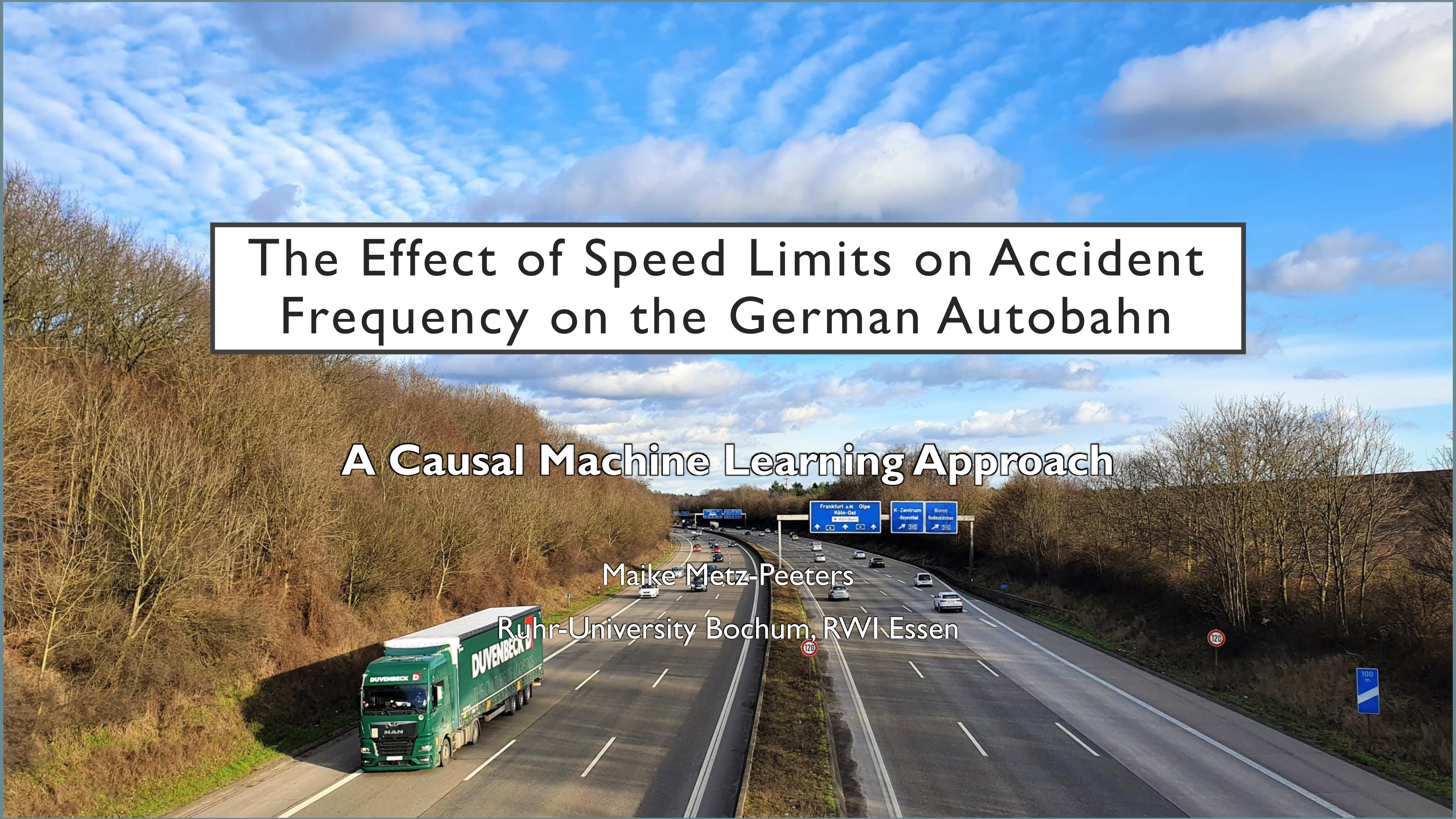


The Effect of Speed Limits on Accident Frequency on the German Autobahn

A Causal Machine Learning Approach

Maike Metz-Peeters

Ruhr-University Bochum, RWI Essen

Background

- No binding speed limit on ~70% of the German motorway network.
    - Vital public debate - Central question: **effects of speed limits on crash frequency**?
- Last large-scale experiment on 130km/h from1974-1976: -10% of all crashes, -20% of severely and fatally injured (BASt, 1977).
- Conducting new experiments politically not feasible.
- Since then, major changes in roadway design and vehicle capabilities etc., no comparable environment in other countries.

➢ Learn about relationship through observational study.

➢ Challenges:

- No readily available dataset → Combine various geospatial data sources.
- Unobserved heterogeneity → Introduce upper bound assumption.

➢ Causal forest: mild assumptions about data generating process (DGP) → spatial overfitting?

Data

- Combine various geospatial data sources:
  - **Federal statistical offices**: injury crash locations,
  - **Open Street Maps**: design & geometric characteristics of motorways,
  - **BASt**: automatic vehicle counting station data (interpolate), network shape, road surface condition measures
  - **NASA**: elevation profile
  - **DWD**: weather data
  - **BBSR**: regional socio-demographic characteristics

  Reduce to 4 principal components

  ➢ Include **spatial lags**.

➢ Fully automated with ArcGIS & Python.

➢ Almost all open data: open version of dataset (& code) planned.

➢ Dataset with characteristics of 500m segments for 13,000 km (50%) of network, 2017-2019.

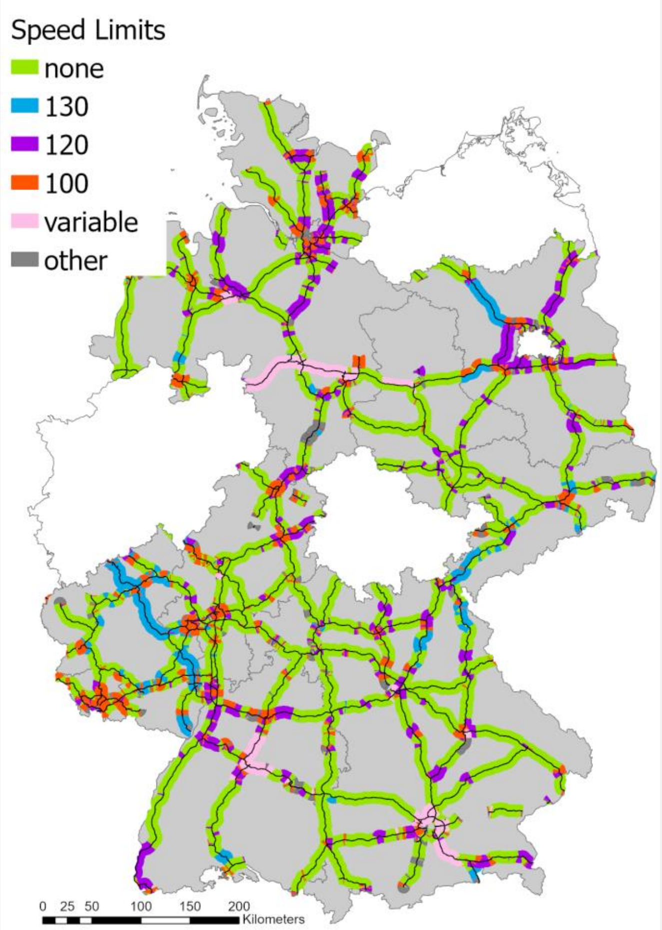


Figure 1: speed limits in 2019

Methodology: Causal Forest

(Athey, Tibshirani & Wager, 2019)

- Based on random forests (RF) by Breiman (2001) - Extension for treatment effect (TE) estimation.
- Ensemble of (de-correlated) regression or classification trees.
- ~ **Adaptive** form of **k-nearest neighbor matching**:
  - Compare segments with and without speed limits, similar w.r.t. relevant observed characteristics.
  - Let neighborhoods get wide along irrelevant dimensions → Overcome *curse of dimensionality.*

- **Potential outcomes framework**: requires *conditional independence assumption* – feasibility?

➢ Instead: **upper bound assumption →** *Uncontrolled factors increasing (decreasing) crash frequency also increase (decrease) speed limit probability*.

  ➢ Estimated effects upward biased: negative upper bound ⇒ larger effects in absolute value.

- Estimates individual treatment effects → Learn about effect heterogeneity.
- Not ideal for strong smooth signals → Use crash rates instead of counts as outcomes.

Methodology: Spatial Prediction - Background

- Robust TE estimation employs *propensity score* & *main effect function* estimates conditional on covariates.
- Default: separate RF for treatment probability & outcome, out-of-bag predictions for each observation.
- Problem: treatment, outcomes & many covariates strongly spatially auto-correlated → **Spatial over-fitting**?
- Solutions (Meyer et al., 2018)?:
  - **Leave-location-out cross validation** (LLO-CV): whole locations into CV-fold, evaluate on *unknown locations*.
  - **Forward-Feature Selection**: Start with all possible 2-variable models, choose best one in LLO-CV, recursively add variable most improving fit on unknown locations, stop if no improvements possible.

Methodology: Spatial Prediction - Experiment

- Estimate following setups:
  a. Random CV with all variables,
  b. Spatial CV with all variables, and
  c. Spatial CV with FFS.

  - LLO-CV improves out-of-location fit for 120 & 130 km/h, not for 100 km/h
  - FFS does not lead to any improvements.

	Oob error			Validation error		
	a	b	c	a	b	c
100	0.037	0.053	0.052	0.049	0.051	0.051
120	0.048	0.12	0.118	0.128	0.125	0.126
130	0.018	0.073	0.072	0.061	0.056	0.057

Table 1: Cross validation results for propensity score estimates: **mean squared error evaluated on** hold-out set with only **unseen location**.

*Note: **cluster robust causal forests** implement LLO-CV.*

Methodology: Spatial Prediction - Reasons

- Random CV performs well at pointing to restricted segments, but not generalizable!
    - ➢ Does not capture important features of DGP.
    - ➢ Spatial auto-correlation causes "leak" in the algorithm: Limit on a segment increases probability of limit on neighboring segment (political reasons, traffic planning, etc.).
- Possible reasons for no improvement for 100 km/h:
    - Less spatial clustering, stronger signal.
- For main effect, no improvements through spatial CV & FFS
    - ➢ Spatial clustering in crash rates only through covariates?

➢ ML methods on spatial data require explicit consideration of spatial nature, especially when:

- Outcome in one location affects outcome in neighboring location **and**
- **Aim** is estimating generalizable function.

➢ Spatial CV implemented with cluster robust causal forests.

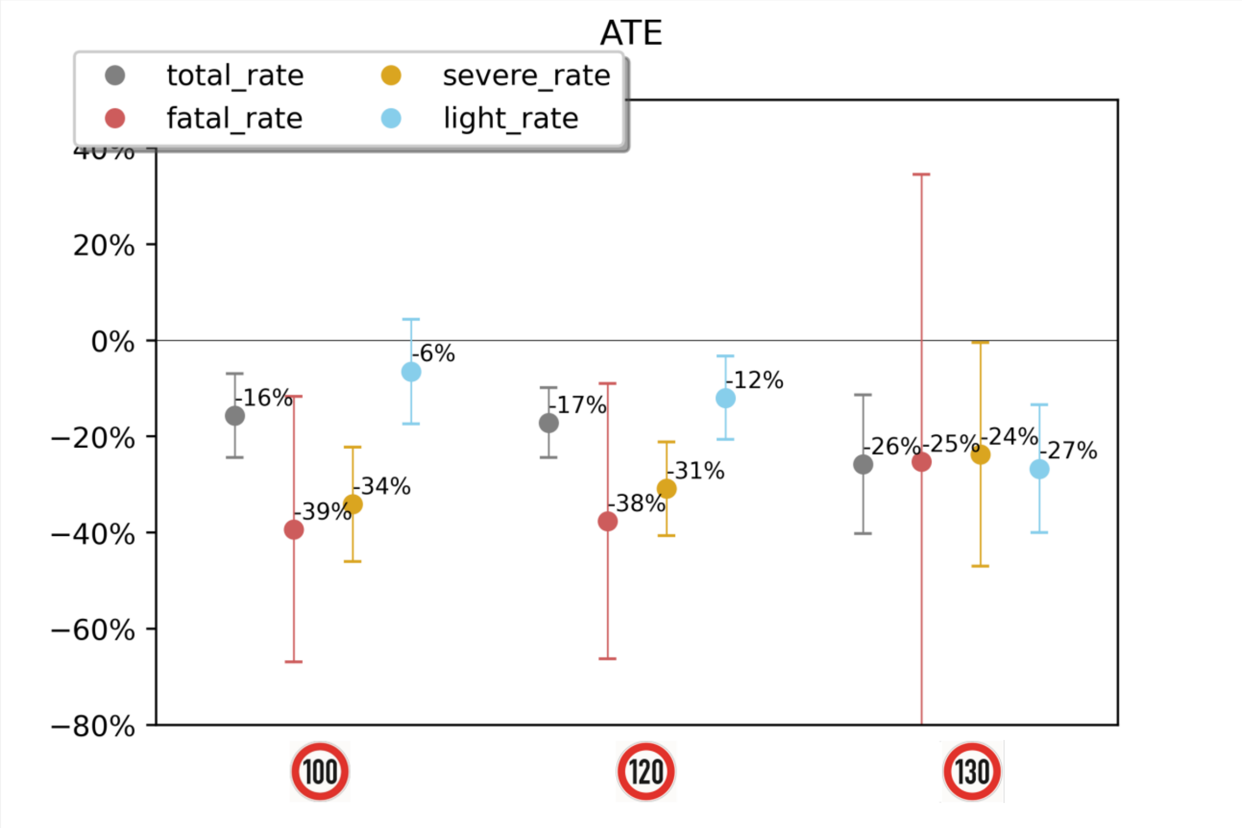
Results

- Large and strongly significant effects of all limits on total crash rate.
- Larger effects on fatal and also on severe crash rate.
- Larger effects on at least severe crashes for more restrictive limits.

Note:

- Largest unobserved heterogeneity for 100 km/h
- Strong spatial clustering of 130 km/h limit → Enough variance in X to capture main features of DGP?

➢ Most reliable results: 120 km/h



N =	21,066	23,038	21,259
$\overline{W}$ =	0.066	0.146	0.074

Figure 2: ATEs estimated with cluster robust causal forests, with 15,000 trees.

Effect Heterogeneity

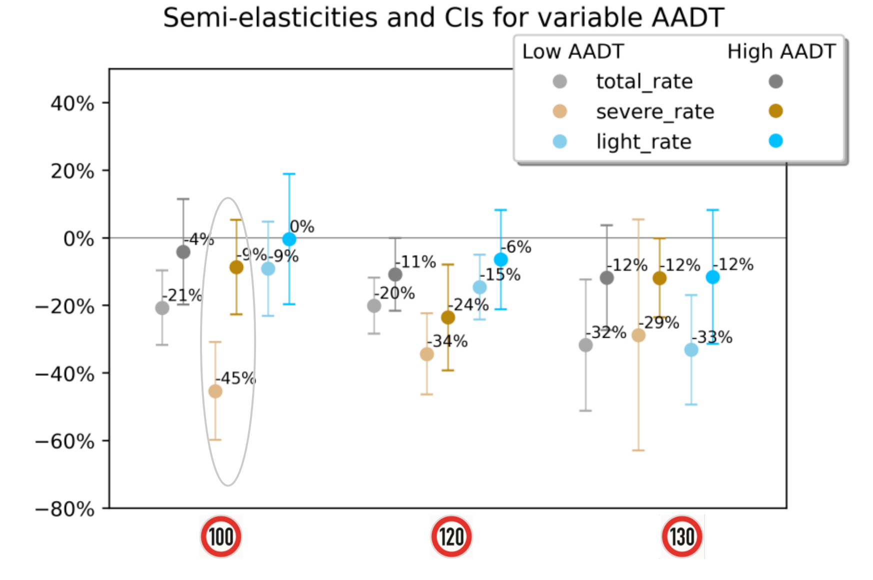


Figure 3: CATEs for segments with low and high AADT

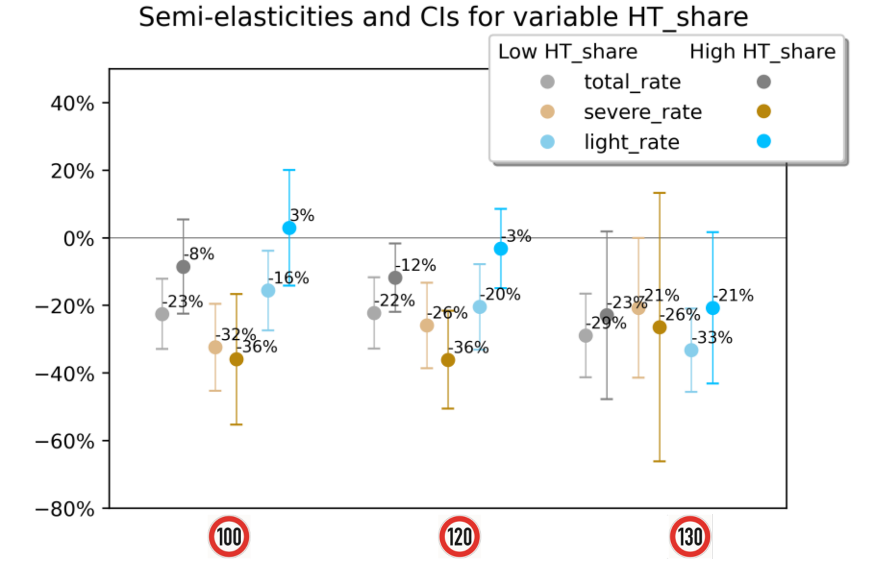


Figure 5: CATEs for segments with low and high shares of heavy traffic

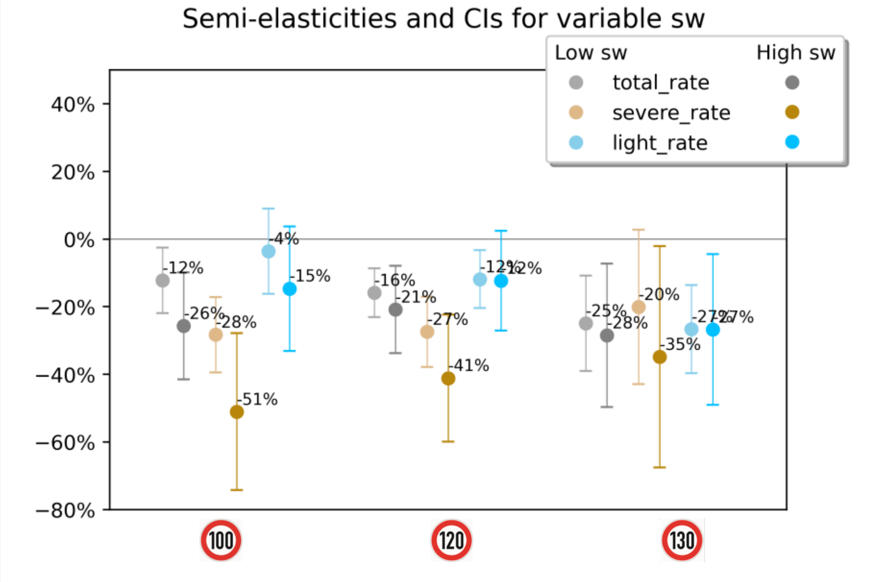


Figure 4: CATEs for segments with and without access and exit ramps.

- Larger effects on segments with larger speed variance.
- Note: Weak statistical significance of differences.

Thank you for your attention!

Contact: Maike.Metz-Peeters@rub.de

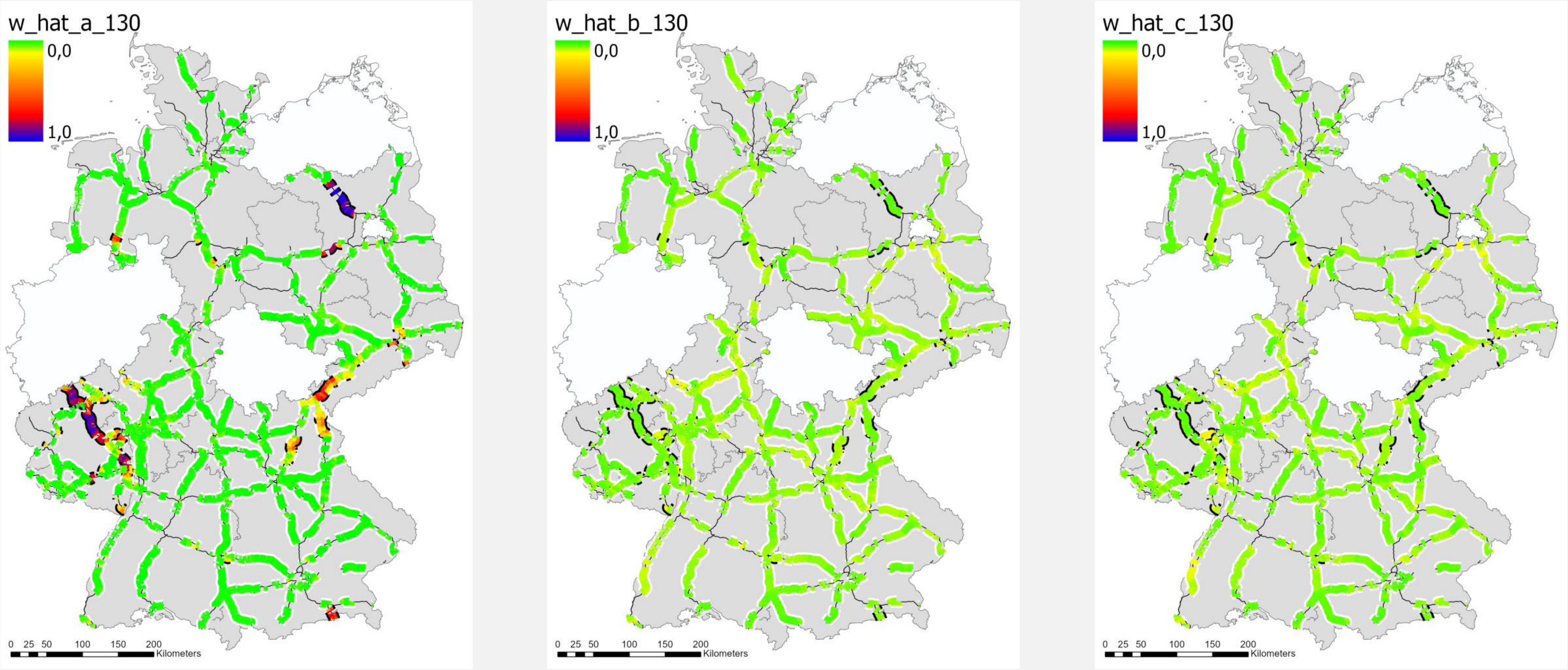


Figure 6, a-c: Propensity scores for speed limit of 130, estimated according to setup a, b, and c. Actually restricted segments displayed with a black outline.

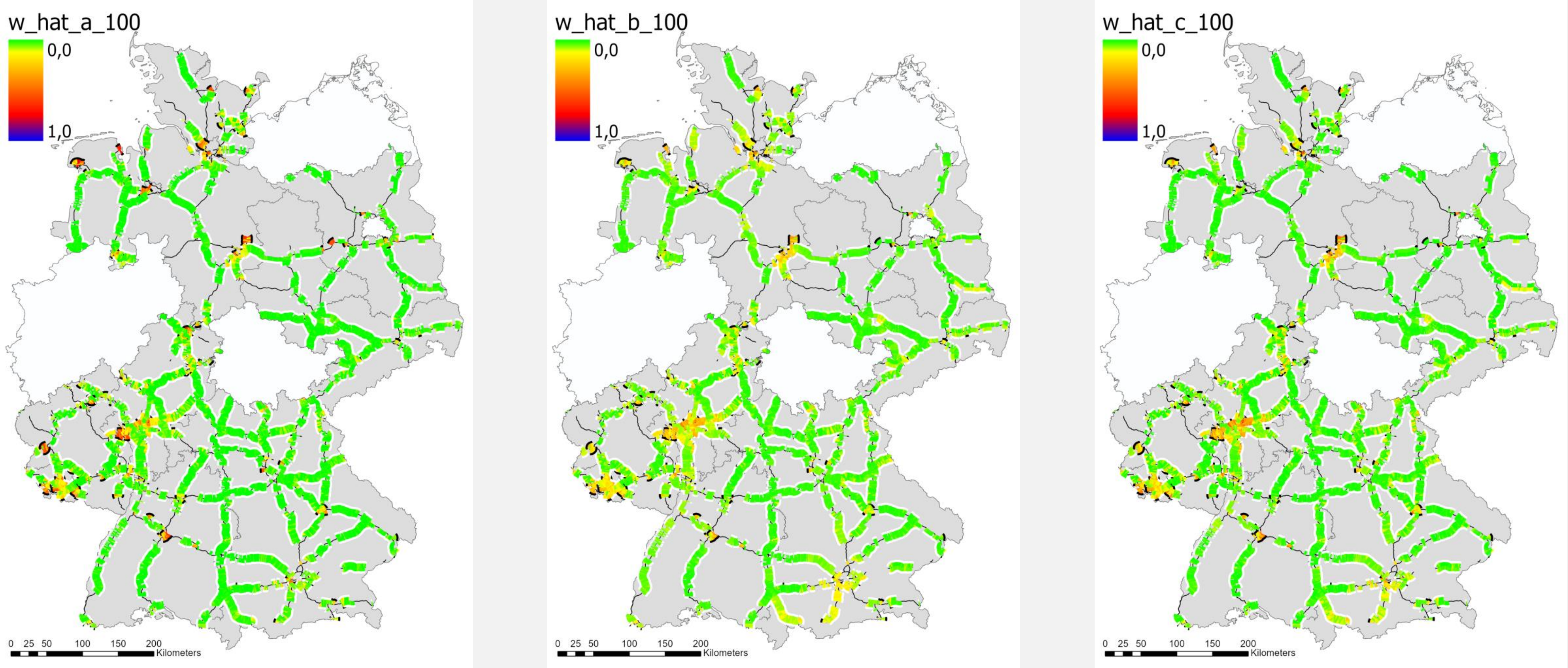


Figure 7, a-c: Propensity scores for speed limit of 100, estimated according to setup a, b, and c.
Actually restricted segments displayed with a black outline.