligibility criteria

    → ...

Counterfactuals

- Observing allows to predict $\bar{Y}$ in the same setting

- Studying interventions allows to predict effects of the same intervention

- Counterfactuals allow to predict what will happen for *some* intervention

Literature

- Judea Pearl and Dana Mackenzie (2018). The Book of Why

- Judea Pearl (2009). Causality. Cambridge University Press.