FAIR CAUSAL INFERENCE FOR FUNCTIONAL DATA

Computational Statistics Conference; Bologna, 2022

Tim Mensinger & Dominik Liebl

University of Bonn

Who Am I?

- Tim
- PhD Candidate in Economics, University of Bonn
- Focus:
  - Econometrics & Statistics
  - Causal Inference
  - Programming

Motivation

- Foot striking patterns:
  - forefoot vs heel

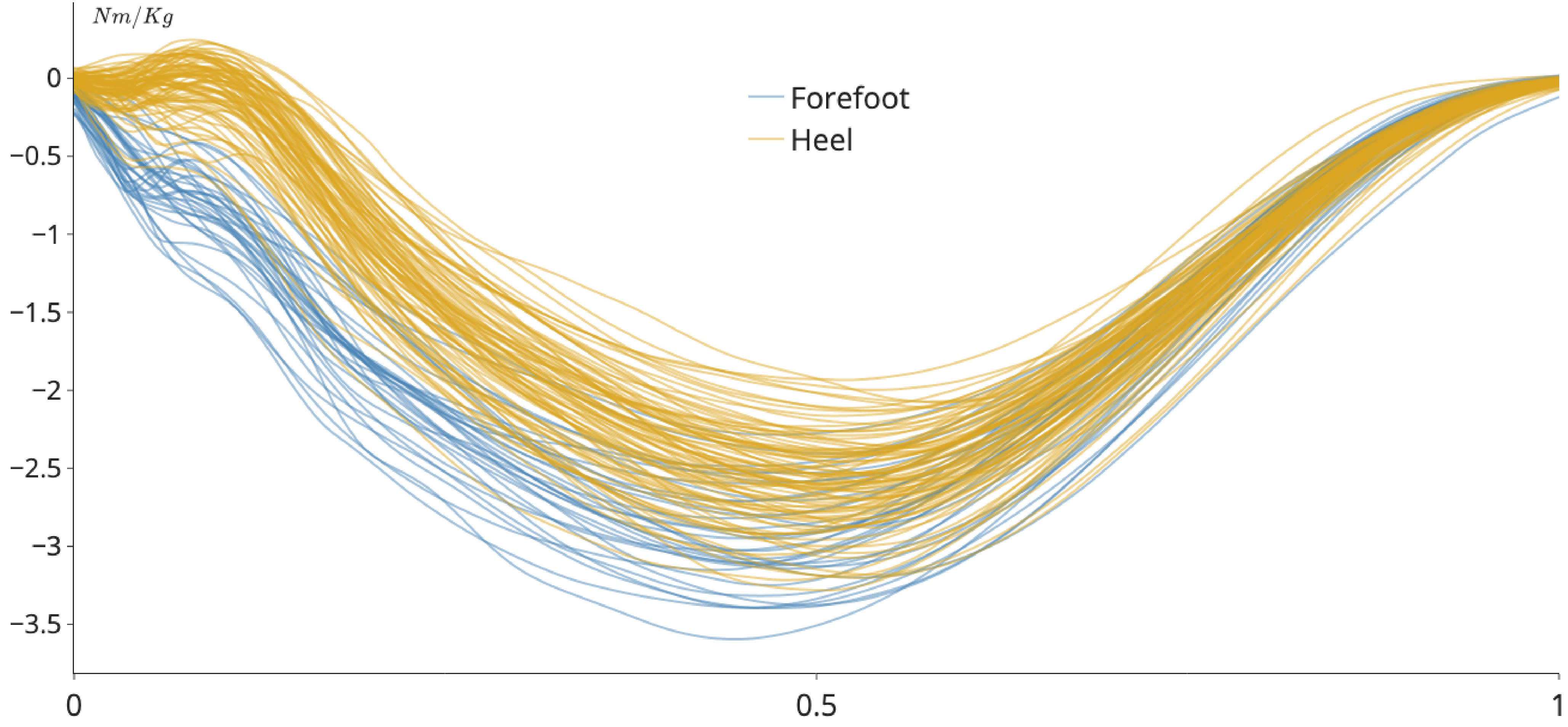
Motivation

- Foot striking patterns:
    - forefoot vs heel
- Consider one metric: Force on ankle joints

Motivation

- Foot striking patterns:
  - forefoot vs heel
- Consider one metric: Force on ankle joints
- ***What's the (causal) effect of forefoot running on ankle joint loading?***

Data



Data Structure

Outcomes	Controls	Treatment
$Y_i \in C^1[0,1]$	$X_i \in \mathbb{R}^p$	$W_i \in \{0,1\}$

Data Structure

Outcomes	Controls	Treatment
$Y_i \in C^1[0,1]$	$X_i \in \mathbb{R}^p$	$W_i \in \{0,1\}$

- Potential Outcomes:
  - $Y_i(1), Y_i(0) \in C^1[0,1]$

Data Structure

Outcomes	Controls	Treatment
$Y_i \in C^1[0,1]$	$X_i \in \mathbb{R}^p$	$W_i \in \{0,1\}$

- Potential Outcomes:
  - $Y_i(1), Y_i(0) \in C^1[0,1]$
  - $Y_i = W_i Y_i(1) + (1 - W_i) Y_i(0)$

Data Structure

Outcomes	Controls	Treatment
$Y_i \in C^1[0,1]$	$X_i \in \mathbb{R}^p$	$W_i \in \{0,1\}$

- Potential Outcomes:
  - $Y_i(1), Y_i(0) \in C^1[0,1]$
  - $Y_i = W_i Y_i(1) + (1 - W_i) Y_i(0)$
  - SUTVA: $Y_i = Y_i(W_i)$

Object of Interest

- Average treatment effect function:

$$\tau(t) = \mathbb{E}[Y_i(1)(t) - Y_i(0)(t)]$$

for $t \in [0, 1]$

Object of Interest

- Average treatment effect function:

$$\tau(t) = \mathbb{E}[Y_i(1)(t) - Y_i(0)(t)]$$

  for $t \in [0, 1]$

- Identification under **unconfoundness** and **overlap**:
  - $(Y_i(1), Y_i(0)) \perp\!\!\!\perp W_i | X_i$

Object of Interest

- Average treatment effect function:

$$\tau(t) = \mathbb{E}[Y_i(1)(t) - Y_i(0)(t)]$$

  for $t \in [0, 1]$

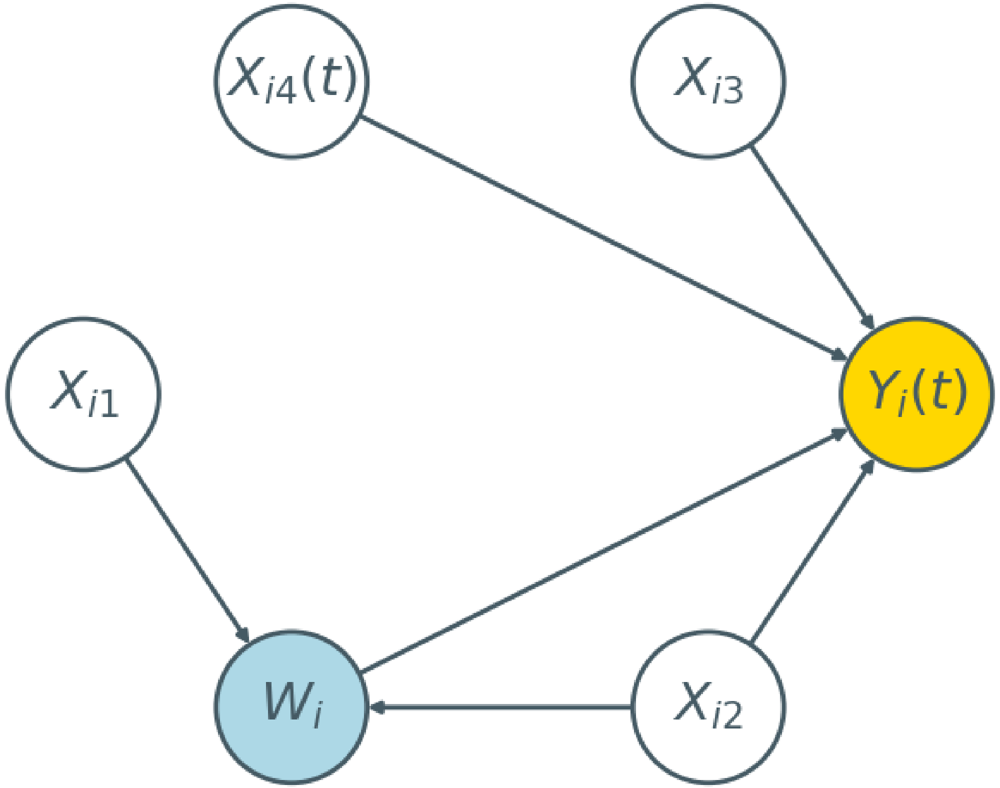
- Identification under **unconfoundness** and **overlap**:
  - $(Y_i(1), Y_i(0)) \perp\!\!\!\perp W_i | X_i$
  - $\eta < \mathbb{P}[W_i = 1 | X_i] < 1 - \eta$

Plan

1. **Find relevant control variables**
   - Utilize causal graphs from *causal inference* literature
2. **Choose a suitable estimator**
   - Utilize methods from *econometrics* literature
3. **Construct confidence bands**
   - Utilize results from *functional data* literature

Find relevant control variables

Directed Acyclical Graph



- For $t \in [0, 1]$
- Structure may change with t
- Set of variables used for prediction of outcome and treatment may differ

Choose a suitable estimator

Augmented Inverse Propensity Score Weighting

$$\hat{\tau}(t) = \frac{1}{n}\sum_{i=1}^{n} \hat{\mathbb{E}}[Y_i(t)|X_i, W_i = 1] - \hat{\mathbb{E}}[Y_i(t)|X_i, W_i = 0]$$

$$+ \frac{1}{n}\sum_{i=1}^{n} W_i \frac{Y_i(t) - \hat{\mathbb{E}}[Y_i(t)|X_i, W_i = 1]}{\hat{\mathbb{P}}[W_i = 1|X_i]} - (1 - W_i)\frac{Y_i(t) - \hat{\mathbb{E}}[Y_i(t)|X_i, W_i = 0]}{1 - \hat{\mathbb{P}}[W_i = 1|X_i]}$$

- (Non-)parametric estimators of nuisance functions:
  - $\hat{\mathbb{E}}[Y_i(t)|X_i, W_i = w]$
  - $\hat{\mathbb{P}}[W_i = 1|X_i]$

Properties and Requirements

Properties of $\hat{\tau}(t)$:

- Consistent for $\tau(t)$
- Doubly robust
- Semiparametric efficient

Requirements:

- Cross-fitting
- Nuisance functions are estimated at $o_P(n^{-1/4})$ rates

Construct Confidence Bands

Simultaneous Confidence Bands

- **To Show:**
  - Asymptotically Gaussian estimator of τ
  - Uniformly consistent estimator of its covariance kernel c (and its 1st and 2nd partial derivatives)
- Liebl and Reimherr (2022):
  - **Get:** Simultaneous and fair confidence bands
  - *Fairness:* Control balance of false-positive rate over $[0, 1]$

Theorem

Under *regularity conditions* on the continuity and differentiability of functions and distributions of the functional errors

$$\sqrt{n}(\hat{\tau} - \tau) \xrightarrow{d} \mathcal{GP}(0, c).$$

And, we can construct an estimator of c and its partial derivatives that is uniformly consistent.

Proof Outline

Proof Outline

- Define oracle estimator $\hat{\tau}^*$ that uses true nuisance functions $\mathbb{E}[Y_i(t)|X_i, W_i = w]$ and $\mathbb{P}[W_i = 1|X_i]$

Proof Outline

- Define oracle estimator $\hat{\tau}^*$ that uses true nuisance functions $\mathbb{E}[Y_i(t)|X_i, W_i = w]$ and $\mathbb{P}[W_i = 1|X_i]$
- Show that $\hat{\tau}^*$ is asymptotically Gaussian with kernel c
  - Structure of c can be derived from the structure of $\hat{\tau}^*$

Proof Outline

- Define oracle estimator $\hat{\tau}^*$ that uses true nuisance functions $\mathbb{E}[Y_i(t)|X_i, W_i = w]$ and $\mathbb{P}[W_i = 1|X_i]$
- Show that $\hat{\tau}^*$ is asymptotically Gaussian with kernel c
  - Structure of c can be derived from the structure of $\hat{\tau}^*$
- Show that $\sqrt{n}||\hat{\tau} - \hat{\tau}^*||_\infty \to_p 0$

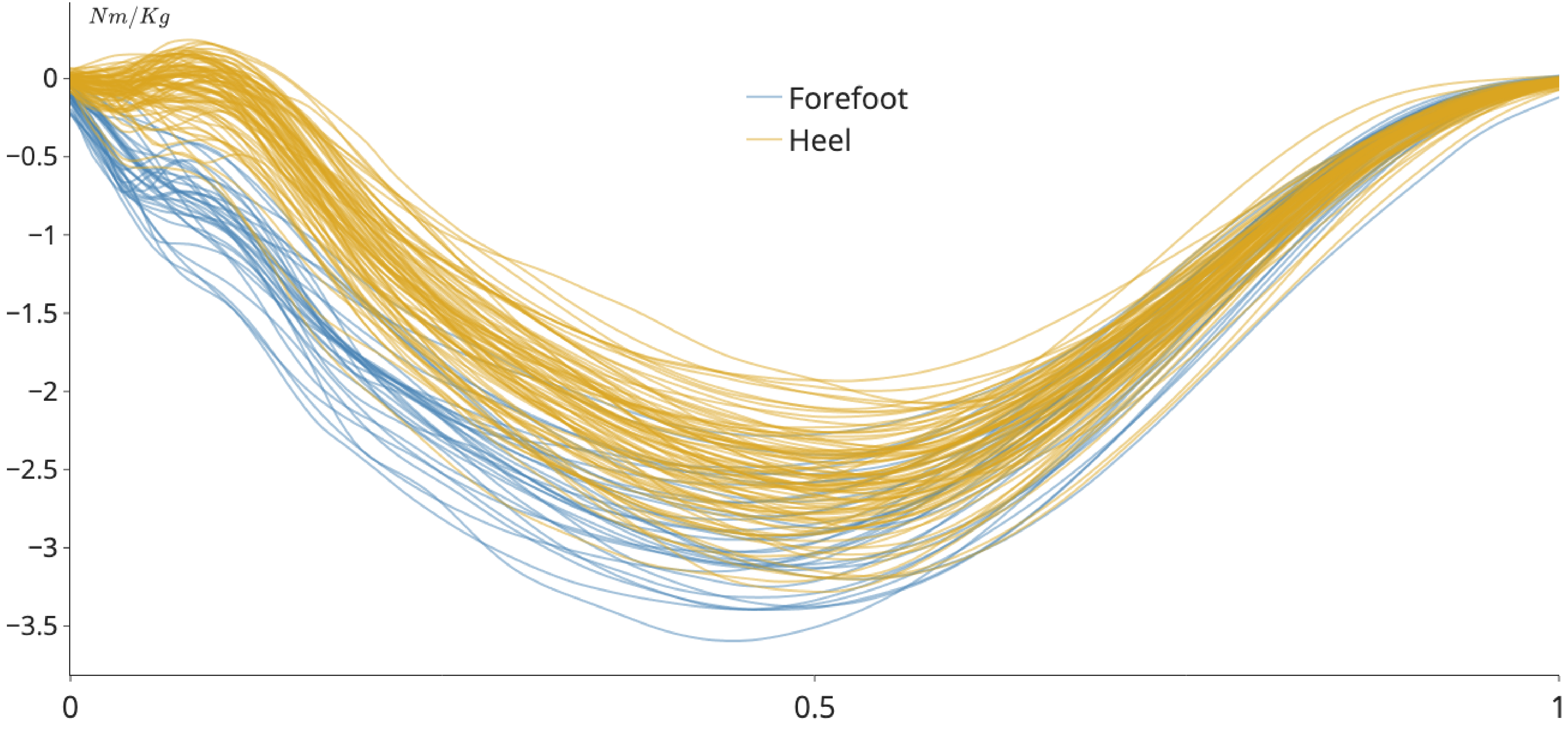
Proof Outline

- Define oracle estimator $\hat{\tau}^*$ that uses true nuisance functions $\mathbb{E}[Y_i(t)|X_i, W_i = w]$ and $\mathbb{P}[W_i = 1|X_i]$
- Show that $\hat{\tau}^*$ is asymptotically Gaussian with kernel c
  - Structure of c can be derived from the structure of $\hat{\tau}^*$
- Show that $\sqrt{n}||\hat{\tau} - \hat{\tau}^*||_\infty \to_p 0$
- Construct sample analogue estimator $\hat{c}$ of c

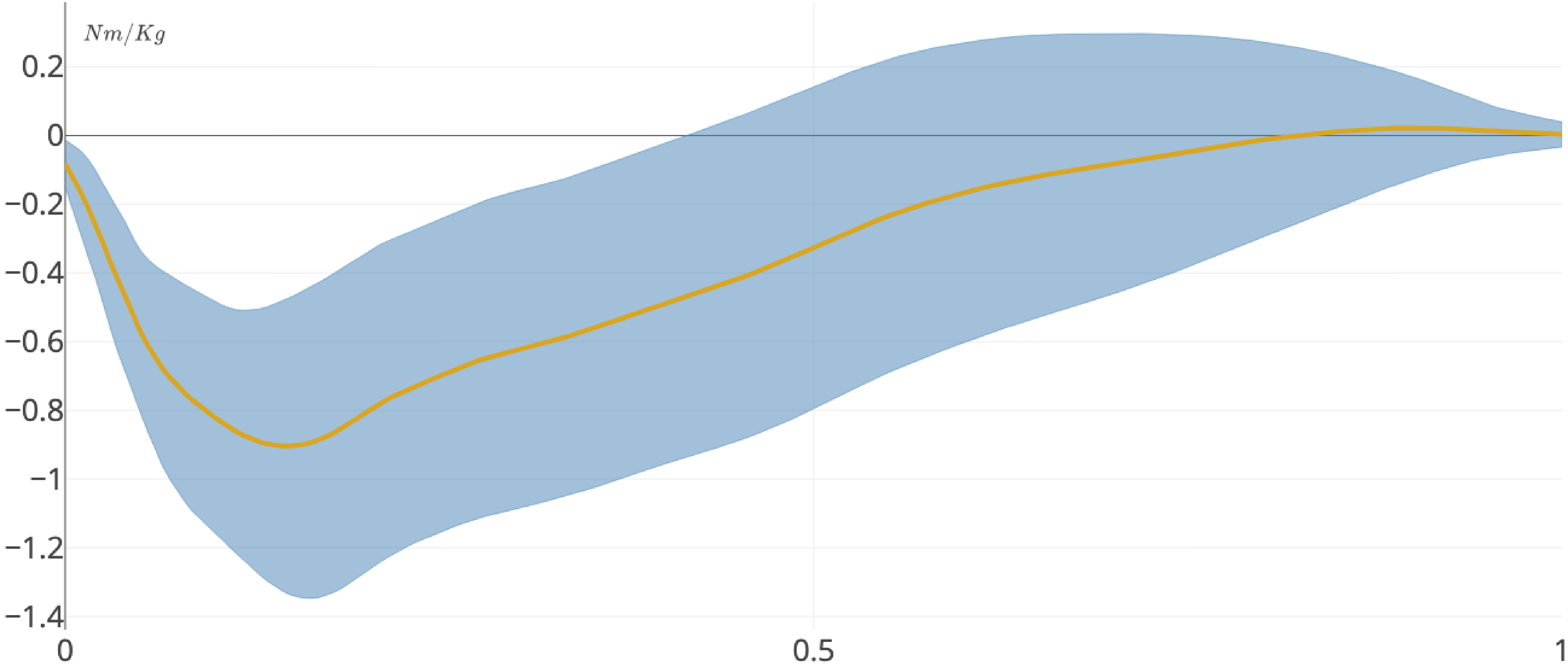
Proof Outline

- Define oracle estimator $\hat{\tau}^*$ that uses true nuisance functions $\mathbb{E}[Y_i(t)|X_i, W_i = w]$ and $\mathbb{P}[W_i = 1|X_i]$
- Show that $\hat{\tau}^*$ is asymptotically Gaussian with kernel c
  - Structure of c can be derived from the structure of $\hat{\tau}^*$
- Show that $\sqrt{n}||\hat{\tau} - \hat{\tau}^*||_\infty \to_p 0$
- Construct sample analogue estimator $\hat{c}$ of c
- Show that $\hat{c}$ and its derivatives converge uniformly

Application



Result



Contact

- Email: tmensinger@uni-bonn.de
- GitHub: [timmens/compstat]()
- Website: [tmensinger.com]()