

Modeling Monotonic Effects of Ordinal Predictors in Regression Models

Paul Bürkner & Emmanuel Charpentier

Linear Regression

Assume that the predictor term η is a linear combination of the predictor variables multiplied by the regression coefficients:

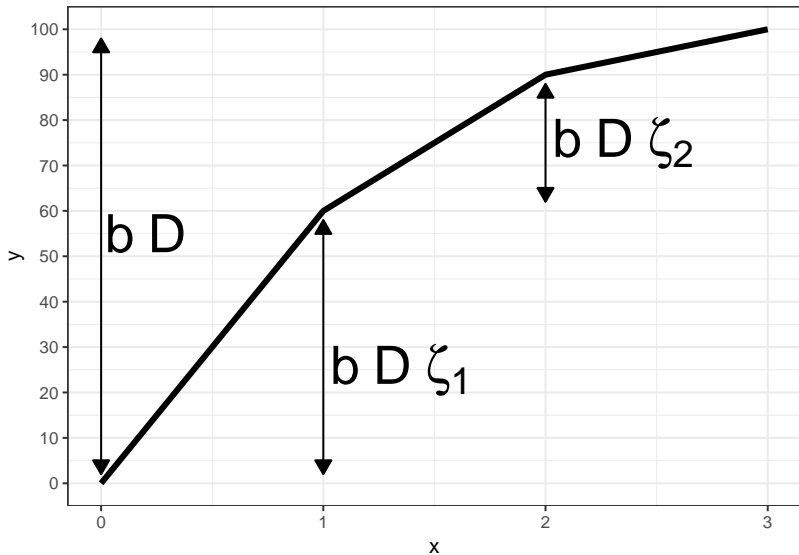
$$\eta = b_0 + \sum_{k=1}^K b_k x_k$$

Predictors x_k may be

- continuous variables
- coded categorical variables

What about ordinal predictors?

Monotonic Effects: Idea



Monotonic Effects: Mathematical Formulation

Monotonic regression of an ordinal predictor $x \in \{0, \dots, D\}$:

$$\eta = b_0 + bD \sum_{i=1}^x \zeta_i$$

- Parameter ζ is a simplex: $\zeta_i \in [0, 1]$ and $\sum_{i=1}^D \zeta_i = 1$
- Parameter b may be any real value

Define the monotonic transform:

$$\text{mo}(x, \zeta) = D \sum_{i=1}^x \zeta_i$$

Monotonic Effects: Interactions

Ordinary Regression model including the interaction of z and x :

$$\eta = b_0 + b_1 z + b_2 x + b_3 z x$$

Generalize to monotonic effects by replacing x with $\text{mo}(x, \zeta)$:

$$\eta = b_0 + b_1 z + b_2 \text{mo}(x, \zeta_{b_2}) + b_3 z \text{mo}(x, \zeta_{b_3})$$

Monotonic Effects in a Bayesian Framework

“If you quantify uncertainty with probability you are a Bayesian.”

Micheal Betancourt

Bayes Theorem:

$$p(\theta | y) = \frac{p(y | \theta) p(\theta)}{p(y)}$$

The monotonic parameters b and ζ are both part of θ

Priors for Monotonic Effects in a Bayesian Framework

Priors on b :

- Any reasonable prior for regression coefficients
- For instance: $b \sim \mathcal{N}(0, s)$ for a fixed standard deviation s

Prior on ζ :

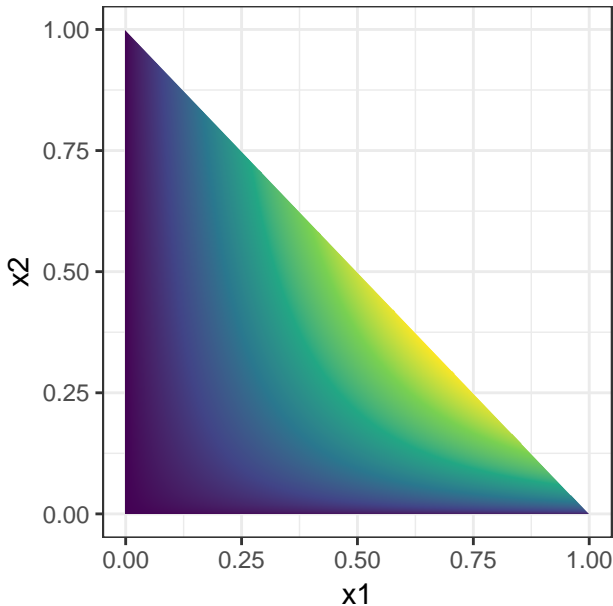
- Dirichlet prior: $\zeta \sim \mathcal{D}(\alpha)$
- α : Concentration parameter of the same length as ζ

Let $\alpha_0 = \sum_{i=1}^D \alpha_i$, then:

$$\mathbb{E}(\zeta_i) = \frac{\alpha_i}{\alpha_0}$$

$$\text{SD}(\zeta_i) = \sqrt{\frac{\alpha_i(\alpha_0 - \alpha_i)}{(\alpha_0^2(\alpha_0 + 1))}}$$

Dirichlet Prior: Visualization for $\alpha = (2, 1.5, 1)$



Monotonic effects in the R package brms

Monotonic effect of x on y :

```
y ~ mo(x)
```

Main effects and interaction of x and z :

```
y ~ mo(x) * z
```

Varying effect of x over group g :

```
y ~ mo(x) + (mo(x) | g)
```

Case Study: Measures of Chronic Widespread Pain (CWP)

Objective: Predict subjective physical health by measures of CWP

Examples for CWP measures:

- Impairments in walking
- Impairments in moving around

Scale from 0 ('no problem') to 4 ('complete problem')

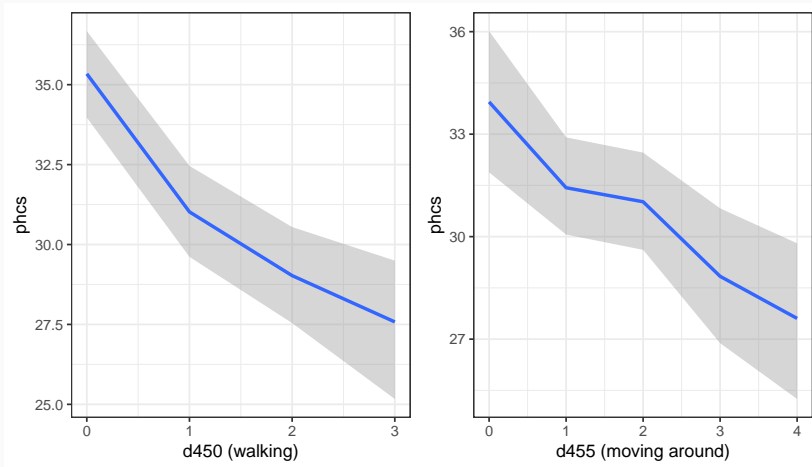
Reference: Gertheiss, Hogger, Oberhauser, & Tutz (2011)

Plausible assumption: CWP measures have **monotonic effects**

Case Study: Model Specification

```
library(brms)
```

```
fit1 <- brm(phcs ~ mo(d450) + mo(d455), data = cwp)
```



Parameter Recovery

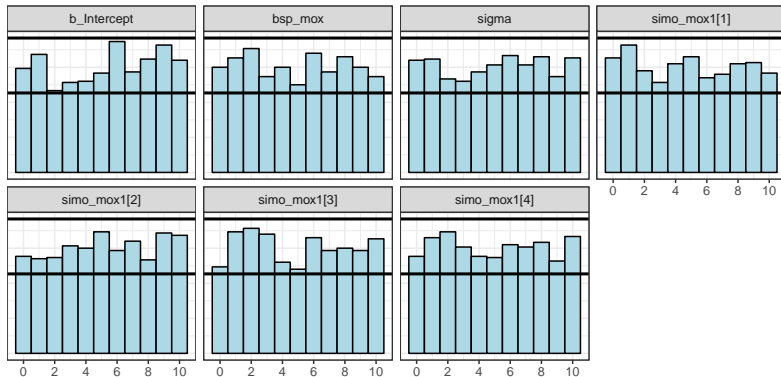
Question: How well can a model recover its own parameters?

Steps of Simulation Based Calibration (SBC):

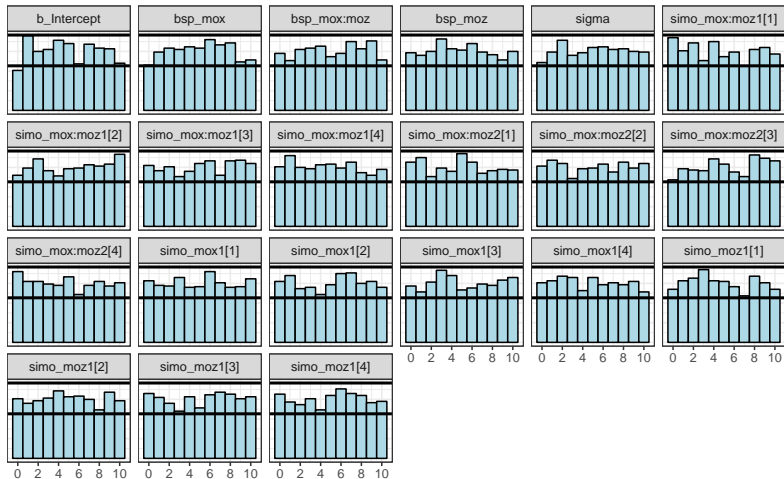
- Sample $\tilde{\theta} \sim p(\theta)$ from the prior
- Sample $\tilde{y} \sim p(y | \tilde{\theta})$ from the likelihood
- Sample $\{\theta_1, \dots, \theta_L\} \sim p(\theta | \tilde{y})$ from the posterior
- Compute the rank statistic $r(\{\theta_1, \dots, \theta_L\} | \tilde{\theta})$
- Repeat the process multiple times
- Plot the rank statistics in a histogram

For well calibrated models the histogram is (approximately) uniform

Parameter Recovery: Monotonic Main Effects



Parameter Recovery: Monotonic Interactions



Other Approaches for Modelling Ordinal Predictors

Categorical isotonic regression:

- Estimate group means of ordinal categories such that $\mu_0 < \mu_1 < \dots < \mu_C$
- Equivalent to monotonic effects in simple cases
- Harder to penalize via priors

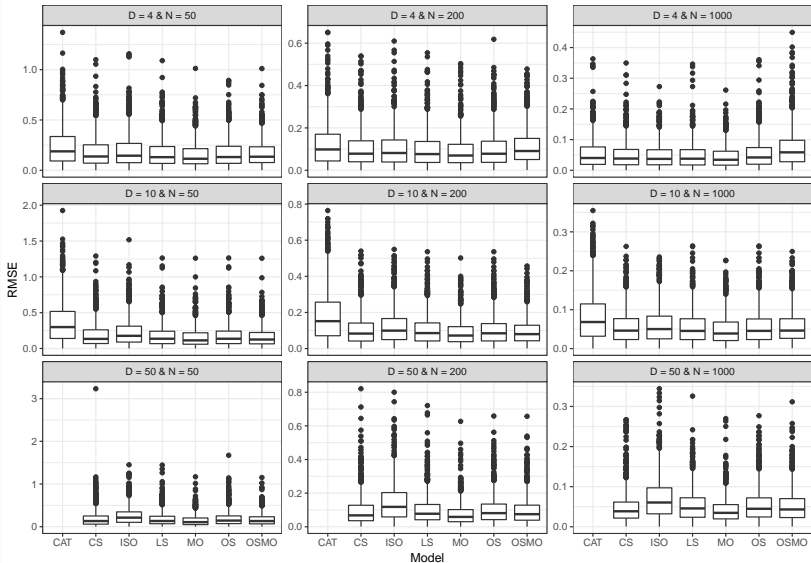
Penalized regression (Gertheiss & Tutz, 2009):

- Apply dummy coding on the ordinal variable
- Penalize larger differences between adjacent categories via

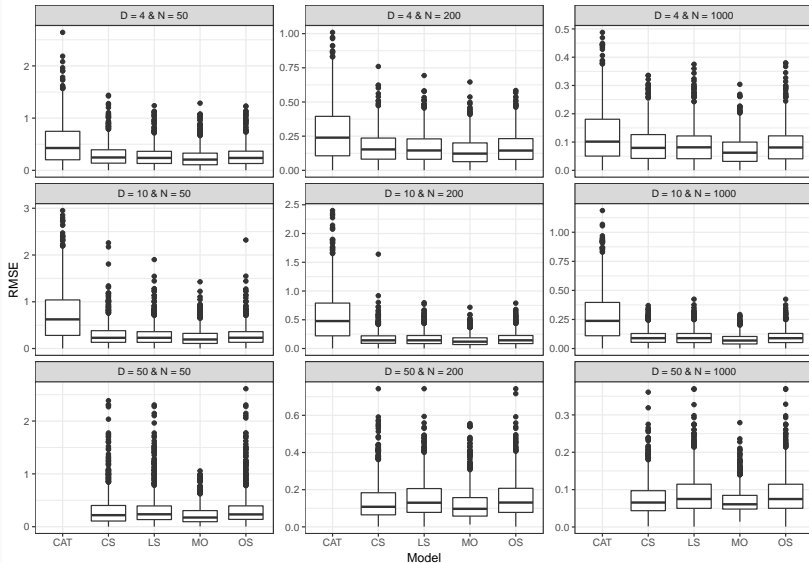
$$J(b) = \sum_{i=1}^D (b_i - b_{i-1})^2$$

- Closely related to regression splines
- No monotonicity constraint

Model Comparison: Monotonic Main Effect



Model Comparison: Monotonic Interactions



Thank you!

References

Bürkner, P., & Charpentier, E. (in review). Modeling Monotonic Effects of Ordinal Predictors in Regression Models.

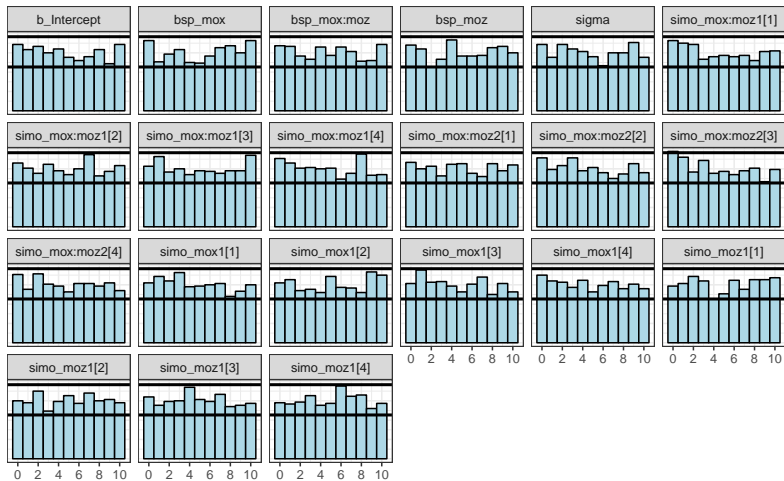
<https://psyarxiv.com/9qkhj/>

Gertheiss, J., Hogger, S., Oberhauser, C., & Tutz, G. (2011). Selection of ordinally scaled independent variables with applications to international classification of functioning core sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 60(3), 377–395.

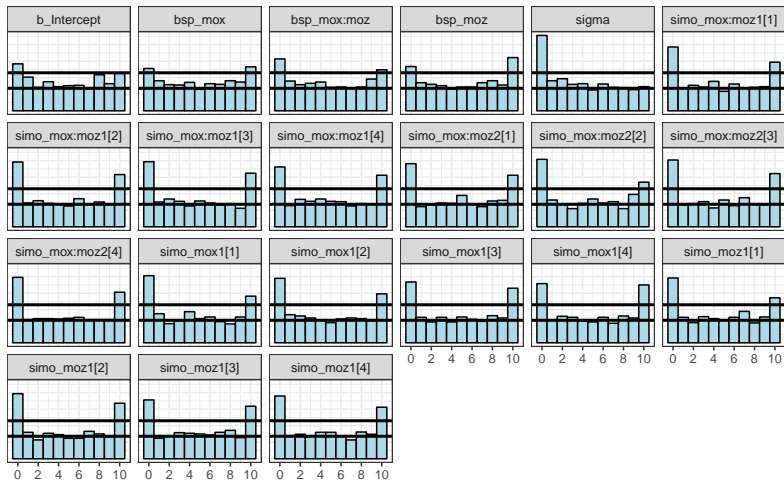
Gertheiss, J., & Tutz, G. (2009). Penalized regression with ordinal predictors. *International Statistical Review*, 77(3), 345–365.

Appendix

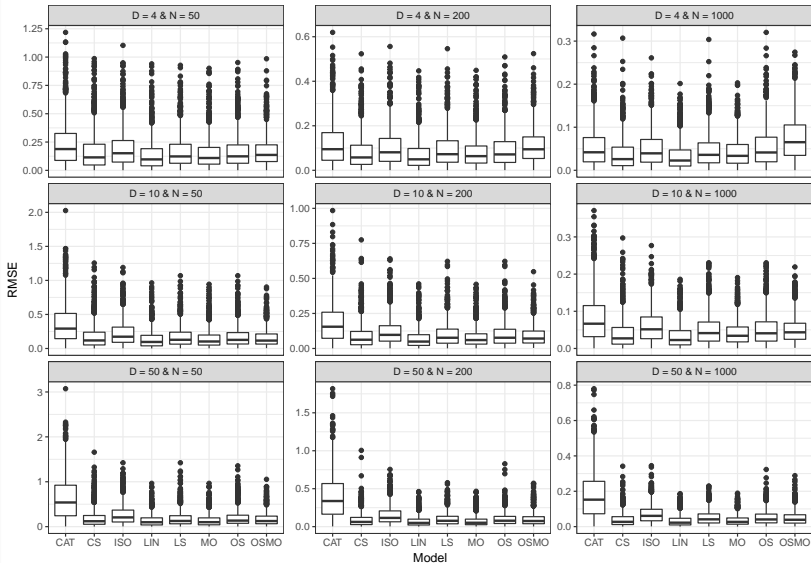
Parameter Recovery: Interactions with many Predictor Categories (1)



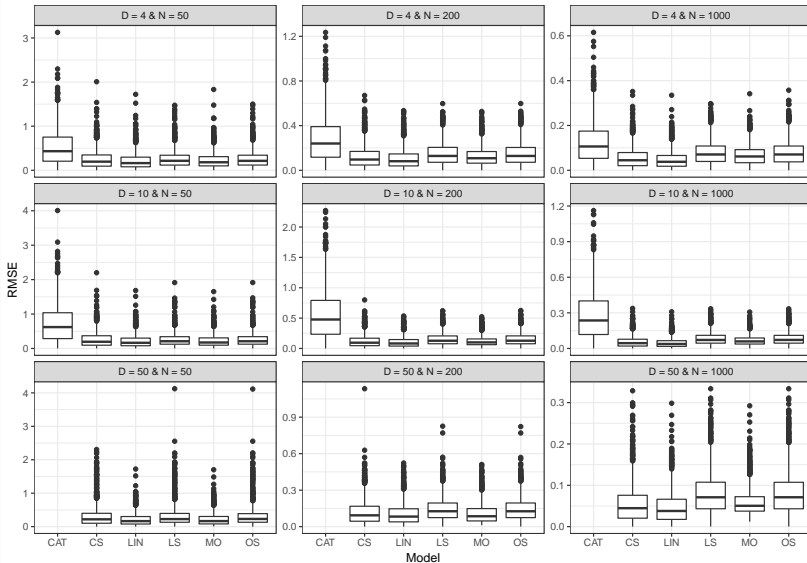
Parameter Recovery: Interactions with many Predictor Categories (2)



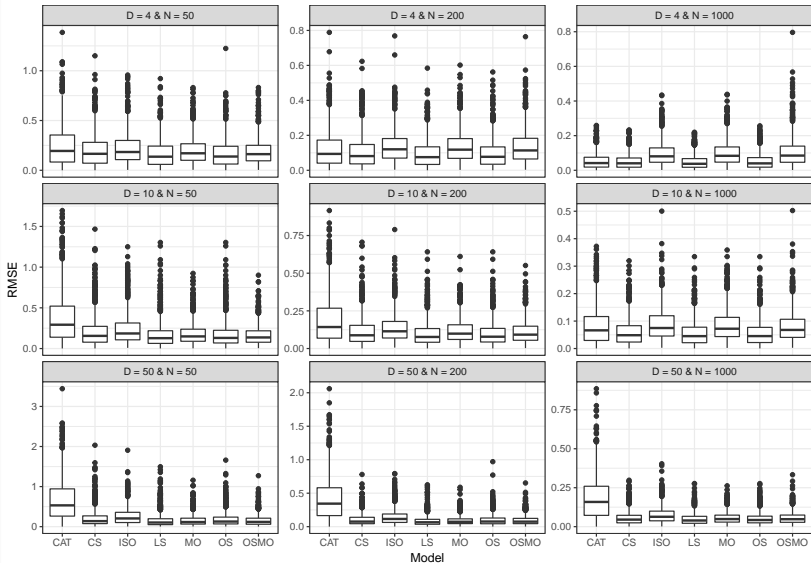
Model Comparison: Linear Main Effects



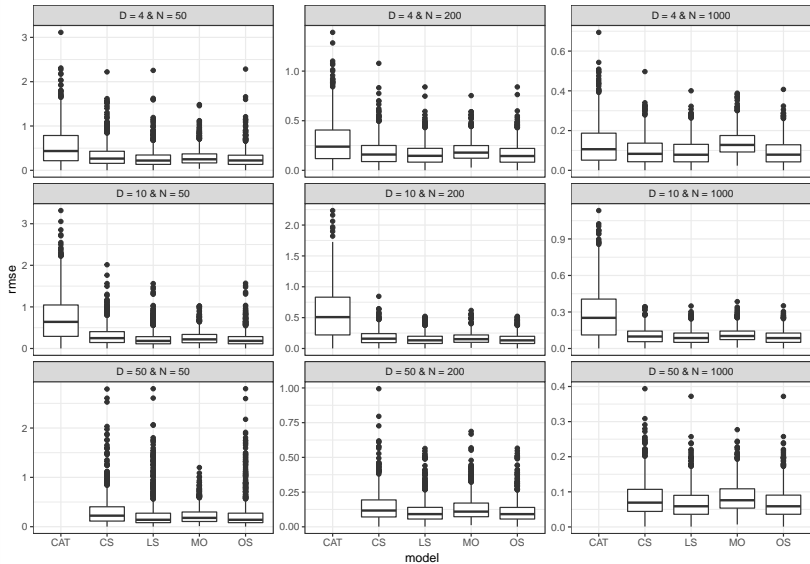
Model Comparison: Linear Interactions



Model Comparison: Categorical Main Effects



Model Comparison: Categorical Interactions



Counter Example to the Conditional Monotonicity

Model: $\eta = b_0 + b_1 z + b_2 \text{mo}(x, \zeta_{b_2}) + b_3 z \text{mo}(x, \zeta_{b_3})$

