Swiss Institute for
Empirical Economic Research

University of St.Gallen

75 NRP

Big Data
National Research Programme

# PI: Michael Lechner
Project title: Causal analysis with Big Data

Presentation of the project

Microeconometrics
+ Causality
- Not Big Data compatible

Supervised
Machine Learning
+ Prediction / Big Data
- No causality

Causal
Machine
Learning
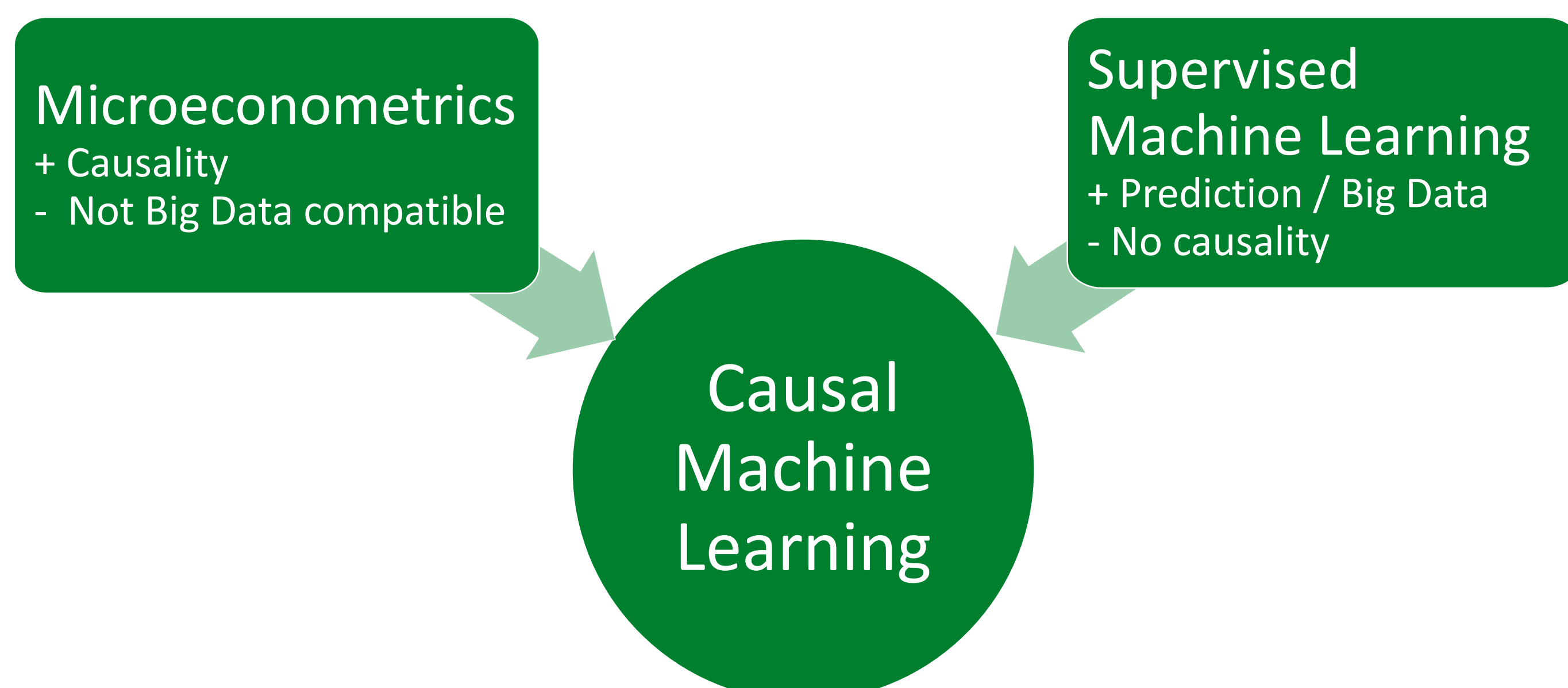
Subprojects

- *Extension* and *evaluation* of methods adapting *machine learning* methods to *causal analysis* to
  1. *… remove impact of confouding variables* in observational studies
  2. … systematically detect *effect heterogeneity*

- Applications of these methods to
  1. evaluation of active *labour market* programmes
  2. hedonic *price formation* of used cars
  3. estimation of the effects of *discrimination in sports*

Aim of the project

$$D \longrightarrow Y$$
$$X$$
$$+$$
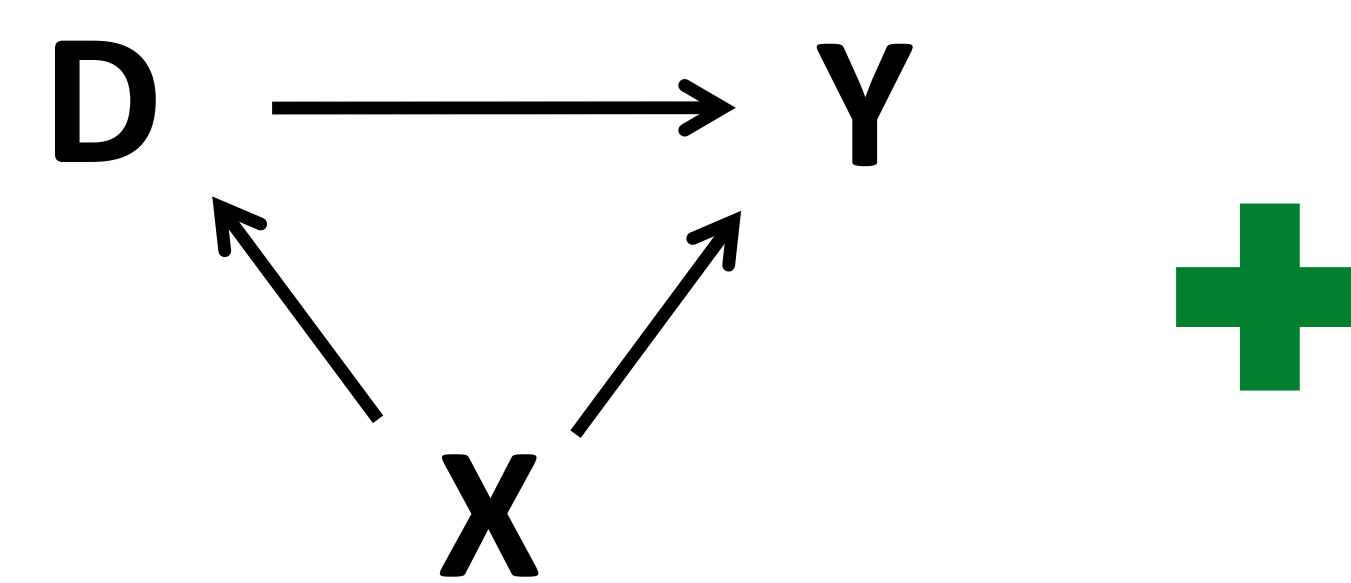
Aims

- *Combine strengths* of two mature literatures to *improve causal analysis* with high-dimensional, large data and with standard settings in general
- Derive *recommendations* about how to set up causal studies for a broad audience of *applied researchers*
- Provide *freely available code* to the research community implementing methods that worked best
- *Inform policy makers* and other relevant *stakeholders* according to the results derived from the applications

Methodology

Microeconometric identification
- Matching
- Instrumental variables
- Difference-in-difference

Machine Learning
- Double Machine Learning
- LASSO-type methods
- Trees, Random Forest & similar
- …

Effect heterogeneity
- Modified covariate method
- Modified outcome method
- Causal trees and forests

Method evaluation
- Theoretical results
- Empirical Monte Carlo study

Relevance of developed methods

- *Decision problems* are intrisically *causal*, answering the «what if» question

- Decision making in …
  - Policy
  - Industry
  - Medicine
  - Daily life
  - …
  … can be improved if they are based on causal effects instead of pure predictions of outcomes
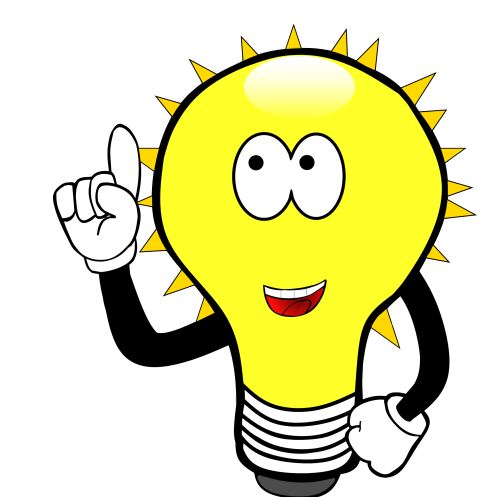
➔ **Project results are useful and relevant far beyond the applications of this project**

We are happy to talk about collaborations