Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

So many choices in Double Machine Learning!? Practical insights from a simulations study

2022 Causal Data Science Meeting

November 8, 2022

Oliver Schacht (University of Hamburg)

Collaborators: Philipp Bach (UHH), Malte Kurz (TUM), Martin Spindler (UHH)

What is Double Machine Learning?

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- **Double/debiased machine learning (DML)** introduced by Chernozhukov et al. (2018)
- General framework based on machine learning tools for causal inference and estimation of treatment effects
- Combines the strength of **machine learning** and **econometrics**
- Resulting estimator has good properties ($\sqrt{N}$-consistency, approx. Gaussian)
- Our object-oriented implementation **DoubleML** provides a general interface for models and methods for DML (in `R` and in `Python`)

The Key Ingredients of DML

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

1. Neyman Orthogonality

   Inference is based on a method-of-moments estimator that obeys the **Neyman orthogonality condition**

2. High-Quality Machine Learning Estimators

   The nuisance parameters are estimated with high-quality (fast-enough converging) machine learning methods

3. Sample Splitting

   To avoid the biases arising from overfitting, a form of **sample splitting** is used at the stage of producing the estimator of the main parameter θ_0

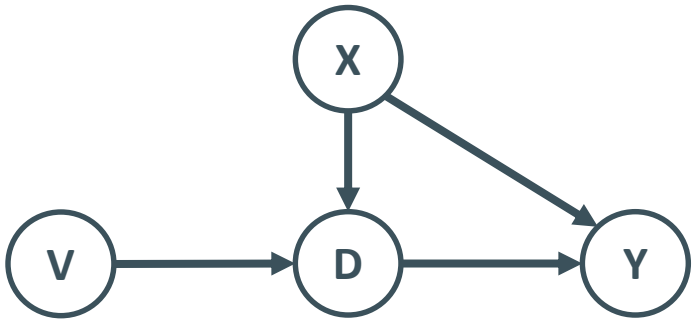
Example: Partially Linear Regression Model

Partially linear regression (**PLR**) model

$$
\begin{aligned}
Y &= D\theta_0 + g_0(X) + \zeta, & \mathbb{E}[\zeta|D, X] &= 0, \\
D &= m_0(X) + V, & \mathbb{E}[V|X] &= 0,
\end{aligned}
$$

with

- outcome variable Y
- Policy or Treatment variable of interest D
- High-dimensional vector of confounding covariates $X = (X_1, \ldots, X_p)$
- Stochastic errors ζ and V



Tuning in Double Machine Learning

- PLR example: To estimate θ_0, the following **Neyman orthogonal score** is used

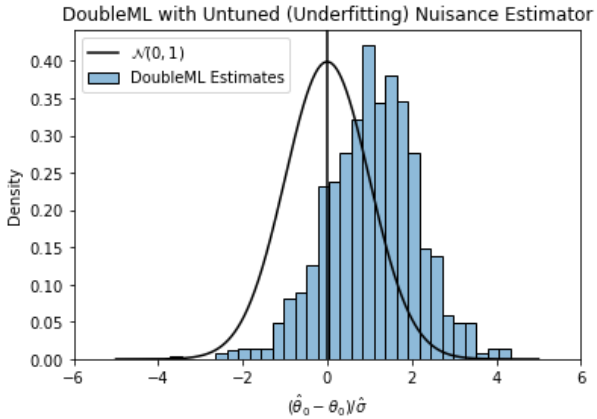
$$\psi(W;\theta,\eta) = \Big(Y - g(X) - \theta\big(D - m(X)\big)\Big)\big(D - m(X)\big)$$

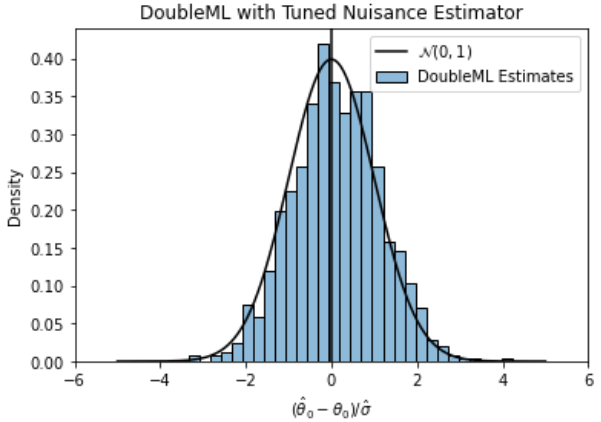
- The nuisance $\eta = (g(X) = \mathbb{E}[\,Y \mid X\,], m(X) = \mathbb{E}[D \mid X])$ is estimated by ML learners
- Double Machine Learning is inherently **robust against small biases** from regularization or overfitting
- **Tuning** DML nuisance predictors is an open question

Tuning in Double Machine Learning

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Using **untuned** ML estimators for nuisance prediction however can lead to a severely biased estimation in the causal parameter of interest





- How do we get to the right estimator?

Our Project

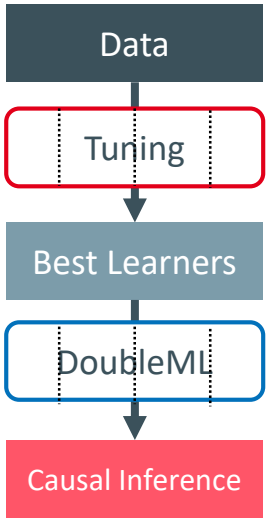
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Carry out large scale simulation study to answer important questions double machine learning users face
  - Which Machine Learning Methods to use?
  - Role of sample splitting in tuning?
  - How to assess the quality of fit?

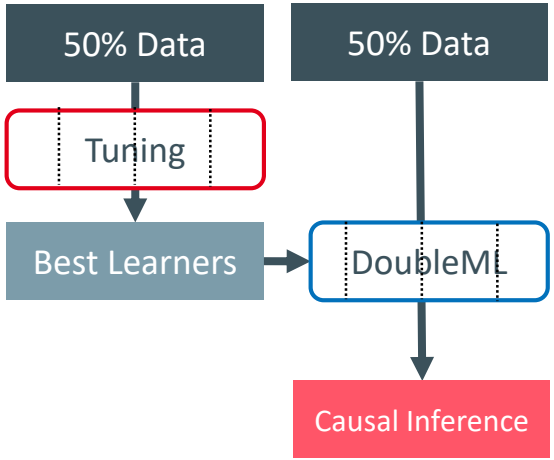
Our Simulation Study

- We use plug-in Automated Machine Learning Estimators from the library `flaml`.
- Three different tuning approaches are tested high-dimensional datasets from the ACIC 2019

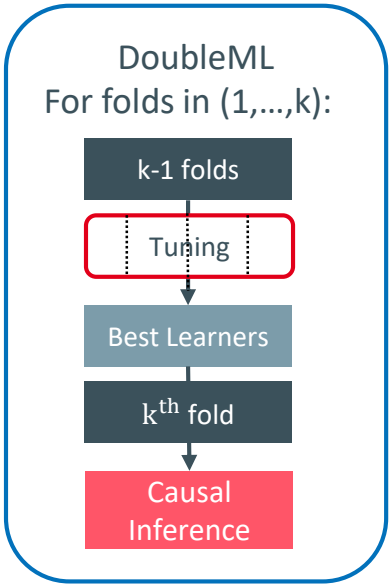
Tune on full sample



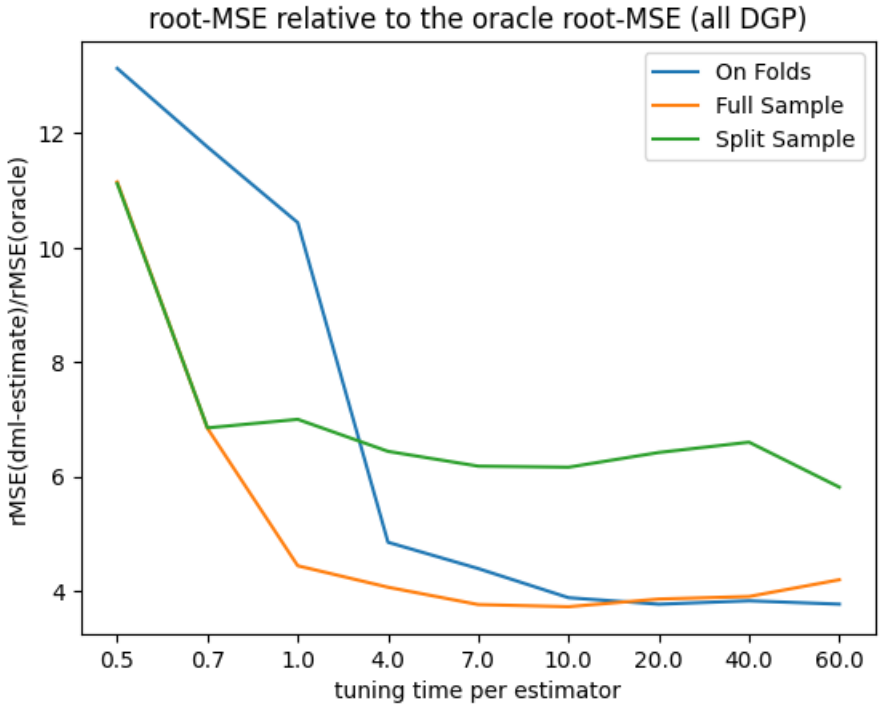
Tune on hold-out sample



Tune on the folds

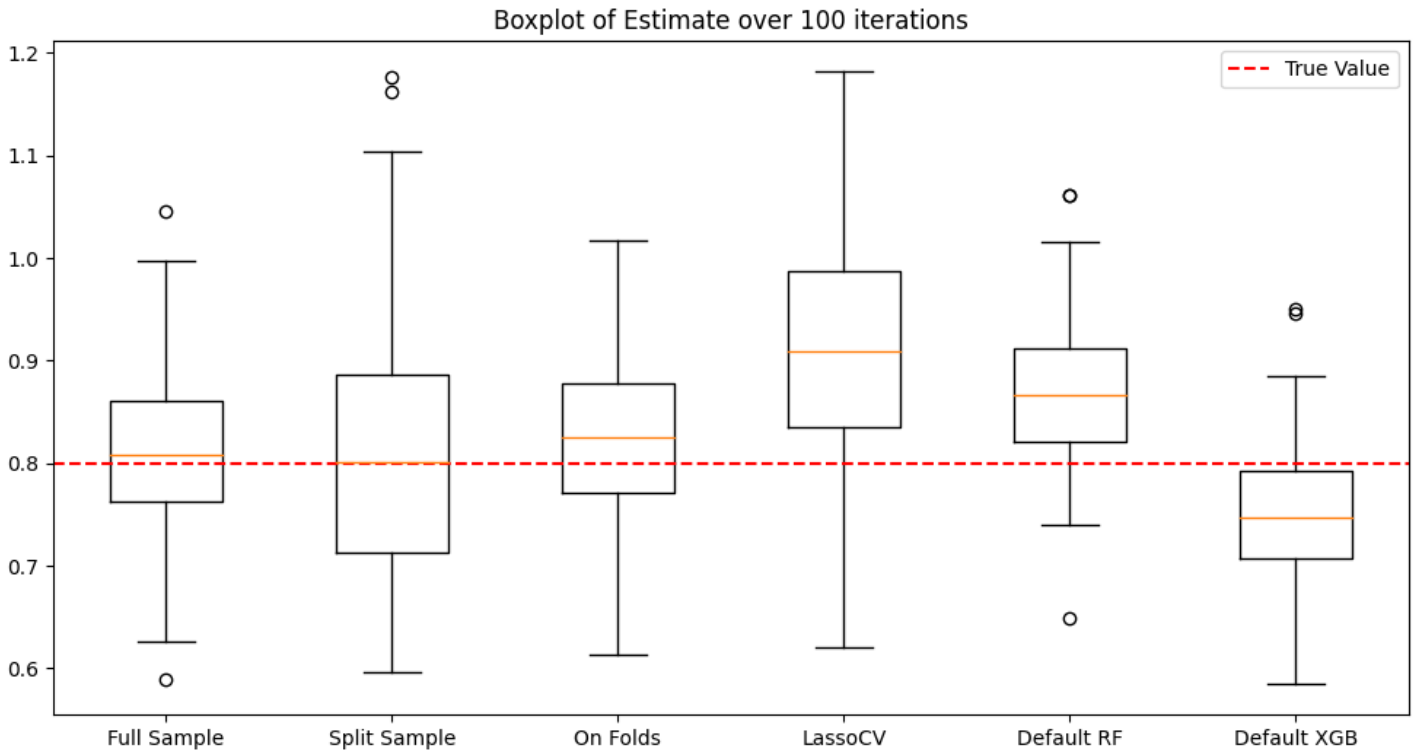


Results



$n = 1000, p = 200$

Results

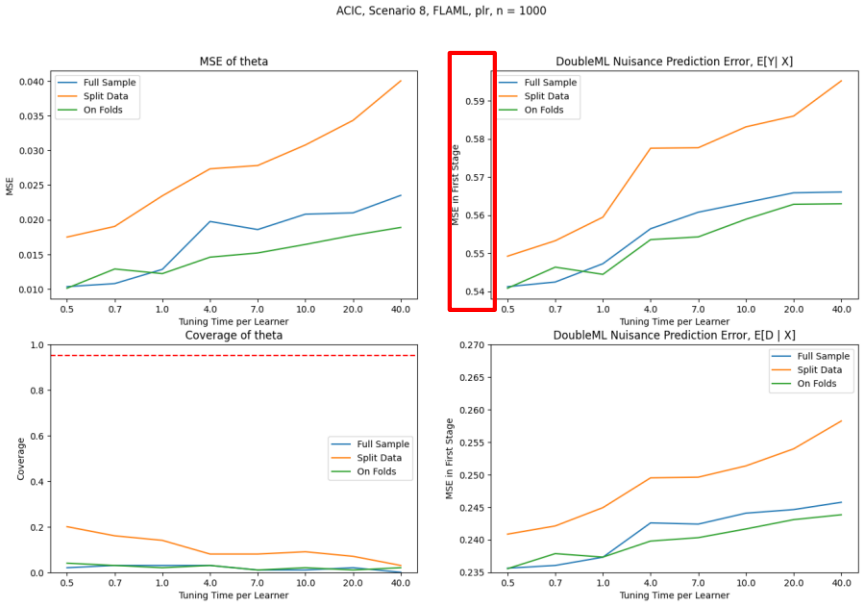

Boxplot of Estimate over 100 iterations
True Value
1.2
1.1
1.0
0.9
0.8
0.7
0.6
Full Sample
Split Sample
On Folds
LassoCV
Default RF
Default XGB

Results

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

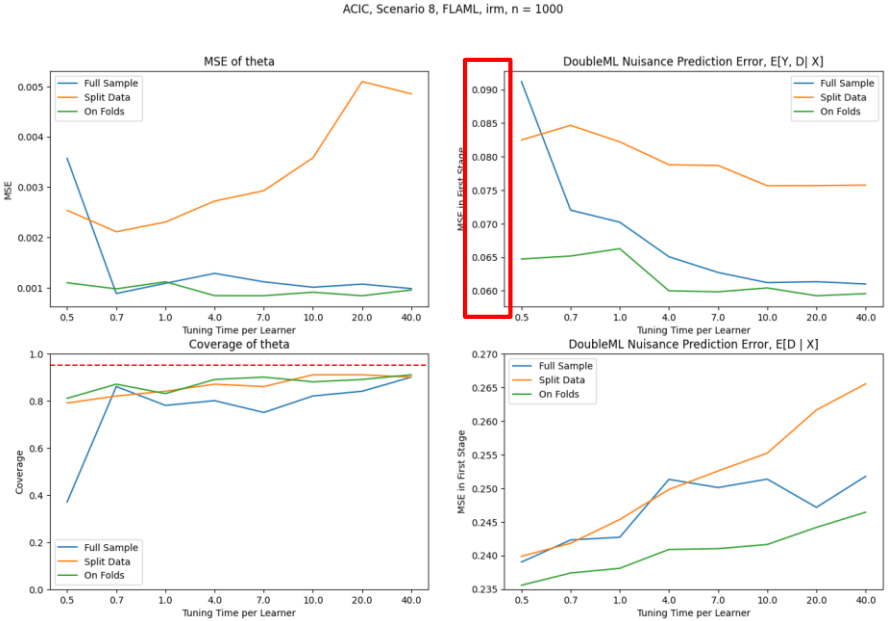
Misspecified Model

(linear, additive effect)



Correct Model

(heterogenous effect)



Key Take-Aways and Outlook

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Tuning the nuisance estimators has influence on inference accuracy in double machine learning
- Plug-in AutoML estimators work well here
- Tuning on hold-out data is in investigated cases not efficient
- We recommend monitoring the nuisance prediction error for assessment of causal inference quality
- Full results for all DGPs to be published / further studies on influence of cross-fitting
- Extension to further AutoML frameworks
- **Extension module** for DoubleML

Thank you for your attention!

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Comments, ideas? **Feel free to reach out!**

✉ oliver.schacht@uni-hamburg.de

More about **DoubleML**:

https://docs.doubleml.org

https://github.com/DoubleML/doubleml-for-py

https://github.com/DoubleML/doubleml-for-r