A Double Machine Learning Approach to Estimate the Effects of Musical Practice on Student's Skills

Michael C. Knaus
University of St. Gallen

Research questions

1. What is the *dose-response effect* of making *music* on cognitive and non-cognitive skills of adolescents?
2. How to *integrate* recently proposed *Double Machine Learning* (DML) into standard causal analysis in *observational studies*?
   - How to investigate *sensitivity* of estimates to *tuning parameter* choices in the machine learning part?
   - How to assess *covariate balancing* in high-dimensional settings?

Motivation

Topic:

- Recent interest in understanding the impact of *extracurricular activities* on *skills*
- *Positive effects* of *musical practice per se* found

Methodological:

- *DML* (Chernozhukov et al., 2018, *Economet J*) *interesting option* for causal inference in observational studies
- However, only illustrative applications in method contributions and *little guidance* for practitioners

Contribution

Topic:

- Investigation of *dose-response* relation between musical practice and skill development
- Observed parental tastes increase *credibility* of *identification*

Methodological:

- Proposal how to address two *practically relevant issues*:
  - Systematic *sensitivity analysis* to the *tuning parameter choice* in the machine learning part
  - Provide *weighted representation* of DML to check *covariate balancing*
- Implemented in R package `dmlmt`

Data

German National Educational Panel Study (NEPS)

- 6,000 students in the 9th grade
- Four intensities of *musical practice* $\{no, low, medium, high\}$
- Objective and subjective cognitive skills, Big Five
- 377 student and parental background characteristics as control variables

Estimation

- Quantity of interest: average potential outcome, $\mu_t = E[Y^t]$, and average treatment effects, $\mu_t - \mu_s$
- Estimated under *conditional independence assumption* using DML method of Farrell (2015, *J Econometrics*)
- ~10,000 potential controls
- *Cross-validated Post-Lasso* used for prediction
- *Post-Lasso* allows balancing checks using w_t from the weighted representation of the DML estimator

$$\hat{\mu}_t = \sum_{i=1}^{N}\left[\mu_t(X_i) + \frac{d_i^t(Y_i - \mu_t(X_i))}{p_t(X_i)}\right] = \sum_{i=1}^{N}\left[Y_i(w_t^Y + w_t^p - w_t^{pY})\right] = \sum_{i=1}^{N}[Y_t w_t]$$

- with treatment dummy d_i^t, conditional outcome $\mu_t(x)$, conditional treatment probability $p_t(x)$
- w_t^Y are weights of outcome prediction, w_t^p are IPW weights, w_t^pY are adjustment weights
- *Cross-validation* allows data-driven sensitivity analysis based on 1SE and 1SE+ rules

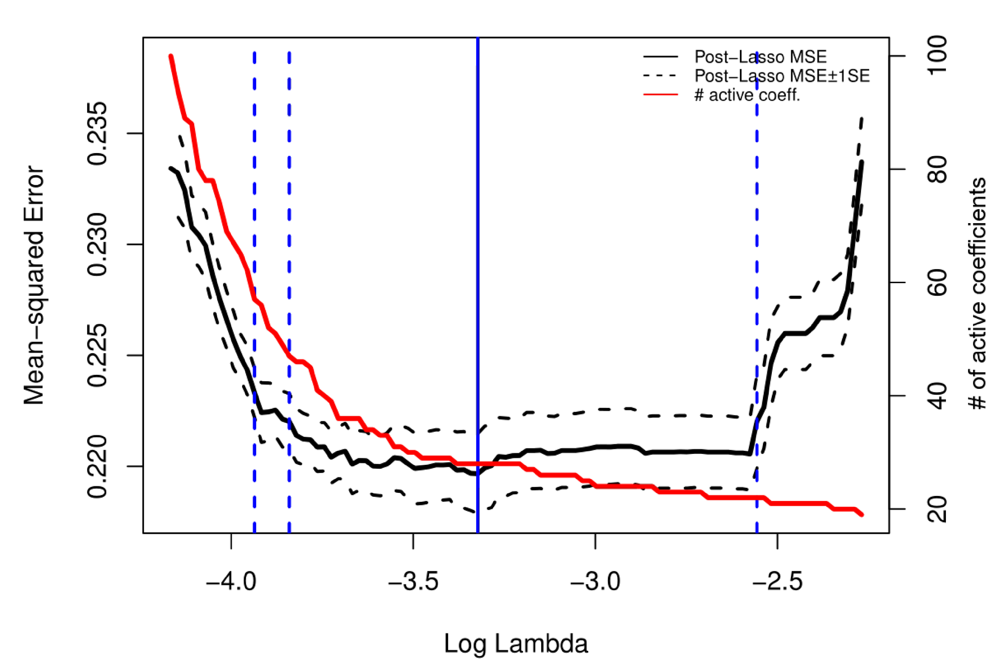

Figure 1: Representative example of cross-validation

Baseline results

Results for *binary music indicator* are in line with previous studies

Cognitive Skills (standardized)				
Science	Math	Vocabulary	Reading	ICT
0.11***	0.08***	0.11***	-0.03	0.12***
(0.02)	(0.02)	(0.02)	(0.03)	(0.02)

Grades (standardized)		
German	Math	Average
0.12***	0.05*	0.09***
(0.03)	(0.03)	(0.03)

Big Five (standardized)				
Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness
0.03	0.11***	-0.04	0.001	0.31***
(0.03)	(0.03)	(0.03)	(0.03)	(0.02)

The data allow to investigate further *different intensities* of musical practice:

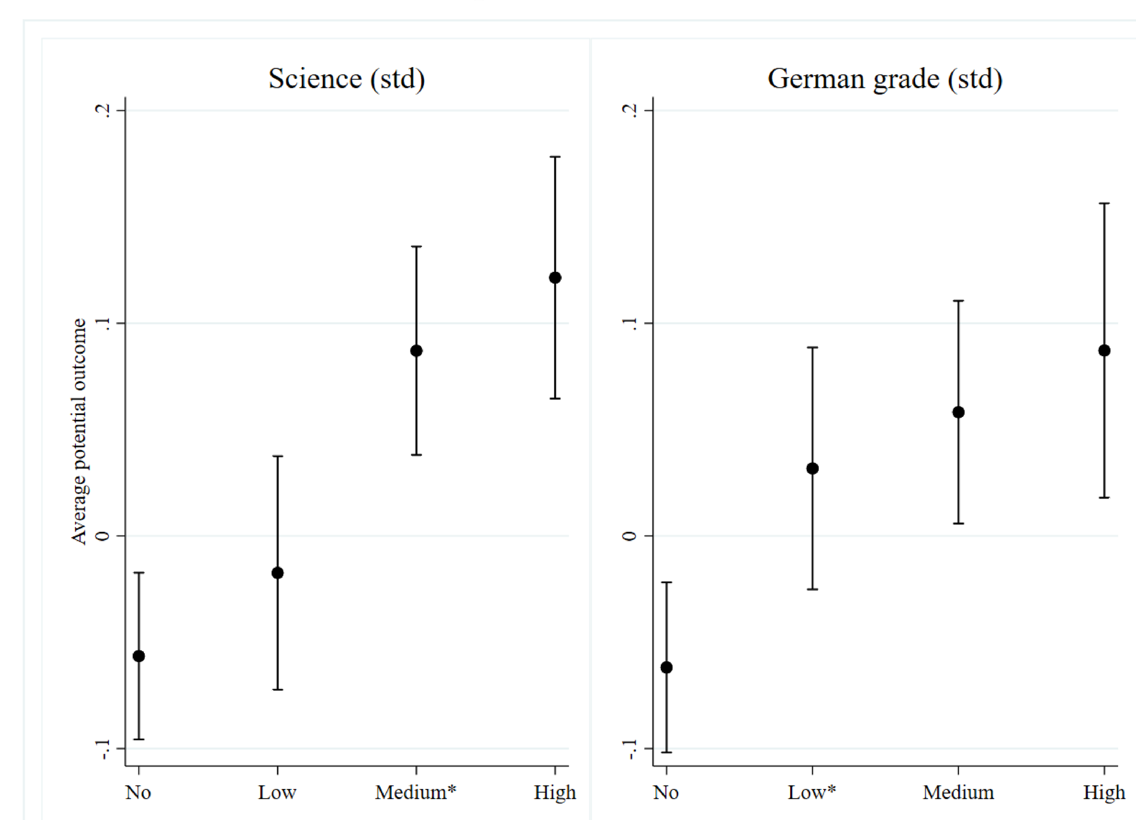

Figure 2: Two examples of estimated potential outcomes

Main Result

- Positive effects on *objectively measured skills* require *at least medium intensity*
- Positive effects on *German grades* already for *low intensity*
- Results are *not sensitive* to the *penalty choice*
- The inclusion of < 30 variables suffices to achieve balancing of the high-dimensional covariates

Sensitivity analysis

Sensitivity to penalty term choice:

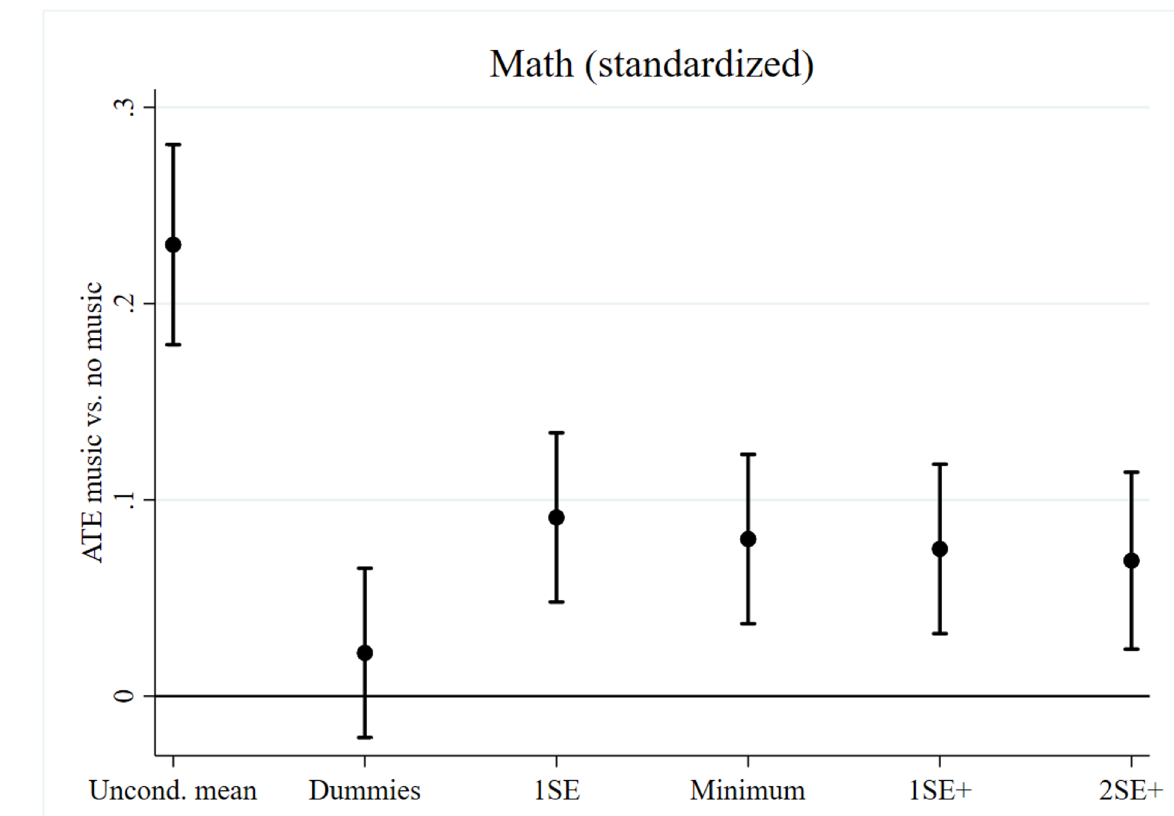

Figure 3: Representative example for sensitivity to penalty choice with binary music indicator

Balancing of all variables assessed via standardized differences

$$SD = \frac{|\bar{X}_1 - \bar{X}_0|}{\sqrt{(Var(X_1) + Var(X_0))/2}} \cdot 100$$

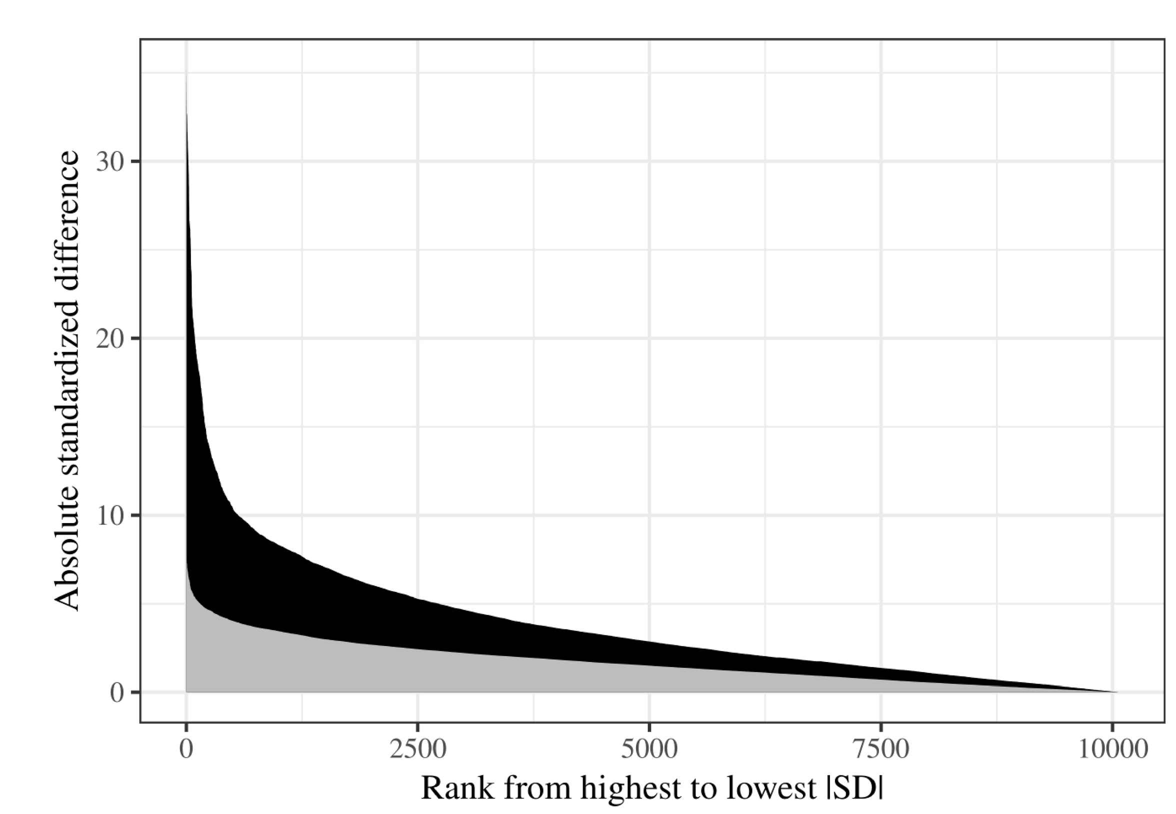

Figure 4: Balancing before (black) and after (grey) covariate adjustment

Swiss Institute for Empirical Economic Research
University of St.Gallen
75 NRP
Big Data
National Research Programme