Unmasking Unfair Correlation: Achieving Credible Interpretation for Group Average Treatment Effects

Jana Mareckova

Applied Machine Learning Days 2024
EPFL Lausanne
March 25, 2024

Potential Outcome Framework

Evaluation of the effect of programs for unemployed:

- **Different Programs:** Basic Courses, Training Courses, Temporary Earning Subsidies, etc.
- **Different Outcomes**: Probability of employment, Level of earnings

Potential Outcome Framework:

- $X \in \mathbb{R}^p$ ... vector of features,
- $D \in \{0, 1, ..., M\}$ ... treatment / program assignment,
  - $d = 0$ ... no program (control group)
  - $d > 0$ ... participation in program (treated group)
- Y^d ... potential outcome under assignment $D = d$.

Parameters of Interest

Average Treatment Effect (ATE):

$$\tau_d = E[Y^d - Y^0]$$

Conditional Average Treatment Effect (CATE):

- $X \rightarrow$ Individualized Average Treatment Effect (IATE)

$$\tau_d(x) = E[Y^d - Y^0 | X = x]$$

- Low-dimensional subset of X, $Z \subset X \rightarrow$ Group Average Treatment Effect (GATE)

$$\tau_d(z) = E[Y^d - Y^0 | Z = z]$$

Identification: Selection-on-Observables

Under certain assumptions ("selection-on-observables"), the following and other identification results are possible:

- $E[Y^d - Y^0|X = x] = E[Y|X = x, D = d] - E[Y|X = x, D = 0]$
- $Y = \sum_{d=0}^{M} 1(D = d)Y^d$
- similar results for GATE and ATE

Causal Machine Learning - IATE

Options for estimation:

- Estimation of a Double Robust Score via Double Machine Learning (DML) (Chernozhukov et al., 2018) combined with a regression step (Foster and Syrgkanis, 2023; Kennedy, 2023)
- Causal Forests (CF)
  - Generalized Regression Forests (GRF) by Athey et al. (2019)
  - Modified Causal Forest (MCF) by Lechner (2018), Lechner & Mareckova (2022)

Advantage:

- Granular effects

Disadvantage:

- Black-box nature $\rightarrow$ lack of interpretability

Causal Machine Learning - GATE

Options for Estimation:

- Estimation of a Double Robust Score via DML or GRF combined with a regression step (Semenova & Chernozhukov, 2021 - linear regression)
- Aggregation of IATEs from the MCF (Lechner, 2018)

Advantage:

- Easier interpretability
- Differences between groups

  $E[Y^1(X) - Y^0(X)|Z = male] - E[Y^1(X) - Y^0(X)|Z = female]$

Disadvantage:

- Prone to be contaminated by correlation with other features when looking at differences among groups

Balanced GATE

- Introduced in Bearth and Lechner (2024)
- Looking at the difference of two GATEs, as if the distribution of other features $W \subset X$ was the same among the groups

$$\tau_d^B(z) = E[E_B[Y^d - Y^0 | Z = male, W] - E_B[Y^d - Y^0 | Z = female, W]]$$

- DML and CF version

Goal:

- Account for variation in other features

Application: Interventions in Switzerland

- Mascolo, Bearth, Muny, Lechner, & Mareckova (2024)
- Swiss administrative data from 2004-2018
- Unemployed individuals eligible for benefits between January 2014 and June 2015
- 5 programs:
  - No program
  - Basic Courses
  - Training Courses
  - Temporary Employment
  - **Temporary Earnings ("Zwischenverdienst", "Le gain intermédiaire")**
- 2 outcomes in the 3rd year after the start of the program:
  - Probability of Employment
  - **Level of Earnings**

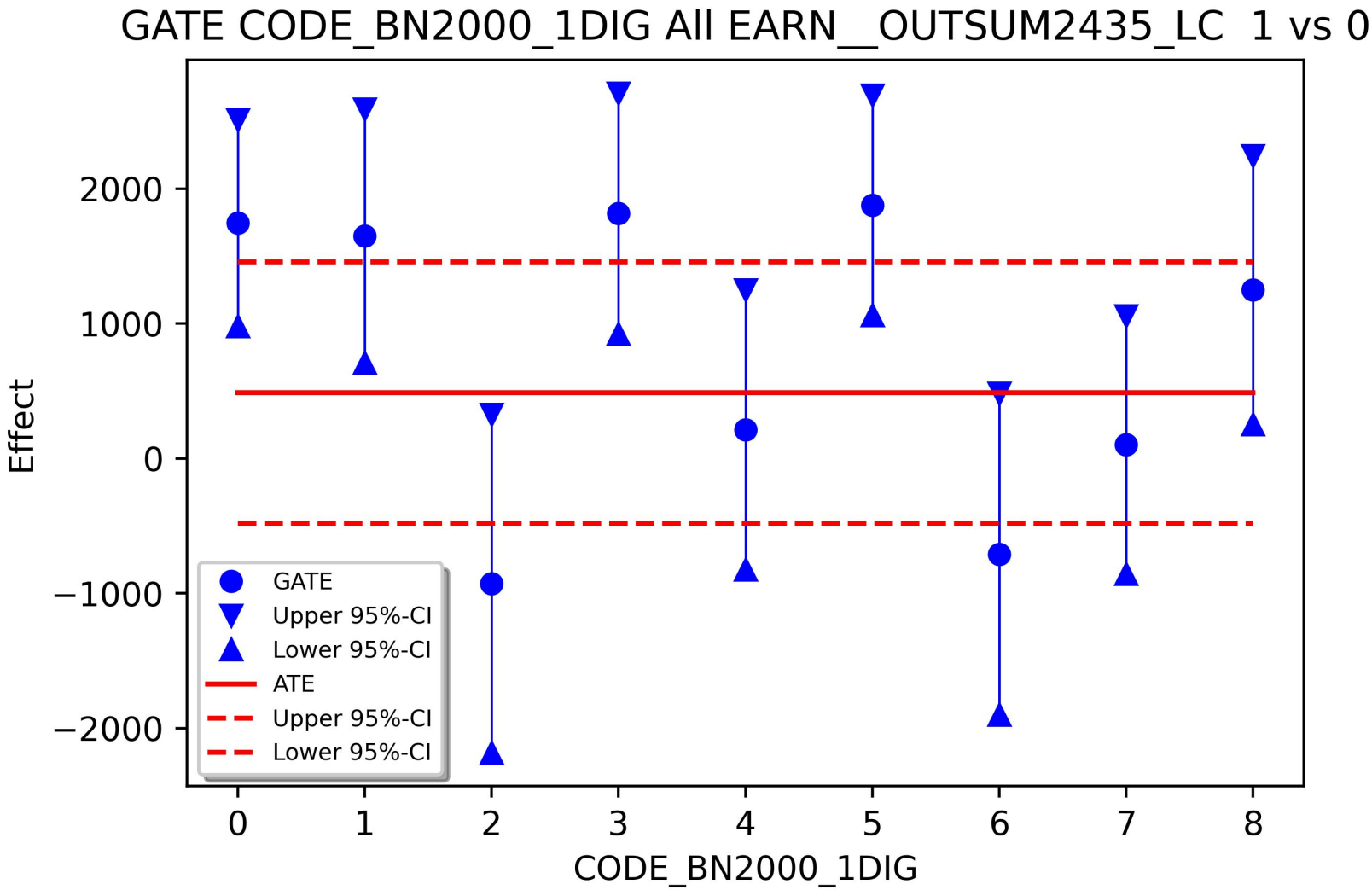
Application: Features

- socio-demographic characteristics (age, gender, civil status, nationality, language proficiency, work permit type, etc.)
- prior employment history
- prior unemployment history
- macroeconomic indicators (regional unemployment rate, etc.)

Application: Method

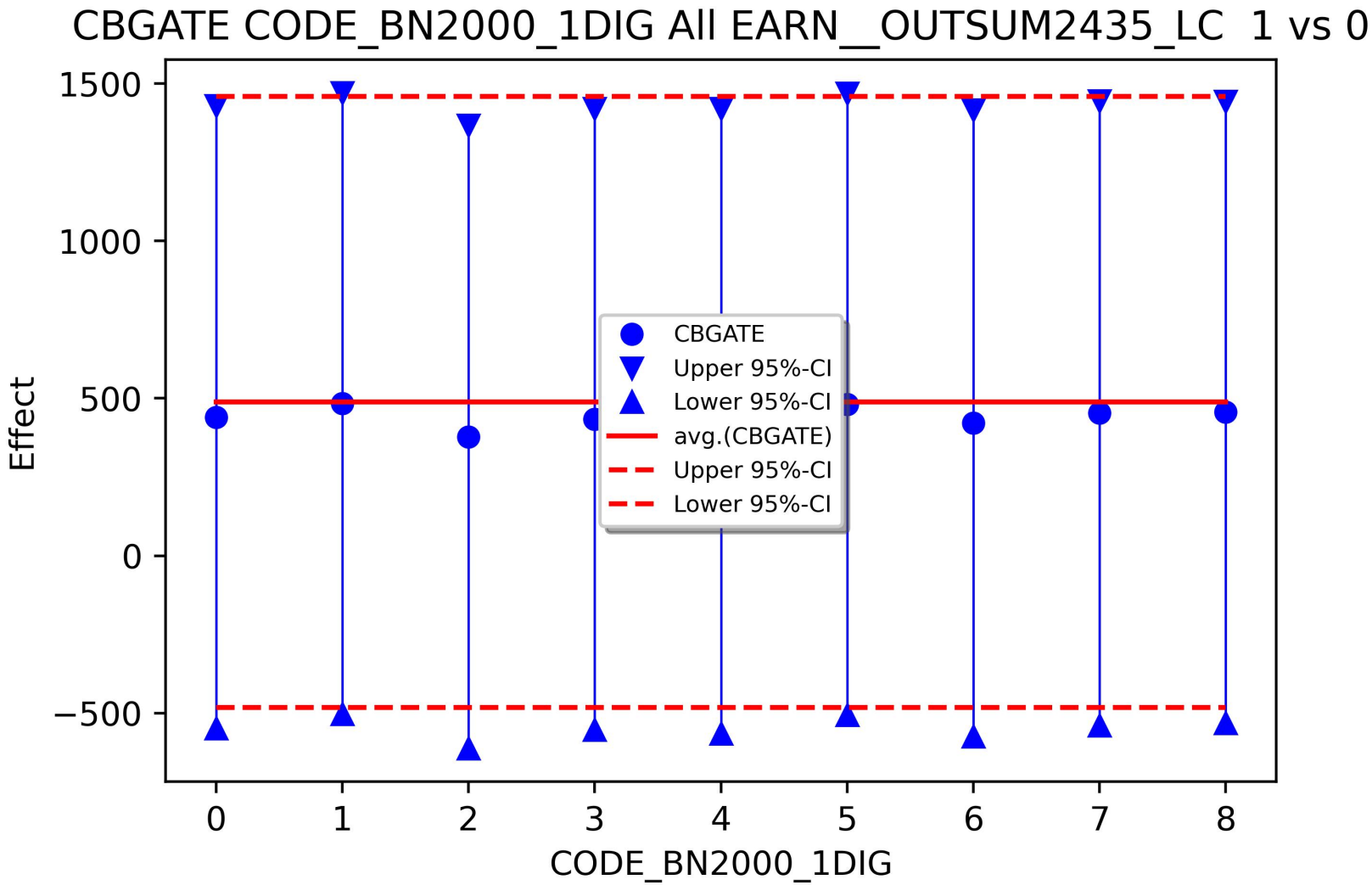
- MCF
- available as Python package

Application: GATE - sector



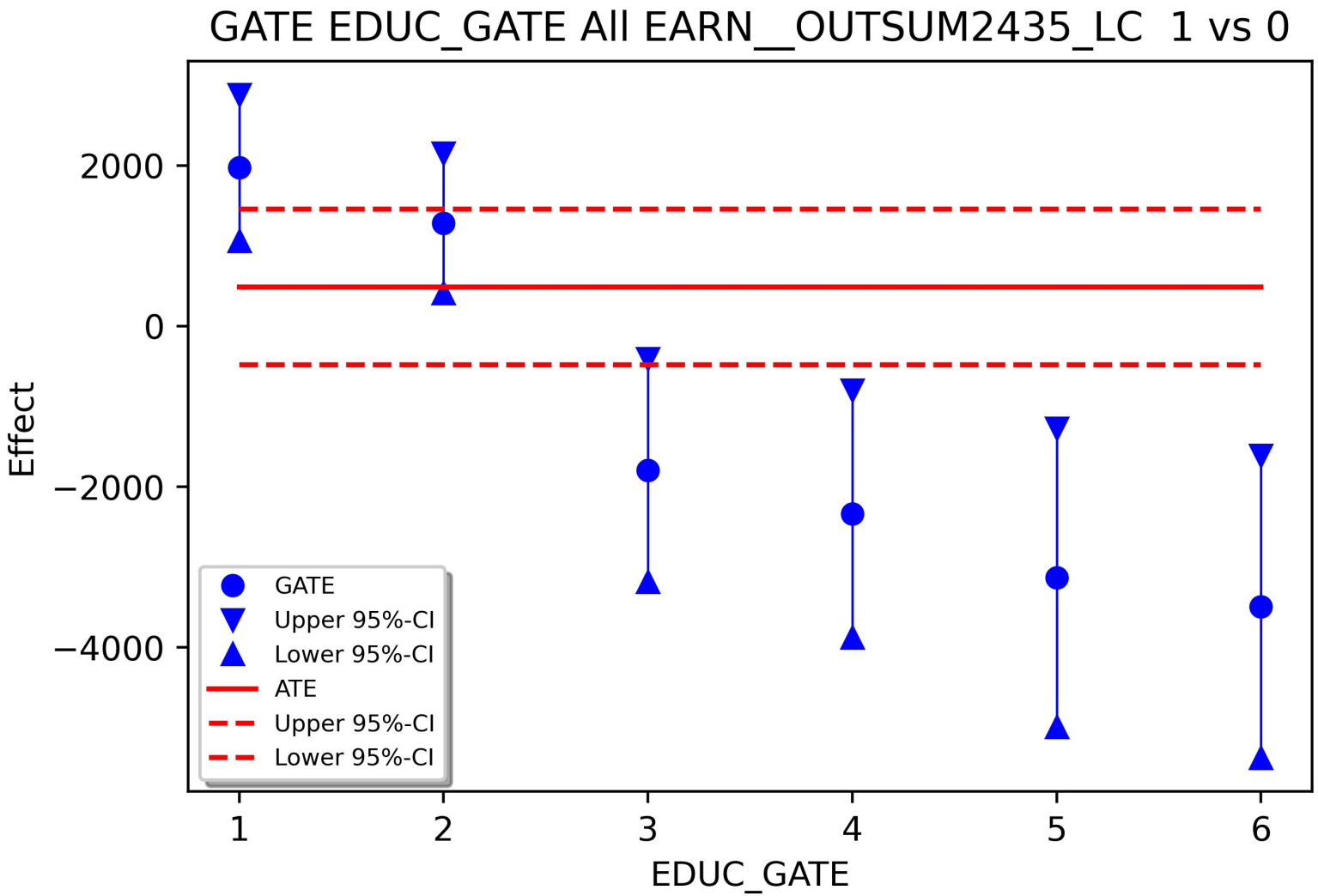
0 - Agriculture, 1 - Industry and trade, 2 - IT, 3 - Construction, 4 - Transport, 5 - Hospitality, 6 - Banking and law, 7 - Health and science, 8 - Rest

Application: BGATE - sector



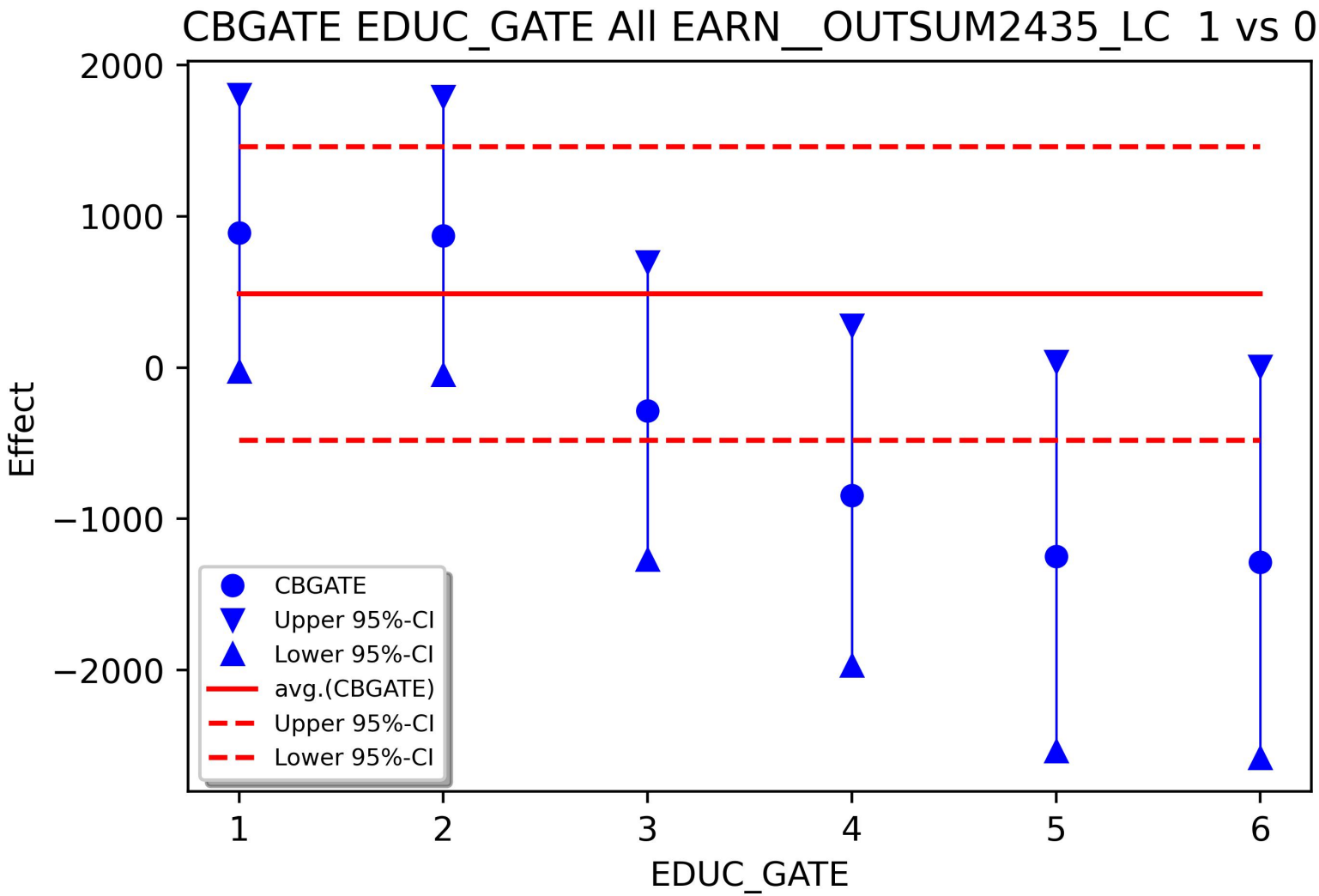
0 - Agriculture, 1 - Industry and trade, 2 - IT, 3 - Construction, 4 - Transport,
5 - Hospitality, 6 - Banking and law, 7 - Health and science, 8 - Rest

Application: GATE - education



1 - Compulsory, 2 - Secondary, 3 - Vocational, 4 - BA, 5 - MA, 6 - PhD

Application: BGATE - education



1 - Compulsory, 2 - Secondary, 3 - Vocational, 4 - BA, 5 - MA, 6 - PhD

Conclusion

CML enables estimation of IATEs and GATEs:

- more granular predictions
- better indications how policy works
- BGATE can help to unmask correlations to other features for better understanding

References I

- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). **Double/debiased machine learning for treatment and structural parameters.** The Econometrics Journal 21(1), 1-68.
- Foster, D. J., and V. Syrgkanis (2023). **Orthogonal statistical learning.** The Annals of Statistics, 51(3), 879-908.
- Kennedy, E. H. (2023). **Towards optimal doubly robust estimation of heterogeneous causal effects.** Electronic Journal of Statistics, 17 (2), 3008-3049.
- Athey, S., Tibshirani, J., & Wager, S. (2019). **Generalized random forests.** The Annals of Statistics, 47(2), 1148.
- Lechner, M. (2018). **Modified causal forests for estimating heterogeneous causal effects.** arXiv preprint arXiv:1812.09487.

References II

- Lechner, M., & Mareckova, J. (2022). **Modified causal forest.** arXiv preprint arXiv:2209.03744.
- Semenova, V., & Chernozhukov, V. (2021). **Debiased machine learning of conditional average treatment effects and other causal functions.** The Econometrics Journal, 24(2), 264-289.
- Bearth, N., & Lechner, M. (2024). **Causal Machine Learning for Moderation Effects.** arXiv preprint arXiv:2401.08290.
- Mascolo, F., Bearth, N., Muny, F., Lechner, M., & Mareckova, J. (2024) **The Heterogeneous Effects of Active Labour Market Policies.** Working Paper.