Max Planck ETH Center for Learning Systems

Aptitude colloquium

Vivian Y. Nastl
November 14 2023
ETH Zürich

Chairperson:
Prof. Dr. Peter L. Bühlmann

Supervisor :
Prof. Dr. Nicolai Meinshausen

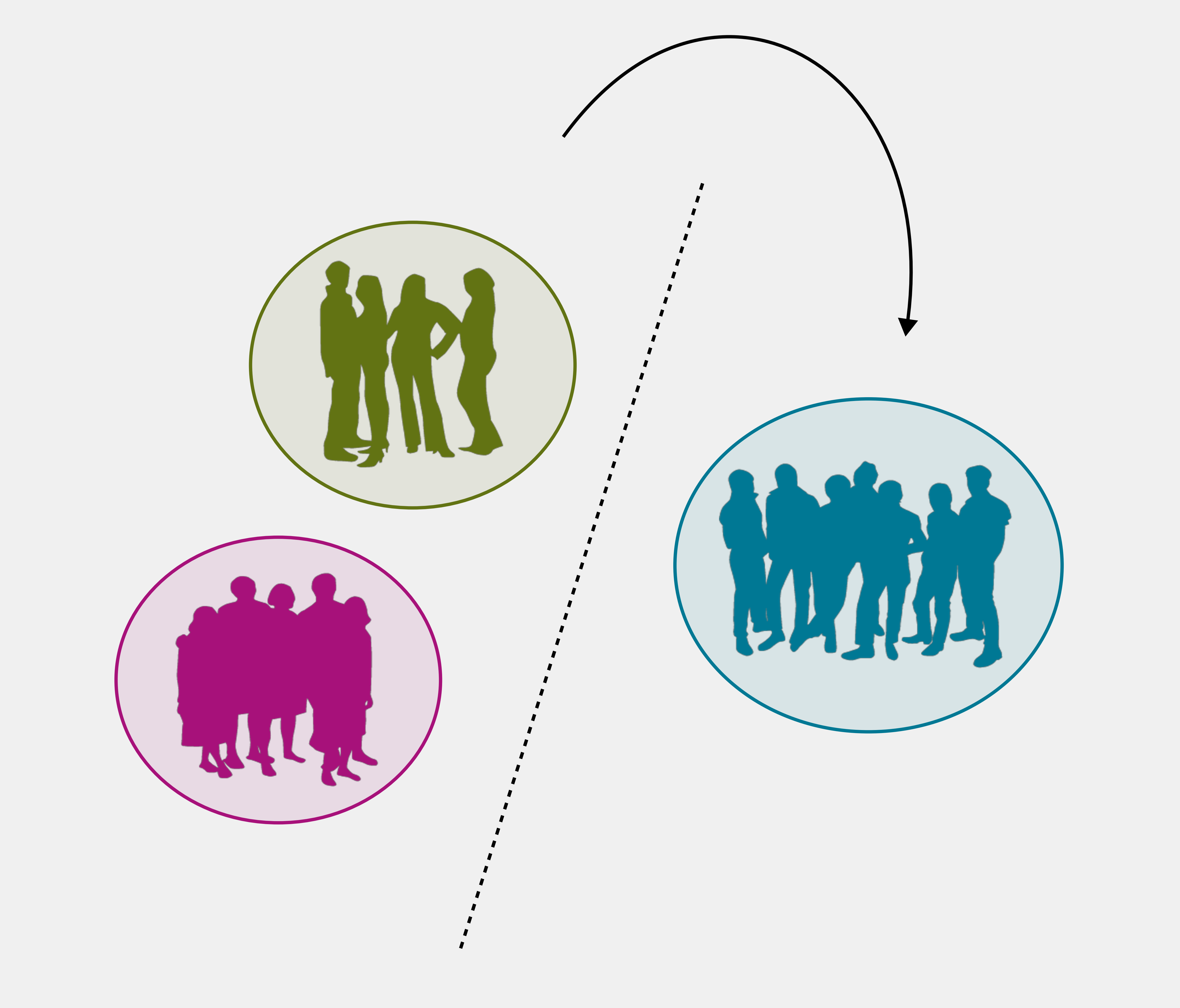
Second Advisor :
Dr. Moritz Hardt

Agenda

1. Project 1 with Moritz Hardt
2. Project 2 with Ana-Andreea Stoica and Moritz Hardt
3. Time Schedule
4. Questions & Discussions

Project 1 with Moritz Hardt

November 2022 - February 2023



Motivation

Invariance of causal models: Assume a SCM over (Y, X, E) with $Y \leftarrow f(X_{S*}) + \varepsilon$ and $S*$ indicating the causal parents of Y. Let E be neither descendants nor parents of Y. Then,

$$Y \perp\!\!\!\perp E \,|\, X_{S*}$$

Learning the causal parents $S*$

Invariant Causal Prediction

Enabling out-of-distribution generalisation of ML algorithms

Invariant Feature Learning

- Invariant Risk Minimization
- Risk Extrapolation
- ...

Arjovsky, Bottou, Gulrajani, & Lopez-Paz, 2019
Heinze-Deml, Peters, & Meinshausen, 2018
Krueger, et al., 2021
Peters, Bühlmann, & Meinshausen, 2016
Schölkopf, et al., 2012

Does theoretical **invariance** of causal models enables **out-of-domain generalisation** in real-world **social applications**?

Method

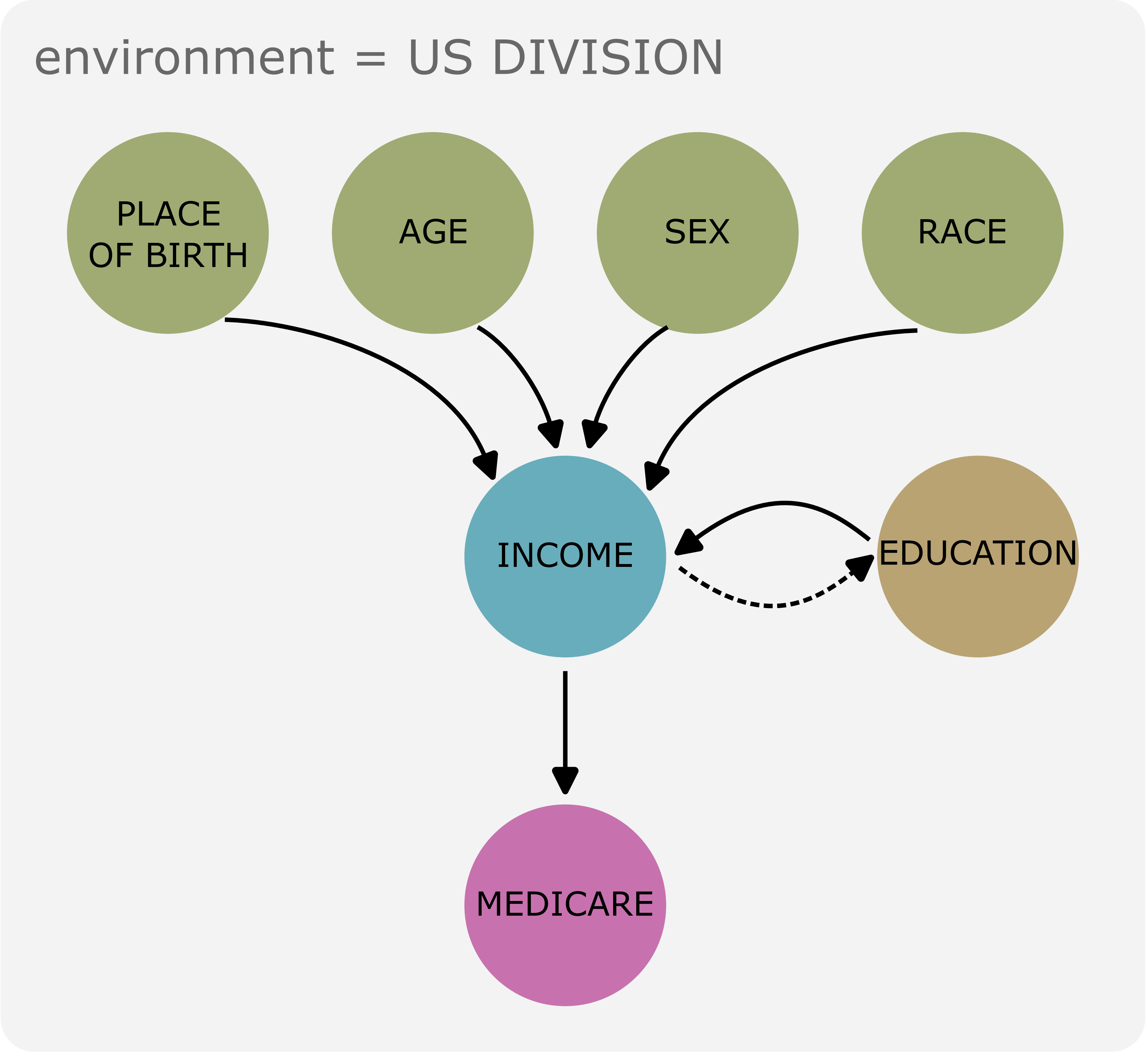
1. Representative real-world social applications with natural domain shift

2. Pre-specified subset of *causal* parents $X_{\hat{S}}$ of Y, i.e., it is reasonable to assume that

   - X_i has causal effect on Y

   - X_i is a root node in the corresponding causal graph

3. Compare in-domain and out-of-domain performance of causal parents $X_{\hat{S}}$ and all features X



TableShift

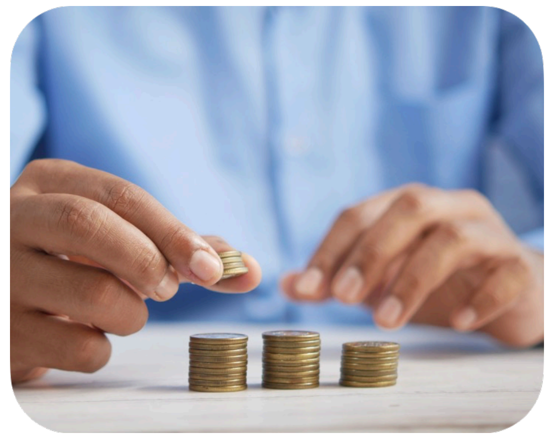
- Domain generalisation benchmark for tabular data
- 15 binary classification task with associated shifts
- Diverse targets in public policy, health care, civic participation, education
- **Definition (Shift gap):**

$$\Delta_{\text{Acc}} = \text{Acc}(f_\theta, \mathcal{D}_{\text{test}}) - \text{Acc}(f_\theta, \mathcal{D}_{\text{train}})$$

with f_θ a fixed classifier, $\mathcal{D}_{\text{test}}$ testing domains and $\mathcal{D}_{\text{train}}$ training domains

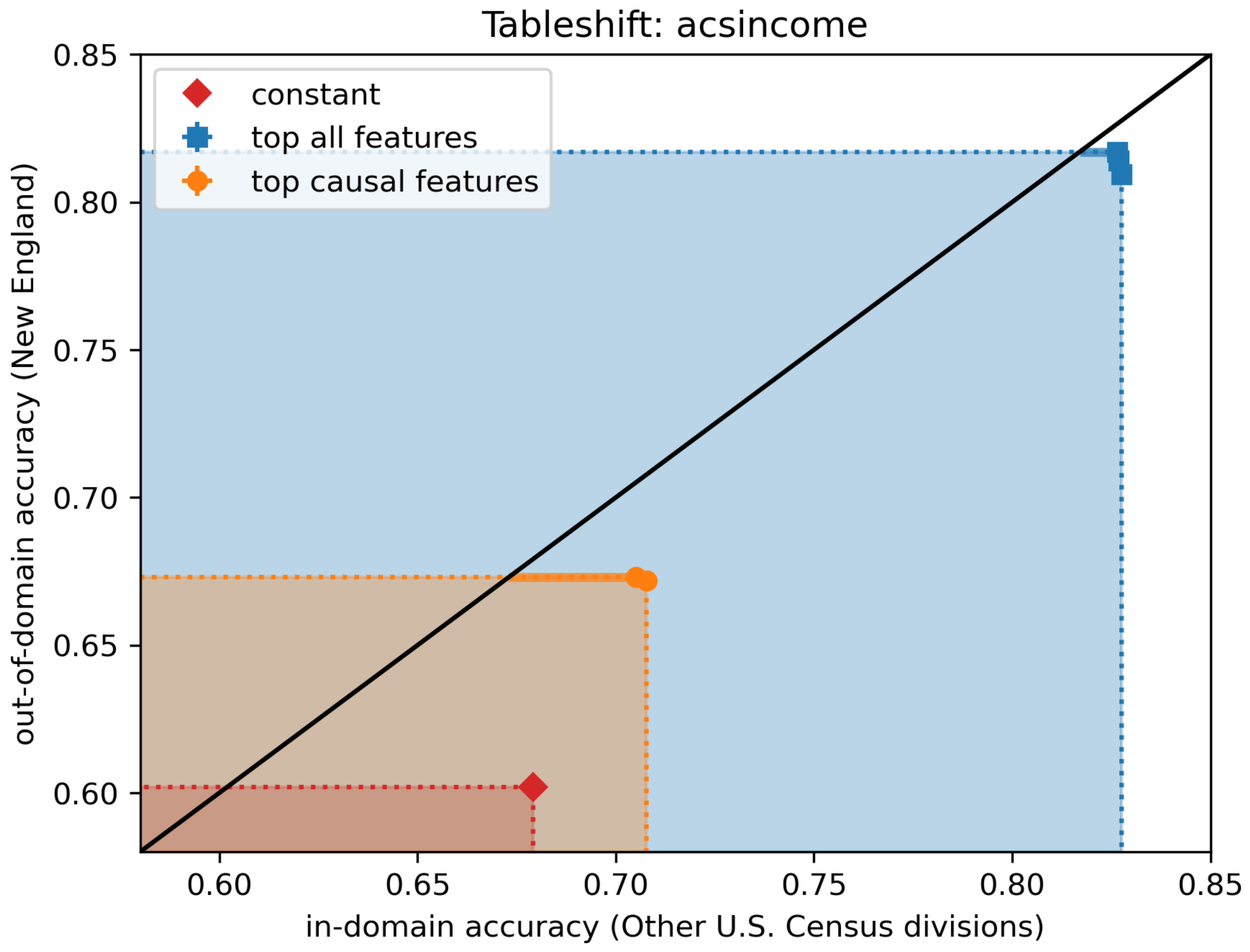
Datasets in TableShift

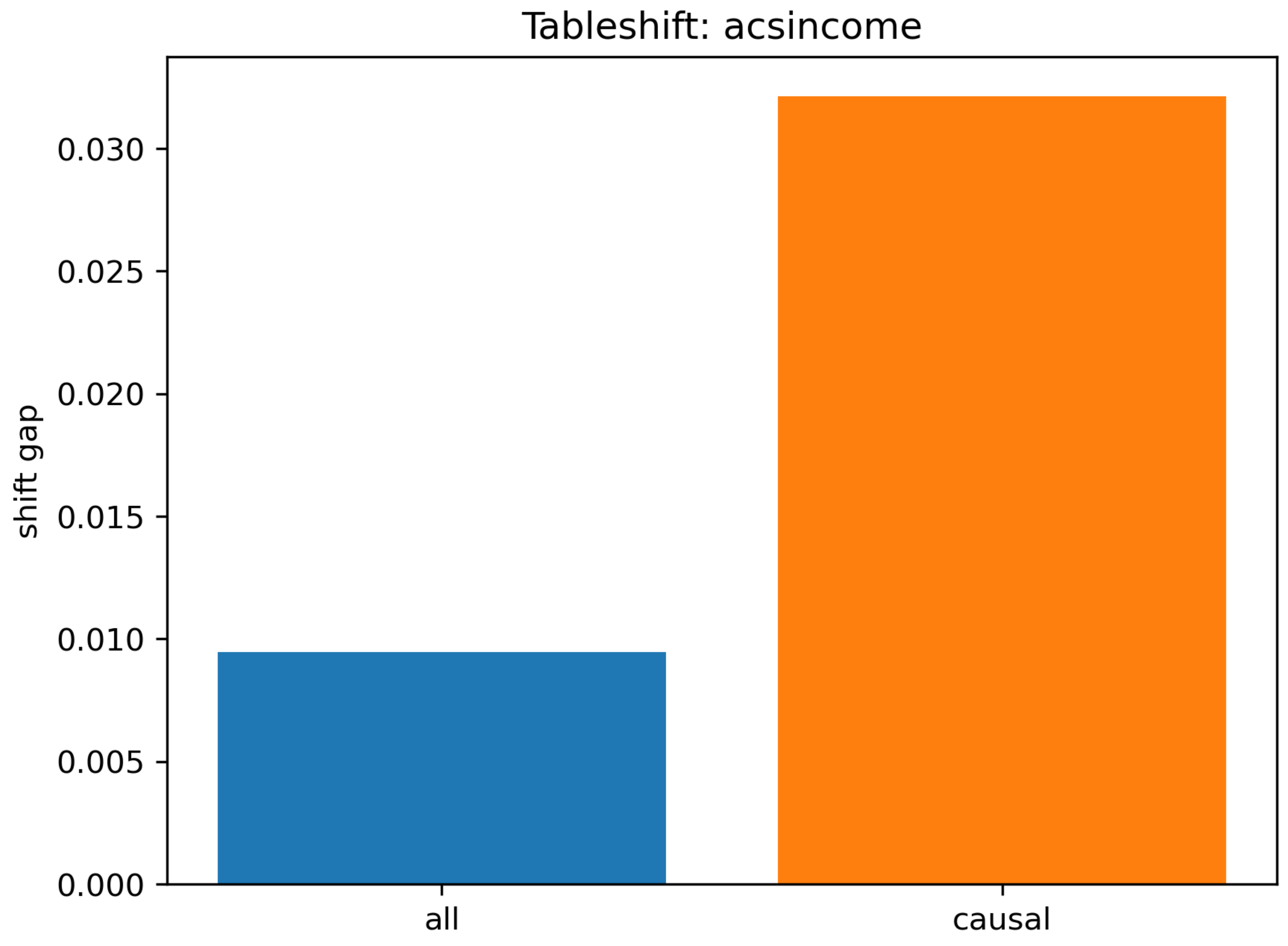
Examples

	Target	Shift	Source
	Income >= 56k for employed adults	Geographic Region (Division)	American Community Survey (via folktables)
	Voted in US presidential election	Geographic Region (Region)	American National Election Studies (ANES)
	Length of stay >= 3 hrs in ICU	Insurance Type	MIMIC-III Clinical Database

Ding, et al., 2023
Gardner, Popovi, & Schmidt, 2023

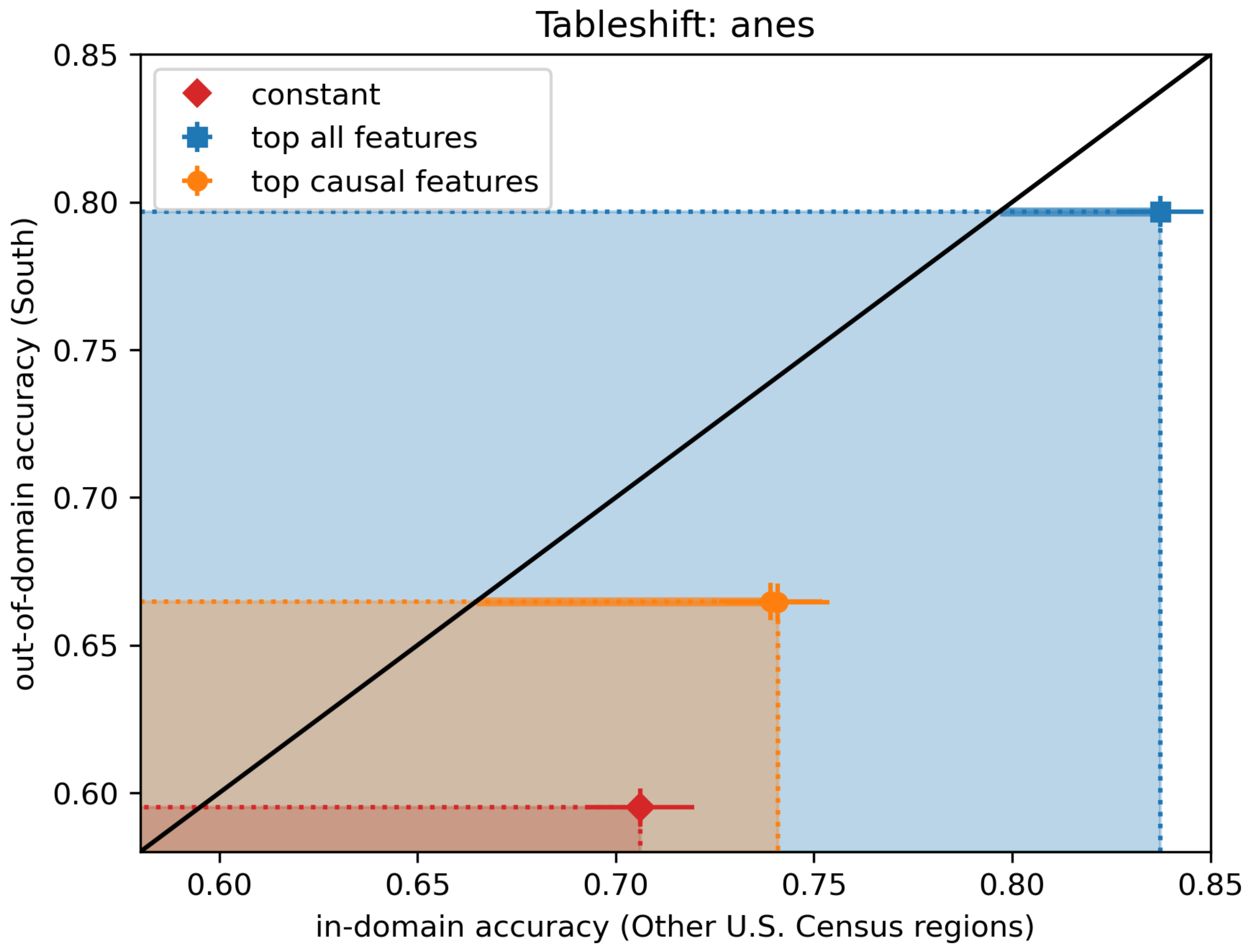
Preliminary Results: Income

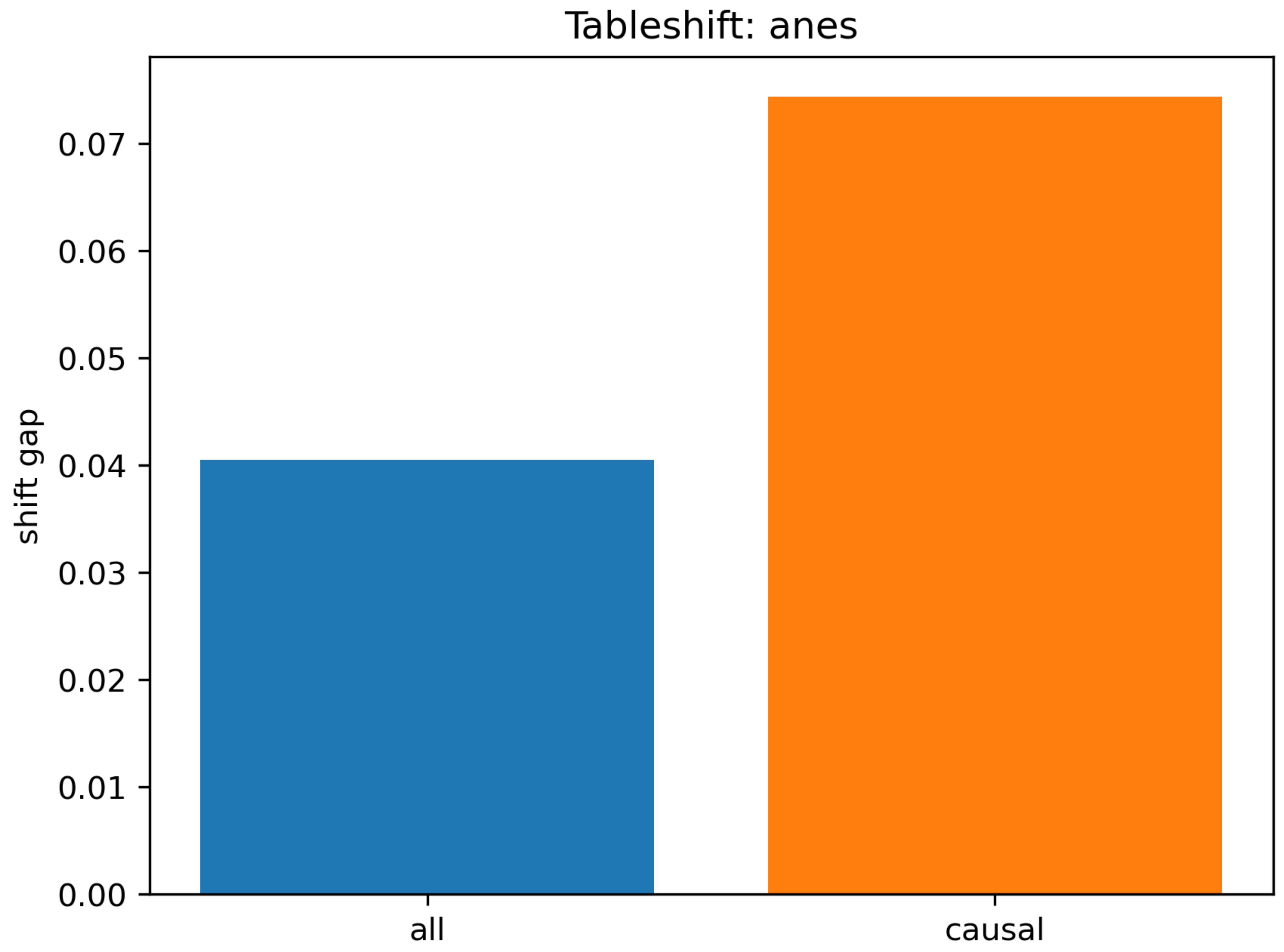




Causal features: age, sex, race, place of birth

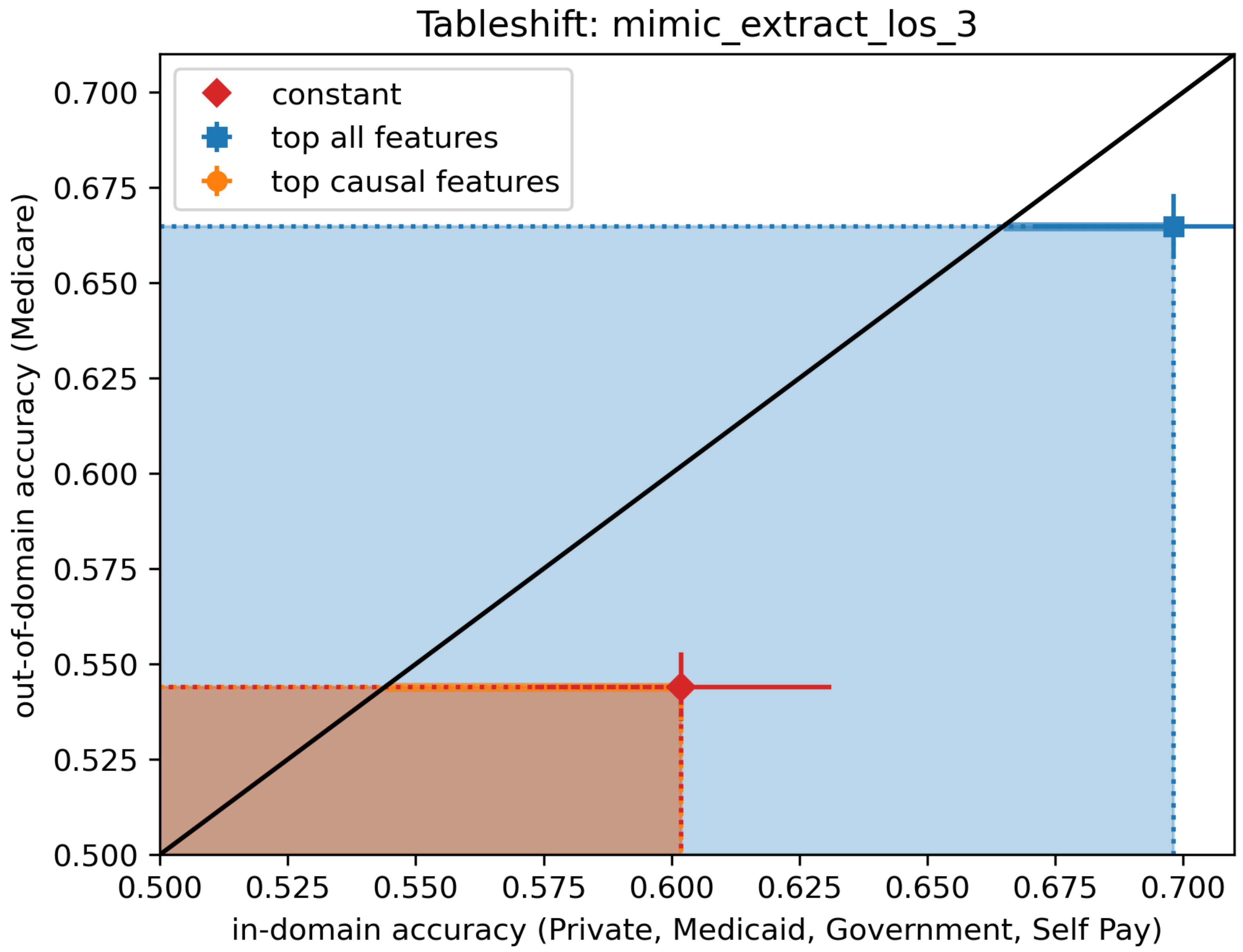
Preliminary Results: Voting

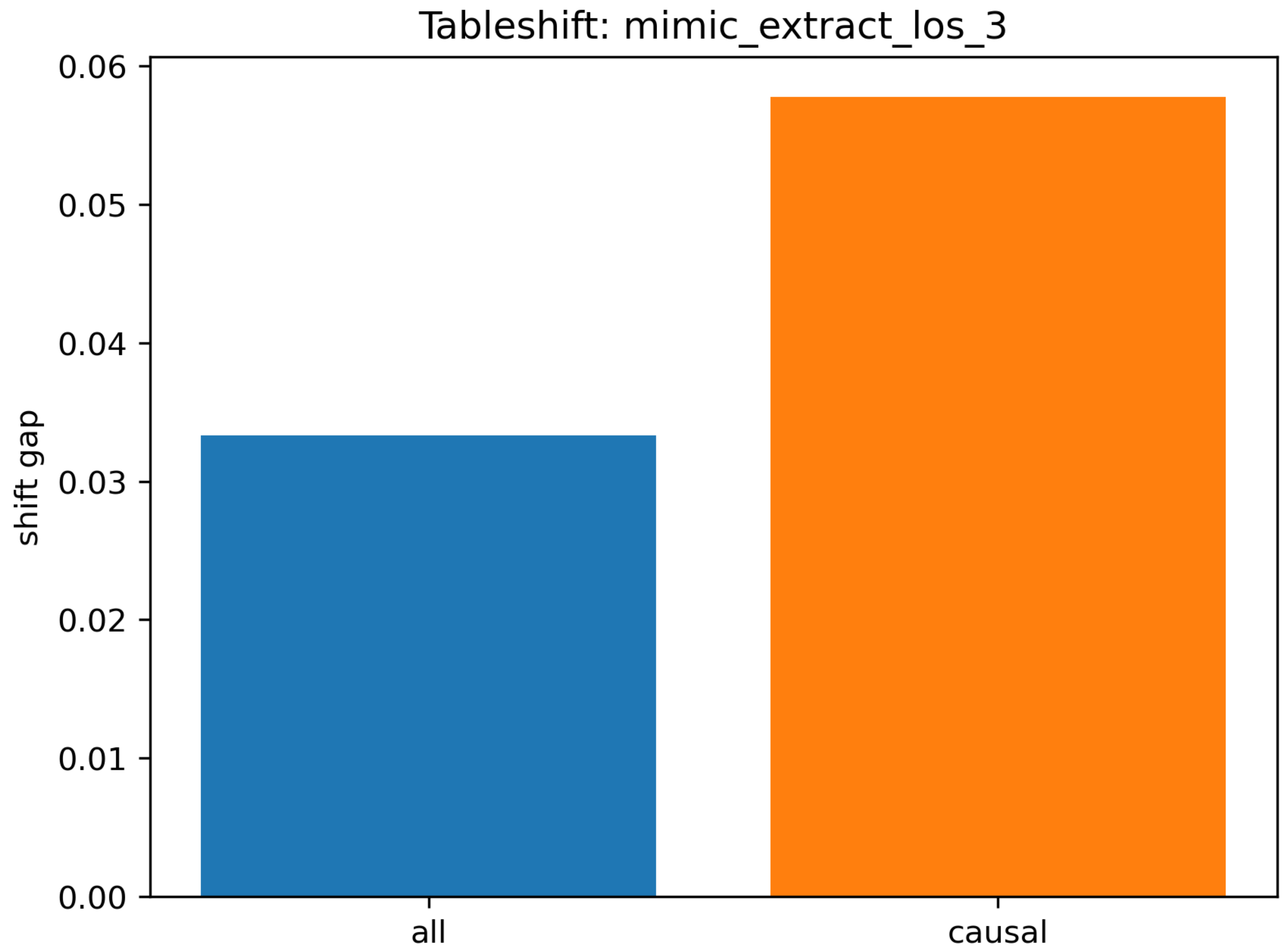




Causal features: election year, state, registered to vote pre-election, age, gender, race

Preliminary Results: Stay in ICU





Causal features: gender, age, ethnicity, height, weight

Summary & Outlook

1. The performance of causal features is **Pareto-dominated** by the performance of all features, w.r.t. in-domain and out-of-domain performance

2. In **some task**, the current selection of causal features have **no predictive power**

3. The **shift gap** of the causal features is **larger** than the shift gap of all features

Relax the current classification of causal features and analyse performance

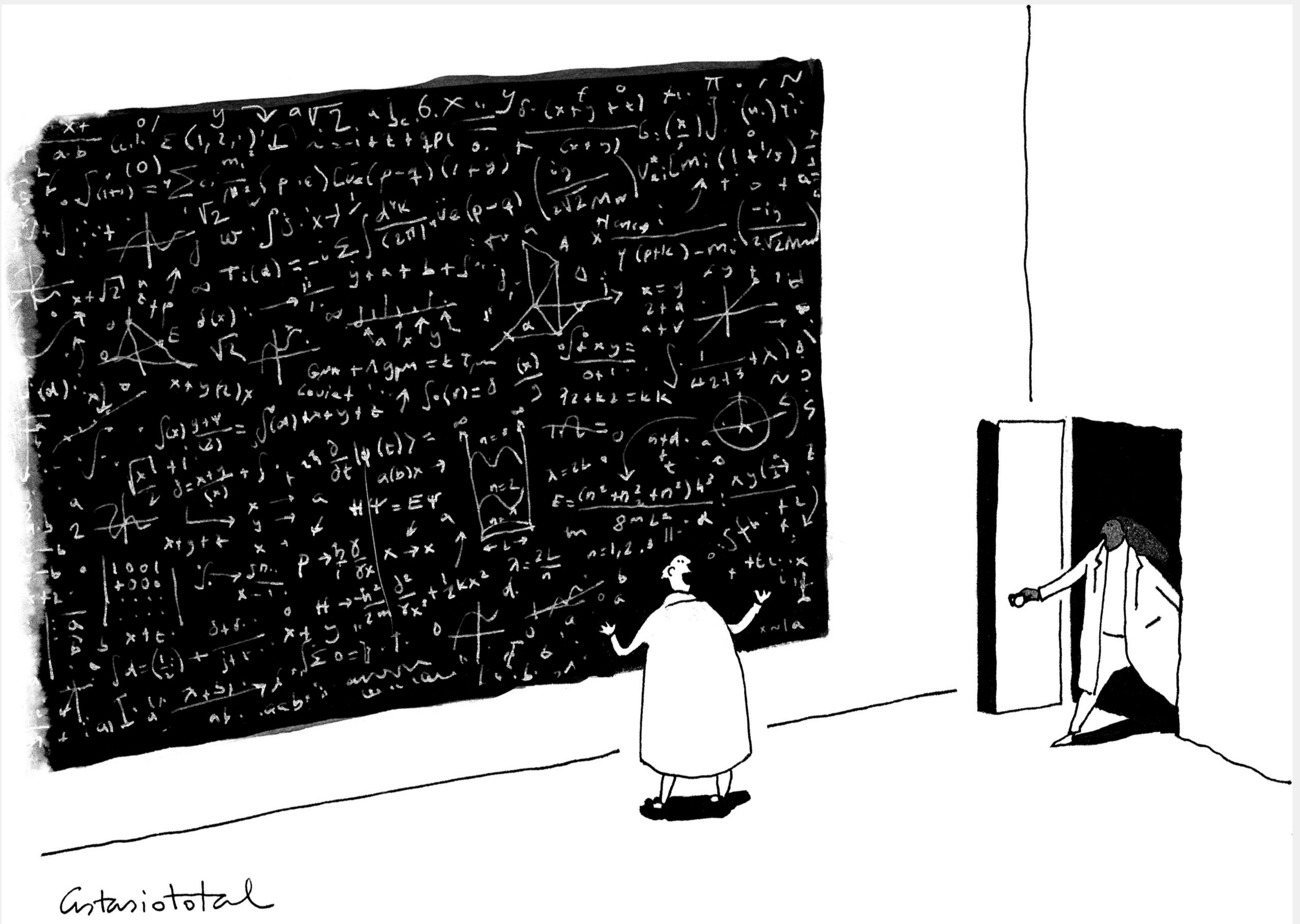
Example: education is ‘causal’ for income

Project 2 with Ana-Andreea Stoica and Moritz Hardt

April 2023 - February 2023

THE NEW YORKER

CARTOON CAPTION CONTEST

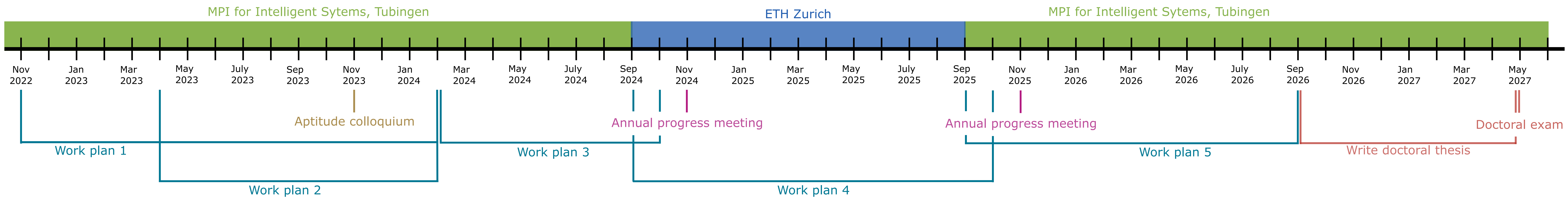


“I do this to scare the students, I actually have no idea what it means.”

Time Schedule

Time schedule

- **Exchange year:** September 01, 2024 - August 31, 2025
- **Annual progress meeting (CLS):** November 2024 & November 2025
- Aptitude colloquium counts for annual progress meeting in 2023



Questions & Discussions

References

Arjovsky, M., Bottou, L., Gulrajani, I., & Lopez-Paz, D. (2019). Invariant risk minimization. *CoRR*, http://arxiv.org/abs/1907.02893.

Ding, F., Hardt, M., Miller, J., & Schmidt, L. (2021). Retiring adult: New datasets for fair machine learning. *Advances in neural information processing systems*, *34*, 6478-6490.

Gardner, J. P., Popovi, Z., & Schmidt, L. (2023). Benchmarking Distribution Shift in Tabular Data with TableShift. *NeurIPS 2023 Datasets and Benchmarks Poster*, preprint.

Heinze-Deml, C., Peters, J., & Meinshausen, N. (2018). Invariant Causal Prediction for Nonlinear Models. *Journal of Causal Inference, 6(2)*, https://doi.org/10.1515/jci-2017-0016.

Peters, J., Bühlmann, P., & Meinshausen, N. (2016). Causal inference using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (with discussion), 78(5)*, 947–1012.

Schölkopf, B., Janzing, D., Peters, J., Sgouritsa, E., Zhang, K., & Mooij, J. (2012). On causal and anticausal learning. *Proceedings of the 29th International Conference on Machine Learning (ICML)*, 1255–1262.