Faculty of
Economics and Statistics

Modeling multiplicative interaction effects in Gaussian structured additive regression models

Philipp Aschersleben, Julian Granna, Thomas Kneib, Stefan Lang, Nikolaus Umlauf, and Winfried Steiner

Working Papers in Economics and Statistics

2024-01

University of Innsbruck
https://www.uibk.ac.at/EconStat/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
Faculty of Economics and Statistics
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 96136
E-mail: Dean-EconStat@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/fakultaeten/volkswirtschaft_und_statistik/forschung/wopec/

For a list of recent papers see the backpages of this paper.

Modeling multiplicative interaction effects in Gaussian structured additive regression models

Philipp Aschersleben[1], Julian Granna[2], Thomas Kneib[3], Stefan Lang[4], Nikolaus Umlauf[5], and Winfried Steiner[6]

[1] *Clausthal University of Technology, Julius-Albert-Straße 2, 38678 Clausthal-Zellerfeld, Germany. e-mail: philipp.aschersleben@tu-clausthal.de*
[2] *University of Innsbruck, Universitätsstr. 15, 6020 Innsbruck, Austria. e-mail: julian.granna@posteo.org*
[3] *University of Göttingen, Humboldtallee 3 ,37073 Göttingen, Germany, e-mail: tkneib@uni-goettingen.de*
[4] *University of Innsbruck, Universitätsstr. 15, 6020 Innsbruck, Austria. e-mail: stefan.lang@uibk.ac.at*
[5] *University of Innsbruck, Universitätsstr. 15, 6020 Innsbruck, Austria. e-mail: nikolaus.umlauf@uibk.ac.at*
[6] *Clausthal University of Technology, Julius-Albert-Straße 2, 38678 Clausthal-Zellerfeld, Germany. e-mail: winfried.steiner@tu-clausthal.de*

February 23, 2024

Abstract

Gaussian Structured Additive Regression provides a flexible framework for additive decomposition of the expected value with nonlinear covariate effects and time trends, unit- or cluster-specific heterogeneity, spatial heterogeneity, and complex interactions between covariates of different types. Within this framework, we present a simultaneous estimation approach for highly complex multiplicative interaction effects. In particular, a possibly nonlinear function $f(z)$ of a covariate z may be scaled by a multiplicative effect of the form $exp(\tilde{\eta})$, where $\tilde{\eta}$ is another possibly structured additive predictor. Inference is fully Bayesian and based on highly efficient Markov Chain Monte Carlo (MCMC) algorithms. We investigate the statistical properties of our approach in extensive simulation experiments. Furthermore, we apply and illustrate the methodology to an analysis of asking prices for 200000 dwellings in Germany.

Keywords: IWLS proposals, MCMC, multiplicative interaction effects, structured additive predictor

1 Introduction

Structured additive predictor models (e.g. Fahrmeir et al., 2022) provide a very broad and rich framework for complex regression modelling. They can deal simultaneously with nonlinear covariate effects and time trends, unit- or cluster-specific heterogeneity, spatial heterogeneity, and complex interactions between covariates of different types.

This paper deals with interactions between a continuous covariate, say z, and other covariates of potentially different types, see e.g. Wood (2017) for an overview of standard approaches for interaction modelling. Here, we assume that the possibly nonlinear effect $f(z)$ of z is scaled by a scaling factor $\exp(\tilde{\eta})$, where $\tilde{\eta}$ is another structured additive predictor such that $f(z)$ may be scaled in a rather complex way. The rationale behind this type of modelling is the assumption of a homogeneous functional form across observations, but possibly heterogeneous functional size with respect to the interacting covariates in the scaling term. A typical application is clustered data where we assume a homogeneous functional form $f(z)$ across all observations but heterogeneous effect size across clusters, see for example Razen and Lang (2020) who model residential prices in Germany where nonlinear covariate effects are scaled according to regional clusters. A key aspect of hedonic real estate modelling of house or apartment prices is the appropriate estimation of the time trend. This is particularly relevant, for example, in the context of hedonic index construction, which requires unbiased underlying estimates (e.g. Granna, Brunauer, and Lang, 2022). A challenge in this context is that the effect of time may not be homogeneous across locations, but also with respect to other covariates in the model. We therefore provide in the application Section 4 of our paper a careful analysis of apartment prices where the possibly nonlinear time trend is not assumed to be homogeneous across observations, but is scaled with respect to the other covariates (e.g. location, area, age of the building) in the model.

This paper stands in the tradition of a number of related papers, in particular Brunauer et al. (2010), Lang et al. (2015) and Razen and Lang (2020). These papers deal with cluster-specific interaction terms of the form $(1 + \alpha_c)f(z)$, where c is a cluster index and $1 + \alpha_c$ is a cluster-specific scaling factor that scales the nonlinear function $f(z)$. Brunauer et al. (2010) and Lang et al. (2015) apply a heuristic two-stage procedure to estimate $f(z)$ and the scaling factor $1+\alpha_c$, while Razen and Lang (2020) provide a simultaneous estimation of the nonlinear function and the cluster-specific scaling factor based on Markov chain Monte Carlo (MCMC) simulation techniques. Compared to the existing literature, the approach of this paper has two distinct improvements:

- First, the exponential scaling introduced in this paper allows us, for the first time, to ensure strictly positive scaling factors. In all previous approaches, negative scaling factors were possible but had no meaningful interpretation.
- Second, the scaling term allows for a full and unconstrained structured additive predictor, possibly composed of all available covariates, and is therefore not limited

to simple cluster-specific heterogeneity terms.

The remainder of the paper is organised as follows: Section 2 introduces the statistical model and provides details on Bayesian inference based on Markov chain Monte Carlo simulation techniques. simulation techniques. Section 3 presents results from extensive simulation experiments, while Section 4 is devoted to our case study of apartment prices described above. The final Section 5 summarises the paper and discusses some ideas for future research.

2 Methodology

2.1 Gaussian structured additive regression

Suppose we are given data on n observations in the form $(y_i, \boldsymbol{z}_i, \boldsymbol{x}_i)$, $i = 1, \ldots, n$, with a continuous response y and a number of covariates $\boldsymbol{z}$ and $\boldsymbol{x}$. Bayesian structured additive regression, as outlined e.g. in Fahrmeir et al. (2022), assumes a Gaussian distribution of the response y, given the covariates, with expected value η and (homoscedastic) variance σ^2. The expected value is defined in terms of the covariates as

$$E(y|\boldsymbol{z}, \boldsymbol{x}) = \eta = f_1(\boldsymbol{z}_1) + \ldots + f_q(\boldsymbol{z}_q) + \boldsymbol{x}^\top \boldsymbol{\gamma}, \tag{1}$$

where the functions f_j, $j = 1, \ldots, q$, are nonlinear functions of the (possibly multivariate) covariates $\boldsymbol{z}_j$ and the term $\boldsymbol{x}^\top \boldsymbol{\gamma}$ comprises the linear effects of the model.

Using known basis functions B_k, a particular effect f can be approximated by

$$f(\boldsymbol{z}) = \sum_{k=1}^{K} \beta_k B_k(\boldsymbol{z}),$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)'$ is a vector of unknown regression coefficients to be estimated. A standard choice for continuous covariates are B-spline basis functions, see below.

Defining the $n \times K$ design matrix $\boldsymbol{Z}$ with elements $\boldsymbol{Z}[i, k] = B_k(\boldsymbol{z}_i)$, the vector $\boldsymbol{f} = (f(\boldsymbol{z}_1), \ldots, f(\boldsymbol{z}_n))'$ of function evaluations can be written in matrix notation as $\boldsymbol{f} = \boldsymbol{Z}\boldsymbol{\beta}$. Accordingly, the predictors in (1) can be written as

$$\boldsymbol{\eta} = \boldsymbol{Z}_1 \boldsymbol{\beta}_1 + \ldots + \boldsymbol{Z}_q \boldsymbol{\beta}_q + \boldsymbol{X}\boldsymbol{\gamma}.$$

In a Bayesian framework, overfitting of a particular function f is usually avoided by employing a suitable smoothness prior for the regression coefficients $\boldsymbol{\beta}$, see e.g. Fahrmeir et al. (2022). A standard choice is a (possibly improper) Gaussian prior of the form

$$p(\boldsymbol{\beta}|\tau^2) \propto \left(\frac{1}{\tau^2}\right)^{\mathrm{rk}(\boldsymbol{K})/2} \exp\left(-\frac{1}{2\tau^2} \boldsymbol{\beta}^\top \boldsymbol{K} \boldsymbol{\beta}\right) \cdot I(\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{0}), \tag{2}$$

where $I(\cdot)$ is the indicator function. The key components of the prior are the penalty matrix $\boldsymbol{K}$, the variance parameter τ^2 and the constraint $\boldsymbol{A\beta} = \boldsymbol{0}$. Usually the penalty matrix is rank deficient, i.e. $\text{rk}(\boldsymbol{K}) < K$, resulting in a partially improper prior. The specific structure of $\boldsymbol{K}$ depends on the covariate type and on prior assumptions about the smoothness of f.

We apply, for example, a Bayesian version of P-splines when modeling a smooth function f that depends on a continuous covariate z, see Lang and Brezger (2004). Here, the columns of the design matrix $\boldsymbol{Z}$ are given by B-spline basis functions evaluated at the observations z_i and we use first or second order random walks as smoothness priors for the regression coefficients, i.e. $\beta_k = \beta_{k-1} + u_k$, or $\beta_k = 2\beta_{k-1} - \beta_{k-2} + u_k$, with Gaussian errors $u_k \sim N(0, \tau^2)$ and diffuse priors $p(\beta_1) \propto const$, or $p(\beta_1)$ and $p(\beta_2) \propto const$, for initial values. This prior is of the form (2) with penalty matrix given by $\boldsymbol{K} = \boldsymbol{D}^\top \boldsymbol{D}$, where $\boldsymbol{D}$ is a first or second order difference matrix.

Further examples regarding priors for f, in particular priors for spatial covariates, cluster indicators or for modeling interactions are outlined e.g. in Fahrmeir et al. (2022).

The amount of smoothness is governed by the variance parameter τ^2. A conjugate inverse Gamma prior is employed for τ^2, i.e. $\tau^2 \sim IG(a, b)$ with small values for the hyperparameters a and b resulting in an uninformative prior on the log scale. As a default we choose $a = b = 0.001$.

The term $I(\boldsymbol{A\beta} = \boldsymbol{0})$ imposes required identifiability constraints on the parameter vector. A straightforward choice is $\boldsymbol{A} = (1, \ldots, 1)$, i.e. the regression coefficients are centered around zero.

2.2 Multiplicative interaction effects

As outlined in the introduction, in Section 4 of this paper we analyse real estate data of apartment prices as a function of explanatory variables. Particular emphasis is placed on the appropriate modelling of the time trend. More specifically, we will assume a homogeneous functional form of the trend across observations, but heterogeneity with respect to the size of the effect, i.e. we assume that the trend is scaled in a complex way, potentially depending on the other available covariates (e.g. location, area, age of building) in the model.

Indeed, in many applications it is reasonable to assume the effect of covariates to have the same functional form but to vary with respect to the overall size of the function. To account for this type of interaction effect, we allow covariate dependent scaling factors for some or all of the nonlinear functions f_j in (1). This results in predictors of the form

$$\eta = \ldots + \exp(\widetilde{\eta}_j)\, f_j(z_j) + \ldots, \tag{3}$$

where $\exp(\widetilde{\eta}_j)$ is a (positive) scaling factor that scales the nonlinear function f_j of the continuous covariate z_j. The scaling factor may depend through the predictor $\widetilde{\eta}_j$ on covariates where $\widetilde{\eta}_j$ is possibly another structured additive predictor of the form (1).

As already mentioned in the introduction, an alternative approach has been proposed in Brunauer et al. (2010) and further developed in Lang et al. (2015) and Razen and Lang (2020). The approach assumes that f_j in 3 is scaled by the scaling factor $1+\alpha_c$ rather than $\exp(\widetilde{\eta}_j)$. Here, α_c is a cluster specific random effect with respect to the cluster variable c that possibly depends on further cluster specific (nonlinear) terms. The disadvantage compared to the proposal of this paper is that only cluster specific terms are possible. Moreover, negative scaling factors are possible.

2.3 Bayesian inference

We describe Markov chain Monte Carlo updating for interaction terms of the form in (3).

The scaled function f_j in (3) is updated with a Gibbs sampling step treating $\exp(\widetilde{\eta}_j)$ as the effect modifier of a varying coefficient term, see Brezger and Lang (2006) for details.

Updates for an effect $g(\tilde{z})$ that is part of the predictor $\widetilde{\eta}_j$ are nonstandard and have to be done with an IWLS sampler using the following ingredients:

- Log-Likelihood contribution of the i-th observation, $i = 1, \dots, n$:

$$l_i = -\frac{1}{2\sigma^2}(y_i - \eta_i)^2 = -\frac{1}{2\sigma^2}\left(y_i - [\cdots + \exp(\widetilde{\eta}_{ij}) \cdot f_j(z_{ij}) + \cdots]\right)^2 .$$

- First derivative:

$$v_i = \frac{\partial}{\partial \widetilde{\eta}_{ij}} l_i = \frac{1}{\sigma^2}(y_i - \eta_i)\exp(\widetilde{\eta}_{ij}) f_j(z_{ij})$$

- Second derivative:

$$\begin{aligned} \frac{\partial}{\partial \widetilde{\eta}_{ij}} v_i &= \frac{f_j(z_{ij})}{\sigma^2}\left[-\exp(\widetilde{\eta}_{ij}) f_j(z_{ij})\exp(\widetilde{\eta}_{ij}) + (y_i - \eta_i)\exp(\widetilde{\eta}_{ij})\right] \\ &= \frac{f_j(z_{ij})\exp(\widetilde{\eta}_{ij})}{\sigma^2}\left[-f_j(z_{ij})\exp(\widetilde{\eta}_{ij}) + (y_i - \eta_i)\right] \end{aligned}$$

- Expected second derivative:

$$E\left(\frac{\partial}{\partial \widetilde{\eta}_{ij}} v_i\right) = -\frac{(f_j(z_{ij})\exp(\widetilde{\eta}_{ij}))^2}{\sigma^2}.$$

- Working weight (based on observed Fisher information):

$$w_i = \frac{f_j(z_{ij})\exp(\widetilde{\eta}_{ij})}{\sigma^2}\left[f_j(z_{ij})\exp(\widetilde{\eta}_{ij}) - (y_i - \eta_i)\right]$$

- Working weight (based on expected Fisher information):

$$w_i = \frac{(f_j(z_{ij})\exp(\widetilde{\eta}_{ij}))^2}{\sigma^2}.$$

- Working response:

$$\widetilde{y}_i = \widetilde{\eta}_{ij} + \frac{v_i}{w_i}$$

An IWLS proposal for the regression coefficients of $g(\tilde{z})$ is then obtained by sampling from a multivariate normal distribution with expected value

$$\boldsymbol{m} = \left(\tilde{\boldsymbol{Z}}'\boldsymbol{W}\tilde{\boldsymbol{Z}} + \frac{1}{\tau^2}\boldsymbol{K}\right)^{-1}\tilde{\boldsymbol{Z}}'\boldsymbol{W}(\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{\eta}}_{-\tilde{z}})$$

and covariance matrix

$$\boldsymbol{\Sigma} = \left(\tilde{\boldsymbol{Z}}'\boldsymbol{W}\tilde{\boldsymbol{Z}} + \frac{1}{\tau^2}\boldsymbol{K}\right)^{-1},$$

where $\tilde{\boldsymbol{Z}}$ and $\boldsymbol{K}$ are the design and penalty matrices corresponding to $g(\tilde{z})$, $\boldsymbol{W} = diag(w_1, \ldots, w_n)$ is a diagonal marix of working weights, $\tilde{\boldsymbol{y}}$ is the vector of working responses. Details on IWLS updates can be found in Brezger and Lang (2006).

Finally, note on identifiability: To render the model identifiable, we assume that $\widetilde{\eta}_{ij}$ does not comprise an intercept and all effects in $\widetilde{\eta}_{ij}$ are centered since otherwise changes in the level of $\widetilde{\eta}_{ij}$ be compensated by rescaling $f_j(z_{ij})$.

The estimation approach is fully embedded in the software package BayesX, see Brezger, Kneib, and Lang (2005) and Umlauf et al. (2014).

3 Simulation study

3.1 Simulation Design

We carry out a simulation study based on a Gaussian structured additive regression model with predictor

$$\eta = f_1(t) + f_2(x) \cdot exp\left(\gamma_{0i} + f_3(z)\right) + f_4(lon,\, lat)$$

and a standard deviation σ^2 such that the signal to noise ratio σ_η/σ is 2. For the nonlinear functions we assume

- $f_1(t) = cos(2\pi(t-1)/100)$ with t being a sequence from 1 through 100,
- $f_2(x) = sin(x)$, where x is uniformly distributed on $[\pi/2,\, 3\pi/2]$,
- $f_3(z) = 0.7\, cos(2\pi(z{-}1)/100)$ where z is uniformly distributed on the interval $[1,\, 100]$,

- $f_4(lon, lat) = sin(2lon + lat)$, where lon and lat correspond to the longitude and latitude coordinates of the centroids of the 3-digit postcode regions.

A visualisation of the functions $f_1 - f_4$ can be found in Figures 3 and 2. Finally, γ_{0i} is a normally distributed cluster-specific random effect with mean 0 and standard deviation 0.49.

We simulate 250 repetions from the model desribed above. In the next Section 3.2, we compare results of the estimated models using the average estimated functions for all 250 repetitions, i.e.

$$\bar{f}(x) = \frac{1}{250} \sum_{j=1}^{250} \hat{f}^j(x)$$

for a particular function $f(x)$ in η and $\hat{f}^j$ being the estimator of the j-th repetition. We further compute the square root of the relative RMSEs for replication $j = 1, \ldots, 250$

$$RMSE(j) = \sqrt{\frac{\sum_x \left(\hat{f}^j(x) - f(x)\right)^2}{\sum_x f(x)^2}}.$$

Regarding the scaling of the functions, are straightforward comparison of the MSEs is impractical. For each of the 113 clusters that we generate, we obtain a different scaling term and thus, a different function of x. Because a comparison would be impracticable for all 250 repetitions, we provide an RMSE for all scaled average functions for clusters $c = 1$, $\ldots$, 113

$$RMSE_{scaled}(c) = \sqrt{\frac{\sum_x \left(\bar{f}_c \hat{s}_c - f(x) s_c\right)^2}{\sum_x (f(x) s_c)^2}},$$

where $\bar{f}_c$ is the average estimated function for term $f_2(x)$ in cluster $c \in 1, \ldots, 113$, and s_c is the scaling term for cluster c.

The following estimation results are based on P-spline priors with 20 knots and second order difference penalty for the functions f_1, f_2, f_3. In case of the spatial function f_4 a Markov random field prior (e.g. Fahrmeir et al., 2022) is assumed. Throughout for the variance parameters τ^2 we assume inverse gamma priors with hyperparameters $a = 0.001$ and $b = 0.002$. Results are based on a Markov chain with 105,000 iterations, a burnin period of 5,000 iterations. Every 100th sample is later used for estimation.

3.2 Results

Results for a single representative replication are provided in Figures 1-4. We present estimates for the approach proposed in this paper as well as for the alternative (competing) approach briefly outlined at the end of Section 2.2. The results are also contrasted with the true effects. Figure 1 shows estimated cluster-specific random effects γ_{0i}, Figures 2 and

3 display results for the functions f_1-f_4, the scaled function $f_2(x)\, exp(\gamma_{0i} + f_3(z))$ based on the median scaling term can be found in Figure 4. For animated graphs showing the effects for all 250 repetitions, see this link[1]. The estimates are in good agreement with the true effects, with increased variability for f_3 (the function within the exp term) compared to the other effects. Even the alternative approach seems to work comparably well, bearing in mind that the model is misspecified.

To shed more light on the statistical properties, Figures 5-7 show the average estimated effects over the 250 replications for the random effect γ_{0i} and the functions f_1-f_4. For the spatial effect we additionally show the estimated county-specific bias in Figure 8. Coverage rates of the 95% Bayesian credible intervals are presented in Figures 9-11. Finally, Figure 12 reports the RMSEs. We draw the following conclusions:

- All functions are estimated with a comparably small bias. We observe the typical pattern of (slightly) increased bias near local extremes or at the boundaries of the covariate space.

- The Bayesian credible interval coverage rates are, on average, slightly higher than the nominal level and, in this sense, conservative. The average coverage rates range between 95 % for the spatial effect f_4 and 1 for the function f_3 within the scaling factor of f_2. Pointwise coverage rates are as expected, with a tendency for slightly less coverage near boundaries or local extremes.

- The RMSEs for all functions are in an acceptable range. RMSEs are higher for the scaled function f_2 compared to the unscaled function f_1 and highest for the function f_3 as part of the scaling factor. As expected, the competing approach outlined at the end of Section 2.2 shows higher RMSEs as the model is misspecified.

[1]Website: https://jgranna.github.io/simulations.html; On the website we also provide all the files needed to replicate our results. This includes the file used to obtain our simulation data and the estimation programme.

(a) γ_{0i}, exp-scaling. **(b)** γ_{0i}, $(1+\alpha)$-scaling.

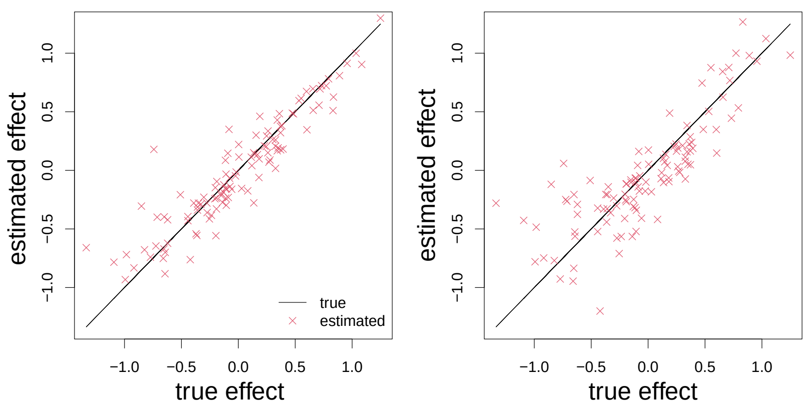


Figure 1: Estimated cluster-specific random effect for random scaling sorted according to effect size.

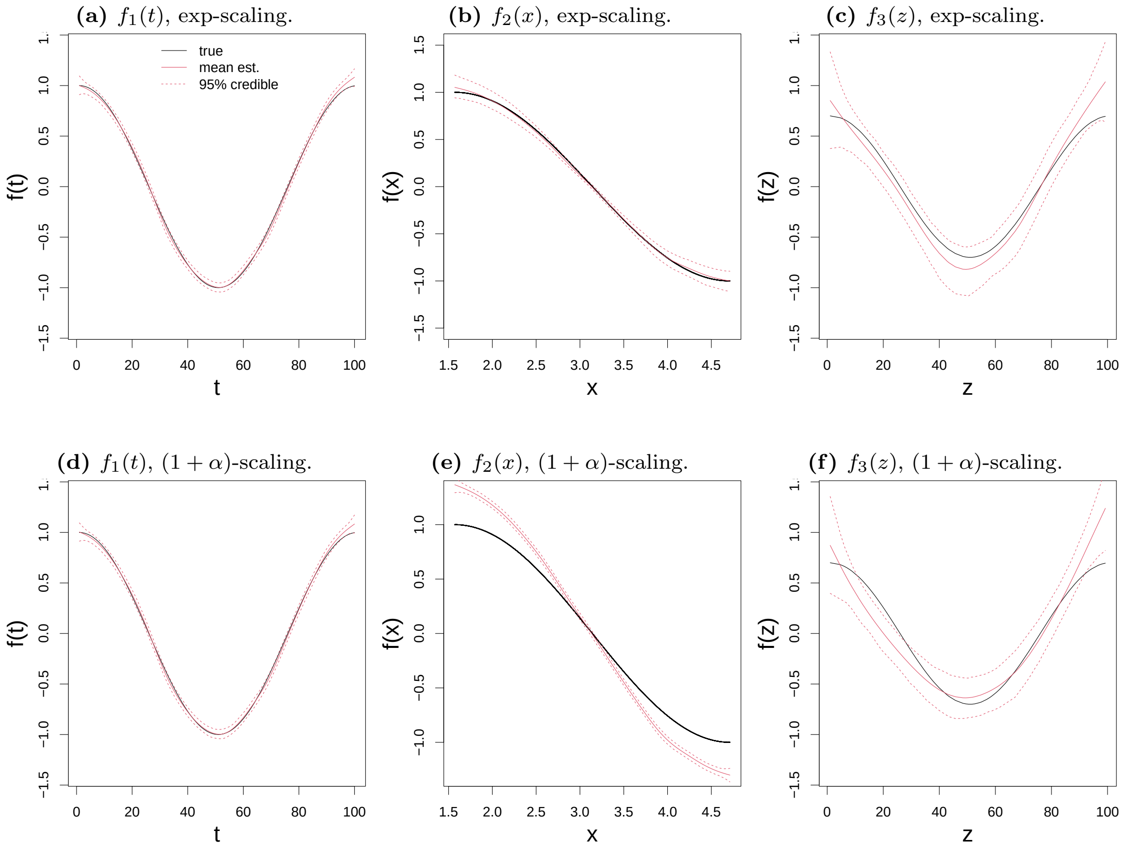


Figure 2: Estimated effects for terms $f_1(t)$, $f_2(x)$, $f_3(z)$. The black solid line corresponds to the true effect, the red solid line corresponds to the mean estimated effect. The red dashed lines are the 95% credible intervals.

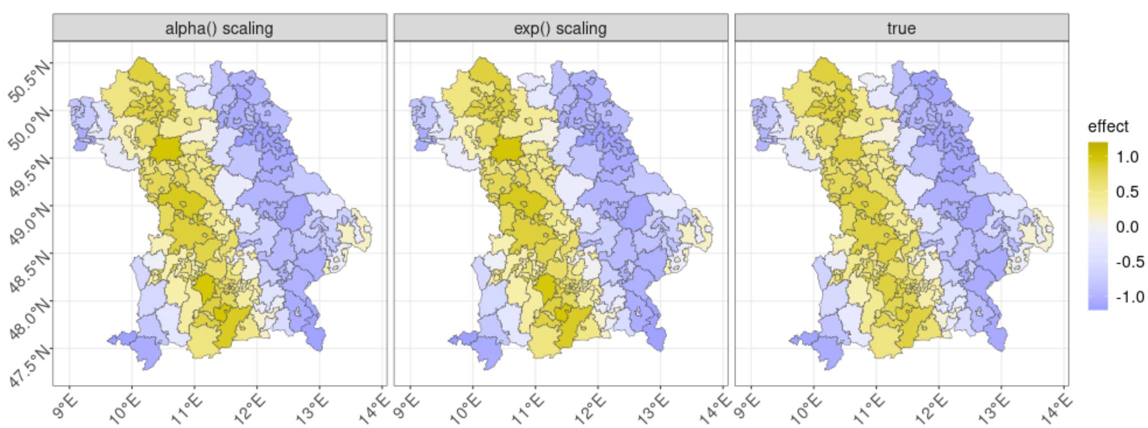


Figure 3: Estimated effects for postcode effect with Markov Random Field prior.

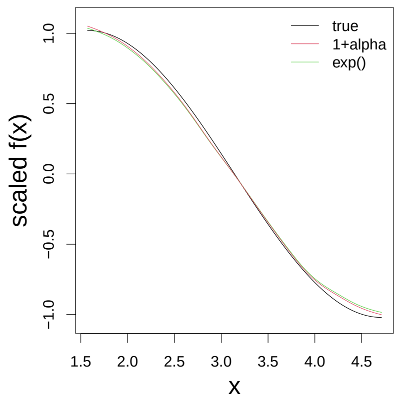


Figure 4: Scaling functions referring to the term $f_2(x)\, s_c$. From all 113 different scaling terms, the median scaling term is reported in this graph.

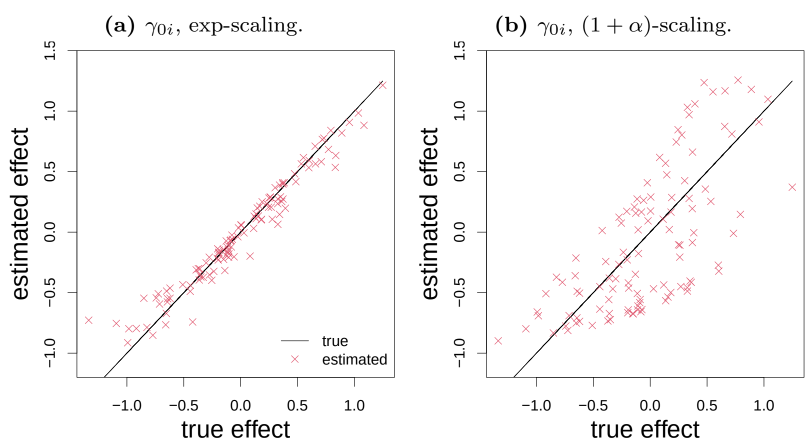


Figure 5: Average estimated effect of γ_{0i}. True values are on the bisecting angle. Dots represent $\hat{\gamma}_{0i}$.

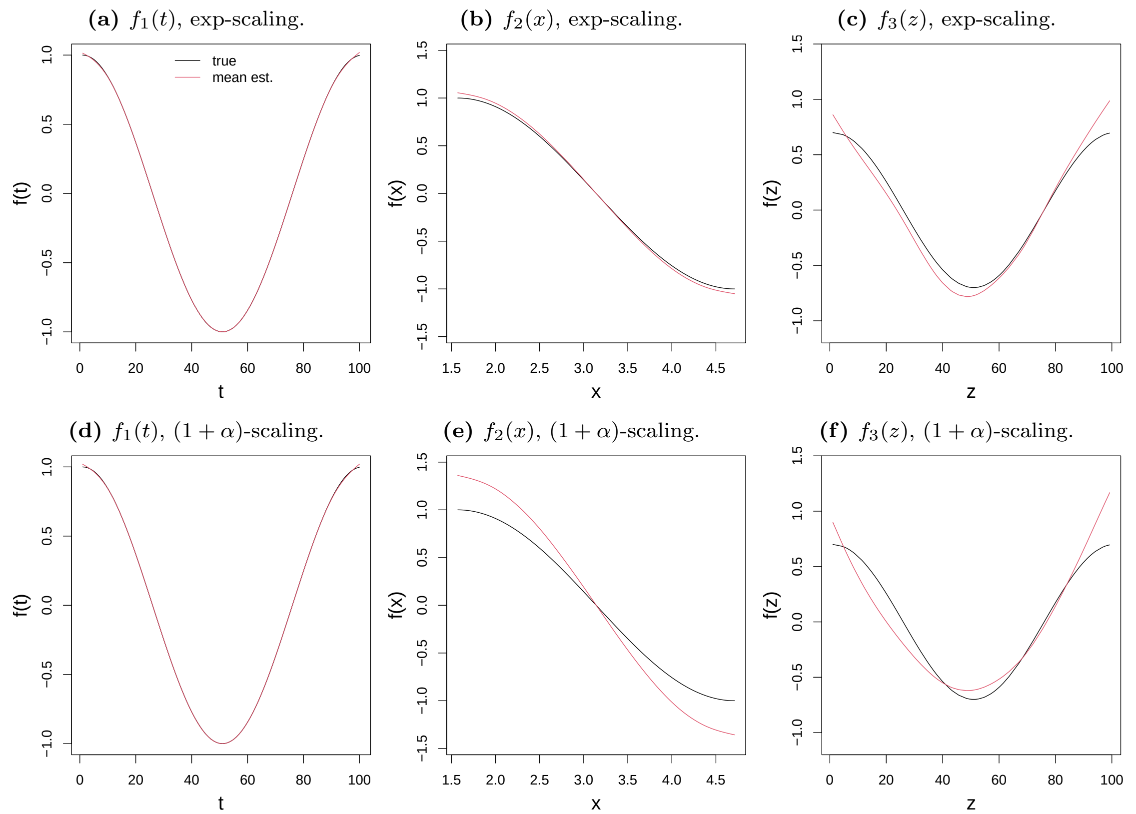


Figure 6: Mean estimated effects for terms $f_1(t)$, $f_2(x)$, $f_3(z)$ (red lines). The black lines are the true curves.

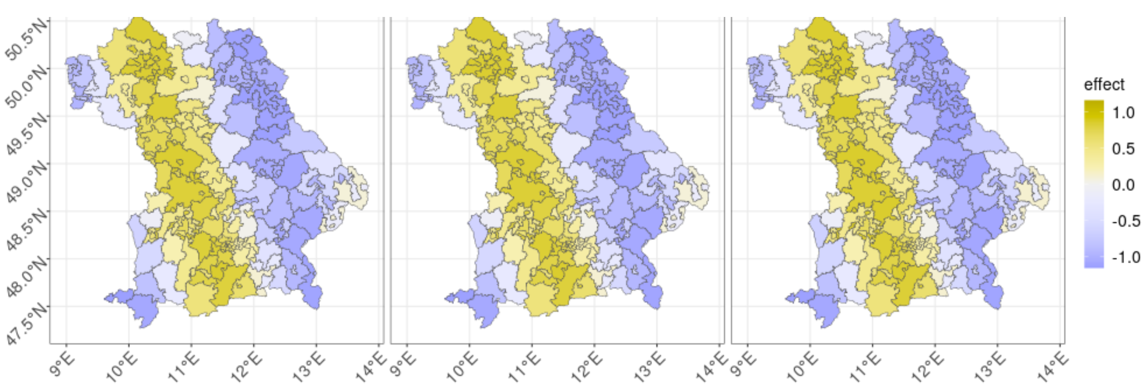


Figure 7: Average effect for plz regions.

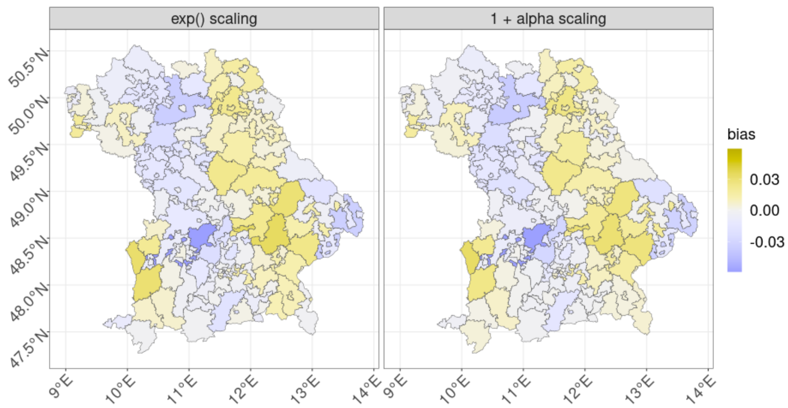


Figure 8: Bias for plz effect.

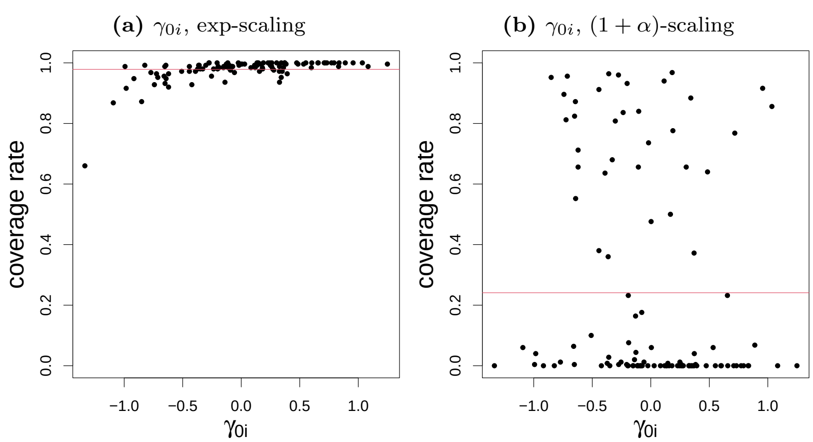


Figure 9: Average coverage rate of 95% credible intervals for all 250 repititions for γ_{0i} term.

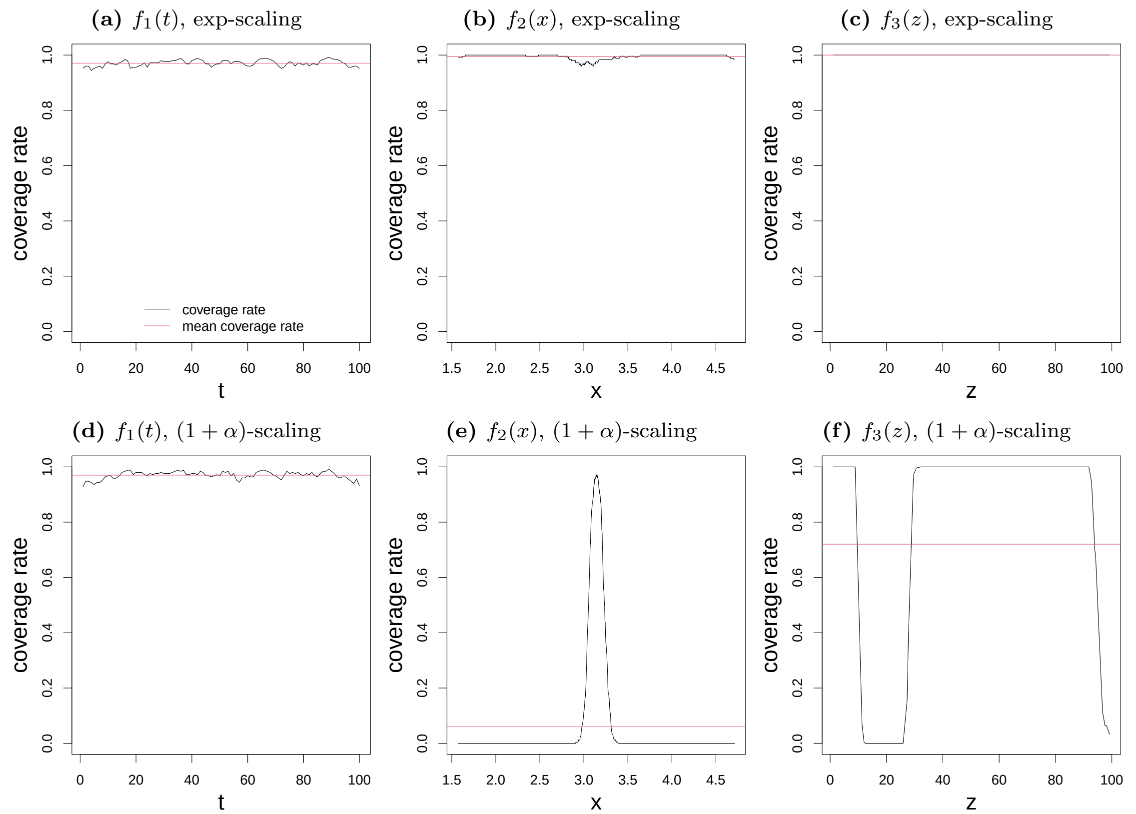


Figure 10: Average coverage rate of 95% credible intervals for all 250 repititions for terms $f_1(t)$, $f_2(x)$, and $f_3(z)$ (rows).

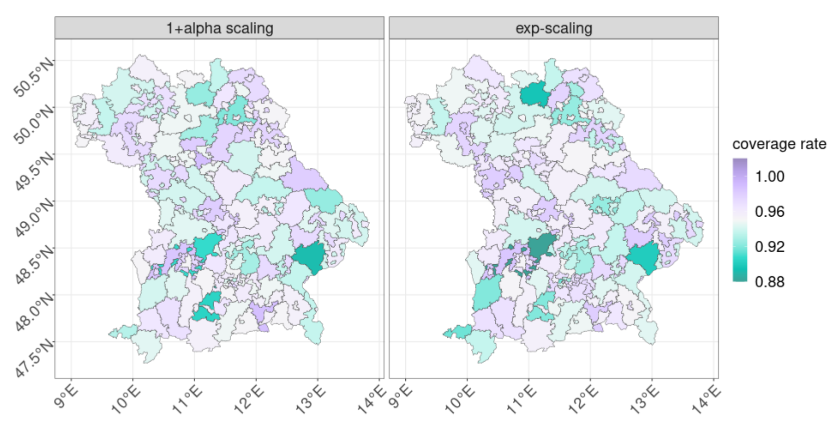


Figure 11: Coverage rate for plz effect. Center of the scale is 95%.

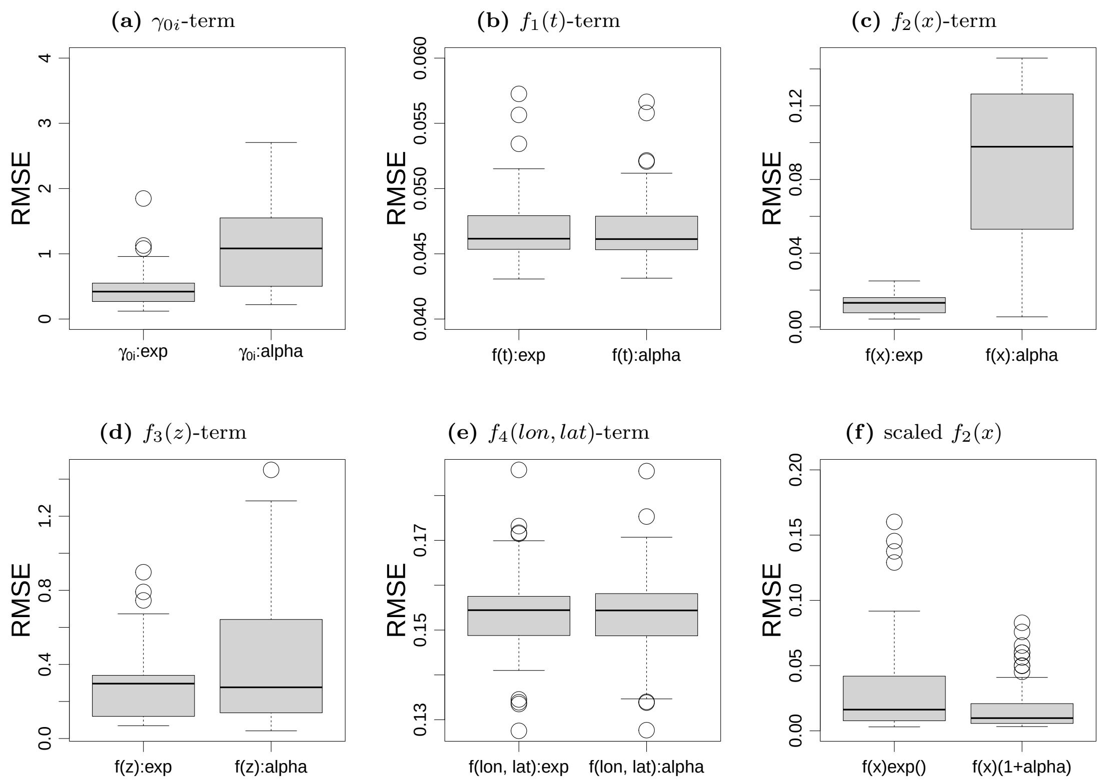


Figure 12: RMSEs for terms γ_{0i}, $f_1(t)$, $f_2(x)$, $f_3(z)$, $f_4(lon, lat)$, and the scaled functions of $f_2(x)$ (rowwise).

4 Application

According to Bondt, Gieseck, and Tujula (2020), housing wealth, i.e. the market value of residential property, constitutes over 55% of total household assets in the euro zone. Tracking real estate prices while accounting for changes in underlying characteristics is hence essential. This is most commonly accomplished in the context of the hedonic method (Rosen, 1974), which interpretes the value of a dwelling as the sum of the prices of its characteristics. These implicit prices are estimated in the context of a regression model, where the price of a unit is regressed on the housing characteristics.

A central aspect of hedonic methods is the appropriate and unbiased estimation of the time trend. This is especially relevant, e.g. in the context of hedonic index construction, which requires unbiased underlying estimates. A challenge in this context is that the effect of time is possibly not homogeneous over location, but also with respect to other covariates in the model. This is discussed in the context of the formation of submarkets by authors like Straszheim (1975), Malpezzi, Ozanne, and Thibodeau (1980), or Helbich et al. (2014).

The exponential scaling introduced in this paper allows us, for the first time, to ensure strictly positive scaling factors. We are also able to consider multiple variables for scaling, which enables us to account not only for spatially heterogeneous time effects, but also with respect to other covariates.

4.1 Data and model specification

For our analysis we use data provided by DataScience Service GmbH, based in Vienna, Austria. The sample includes 200,000 asking prices of dwellings in Germany. The target variable in our application is the price per square metre, or rather its logarithm, since the log transformation offers better distributional properties. In a hedonic distributional regression, we regress the log price on a set of dwelling characteristics to estimate the prices of the corresponding characteristics:
As a continuous regressor, we use the `age` of a dwelling, which is expressed in years as the difference between the last offer for a dwelling and the year of construction. We use age rather than year of construction to ensure better comparability between dwellings. We also consider the `area` of an object, which is the area of the dwelling expressed in m^2, and the `time` at which an object was sold. The latter is given in the corresponding quarter.
The discrete covariates included in our analysis are the binary indicator variables `balcony`, `floor heating` and `furnished`. Finally, we consider the location of the dwelling using the `county` in which the dwelling is located, which corresponds to the so-called "Landkreise" in Germany. There are 400 districts in Germany.
A more detailed description of all included variables is provided in Table 1, which includes summary statistics for all variables.

variable	description	mean / rel. frequency	std. deviation	min	max
`ppqm`	Price per square meter	3292.500	2064.040	50.000	40000.000
`log.ppqm`	log(Price per square meter)	7.930	0.600	3.910	10.600
`age`	Age of flat in years	37.760	30.250	-4.000	119.000
`area`	Area of flat in square meters	80.070	32.380	21.000	240.000
`quarter`	Quarter of last offer	2019.200	2.020	2016.000	2022.500
`county`	county ID				
`balcony`	Whether object has balcony				
	0 = no	0.436			
	1 = yes	0.564			
`floor.heating`	Whether object has floor heating				
	0 = no	0.909			
	1 = yes	0.091			
`furnished`	Whether object is furnished				
	0 = no	0.967			
	1 = yes	0.033			

Table 1: Summary statistics for variables included in the real estate data application. For continuous variables, mean, standard deviation, minimum, and maximum values are provided. For discrete attributes, the relative frequencies of the corresponding categories are shown.

We model dwelling prices using a log Gaussian model. This is the standard distribution in the context of hedonic housing modeling. We relate the location parameter μ to the predictor η as

$$\eta_i = f_1(\texttt{age}_i) + f_2(\texttt{area}_i) + exp(\tilde{\eta}_i) f_3(\texttt{time}_i) + f_4(\texttt{county}_i) + \boldsymbol{X}\boldsymbol{\gamma}, \tag{4}$$

where f_1, f_2, f_3 are possibly nonlinear functions of the included continuous covariates, fitted with P(enalised) splines. f_4 corresponds to the effect of location on the dependent variable modelled as a Markov Random Field. The effects of the `counties` are hence spatially correlated. $\boldsymbol{X}$ is the design matrix with an intercept next to the discrete explanatory variables. $\boldsymbol{\gamma}$ contains the corresponding coefficients of the linear effects. $exp(\tilde{\eta}_i)$ is the positive scaling factor that scales the time effect and includes the structured additive predictor.

$$\tilde{\eta}_i = g_1(\texttt{county}_i) + g_2(\texttt{area}_i) + g_3(\texttt{age}_i) + \tilde{\boldsymbol{X}}\tilde{\boldsymbol{\gamma}}. \tag{5}$$

g_1 is a Markov random field, since we assume spatially correlated scaling factors. g_2, g_3 are possibly nonlinear functions to model the effects of area and age on the scaling term. $\tilde{\boldsymbol{X}}\tilde{\boldsymbol{\gamma}}$ are the linear effects of the discrete variables on the scaling term.

4.2 Results

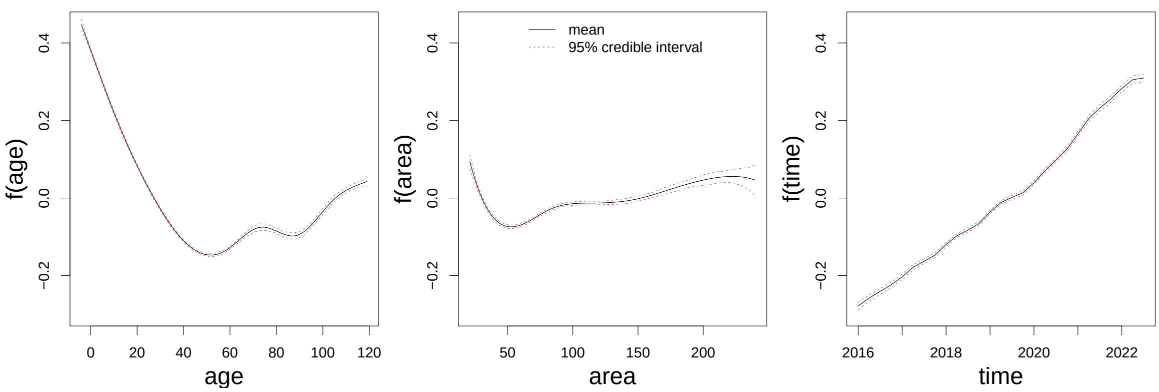
For the effects of continuous covariates, we provide graphs showing the mean, 2.5% and 97.5% quantiles, as the values of the estimated coefficients are not meaningful for interpretation. For linear effects, we provide tables summarising the corresponding estimated coefficients.

4.2.1 Terms outside scaling term

Figure 13 presents the mean effects along with the 95% credible intervals for the continuous covariates. The variables `age` and `area` are not scaled, while `time` is scaled with an exponential scaling term. In line with many findings in the literature, the marginal effect of `age` is U-shaped. Very new dwellings have on average the highest prices per square metre. Dwellings older than 50 years are again increasingly expensive, reflecting buyers' preference for historic buildings.

In terms of the total `area` of apartments, we also find a U-shaped effect on price. As expected, very small homes have the highest average price per square metre. As the `area` of the dwelling increases, the price per square metre decreases until about $50m^2$. This is common in the literature and corresponds to a quantity discount due to the diminishing marginal utility of floor area. Beyond this size, the average price increases monotonically.

Over the `time` horizon covered, the average price per square metre increases monotonically. The scaling of the time effect and the resulting scaled functions for the time effect are examined below.



(a) `age` effect, $f_1($`age`$)$. **(b)** `area` effect, $f_2($`area`$)$. **(c)** `time` effect, $f_3($`time`$)$.

Figure 13: Estimated effects for covariates `age`, `area`, and `time`. Solid line corresponds to the mean, dashed lines refer to 95% credible intervals.

The summary statistics for the estimated linear effect for the categorical covariates `balcony`, `floor heating`, and `furnished` are shown in Table 2.

All effects are statistically significant, as none of the reported credible intervals overlap the value of 0. As expected, and in line with the findings of other authors, all mean effects are greater than 0, implying a positive impact on the price per square metre. The largest average effect is for `furnished`.

	Mean	Sd	2.5%	50%	97.5%
(Intercept)	7.615	0.003	7.611	7.615	7.621
`balcony`	0.045	0.001	0.042	0.045	0.048
`floor heating`	0.042	0.003	0.037	0.042	0.047
`furnished`	0.060	0.004	0.052	0.060	0.068

Table 2: Summary of samples of estimated coefficients for linear effects.

Figure 14 illustrates the effect of location on μ. For the sake of orientation, we also label large cities with more than $400,000$ inhabitants. Two main findings emerge from the graph. First, average prices per square metre are increased around large cities. This is particularly true for Berlin, Frankfurt, Hamburg and Munich. A large area around Munich, stretching towards the Alps in the very south of Germany, is characterised by very high average prices. But areas close to the North Sea and the Baltic Sea are also more expensive than neighbouring `counties`.

Second, rural areas in the former East Germany in particular are less valuable than in the former West Germany. These findings are consistent with many studies such as Razen and Lang (2020).

4.2.2 Scaled time effect

Below, we present the estimated effects within the scaling term together with the scaled time effects to gain insight into the heterogeneity of time across the scaling variables.

Table 3 reports the estimated coefficients for the linear effects included in the scaling term. The credible interval for `balcony` includes 0, indicating that `balcony` does not have a significant effect on the scaling of time.

With regard to the impact of `floor heating`, the negative mean coefficient indicates that dwellings with underfloor heating are subject to a flatter time trend, which means that dwellings without underfloor heating face steeper price appraisals in the time horizon considered.

Finally, `furnished` apartments are associated with a larger scaling term, which corresponds to a steeper scaling effect of time on price. Thus, furnished dwellings were subject to larger price increases than unfurnished dwellings.

The scaled time functions resulting from the effects on the scaling term from Table 3 are plotted in Figure 15. For the plots, only the variable of interest is varied within the scaling term, while the other covariates are held fixed at their means.

The two functions for `balcony` are virtually indistinguishable. This is not surprising given the very small and insignificant effect on the scaling term. The scaled function for apartments equipped with `floor heating` (red line, middle panel) is slightly flatter than that for apartments without `floor heating` (black line). Lastly, the right panel

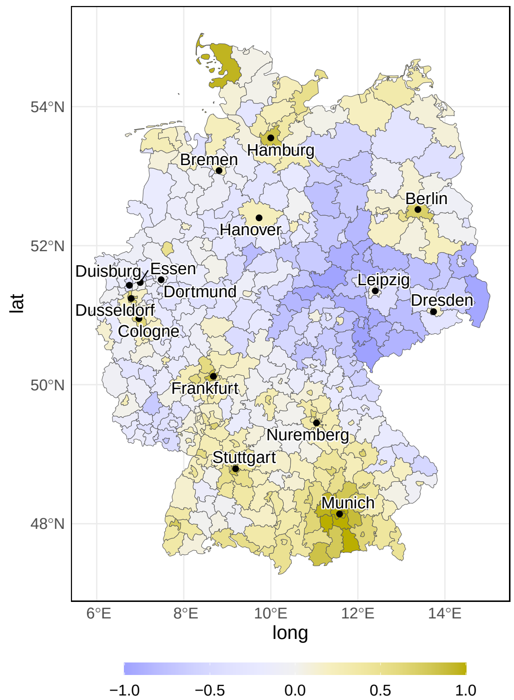


Figure 14: Markov Random Field for the 400 `counties` in Germany corresponding to term f_4. Dark yellow areas correspond to higher average square metre prices, dark blue areas to lower mean prices.

corresponds to the coefficient of the covariate `furnished`. The red line refers to the positive, albeit smaller in absolute value, estimated effect on the scaling, indicating a slightly steeper average price increase for `furnished` dwellings compared to unfurnished ones.

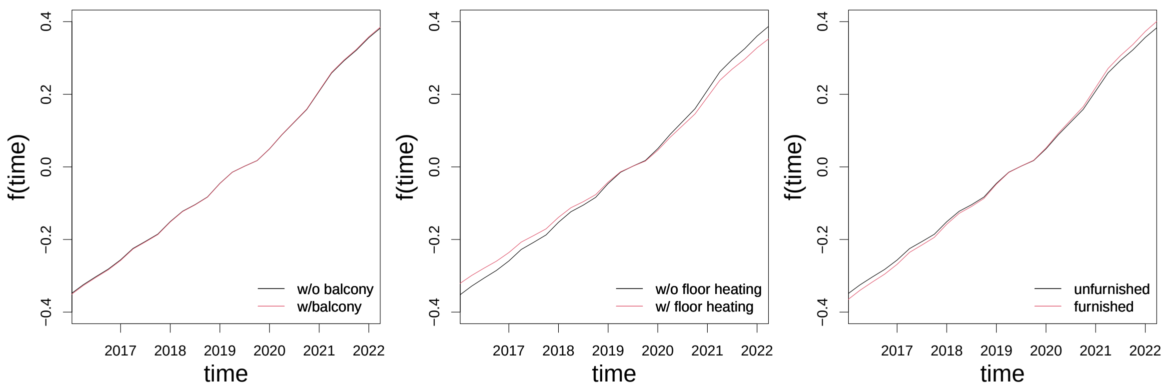
The effect of the continuous covariates on the scaling of `time` along with the corresponding credible intervals is illustrated in Figure 16.

The first panel displays the influence of `age`. The function fluctuates around zero and the credible intervals at least partially overlap the zero. There is no clear association with the heterogeneity of the time effect. This finding translates into the scaled f_3 time functions in the left panel of Figure 17. There is some variation in the functions, but no clear pattern.

The right panel of Figure 16 shows the effect of `area` on the scaling term. The effect is (almost) monotonically decreasing - but with wider credible confidence bands around very high values of `area`. This implies that larger values of `area` translate into smaller scaling of `time`, or in other words, smaller dwellings are subject to steeper price rises between 2016 and 2022 compared to larger dwellings. This finding is illustrated in the middle panel of

	Mean	Sd	2.5%	50%	97.5%
`balcony`	0.007	0.007	-0.007	0.007	0.020
`floor heating`	-0.094	0.015	-0.126	-0.093	-0.064
`furnished`	0.045	0.016	0.014	0.044	0.075

Table 3: Summary statistics for samples of estimated coefficients for discrete covariates with the scaling term $\tilde{\eta}_i$.



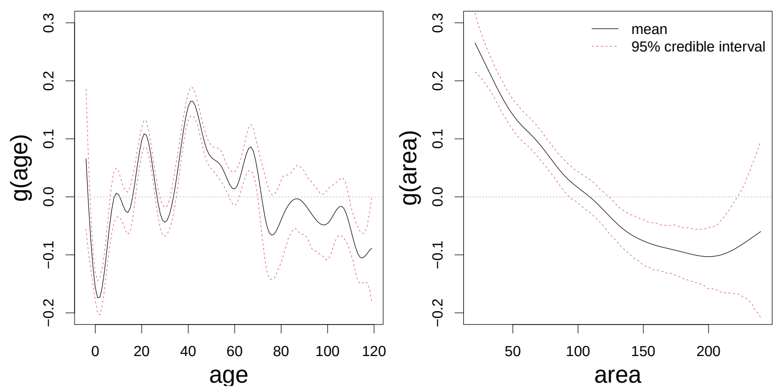
(a) Scaled with `balcony`. **(b)** Scaled with `floor heating`. **(c)** Scaled with `furnished`

Figure 15: Effect of `time` scaled with `balcony`, `floor area`, and `furnished` evaluated at the mean level.

Figure 17.

The effect of location on the scaling term is illustrated in Figure 18. It is worth noting that the patterns are not uniform across Germany, particularly in the north and south. In Bavaria and Baden-Württemberg, `counties` with a high overall price level, as seen in Figure 14, are generally associated with lower price increases compared to areas with lower overall prices per square metre, i.e. rural areas. These findings are reversed in northern Germany. In particular, the `counties` around Berlin in Brandenburg exhibit steeper price appreciation in comparison with rural areas around cities such as Leipzig or Bremen. Whether this effect is - at least in part - induced by the rent cap that was in force in Berlin between 30 January 2020 and 15 April 2021, remains to be investigated. The resulting scaled functions of `time` are illustrated in the right panel in Figure 17.

In terms of its impact on scaling, location has the greatest variability with regard to the range of scaled functions. This is little surprising, as location is also considered to be the most important single contributing characteristic of dwellings on the price. With the exception of `age` and `balcony`, all the covariates considered have a significant impact on the scaling of the `time` effect. Our results underline the need to take into account the heterogeneity of the effect of time on the price per square metre, not only with respect to location, but also with respect to other covariates.



(a) `age` effect, $g_1(\texttt{age})$.

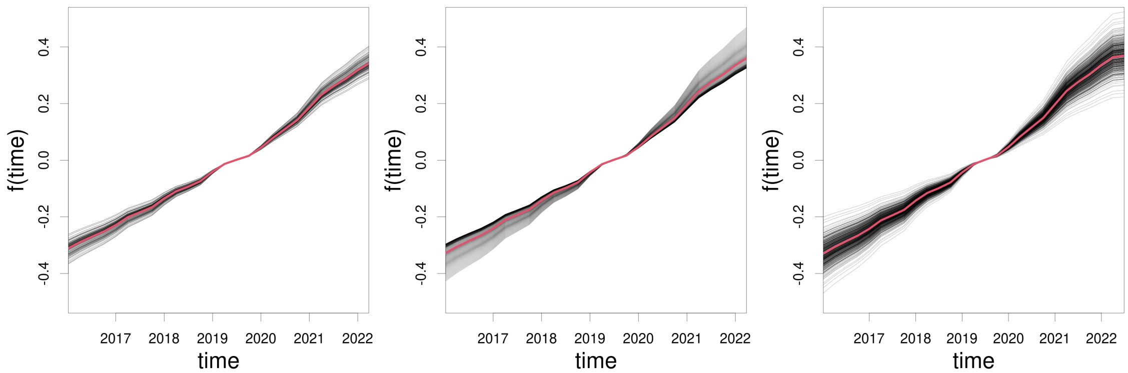
(b) `area` effect, $g_2(\texttt{area})$.

Figure 16: Estimated smooth functions of effects for `age` and `area` within the scaling factor. Solid line corresponds to the mean, dashed red lines refer to 95% credible intervals.

5 Conclusion

This paper presents a simultaneous estimation approach for Bayesian structured additive regression models with possibly complex interactions where a nonlinear function f is scaled with respect to a strictly positive scaling factor $\exp(\tilde{\eta})$ where the additional structured additive predictor $\tilde{\eta}$ may be composed of arbitrary terms of the remaining covariates. Extensive simulation experiments show the validity of the approach. A detailed case study demonstrates the usefulness of the methodology in practice.

So far our approach is restricted to homoscedastic Gaussian responses. For the future we plan to extend the methodology to more general Bayesian distributional regression, see Klein et al. (2015), such that every parameter of a distribution may be modelled in terms of a full structured additive predictor with the interaction terms discussed in this paper.



(a) Scaled with `age`. **(b)** Scaled with `area`. **(c)** Scaled with `counties`.

Figure 17: Effect of time scaled with **`area`**, **`age`**, and **`counties`** evaluated at the mean level. Red lines refer to the unscaled functions.

References

Bondt, Gabe de, Arne Gieseck, and Mika Tujula (2020). "Household wealth and consumption in the euro area". In: *ECB Economic Bulletin* (1).

Brezger, A. and S. Lang (2006). "Generalized Structured Additive Regression based on Bayesian P-splines". In: *Computational Statistics and Data Analysis* 50, pp. 967–991.

Brezger, Andreas, Thomas Kneib, and Stefan Lang (2005). "BayesX: Analyzing Bayesian Structured Additive Regression Models". In: *Journal of Statistical Software* 14.11, pp. 1–22. URL: `http://www.jstatsoft.org/v14/i11/`.

Brunauer, W., S. Lang, P. Wechselberger, and S. Bienert (2010). "Additive Hedonic Regression Models with Spatial Scaling Factors: An Application for Rents in Vienna". In: *Journal of Real Estate Finance and Economics* 41, pp. 390–410.

Fahrmeir, Ludwig, Thomas Kneib, Stefan Lang, and Brian Marx (2022). *Regression. Models, Methods and Applications*. Springer Verlag.

Granna, Julian, Wolfgang A. Brunauer, and Stefan Lang (2022). "Proposing a global model to manage the bias-variance tradeoff in the context of hedonic house price models".

Helbich, Marco, Wolfgang Brunauer, Eric Vaz, and Peter Nijkamp (Feb. 2014). "Spatial Heterogeneity in Hedonic House Price Models: The Case of Austria". In: *Urban Studies* 51 (2), pp. 390–411. ISSN: 0042-0980. DOI: `10.1177/0042098013492234`.

Klein, N., T. Kneib, S. Lang, and A. Sohn (2015). "Bayesian structured additive distributional regression with an application to regional income inequality in Germany". In: *The Annals of Applied Statistics* 9, pp. 1024–1052.

Lang, S. and A. Brezger (2004). "Bayesian P-Splines". In: *Journal of Computational and Graphical Statistics* 13, pp. 183–212.

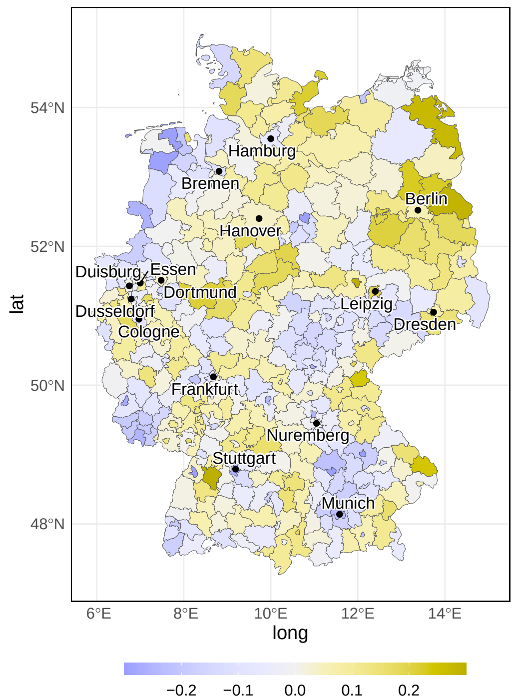


Figure 18: Markov Random Field for term g_1 within the scaling factor for the 400 `counties` in Germany.

Lang, S., W.J. Steiner, A. Weber, and P. Wechselberger (2015). "Accommodating heterogeneity and nonlinearity in price effects for predicting brand sales and profits". In: *European Journal of Operational Research* 246.1, pp. 232–241.

Malpezzi, Stephen, Larry Ozanne, and Thomas Thibodeau (1980). *Characteristic prices of housing in fifty-nine metropolitan areas*. Urban Institute.

Razen, A. and S. Lang (2020). "Random scaling factors in Bayesian distributional regression models with an application to real estate data". In: *Statistical Modelling* 20.4, pp. 347–368.

Rosen, Sherwin (Jan. 1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition". In: *Journal of Political Economy* 82 (1), pp. 34–55. ISSN: 0022-3808. DOI: `10.1086/260169`.

Straszheim, Mahlon R (1975). *An Econometric Analysis of the Urban Housing Market*. NBER. URL: `https://www.nber.org/books-and-chapters/econometric-analysis-urban-housing-market`.

Umlauf, N., Thomas Kneib, Stefan Lang, and Achim Zeileis (2014). "Structured Additive Regression Models: An R Interface to BayesX." In: *Journal of Statistical Software* 63.23. URL: `https://www.jstatsoft.org/article/view/v063i21`.

Wood, Simon N (2017). *Generalized additive models: an introduction with R*. CRC press.

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2024-01 **Philipp Aschersleben, Julian Granna, Thomas Kneib, Stefan Lang, Nikolaus Umlauf, and Winfried Steiner:** Modeling multiplicative interaction effects in Gaussian structured additive regression models

2023-18 **Luke Glowacki, Florian Morath and Hannes Rusch:** High minority power facilitates democratization across ethnic fault lines

2023-17 **Felix Holzmeister, Magnus Johannesson, Robert Böhm, Anna Dreber, Jürgen Huber, Michael Kirchler:** Heterogeneity in effect size estimates: Empirical evidence and practical implications

2023-16 **Changxia Ke, Florian Morath, Sophia Seelos:** Do groups fight more? Experimental evidence on conflict initiation

2023-15 **Loukas Balafoutas, Helena Fornwagner, Rudolf Kerschbamer, Matthias Sutter, and Maryna Tverdostup:** Serving consumers in an uncertain world: A credence goods experiment

2023-14 **Loukas Balafoutas, Helena Fornwagner, Rudolf Kerschbamer, Matthias Sutter, and Maryna Tverdostup:** Diagnostic Uncertainty and Insurance Coverage in Credence Goods Markets

2023-13 **Sarah Lynn Flecke, Rene Schwaiger, Jürgen Huber, Michael Kirchler:** Nature Experiences and Pro-Environmental Behavior: Evidence from a Randomized Controlled Trial

2023-12 **Christian König-Kersting and Stefan T. Trautmann:** Grit, Discounting, & Time Inconsistency

2023-11 **Stefano Piasenti, Marica Valente, Roel van Veldhuizen, Gregor Pfeifer:** Does Unfairness Hurt Women? The Effects of Losing Unfair Competitions

2023-10 **Pascal Kieren, Christian König-Kersting, Robert Schmidt, Stefan Trautmann, Franziska Heinicke:** First-Order and Higher-Order Inflation Expectations: Evidence about Households and Firms

2023-09 **Marica Valente, Timm Gries, Lorenzo Trapani:** Informal employment from migration shocks

2023-08 **Natalie Struwe, Esther Blanco, James M. Walker:** No response to changes in marginal incentives in one-shot public good experiments

2023-07 **Sebastian Bachler, Sarah Lynn Flecke, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Carbon Pricing, Carbon Dividends and Cooperation: Experimental Evidence

2023-06 **Elisabeth Gsottbauer, Michael Kirchler, and Christian König-Kersting:** Climate Crisis Attitudes among Financial Professionals and Climate Experts

2023-05 **Xiaogeng Xu, Satu Metsälampi, Michael Kirchler, Kaisa Kotakorpi, Peter Hans Matthews, Topi Miettinen:** Which income comparisons matter to people, and how? Evidence from a large field experiment

2023-04 **Tanja Hoertnagl, Rudolf Kerschbamer, Regine Oexl, Rudi Stracke, and Uwe Sunde:** Heterogeneity in Rent-Seeking Contests with Multiple Stages: Theory and Experimental Evidence

2023-03 **Melissa Newham, Marica Valente:** The Cost of Influence: How Gifts to Physicians Shape Prescriptions and Drug Costs

2023-02 **Natalie Struwe, Esther Blanco, James M. Walker:** Determinants of Financial Literacy and Behavioral Bias among Adolescents

2023-01 **Marco Aschenwald, Armando Holzknecht, Michael Kirchler, Michael Razen:** Determinants of Financial Literacy and Behavioral Bias among Adolescents

2022-20 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, and Thomas Rittmannsberger:** Beliefs about social norms and (the polarization of) COVID-19 vaccination readiness

2022-19 **Edward I. Altman, Marco Balzano, Alessandro Giannozzi, Stjepan Srhoj:** Revisiting SME default predictors: The Omega Score

2022-18 **Johannes Diederich, Raphael Epperson, Timo Goeschl:** How to Design the Ask? Funding Units vs. Giving Money

2022-17 **Toman Barsbai, Vojtěch Bartoš, Victoria Licuanan, Andreas Steinmayr, Erwin Tiongson, and Dean Yang:** Picture This: Social Distance and the Mistreatment of Migrant Workers

2022-16 **Andreas Steinmayr, Manuel Rossi:** Vaccine-skeptic physicians and COVID-19 vaccination rates

2022-15 **Stjepan Srhoj, Alex Coad, Janette Walde:** HGX: The Anatomy of High Growth Exporters

2022-14 **Martin Obradovits, Philipp Plaickner** Price-Directed Search, Product Differentiation and Competition

2022-13 **Utz Weitzel, Michael Kirchler** The Banker's Oath And Financial Advice

2022-12 **Julian Granna, Wolfgan Brunauer, Stefan Lang:** Proposing a global model to manage the bias-variance tradeoff in the context of hedonic house price models

2022-11 **Christoph Baumgartner, Stjepan Srhoj and Janette Walde:** Harmonization of product classifications: A consistent time series of economic trade activities

2022-10 **Katharina Momsen, Markus Ohndorf:** Seller Opportunism in Credence Good Markets ? The Role of Market Conditions

2022-09 **Christoph Huber, Michael Kirchler:** Experiments in Finance ? A Survey of Historical Trends

2022-08 **Tri Vi Dang, Xiaoxi Liu, Florian Morath:** Taxation, Information Acquisition, and Trade in Decentralized Markets: Theory and Test

2022-07 **Christoph Huber, Christian König-Kersting:** Experimenting with Financial Professionals

2022-06 **Martin Gächter, Martin Geiger, Elias Hasler:** On the structural determinants of growth-at-risk

2022-05 **Katharina Momsen, Sebastian O. Schneider:** Motivated Reasoning, Information Avoidance, and Default Bias

2022-04 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Thomas Rittmannsberger:** How does the vaccine approval procedure affect COVID-19 vaccination intentions?

2022-03 **Robert Böhm, Cornelia Betsch, Yana Litovsky, Philipp Sprengholz, Noel Brewer, Gretchen Chapman, Julie Leask, George Loewenstein, Martha Scherzer, Cass R. Sunstein, Michael Kirchler:** Crowdsourcing interventions to promote uptake of COVID-19 booster vaccines

2022-02 **Matthias Stefan, Martin Holmén, Felix Holzmeister, Michael Kirchler, Erik Wengström:** You can't always get what you want–An experiment on finance professionals' decisions for others

2022-01 **Toman Barsbai, Andreas Steinmayr, Christoph Winter:** Immigrating into a Recession: Evidence from Family Migrants to the U.S.

2021-32 **Fanny Dellinger:** Housing Support Policies and Refugees' Labor Market Integration in Austria

2021-31 **Albert J. Menkveld, Anna Dreber, Felix Holzmeister, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusüss, Michael Razen, Utz Weitzel and et al:** Non-Standard Errors

2021-30 **Toman Barsbai, Victoria Licuanan, Andreas Steinmayr, Erwin Tiongson, Dean Yang:** Information and Immigrant Settlement

2021-29 **Natalie Struwe, Esther Blanco, James M. Walker:** Competition Among Public Good Providers for Donor Rewards

2021-28 **Stjepan Srhoj, Melko Dragojević:** Public procurement and supplier job creation: Insights from auctions

2021-27 **Rudolf Kerschbamer, Regine Oexl:** The effect of random shocks on reciprocal behavior in dynamic principal-agent settings

2021-26 **Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon:** Competitive versus cooperative incentives in team production with heterogeneous agents

2021-25 **Anita Gantner, Regine Oexl:** Respecting Entitlements in Legislative Bargaining - A Matter of Preference or Necessity?

2021-24 **Silvia Angerer, E. Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** The formation of risk preferences throughsmall-scale events

2021-23 **Stjepan Srhoj, Dejan Kovač, Jacob N. Shapiro, Randall K. Filer:** The Impact of Delay: Evidence from Formal Out-of-Court Restructuring

2021-22 **Octavio Fernández-Amador, Joseph F. Francois, Doris A. Oberdabernig, Patrick Tomberger:** Energy footprints and the international trade network: A new dataset. Is the European Union doing it better?

2021-21 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Nudging Debtors to Pay Their Debt: Two Randomized Controlled Trials

2021-20 **Daniel Müller, Elisabeth Gsottbauer:** Why Do People Demand Rent Control?

2021-19 **Alexandra Baier, Loukas Balafoutas, Tarek Jaber-Lopez:** Ostracism and Theft in Heterogeneous Groups

2021-18 **Zvonimir Bašić, Parampreet C. Bindra, Daniela Glätzle-Rützler, Angelo Romano, Matthias Sutter, Claudia Zoller:** The roots of cooperation

2021-17 **Silvia Angerer, Jana Bolvashenkova, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Children's patience and school-track choices several years later: Linking experimental and field data

2021-16 **Daniel Gründler, Eric Mayer, Johann Scharler:** Monetary Policy Announcements, Information Schocks, and Exchange Rate Dynamics

2021-15 **Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan:** The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

2021-14 **Jeppe Christoffersen, Felix Holzmeister, Thomas Plenborg:** What is Risk to Managers?

2021-13 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Trust in health care credence goods: Experimental evidence on framing andsubject pool effects

2021-12 **Rene Schwaiger, Laura Hueber:** Do MTurkers Exhibit Myopic Loss Aversion?

2021-11 **Felix Holzmeister, Christoph Huber, Stefan Palan:** A Critical Perspective on the Conceptualization of Risk in Behavioral and Experimental Finance

2021-10 **Michael Razen, Alexander Kupfer:** Can increased tax transparency curb corporate tax avoidance?

2021-09 **Changxia Ke, Florian Morath, Anthony Newell, Lionel Page:** Too big to prevail: The paradox of power in coalition formation

2021-08 **Marco Haan, Pim Heijnen, Martin Obradovits:** Competition with List Prices

2021-07 **Martin Dufwenberg, Olof Johansson-Stenman, Michael Kirchler, Florian Lindner, Rene Schwaiger:** Mean Markets or Kind Commerce?

2021-06 **Christoph Huber, Jürgen Huber, and Michael Kirchler:** Volatility Shocks and Investment Behavior

2021-05 **Max Breitenlechner, Georgios Georgiadis, Ben Schumann:** What goes around comes around: How large are spillbacks from US monetary policy?

2021-04 **Utz Weitzel, Michael Kirchler:** The Banker's Oath And Financial Advice

2021-03 **Martin Holmen, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström:** Economic Preferences and Personality Traits Among Finance Professionals and the General Population

2021-02 **Christian König-Kersting:** On the Robustness of Social Norm Elicitation

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

University of Innsbruck

Working Papers in Economics and Statistics

2024-01

Philipp Aschersleben, Julian Granna, Thomas Kneib, Stefan Lang, Nikolaus Umlauf, and Winfried Steiner

Modeling multiplicative interaction effects in Gaussian structured additive regression models

Abstract
Gaussian Structured Additive Regression provides a flexible framework for additive decomposition of the expected value with nonlinear covariate effects and time trends, unit- or cluster-specific heterogeneity, spatial heterogeneity, and complex interactions between covariates of different types. Within this framework, we present a simultaneous estimation approach for highly complex multiplicative interaction effects. In particular, a possibly nonlinear function f(z) of a covariate z may be scaled by a multiplicative effect of the form exp($\tilde{\eta}$), where $\tilde{\eta}$ is another possibly structured additive predictor. Inference is fully Bayesian and based on highly efficient Markov Chain Monte Carlo (MCMC) algorithms. We investigate the statistical properties of our approach inextensive simulation experiments. Furthermore, we apply and illustrate themethodology to an analysis of asking prices for 200000 dwellings in Germany.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)