University of Innsbruck

Working Papers
in
Economics and Statistics

Cost Drivers of Operation Charges and Variation over Time: An Analysis Based on Semiparametric SUR Models

Wolfgang A. Brunauer, Sebastian Keiler
and Stefan Lang

2010-10

Cost Drivers of Operation Charges and Variation over Time: An Analysis Based on Semiparametric SUR Models

Wolfgang A. Brunauer[a], Sebastian Keiler[*,b], Stefan Lang[c]

[a] *Immobilien Rating GmbH, Taborstr. 1–3, 1020 Wien*

[b] *University of Applied Sciences Kufstein, Andreas Hofer-Str. 7, A-6330 Kufstein*

[c] *University of Innsbruck, Universitätsstr. 15, A-6020 Innsbruck*

Abstract

Although building operating charges have turned out to be a major determinant of profitability for real estate investments, there is a noticeable lack of reports or studies that analyze these costs with state-of-the-art statistical techniques. Specifically, past studies usually assume linear relationships between costs and building attributes, they do not control for cluster-specific or longitudinal effects and do not account for the simultaneous structure of cost categories. Therefore, in this study we provide a novel approach to real estate cost benchmarking: We analyze the effects of building attributes on electricity, heating and maintenance costs for office buildings in Germany in a multivariate structured additive regression (STAR) model *simultaneously*, modeling potentially nonlinear effects as P(enalized)-Splines and controlling for cluster-specific and individual heterogeneity in a three-way random effects structure. This way, we gain insights into how building attributes influence costs, and how cost levels vary across cities, companies and buildings. We furthermore derive quality-adjusted time indices for the two major German submarkets, the former German Democratic Republic and the old West German states. The results obtained can be used to derive portfolio allocation strategies and for planning, constructing, operating and redeveloping real estate.

Key words: benchmarking, operating charges, P-Splines, random effects, seemingly unrelated regression, structured additive regression

---

1. Introduction

In an increasingly more competitive market environment, building operating charges (the costs of building provision and management) have turned out to be a major determinant of profitability of real estate investment. This has drawn the attention to the need of identifying suboptimal cost structures in order to exploit potentials for optimization in asset and portfolio management. For this purpose, real estate professionals in the German speaking area largely rely on benchmarking reports such as the annual *Office Service Charges Analysis Report* (OSCAR), [12]. However, most of these reports evaluate operating costs over single building characteristics or regions without controlling for other systematic influences, which is likely to lead to biased results. Linear multiple regression models, where the cost drivers are estimated in a regression analysis of costs against building characteristics, provide a way to cope with these problems. For example, Stoy and Kytzia [18] explain office electricity consumption mainly by characteristics of technical installation, while Chung and Hui [6] regress energy use intensity in office buildings on a set of predictors such as the age of the building, energy system parameters and floor space.

However, the assumption of linearity seems to be too restrictive in many cases, which advocates the use of more flexible,

---

[*] Corresponding author
Email address: `sebastian.keiler1@gmail.com` (Sebastian Keiler)

namely non- and semiparametric regression models. A particularly broad and rich framework of semiparametric models is provided by structured additive regression (STAR) models introduced by Fahrmeir et. al. [10] and Lang and Brezger [5]. In STAR-models, continuous covariate effects are modeled as P(enalized) splines as introduced by Eilers and Marx [9]. Furthermore, such models are also capable of accounting for (possibly correlated) random effects for spatial or cluster indexes, which makes them particularly suitable for our regression situation, where unobserved (a) building-, (b) company-, and (c) city-specific influences may affect cost structures considerably.

Yet, some of the cost categories are likely to be subject to the *same* unobserved influences (e.g. due to the overall building condition, the market environment or public policy). Since neglecting existing association between response variables can lead to inefficient estimation of covariate effects, this article seeks to acquire information about the structure and causes of three major parts of building operating charges *simultaneously* using seemingly unrelated regression (SUR) models. For multivariate Gaussian response, the parametric SUR model (introduced by Zellner [22]) is a standard tool in econometrics. SUR models improve estimation results if disturbance terms are correlated, and if the explanatory covariates differ to some extent between the equations. Astonishingly, there is a distinct lack of non- and semiparametric SUR models. Notable exceptions are the approaches in Smith and Kohn [17] and Lang et al. [13]. In this paper, we rely on the latter, which presents structured additive seemingly unrelated regression modeling. Thus, we allow for simultaneous nonparametric estimation of covariate effects as well as spatial and cluster specific effects, offering an efficient method of identifying the way and degree in which building attributes and cluster specific effects influence operating costs.

As a basis of our study, we use a sample of 699 observations in 56 German cities, collected for benchmarking purposes by CREIS Real Estate Solutions from 2000 to 2005. The cost categories heating, maintenance and electricity are explained by various (continuous and categorical) building attributes as well as random effects for owning companies, cities and repeatedly observed buildings.

The results provide building owners and operators with unbiased benchmarks for operating cost drivers and helps them to identify potentials for optimization in their portfolio. Furthermore, we describe cost developments over time and geographical regions as quality-adjusted price indexes.

The remainder of the paper is organized as follows: In section 2 we describe our structured additive SUR model with respect to our application. In section 3, the data set is described and we set up the working model. Next, we present the results in section 4, and draw some conclusions in the last part of this article.

2. Semiparametric SUR models

As discussed in section 1, we have to consider the following regression situation: We want to explain three response variables simultaneously in a multivariate regression framework with possibly nonlinear effects of continuous covariates. Furthermore, buildings are clustered in cities and companies, and some of them are repeatedly measured. We will discuss the methodological approaches applied for this situation in the following subsections.

2.1. Structured additive regression

The dataset used for this article consists of observations $\mathbf{y}_i = (y_{i1}, \ldots, y_{ik})'$, $i = 1, \ldots, n$, on a $k-$variate response $\mathbf{y}$ (in our case electricity, heating and maintenance costs) and on covariates. We distinguish between a vector $\mathbf{x}_{ir} = (x_{ir1}, \ldots, x_{irp_r})'$ of *continuous* covariates (in our application e.g. net floor area, the age of the building or the time trend) or *cluster indicators* whose influence on the r-th cost category of $\mathbf{y}_i$, will be modeled nonparametrically, and a further vector $\mathbf{v}_{ir} = (v_{ir1}, \ldots, v_{irq_r})'$ of *categorical* covariates (e.g. the existence of an elevator or air condition) whose effect is modeled in the usual linear form with parameters $\boldsymbol{\gamma}_r$. We call a covariate a cluster indicator if it provides information in which city a building is located, which company a building belongs to and which building an observation pertains to. For each cost category y_{ir}, $r = 1, \dots, k$, of the response we assume a semiparametric regression model

$$y_{ir} \;=\; \eta_{ir} \;+\; \varepsilon_{ir}, \quad i = 1, \dots, n, \tag{1}$$

with structured additive predictors

$$\eta_{ir} \;=\; f_{r1}(x_{ir1}) \;+\; \dots \;+\; f_{rp_r}(x_{irp_r}) + \mathbf{v}_{ir}'\boldsymbol{\gamma}_r, \tag{2}$$

where η_{ir} is the predictor of equation r and $f_{r1}, \dots, f_{rp_r}$ are possibly nonlinear functions of the continuous covariates or random effects of the cluster indicators, and $\mathbf{v}_{ir}'\boldsymbol{\gamma}_r$ is the usual linear part of the model.

The errors $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \dots, \varepsilon_{ik})$, $i = 1, \dots, n$, are assumed to be i.i.d. multivariate Gaussian with mean zero and a covariance matrix $\boldsymbol{\Sigma}$, i.e. $\boldsymbol{\varepsilon}_i \,|\, \boldsymbol{\Sigma} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. This implies with $\boldsymbol{\eta}_i = (\eta_{i1}, \dots, \eta_{ik})'$ that

$$\mathbf{y}_i \,|\, \boldsymbol{\eta}_i, \boldsymbol{\Sigma} \sim N(\boldsymbol{\eta}_i, \boldsymbol{\Sigma}), \tag{3}$$

where responses $\mathbf{y}_i$ are conditionally independent, given the predictors $\boldsymbol{\eta}_i$.

The nonlinear functions or random effects f_{rj} in (2) are modeled by a basis functions approach, i.e. a particular f of a covariate x is approximated by a linear combination of K basis or indicator functions

$$f(x) = \sum_{k=1}^{K} \beta_k B_k(x).$$

The B_k are known basis functions and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)'$ is a vector of unknown regression coefficients to be estimated. Typically K is a large number to capture the variability of the data. Overfitting, i.e. highly variable estimators, are avoided by a roughness penalty, that is applied to the regression coefficients. We use quadratic penalties of the form $\lambda\boldsymbol{\beta}'\mathbf{K}\boldsymbol{\beta}$ where $\mathbf{K}$ is a penalty matrix. The penalty depends on a smoothing parameter λ that govern the amount of smoothness imposed on the function f.

From a Bayesian point of view the quadratic penalty $\lambda\boldsymbol{\beta}'\mathbf{K}\boldsymbol{\beta}$ corresponds to a Gaussian (improper) prior for the regressions coefficients $\boldsymbol{\beta}$, i.e.

$$p(\boldsymbol{\beta}|\tau^2) \propto \left(\frac{1}{\tau^2}\right)^{rk(\mathbf{K})/2} \exp\left(-\frac{1}{2\tau^2}\boldsymbol{\beta}'\mathbf{K}\boldsymbol{\beta}\right). \tag{4}$$

Smoothness is now controlled by the variance parameter τ^2 which takes the role of an inverse smoothing parameter.

The choice of basis functions $B_1, \dots, B_K$ and penalty matrix $\mathbf{K}$ depends on our prior assumptions about the smoothness of f as well as the type and dimension of x. In the next subsections we will give specific examples for modeling the f_{rk} or in other words for the choice of basis functions and penalty matrices. We restrict ourselves to examples used in the subsequent analysis. More examples can be found e.g. in Belitz and Lang [2].

2.2. P-Splines

Suppose that a particular covariate x is continuous as is the case for the floor area of the building. We apply the P-splines approach for modeling nonlinear effects of continuous covariates. P-splines have been introduced in a frequentist setting by Eilers and Marx [9] and in a Bayesian framework by Lang and Brezger [14]. P-splines approximate the nonlinear function f by a polynomial spline of degree l with equally spaced knots

$$x_{min} = \kappa_0 < \kappa_1 < \ldots < \kappa_{m-1} < \kappa_m = x_{max}$$

over the domain of x. A spline has the following two properties:

- In each of the intervals $[\kappa_j, \kappa_{j+1}]$, $j = 0, \ldots, m-1$ the spline f is a polynomial of degree l, and
- at the *knots* κ_j (the interval boundaries) the spline is $l-1$ times continuously differentiable.

A spline can be written in terms of a linear combination of $K = m + l$ basis functions, see DeBoor [7]. The most widely used bases are the truncated power series basis and the B-spline basis. Using a truncated power series basis the function f is

$$f(x) = \beta_0 + \beta_1 x + \ldots + \beta_l x^l + \sum_{j=1}^{m-1} \beta_{l+j} t_j(x, l), \tag{5}$$

where

$$t_j(x, l) = (x - \kappa_j)_+^l = \begin{cases} (x - \kappa_j)^l & x > \kappa_j \\ 0 & \text{else.} \end{cases} \tag{6}$$

In a simple regression spline approach, the regression coefficients β_k are estimated using standard inference techniques for linear models. The crucial problem with such regression splines is the choice of the number and the position of the knots. A small number of knots may result in a function space which is not flexible enough to capture the variability of the data. A large number may lead to overfitting. As a remedy Eilers and Marx [9] propose to define a fairly large number of knots (usually between 10 and 40) to ensure enough flexibility. Sufficient smoothness of the fitted curve is achieved through a roughness penalty on the regression coefficients.

Using a truncated power series basis, overfitting is prevented using a quadratic ridge type penalty

$$P(\lambda) = \lambda \sum_{j=1}^{m-1} \beta_{l+j}^2, \tag{7}$$

leading to the penalized least squares criterion

$$PLS(\boldsymbol{\beta}, \lambda) = \sum_{i=1}^{n} (y_i - f(x_i))^2 + \lambda \sum_{j=1}^{m-1} \beta_{l+j}^2 \tag{8}$$

to be minimized with respect to $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_{K-1})'$.

Despite their simplicity P-splines based on a truncated power series basis in combination with penalty (7) are rarely used in practice, due to the numerical instability of the highly collinear basis functions. Instead a local B-splines basis is applied. There is a close relationship between B-splines and truncated polynomials as B-splines can be computed as differences of truncated powers (see Eilers and Marx [9]). For instance B-spline basis functions of degree one are computed as

$$B_j(x, 1) = t_{j-2}(x, 1) - 2t_{j-1}(x, 1) + t_j(x, 1) = \Delta^2 t_j(x, 1),$$

with t_j defined in (6). Note that extra knots $\kappa_{-l}, \ldots, \kappa_{-1}$ left to κ_0 and $\kappa_{m+1}, \ldots, \kappa_{m+l}$ right to κ_m are required, so that the truncated polynomials in the above formula are properly defined to compute all basis functions B_j close to the left and right borders. Now the spline f may be written as

$$f(x) = \sum_{k=1}^{K} \beta_k B_k(x, l).$$

The local basis also gives rise to alternative penalization. The widely used approach by Eilers and Marx [9] penalizes the sum of squared d-th order differences

$$P(\lambda) = \lambda \sum_{k=d+1}^{K} \left(\Delta^d \beta_k\right)^2 = \lambda \boldsymbol{\beta}' \mathbf{K} \boldsymbol{\beta} \tag{9}$$

were Δ^d is the difference operator of order d. The penalty matrix is given by $\mathbf{K} = \mathbf{D}'\mathbf{D}$ where $\mathbf{D}$ is a d-th order difference matrix. The default for d in most implementations (e.g. `mgcv` in R or the software package BayesX, see below) is $d = 2$.

2.3. Varying coefficients

Of course, we can also model interaction effects in this framework. Specifically, we are interested in heterogeneous nonlinear time trends for the two main German submarkets, the old West German states (henceforth Western Germany) and the former German Democratic Republic (henceforth GDR). Suppose that x_{i1} is a time index and v_{i1} is an indicator for GDR in effect coding, i.e. $v_{i1} = 1$ if the i-th observed building is located in GDR and $v_{i1} = -1$ else. We use effect rather than dummy coding for technical reasons. Now we arrive at a model of the form

$$y_i \;=\; f_1(x_{i1}) \;+\; \ldots \;+\; f_{x_1|v_1}(x_{i1})v_{i1} \;+\; \ldots \;+\; v_{i1}\gamma_1 \;+\; \ldots, \tag{10}$$

where $f_1(x_{i1})$ and $\gamma_1 v_1$ are the main effects of the time index x_{i1} and the effect of GDR v_{i1} and $f_{x_1|v_1}$ is the smooth nonparametric function of the interaction term (we drop the equation index for simplicity). The function f_1 can be interpreted as an average time trend of x_{i1}, and $f_{x_1|v_1} + \gamma_1$ respectively $-f_{x_1|v_1} - \gamma_1$ is now the deviation from f_1 for $v_{i1} = 1$ (building is located in the former GDR) and $v_{i1} = -1$ (building is located in Western Germany). Summarizing we obtain the following decomposition of the time trend:

$$f(x_{i1}) \;=\; \begin{cases} f_1(x_{i1}) \;+\; f_{x_1|v_1}(x_{i1}) + \gamma_1 & \text{if building } i \text{ is situated in the former GDR} \\ f_1(x_{i1}) \;-\; f_{x_1|v_1}(x_{i1}) - \gamma_1 & \text{if building } i \text{ is situated in Western Germany.} \end{cases} \tag{11}$$

2.4. Cluster Effects

Suppose now that x is a unit- or cluster index variable such as the company, city or building index in our analysis. There is a vast literature on modeling unit- or cluster specific heterogeneity, see e.g. Verbeke and Molenberghs [19]. An important special case arises for longitudinal data where individuals are repeatedly observed over time (Diggle et al. [8]). Typically, unit- or cluster specific random effects are introduced to account for heterogeneity. In its simplest form an $i.i.d.$ random intercept β_x with

$$\beta_x \sim N(0, \tau^2) \tag{12}$$

is introduced. Here, $x \in \{1, \ldots, K\}$ is an index variable that denotes the cluster a particular observation pertains to. This is equivalent to a penalized least squares approach with function $f(x) = \beta_x$ and penalty matrix $\mathbf{I}$ or a Bayesian approach with prior (12).

Note that more than one random intercept with respect to different cluster variables are possible, as in our application (we specify a company-, city- and building-specific random effect).

2.5. Inference and Software

The semiparametric SUR model is estimated in a Bayesian framework using Markov chain Monte Carlo (MCMC) simulation techniques for inference. Details can be found in Lang et al. [13]. The approach is implemented in the software package BayesX which is publicly available at

http://www.stat.uni-muenchen.de/ bayesx/bayesx.html,

see Brezger et al. [3] and [4]. The homepage of BayesX contains also a number of tutorials.

The following code fragment exemplifies the usage of BayesX in the context of semiparametric SUR modeling:

```
dataset costs;
costs.infile using c:\data\costs.raw;
bayesreg sur_model;

sur_model.mregress
ln_elect_psqm =  age(psplinerw2) + ... + company(random) + ... + quality_h + quality_m:
ln_heat_psqm = age(psplinerw2) + ... + company(random) + ... + quality_h + quality_m:
ln_maintenance_psqm = age(psplinerw2) + ... + company(random) + ... + quality_h + quality_m,
       family=multgaussian iterations=12000 step=10 burnin=2000 using costs;
```

Having defined the dataset object `costs`, we read the cost data from an ASCII file. We then define the `bayesreg`-object and apply the method `mregress` to fit our multivariate gaussian (`multgaussian`) model with structured additive predictor. In all of our three equations we have P-splines with second order difference penalties (`psplinerw2`), unstructured random company-, city- and building-specific effects (`random`) and linear effects. Furthermore, we define the number of MCMC-`iterations` and the number of `burnin`-iterations thereof as well as the thinning parameter for the MCMC simulation, `step`. See the tutorial on Bayesian inference via MCMC simulation available at the BayesX homepage for details.

3. Data description and model specification

3.1. Data

The data this research is based on, has been collected for the annual OSCAR report [12]. For this report, data concerning floor areas, technical quality and equipment and operating charges of office buildings are carefully captured and presented in great detail for comparison purposes on a yearly basis. Total operating charges consist of 10 cost categories, which are depicted in figure 1. We have chosen to analyze electricity, heating and maintenance costs in this article (a total of 699 observations dispose of these cost categories) mainly for the following reasons:

- As these cost categories mainly depend on the same (possibly unobserved) "technical" influences, they are likely to be correlated, which makes a SUR model appropriate (which is not necessarily true for the other cost categories). The other cost categories do not depend on the same influences as electricity, heating and maintenance costs and therefore do not add to the quality of a SUR-model. Specifically the costs for administration, security, janitors and cleaning are largely dependent on staff costs and driven by different shocks than the chosen categories.
- Electricity, heating and maintenance costs are relevant in a sense that they account for approximately one third of total operating costs, while security, water and insurance have relatively small shares of total costs (see figure 1).
- Large deviations from the expected costs point to inefficiencies in building equipment, operation or user behavior and are a starting point for an optimization process. In contrast, influencing the cost categories public charges and administration is not that straightforward and therefore of less practical interest.

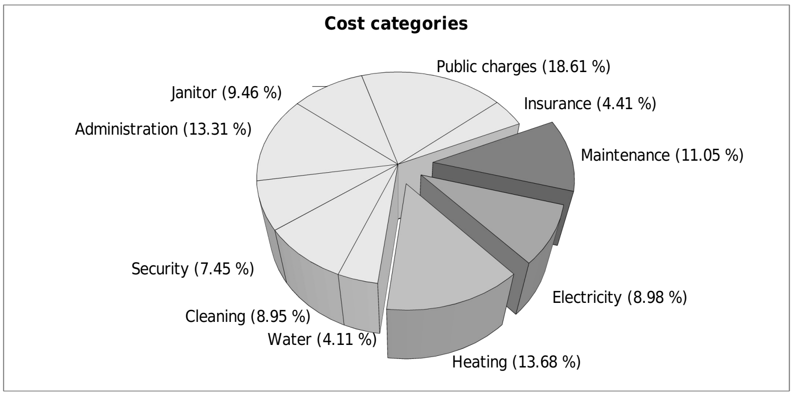

Figure 1: *All ten cost categories and their share of total costs. The three analyzed categories are highlighted.*

The data were collected for the years 2000 to 2005, and some of the buildings have been observed repeatedly. As table 1 shows, one out of four buildings has been observed more than once, one building even over the whole period. This leads to a panel structure of 501 buildings. Furthermore, observations are situated in cities on the one hand and owned by companies on the other hand, both of which are likely to affect cost structures. In table 2, the numbers of unique values for these cluster categories are displayed together with minimum, 25 % quantile, median, 75 % quantile and maximum of observations per cluster. Note that for both categories, the smallest group contains only one observation, while there is one city that contains 88 observation and one company to which 277 observations (nearly half of all observations) pertain.

No. of years	No. of buildings	Freq.
1	365	72.85
2	94	18.76
3	29	5.79
4	7	1.4
5	5	1
6	1	0.2
Total	501	100

Table 1: *Multiply observed objects. Most of the buildings (73 %) have been observed only once, while one building has been observed over the whole period.*

	unique values	min	p25	median	p75	max
Cities	56	1	2	4.5	11	88
Companies	21	1	4	9	31	277

Table 2: *Description of cluster categories* cities *and* companies*: Number of groups (unique values) per category along with minimum, 25 % quantile, median, 75 % quantile and maximum of observations* per group *that are observed in the respective category. The size of the groups varies considerably, from only one observation to 88 observations for cities, and one to 277 observations for companies.*

All buildings are characterized by continuous covariates such as the net floor area of the building (*net*), integer variables such as the age at the time of data collection (*age*), the time of measurement (*time*) and the number of floors of the building (*floors*), as well as categorical variables such as the quality of the building, the existence and quality of air condition, an indicator for a past refurbishment, identifying whether the building has been built in row with other buildings and the existence of a garage and an elevator. Furthermore, in order to capture location specific effects, the data has been matched with data on climate conditions (heating degree days, *hdd*—obtained from the German Weather Service; heating degree days are computed as the sum of days with an outside temperature below a threshold of 15°C, weighted with the difference between the actual outside temperature and this threshold; see the respective norm from the German Institute for Standardization, DIN 4108-6 [1] and the guideline from the Association of German Engineers, VDI 3807 [20]) as well as data on economic strength (standardized purchasing power, *pps* (provided form the Michael Bauer Research GmbH). Table 5 in appendix **??** gives an overview over the covariates employed in our multivariate regression model.

3.2. Model

The model we set up using the covariates described in section 3 can thus be described as follows:

$$\begin{aligned}
ln(elect) &= f_1\,(age) + f_2\,(floors) + f_3\,(time) + f_4\,(time)\,e\_w + \\
&\quad f_5\,(company) + f_6\,(city) + f_7\,(building) + \boldsymbol{v}_1'\boldsymbol{\gamma}_1 + \varepsilon_1 \qquad (13)\\
ln(heat) &= f_1\,(age) + f_2\,(floors) + f_3\,(net) + f_4\,(hdd) + f_5\,(time) + f_6\,(time)\,e\_w + \\
&\quad f_7\,(company) + f_8\,(city) + f_9\,(building) + \boldsymbol{v}_2'\boldsymbol{\gamma}_2 + \varepsilon_2 \qquad (14)\\
ln(maint) &= f_1\,(age) + f_2\,(floors) + f_3\,(net) + f_4\,(pps) + f_5\,(time) + f_6\,(time)\,e\_w + \\
&\quad f_7\,(company) + f_8\,(city) + f_9\,(building) + \boldsymbol{v}_3'\boldsymbol{\gamma}_3 + \varepsilon_3 \qquad (15)
\end{aligned}$$

In all equations, continuous and integer covariates are modeled as P-Splines, the categorical covariates describing the quality and condition of the building are encoded as dummy variables and subsumed in $\boldsymbol{v}_r$ with estimated parameters $\boldsymbol{\gamma}_r$. Furthermore, we included time trends with possible interactions with the two German submarkets (former GDR, Western Germany) as described in section 2.3.

We transform the response variables logarithmically, as we expect building characteristics to have multiplicative effects on operating cost categories (this topic has been discussed extensively in the context of hedonic price theory, see Malpezzi [15] for an overview). Due to these transformations, the estimated effects can be approximately interpreted as semi-elasticities (i.e. the percentage change of price by the absolute change of the covariate, see Greene [11]). However, for large values, this approximation becomes too rough, so we calculate some effects explicitly (specifically, the dummy effects of quality, air condition, elevator and the indicator for the former GDR, see table 4). The percentage cost effect of dummy v_j is then $\Delta_j(costs) = [exp(\gamma_j) - 1] \times 100$, where γ_j is the estimated parameter.

Based on theoretical considerations, we expect the following results: For all of the cost categories examined, we expect the age of the building to have an increasing effect due to the decreasing technical and structural condition. Also an increasing number of floors is likely to lead to higher electricity and maintenance costs, considering the additional equipment for higher buildings (e.g. elevators, pumps, fans etc.) and heating costs (high buildings tend to have unfavorable ratios of envelope to volume). The net floor area is only relevant for heating and maintenance costs (equations (14) and (15)), where we assume that large buildings may benefit from economies of scale, leading to a negative effect in both equations. The heating degree days only appear in equation (14), where a larger number of heating degree days should lead to an increase in heating energy consumption. We employ the purchase power as a predictor for maintenance costs in equation (15), as in cities with a high purchase power labor costs are also high, which is likely to affect this labor intensive cost category. For all cost categories, we expect costs to increase over the years.

Turning to the linear effects, the existence of air condition should be a major predictor both for energy and maintenance costs. Furthermore, refurbished buildings may tend to have lower costs in general. Buildings in a row should exhibit lower heating costs, as the outside envelope area is reduced, and the existence of an elevator most likely leads to higher electricity and maintenance costs. We do not have any prior assumptions concerning the quality of the building and the existence of a garage. We expect the effect of the indicator for the former GDR to capture lower factor prices in this part of Germany, leading to lower costs.

4. Results

In this section we present the results for the model specified in section 3.2. First, effects are presented for each cost category separately. Next, we will analyze cluster specific effects, and finally the variation over time is explored. Table 3 presents the estimated error correlation structure of our three equation model. The strongest correlation is observed for the energy dominated categories heating and electricity (although there is also a notable correlation between the other equations).

	elect	heat	maint
elect	1		
heat	0.210	1	
maint	0.183	0.124	1

Table 3: *Estimated error correlations between the equations.*

	Electricity	*Heating*	*Maintenance*
intercept	2.473	1.754	1.300
	***(0.2877)	***(0.0378)	***(0.2887)
refurbished	−0.054	−0.037	0.052
	(0.0631)	(0.0399)	(0.0667)
garage	0.032		0.020
	(0.0619)		(0.0656)
quality low	−0.566	−0.130	−0.353
	***(0.1071)	***(0.0578)	***(0.1054)
quality medium	−0.257	−0.103	−0.138
	***(0.0725)	***(0.0397)	***(0.0682)
no air condition	−0.747		−0.289
	***(0.0959)		***(0.0937)
partially air conditioned	-0.342		−0.201
	***(0.0931)		***(0.0933)
elevator	−0.389		0.390
	**(0.2126)		***(0.2106)
vacancy	−0.031	−0.009	−0.038
	(0.0546)	(0.0309)	(0.0540)
row		−0.036	
		(0.0355)	
e_w_dummy	0.118	0.155	0.129
	(0.1739)	**(0.091)	(0.2266)
outlier	−1.329		
	***(0.5356)		

Table 4: *Results for binary covariates of the three equations. The table shows the posterior means along with the posterior standard deviation (in brackets). *** indicate a credible interval of 95 % not including zero, ** indicate a credible interval of 80 % not including zero. The reference is a high quality fully air conditioned building.*

4.1. Electricity costs

The linear effects for the electricity cost equation are displayed in the first column of table 4. Electricity costs per square meter rise on average by 100 % for buildings with a full air condition while partly air conditioned buildings still have 40 % higher costs, both compared to buildings without air condition.[1] This magnitude is largely in line

---

[1] Note that we employed the $\mathit{\Delta}_{ij}(costs) = [exp(\gamma_j \times v_{ij}) - 1] \times 100$ approximation for obtaining the percentage value of the effect.

with former publications described in the VDI 3807 [21]. The negative sign for the binary variable elevator may be counterintuitive at first, as additional technical equipment should lead to higher costs. Yet, considering the fact that all but nine buildings in the sample have an elevator, this may rather be seen as a proxy for sub-standard objects with—in this case—clearly higher electricity costs. There is no evidence that the presence of a garage has any influence on the electricity costs per square meter. Furthermore neither vacancy within the building nor a previous refurbishment seem to have any effect on these costs. Note that the indicator for buildings in the former GDR is not significant. We will explore this issue further in section 4.5.

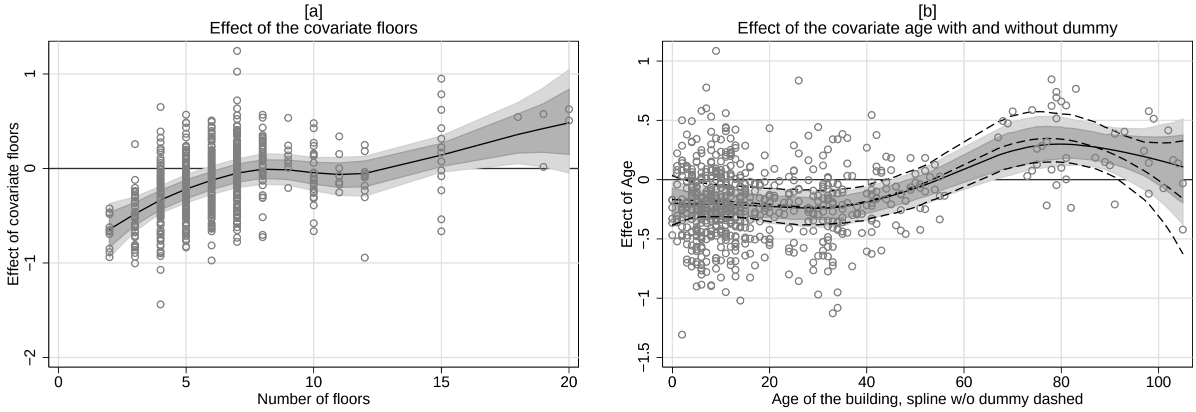

Figure 2: *Effects of continuous covariates for electricity costs, overlaid with partial residuals. Panel [a] variable floors, panel [b] variable age. Note that the dashed line indicates the posterior mean as well as the 95 % credible interval for the case without the dummy for building number 58.*

Figure 2 depicts the effects of nonparametrically modeled continuous covariates (centered about zero for identifiability). Both, the number of floors and the age of the building have a positive effect on electricity costs per square meter. The effect of the number of floors is strongly upward sloping with a consolidation phase between the eighth and the twelfth floor; for buildings with more than twelve floors, there is again an increasing effect (interestingly, this pattern corresponds with the German building code definition of high-rise buildings). In total, heating costs *ceteris paribus* differ by 100 % between a two and a 20 storey building (the shape of this effect is in line with our prior assumptions, see section 3.2). The age of the building in the right panel of this figure has a negligible effect on electricity costs up to an age of approximately 40 years, followed by a steady increase up to 70 years. This suggests that old buildings tend to exhibit an inferior building equipment standard, resulting in sub-optimal energy efficiency and thus higher energy consumption. However, there is a notable decrease of this effect in the very high range of this covariate. The reason for this surprising pattern is that the oldest building in the dataset, which happens to exhibit rather low energy costs, has been observed in three consecutive years. As data in this age group is rather sparse, this leads to a substantial drop in the age effect towards the end of the spline indicated by the dashed lines in figure 2 panel [b]. We control for this outlier beyond the random effect by introducing a dummy variable (alternatively, we could have just dropped it from the dataset). This considerably changes the cost gradient and yields a more reliable estimate, as depicted in the right panel of figure 2.

4.2. Heating costs

Heating costs are largely driven by the building's (technical and constructive) quality, which becomes obvious from the second column of table 4, although the effect is considerably smaller than in the case of electricity costs (in this case minus 13 % and minus 10 % respectively). Neither a refurbishment nor a detached construction type have a significant impact on heating costs.

Figure 3 again illustrates the effects of nonparametrically modeled covariates. Although the shape of the effect of covariate *floor* in panel [a] resembles the one in the case of electricity costs, its impact is smaller in this case. Panel

[b] shows the effect of the age of the building continuously increasing up to 45 years and a constant effect afterwards, indicating that new buildings *ceteris paribus* have 56 % lower heating costs than a buildings built during the 1960s. This gradient portrays the changing technical standards and building codes. While improved construction methods allowed for thin walls in the sixties and beyond, buildings before this time period needed relatively thick walls, which (unintendedly) contribute to thermal performance. The net floor area, depicted in panel [c], has a negative effect on costs, suggesting the presence of economies of scale (although with a decreasing marginal effect). This is again in line with theoretical assumptions, since larger heating systems usually work more efficiently. The effect of the climate related covariate heating degree days (panel [d]) is not significant in the sense that pointwise confidence intervals cover the 0-line over the whole range of this covariate. This may be due to the fact that buildings in cooler areas are better insulated, thus endogenously removing any effect resulting in this observed pattern.

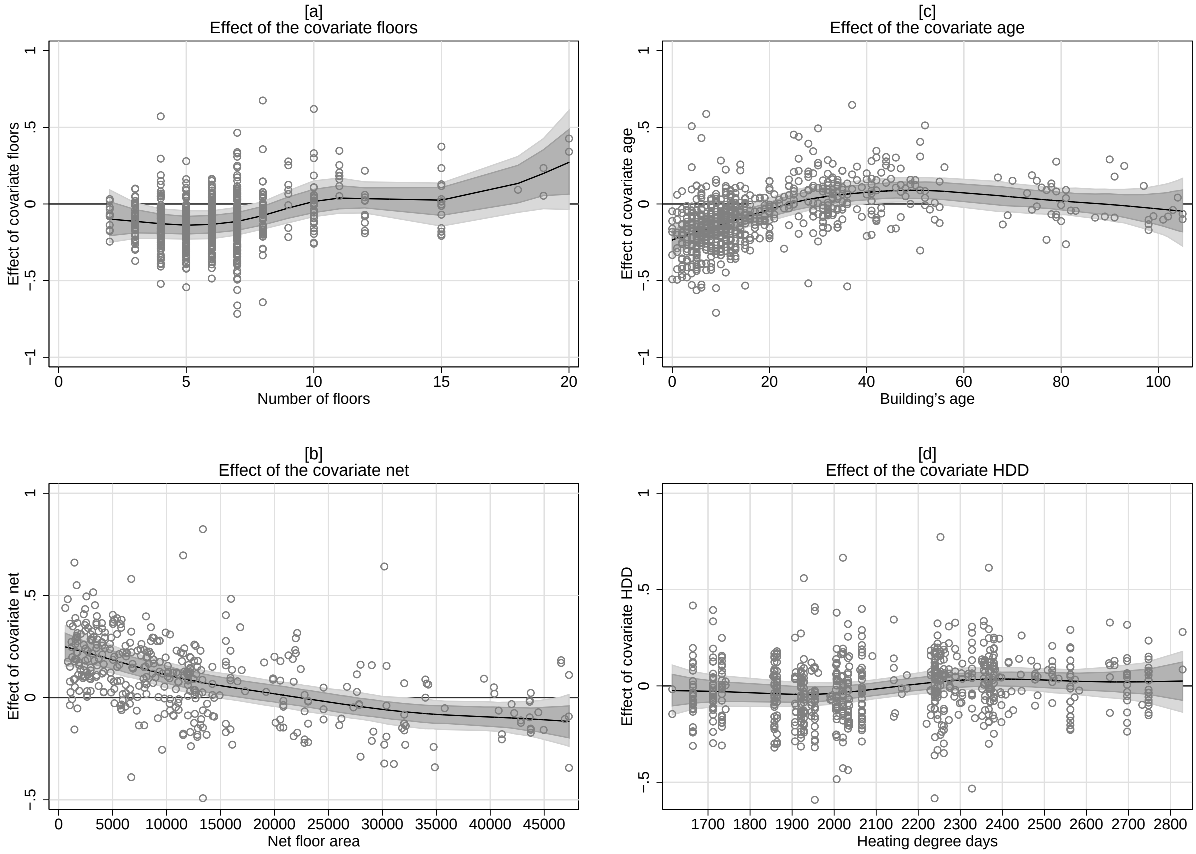

Figure 3: *Effects of continuous covariates of category heating costs, overlaid with partial residuals. Panel [a] shows the effect of the number of floors, panel [b] plots the effect of the variable age, panel [c] illustrates the gross internal floor area effect. Panel [d] shows the spline for the heating degree days.*

4.3. Maintenance costs

Also for maintenance costs, technical equipment (air condition) as well as building quality are highly relevant. Maintenance costs rise by 25 % for fully air conditioned buildings compared to no air condition and 18 % compared to partial air condition. A high building quality results in 29 % higher maintenance costs compared to low and 13 % compared to medium quality. The significant positive effect for the elevator dummy is in line with theoretical considerations.

The nonparametric effects in figure 4 reveal the rather strong effect of the covariate *floors* (panel [a]) with 34 % difference between the highest and the lowest building in the sample. Furthermore, almost a linear gradient for covariate net floor area (panel [b]) is observable, again indicating considerable economies of scale (an increase of floor space by 10,000 square meters leads to a reduction in maintenance costs of roughly 7 %). In panel [c], the effect of

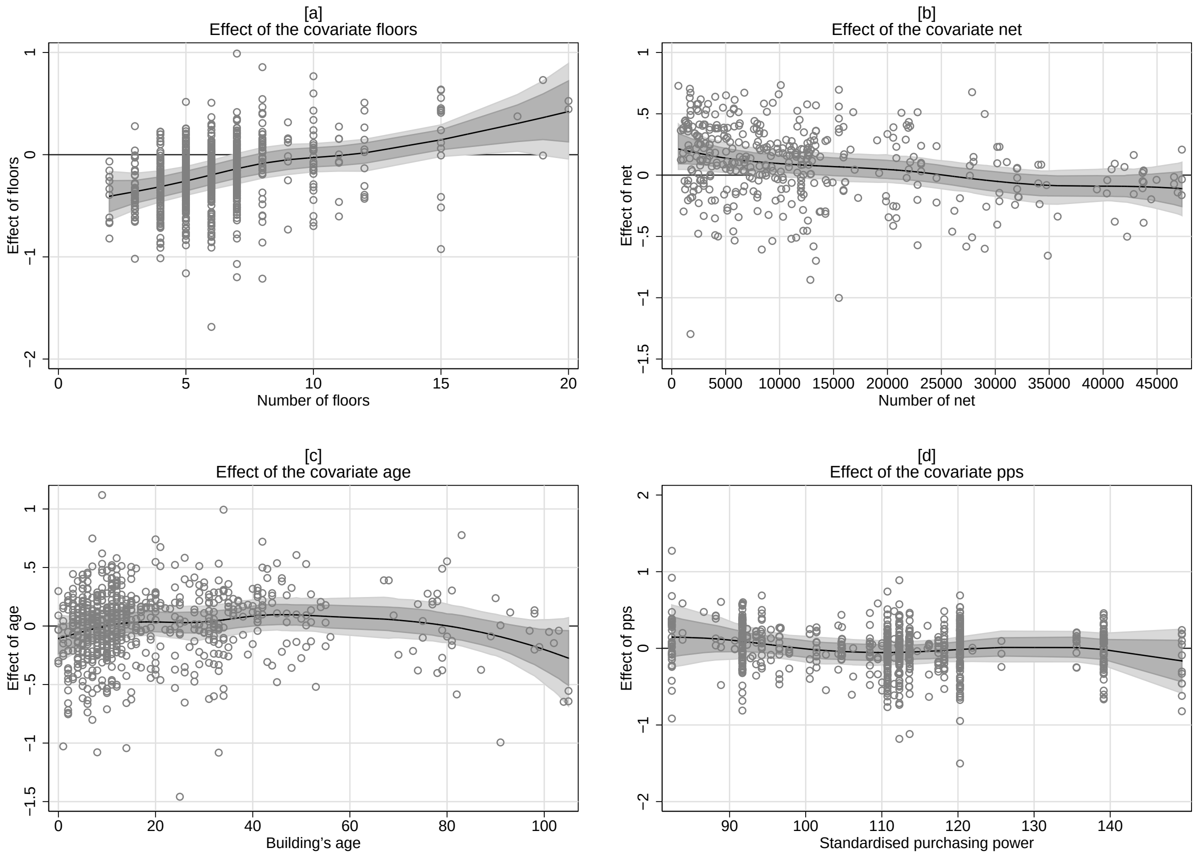

Figure 4: *Effects of continuous covariates of category maintenance costs, overlaid with partial residuals. Panel [a] shows the effect of the number of floors, panel [b] illustrates the gross internal floor area effect, panel [c] plots the effect of the variable age. Panel [d] shows the spline for the standardized purchasing power.*

covariate *age* is insignificant over the whole range of values. However, the shape suggests a (weak) quadratic cost gradient. Reasons for this may be that either there is a lesser amount of maintainable equipment in older buildings in general, or older buildings have already been equipped with new technique as the technical lifespan of the original equipment has already expired. In order to account for local market structures, the purchasing power index on NUTS-3 level has been introduced into this equation (differences in economic power affect wages and therefore maintenance costs). However, as can be seen in panel [d], this effect is insignificant as well, indicating a common price level in all areas. This may be due to an increasing competition and a concentration of service suppliers (one notable example is the fusion of STRABAG and the Facility Management department of the German Telekom).

4.4. Random effects

Figure 5 illustrates the random effects for all three models. Recall that there are three types of random effects, the effect of the city, the company and a longitudinal effect for repeatedly observed buildings. Figure 5 shows the results graphically, where each panel depicts the effects of one cost category. In order to get an impression of the dimension of the random effects, they are contrasted with the residuals.

In general, city specific random effects account for a relatively small difference, while the building specific effects are quite pronounced, indicating the large relevance of user behavior: Recent analyses e. g. by Messerschmidt [16] estimate user induced fluctuation in a building's energy demand at around 50 %. However, also the company specific random effects indicate optimization potentials on a company level. In total, the random effects for category heating costs (panel [b]) are smallest (we also observe the smallest standard error of regression for this equation).

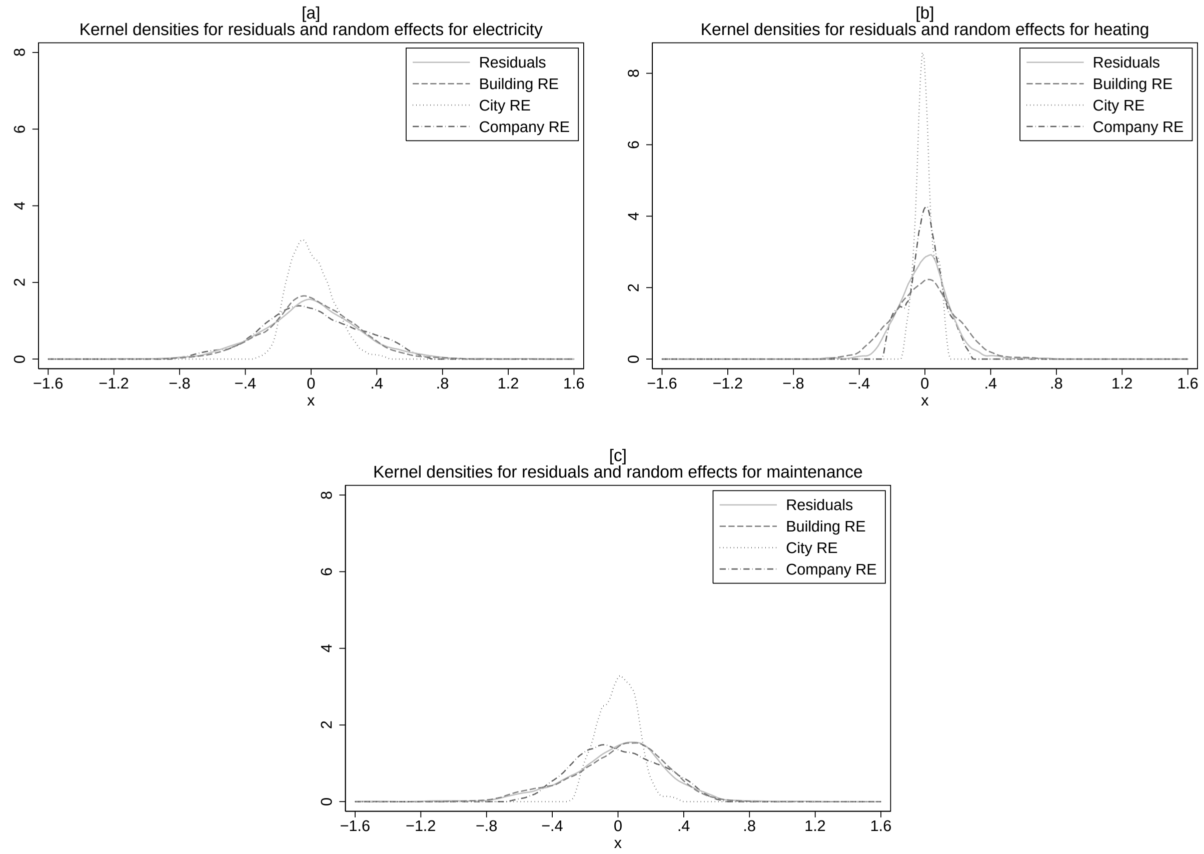

Figure 5: *Kernel density estimates for the models' residuals and the random effects. The company specific random effects are illustrated as a histogram. The panels show the models electricity [a], heating [b] and maintenance [c].*

4.5. Variation over Time

One of the major goals of real estate benchmarking is the revelation of developments over time. Thefore, we present quality adjusted indexes for the cost categories under consideration, reflecting the heterogeneous market structure in Germany by splitting the effect between Western Germany and the former GDR as explained in section 2. Figure 6 illustrates the results of this procedure. Note that in 2002, electricity costs in the former GDR started decreasing considerably (panel [a]), while heating and maintenance costs continued to rise, although more slowly (panels [b] and [c]). Interestingly, this pattern is also present in the German Gross Domestic Product (panel [d]). The fact that price increases in Western Germany are not affected by this in any way might indicates systematic differences in market structure between the former GDR and Western Germany: The pattern observed may give evidence for a higher price elasticity in the former GDR during phases of reduced growth.

5. Conclusions

The main goal of this article is to provide a cutting edge approach to real estate cost benchmarking for electricity, heating and maintenance costs that still provides useful results for practitioners. There are three challenges that make this regression problem particularly difficult: First, possibly nonlinear covariate effects prevent the use of purely linear models. Second, unobserved cluster-specific or individual heterogeneity (specifically, for the companies owning the buildings, the cities where they are located and longitudinal effects) must be accounted for in order to get unbiased results. And finally, as the three cost categories analyzed are likely to be simultaneously influenced by common shocks, a Seemingly Unrelated Regression (SUR) model should be applied, providing more efficient estimation results. Therefore,

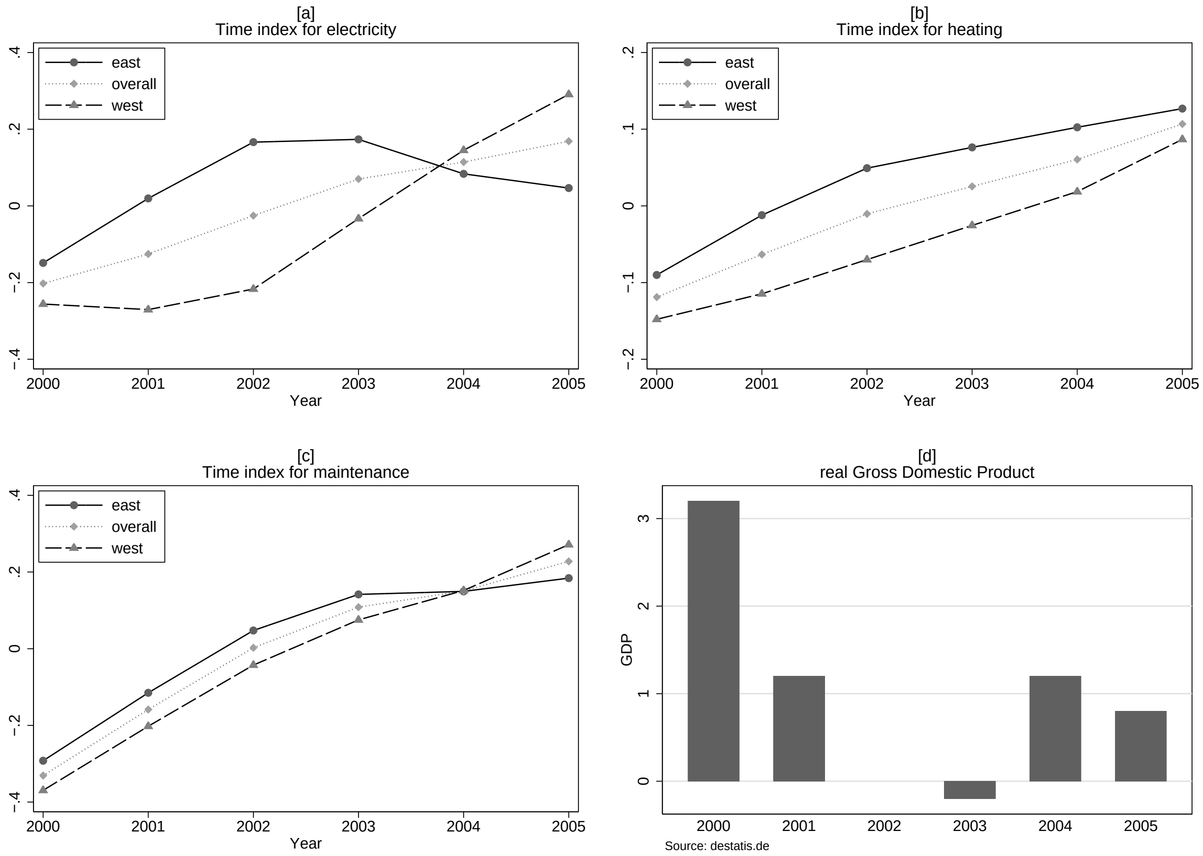

Figure 6: *The variation over time split up for the former GDR (solid) and Western Germany (dashed). The dotted line indicates the aggregated index. Panels [a–c] show the three cost categories, panel [d] shows the German Gross Domestic Product (GDP) for the time span.*

we propose a novel approach based on three components: Potentially nonlinear effects of cost drivers are modeled as P(enalized) Splines; a three-way random effects structure accounts for unobserved cluster-specific heterogeneity, leading to unbiased results and providing the basis for further exploration of the data; and finally, contemporaneous building-specific correlation between cost categories is modeled by a SUR-approach. This concept proves to be superior to conventional benchmarking attempts, providing insights into cost structures and cost developments over time and between markets. Our findings may help to anticipate operating charges *ex ante*, which can help to derive investment strategies and identify opportunities. On the other hand, the costs of individual buildings can be benchmarked *ex post*, providing insights into potentials for cost reductions.

Acknowledgements

The authors gratefully acknowledge financial support by the *Tiroler Wissenschaftsfonds* (Tyrolean Science Fund). Parts of this study result from the *Haus der Zukunft* (Future Buildings) project for Life Cycle Cost prediction at the Kufstein University of Applied Sciences.

References

[1] Deutsches Institut für Normung eV, *DIN 4108 T6 Wärmeschutz in Gebäuden.*

[2] C. Belitz and S. Lang, *Simultaneous selection of variables and smoothing parameters in structured additive regression models*, Computational Statistics and Data Analysis, 53 (2008), pp. 61–81.

[3] A. Brezger, T. Kneib, and S. Lang, *BayesX: Analyzing Bayesian Structural Additive Regression Models*, Journal of Statistical Software, 14 (2005), pp. 1–22.

[4] ——, *BayesX manuals*, 2009.

[5] A. Brezger and S. Lang, *Generalized structured additive regression based on Bayesian P-splines*, Computational Statistics & Data Analysis, 50 (2006), pp. 967–991.

[6] W. Chung and Y. Hui, *A study of energy efficiency of private office buildings in Hong Kong*, Energy and Buildings, 41 (2009), pp. 696–701.

[7] C. DeBoor, *A Practical Guide to Splines (Applied Mathematical Sciences)*, Springer, 2001.

[8] P. J. Diggle, P. Heagerty, K.-L. Liang, and S. L. Zeger, *Analysis of Longitudinal Data*, Oxford University Press, Oxford, 2 ed., 2002.

[9] P. H. Eilers and B. D. Marx, *Flexible smoothing with B-splines and penalties*, Statistical Science, 11 (1996), pp. 89–121.

[10] L. Fahrmeir, T. Kneib, and S. Lang, *Penalized additive regression for space–time data a bayesian perspective*, Statistica Sinica, 14 (2004), pp. 731–761.

[11] W. H. Greene, *Econometric analysis*, Prentice Hall Internat., London, 5. ed., in ed., 2003.

[12] Jones Lang LaSalle, *OSCAR Service Charges Analysis Report*, 2010.

[13] S. Lang, S. Adebayo, L. Fahrmeir, and W. Steiner, *Bayesian Geoadditive Seemingly Unrelated Regression*, Computational Statistics, 18 (2003), pp. 263–292.

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

[15] S. Malpezzi, *Hedonic Pricing Models: A Selective and Applied Review*, in Housing Economics and Public Policy: Essays in Honor of Duncan Maclennan, T. O'Sullivan and K. Gibb, eds., Blackwell Science Ltd, 2003, pp. 67–89.

[16] J. Messerschmidt, *Sparer und Verschwender gibt es überall. Einfluss des Gebäudestandards und des Nutzerverhaltens auf die Heizkosten*, TGA Fachplaner, 3 (2004), pp. 18–21.

[17] M. Smith and R. Kohn, *Nonparametric Seemingly Unrelated Regression*, Journal of Econometrics, 98 (2000), pp. 257–281.

[18] C. Stoy and S. Kytzia, *Benchmarking electricity consumption*, Construction Management and Economics, 24 (2006), pp. 1083–1089.

[19] G. Verbeke and G. Molenberghs, *Linear Mixed Models for Longitudinal Data*, Springer, New York, 2000.

[20] Verein Deutscher Ingenieure, *VDI 2067 Raumheizung: Berechnung der Kosten von Wärmeversorgungsanlagen.*

[21] ——, *VDI 3807 Energie- und Wasserverbrauchskennwerte für Gebäude – Grundlagen.*

[22] A. Zellner, *An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias*, Journal of The American Statistical Association, 57 (1962), pp. 348–368.

A. Variable definition

Name	Explanation	elect	heat	maint
refurbished	building has been refurbished (binary)	x	x	x
quality	Building's quality in three categories: high (*reference category*), medium or low	x	x	x
air condition	Building's type air condition in three categories: no air condition, partially and fully (*reference category*) air conditioned	x		x
row	building is detached to another building (binary)		x	
garage	building has a garage (binary)	x		
e_w_dummy	building is located in the former GDR (binary)	x	x	x
vacancy	building has reported vacancy (binary)	x	x	x
elevator	building has an elevator (binary)	x		x
net	Net floor area		x	x
floors	Number of floors	x	x	x
age	Building's age	x	x	x
pps	Purchasing Power Standardized of the city			x
hdd	Heating degree days of the region		x	

Table 5: *Variable description and model summary.*

University of Innsbruck – Working Papers in Economics and Statistics
Recent papers

2010-10 **Wolfgang A. Brunauer, Sebastian Keiler and Stefan Lang:** Trading strategies and trading profits in experimental asset markets with cumulative information

2010-09 **Thomas Stöckl and Michael Kirchler:** Trading strategies and trading profits in experimental asset markets with cumulative information

2010-08 **Martin G. Kocher, Marc V. Lenz and Matthias Sutter:** Psychological pressure in competitive environments: Evidence from a randomized natural experiment: Comment

2010-07 **Michael Hanke and Michael Kirchler:** Football Championships and Jersey Sponsors' Stock Prices: An Empirical Investigation

2010-06 **Adrian Beck, Rudolf Kerschbamer, Jianying Qiu and Matthias Sutter:** Guilt from Promise-Breaking and Trust in Markets for Expert Services - Theory and Experiment

2010-05 **Martin Gächter, David A. Savage and Benno Torgler:** Retaining the Thin Blue Line: What Shapes Workers' Intentions not to Quit the Current Work Environment

2010-04 **Martin Gächter, David A. Savage and Benno Torgler:** The relationship between Stress, Strain and Social Capital

2010-03 **Paul A. Raschky, Reimund Schwarze, Manijeh Schwindt and Ferdinand Zahn:** Uncertainty of Governmental Relief and the Crowding out of Insurance

2010-02 **Matthias Sutter, Simon Czermak and Francesco Feri:** Strategic sophistication of individuals and teams in experimental normal-form games

2010-01 **Stefan Lang and Nikolaus Umlauf:** Applications of Multilevel Structured Additive Regression Models to Insurance Data

2009-29 **Loukas Balafoutas:** How much income redistribution? An explanation based on vote-buying and corruption. *Revised version forthcoming in Public Choice.*

2009-28 **Rudolf Kerschbamer, Matthias Sutter and Uwe Dulleck:** The Impact of Distributional Preferences on (Experimental) Markets for Expert Services

2009-27 **Adrian Beck, Rudolf Kerschbamer, Jianying Qiu and Matthias Sutter:** Car Mechanics in the Lab - Investigating the Behavior of Real Experts on Experimental Markets for Credence Goods

2009-26 **Michael Kirchler, Jürgen Huber and Thomas Stöckl:** Bubble or no Bubble - The Impact of Market Model on the Formation of Price Bubbles in Experimental Asset Markets

2009-25 **Rupert Sausgruber and Jean-Robert Tyran:** Tax Salience, Voting, and Deliberation

2009-24 **Gerald J. Pruckner and Rupert Sausgruber:** Honesty on the Streets - A Natural Field Experiment on Newspaper Purchasing

2009-23 **Gerlinde Fellner, Rupert Sausgruber and Christian Traxler:** Testing Enforcement Strategies in the Field: Legal Threat, Moral Appeal and Social Information

2009-22 **Ralph-C. Bayer, Elke Renner and Rupert Sausgruber:** Confusion and Reinforcement Learning in Experimental Public Goods Games

2009-21 **Sven P. Jost:** Transfer Pricing Risk Awareness of Multinational Corporations - Evidence from a Global Survey

2009-20 **Andrea M. Leiter and Engelbert Theurl:** The Convergence of Health Care Financing Structures: Empirical Evidence from OECD-Countries

2009-19 **Francesco Feri and Miguel A. Meléndez-Jiménez:** Coordination in Evolving Networks with Endogenous Decay

2009-18 **Harald Oberhofer:** Firm growth, European industry dynamics and domestic business cycles

2009-17 **Jesus Crespo Cuaresma and Martin Feldkircher:** Spatial Filtering, Model Uncertainty and the Speed of Income Convergence in Europe

2009-16 **Paul A. Raschky and Manijeh Schwindt:** On the Channel and Type of International Disaster Aid

2009-15 **Jianying Qiu:** Loss aversion and mental accounting: The favorite-longshot bias in parimutuel betting

2009-14 **Siegfried Berninghaus, Werner Güth, M. Vittoria Levati and Jianying Qiu:** Satisficing in sales competition: experimental evidence

2009-13 **Tobias Bruenner, Rene Levinský and Jianying Qiu:** Skewness preferences and asset selection: An experimental study

2009-12 **Jianying Qiu and Prashanth Mahagaonkar:** Testing the Modigliani-Miller theorem directly in the lab: a general equilibrium approach

2009-11 **Jianying Qiu and Eva-Maria Steiger:** Understanding Risk Attitudes in two Dimensions: An Experimental Analysis

2009-10 **Erwann Michel-Kerjan, Paul A. Raschky and Howard C. Kunreuther:** Corporate Demand for Insurance: An Empirical Analysis of the U.S. Market for Catastrophe and Non-Catastrophe Risks

2009-09 **Fredrik Carlsson, Peter Martinsson, Ping Qin and Matthias Sutter:** Household decision making and the influence of spouses' income, education, and communist party membership: A field experiment in rural China

2009-08 **Matthias Sutter, Peter Lindner and Daniela Platsch:** Social norms, third-party observation and third-party reward

2009-07 **Michael Pfaffermayr:** Spatial Convergence of Regions Revisited: A Spatial Maximum Likelihood Systems Approach

2009-06 **Reimund Schwarze and Gert G. Wagner:** Natural Hazards Insurance in Europe – Tailored Responses to Climate Change Needed

2009-05 **Robert Jiro Netzer and Matthias Sutter:** Intercultural trust. An experiment in Austria and Japan

2009-04 **Andrea M. Leiter, Arno Parolini and Hannes Winner:** Environmental Regulation and Investment: Evidence from European Industries

2009-03 **Uwe Dulleck, Rudolf Kerschbamer and Matthias Sutter:** The Economics of Credence Goods: On the Role of Liability, Verifiability, Reputation and Competition. *Revised version forthcoming in American Economic Review.*

2009-02 **Harald Oberhofer and Michael Pfaffermayr:** Fractional Response Models - A Replication Exercise of Papke and Wooldridge (1996)

2009-01 **Loukas Balafoutas:** How do third parties matter? Theory and evidence in a dynamic psychological game.

2008-27 **Matthias Sutter, Ronald Bosman, Martin Kocher and Frans van Winden:** Gender pairing and bargaining – Beware the same sex! *Revised version published in Experimental Economics, Vol. 12 (2009): 318-331.*

2008-26 **Jesus Crespo Cuaresma, Gernot Doppelhofer and Martin Feldkircher:** The Determinants of Economic Growth in European Regions.

2008-25 **Maria Fernanda Rivas and Matthias Sutter:** The dos and don'ts of leadership in sequential public goods experiments.

2008-24 **Jesus Crespo Cuaresma, Harald Oberhofer and Paul Raschky:** Oil and the duration of dictatorships.

2008-23 **Matthias Sutter:** Individual behavior and group membership: Comment. *Revised Version published in American Economic Review, Vol.99 (2009): 2247-2257.*

2008-22 **Francesco Feri, Bernd Irlenbusch and Matthias Sutter:** Efficiency Gains from Team-Based Coordination – Large-Scale Experimental Evidence. *Revised and extended version forthcoming in American Economic Review.*

2008-21 **Francesco Feri, Miguel A. Meléndez-Jiménez, Giovanni Ponti and Fernando Vega Redondo:** Error Cascades in Observational Learning: An Experiment on the Chinos Game.

2008-20 **Matthias Sutter, Jürgen Huber and Michael Kirchler:** Bubbles and information: An experiment.

2008-19 **Michael Kirchler:** Curse of Mediocrity - On the Value of Asymmetric Fundamental Information in Asset Markets.

2008-18 **Jürgen Huber and Michael Kirchler:** Corporate Campaign Contributions as a Predictor for Abnormal Stock Returns after Presidential Elections.

2008-17 **Wolfgang Brunauer, Stefan Lang, Peter Wechselberger and Sven Bienert:** Additive Hedonic Regression Models with Spatial Scaling Factors: An Application for Rents in Vienna.

2008-16 **Harald Oberhofer, Tassilo Philippovich:** Distance Matters! Evidence from Professional Team Sports. *Extended and revised version forthcoming in Journal of Economic Psychology.*

2008-15 **Maria Fernanda Rivas and Matthias Sutter:** Wage dispersion and workers' effort.

2008-14 **Stefan Borsky and Paul A. Raschky:** Estimating the Option Value of Exercising Risk-taking Behavior with the Hedonic Market Approach. *Revised version forthcoming in Kyklos.*

2008-13 **Sergio Currarini and Francesco Feri:** Information Sharing Networks in Oligopoly.

2008-12 **Andrea M. Leiter:** Age effects in monetary valuation of mortality risks - The relevance of individual risk exposure.

2008-11 **Andrea M. Leiter and Gerald J. Pruckner:** Dying in an Avalanche: Current Risks and their Valuation.

2008-10 **Harald Oberhofer and Michael Pfaffermayr:** Firm Growth in Multinational Corporate Groups.

2008-09 **Michael Pfaffermayr, Matthias Stöckl and Hannes Winner:** Capital Structure, Corporate Taxation and Firm Age.

2008-08 **Jesus Crespo Cuaresma and Andreas Breitenfellner:** Crude Oil Prices and the Euro-Dollar Exchange Rate: A Forecasting Exercise.

2008-07 **Matthias Sutter, Stefan Haigner and Martin Kocher:** Choosing the carrot or the stick? – Endogenous institutional choice in social dilemma situations. Revised version forthcoming in *Review of Economic Studies*.

2008-06 **Paul A. Raschky and Manijeh Schwindt:** Aid, Catastrophes and the Samaritan's Dilemma.

2008-05 **Marcela Ibanez, Simon Czermak and Matthias Sutter:** Searching for a better deal – On the influence of group decision making, time pressure and gender in a search experiment. *Revised version published in Journal of Economic Psychology, Vol. 30 (2009): 1-10.*

2008-04 **Martin G. Kocher, Ganna Pogrebna and Matthias Sutter:** The Determinants of Managerial Decisions Under Risk.

2008-03 **Jesus Crespo Cuaresma and Tomas Slacik:** On the determinants of currency crises: The role of model uncertainty. *Revised version accepted for publication in Journal of Macroeconomics.*

2008-02 **Francesco Feri:** Information, Social Mobility and the Demand for Redistribution.

2008-01 **Gerlinde Fellner and Matthias Sutter:** Causes, consequences, and cures of myopic loss aversion - An experimental investigation. *Revised version published in The Economic Journal, Vol. 119 (2009), 900-916.*

2007-31 **Andreas Exenberger and Simon Hartmann:** The Dark Side of Globalization. The Vicious Cycle of Exploitation from World Market Integration: Lesson from the Congo.

2007-30 **Andrea M. Leiter and Gerald J. Pruckner:** Proportionality of willingness to pay to small changes in risk - The impact of attitudinal factors in scope tests. *Revised version forthcoming in Environmental and Resource Economics.*

2007-29 **Paul Raschky and Hannelore Weck-Hannemann:** Who is going to save us now? Bureaucrats, Politicians and Risky Tasks.

2007-28 **Harald Oberhofer and Michael Pfaffermayr:** FDI versus Exports. Substitutes or Complements? A Three Nation Model and Empirical Evidence.

2007-27 **Peter Wechselberger, Stefan Lang and Winfried J. Steiner:** Additive models with random scaling factors: applications to modeling price response functions.

2007-26 **Matthias Sutter:** Deception through telling the truth?! Experimental evidence from individuals and teams*. Revised version published in The Economic Journal, Vol. 119 (2009), 47-60.*

2007-25 **Andrea M. Leiter, Harald Oberhofer and Paul A. Raschky:** Productive disasters? Evidence from European firm level data. *Revised version forthcoming in Environmental and Resource Economics.*

2007-24 **Jesus Crespo Cuaresma:** Forecasting euro exchange rates: How much does model averaging help?

2007-23 **Matthias Sutter, Martin Kocher and Sabine Strauß:** Individuals and teams in UMTS-license auctions. *Revised version with new title "Individuals and teams in auctions" published in Oxford Economic Papers, Vol. 61 (2009): 380-394***).**

2007-22 **Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz:** Modelling and Forecasting Oil Prices: The Role of Asymmetric Cycles. *Revised version accepted for publication in The Energy Journal.*

2007-21 **Uwe Dulleck and Rudolf Kerschbamer:** Experts vs. discounters: Consumer free riding and experts withholding advice in markets for credence goods. *Revised version published in International Journal of Industrial Organization, Vol. 27, Issue 1 (2009): 15-23.*

2007-20 **Christiane Schwieren and Matthias Sutter:** Trust in cooperation or ability? An experimental study on gender differences*. Revised version published in Economics Letters, Vol. 99 (2008): 494-497.*

2007-19 **Matthias Sutter and Christina Strassmair:** Communication, cooperation and collusion in team tournaments – An experimental study. *Revised version published in: Games and Economic Behavior, Vol.66 (2009), 506-525.*

2007-18 **Michael Hanke, Jürgen Huber, Michael Kirchler and Matthias Sutter:** The economic consequences of a Tobin-tax – An experimental analysis. *Revised version forthcoming in Journal of Economic Behavior and Organization.*

2007-17 **Michael Pfaffermayr:** Conditional beta- and sigma-convergence in space: A maximum likelihood approach. *Revised version forthcoming in Regional Science and Urban Economics.*

2007-16 **Anita Gantner:** Bargaining, search, and outside options. *Published in: Games and Economic Behavior, Vol. 62 (2008), pp. 417-435.*

2007-15 **Sergio Currarini and Francesco Feri:** Bilateral information sharing in oligopoly.

2007-14 **Francesco Feri:** Network formation with endogenous decay.

2007-13 **James B. Davies, Martin Kocher and Matthias Sutter:** Economics research in Canada: A long-run assessment of journal publications. *Revised version published in: Canadian Journal of Economics, Vol. 41 (2008), 22-45.*

2007-12 **Wolfgang Luhan, Martin Kocher and Matthias Sutter:** Group polarization in the team dictator game reconsidered. *Revised version published in: Experimental Economics, Vol. 12 (2009), 26-41.*

2007-11 **Onno Hoffmeister and Reimund Schwarze:** The winding road to industrial safety. Evidence on the effects of environmental liability on accident prevention in Germany.

2007-10 **Jesus Crespo Cuaresma and Tomas Slacik:** An “almost-too-late” warning mechanism for currency crises. *(Revised version accepted for publication in Economics of Transition***)**

2007-09 **Jesus Crespo Cuaresma, Neil Foster and Johann Scharler:** Barriers to technology adoption, international R&D spillovers and growth.

2007-08 **Andreas Brezger and Stefan Lang:** Simultaneous probability statements for Bayesian P-splines.

2007-07 **Georg Meran and Reimund Schwarze:** Can minimum prices assure the quality of professional services? (*Accepted for publication in European Journal of Law and Economics*)

2007-06 **Michal Brzoza-Brzezina and Jesus Crespo Cuaresma:** Mr. Wicksell and the global economy: What drives real interest rates?.

2007-05 **Paul Raschky:** Estimating the effects of risk transfer mechanisms against floods in Europe and U.S.A.: A dynamic panel approach.

2007-04 **Paul Raschky and Hannelore Weck-Hannemann:** Charity hazard - A real hazard to natural disaster insurance. *Revised version forthcoming in: Environmental Hazards.*

2007-03 **Paul Raschky:** The overprotective parent - Bureaucratic agencies and natural hazard management.

2007-02 **Martin Kocher, Todd Cherry, Stephan Kroll, Robert J. Netzer and Matthias Sutter:** Conditional cooperation on three continents. *Revised version published in: Economics Letters, Vol. 101 (2008): 175-178.*

2007-01 **Martin Kocher, Matthias Sutter and Florian Wakolbinger:** The impact of naïve advice and observational learning in beauty-contest games.

University of Innsbruck

Working Papers in Economics and Statistics

2010-10

Wolfgang A. Brunauer, Sebastian Keiler and Stefan Lang

Cost Drivers of Operation Charges and Variation over Time: An Analysis Based on Semiparametric SUR Models

Abstract

Although building operating charges have turned out to be a major determinant of profitability for real estate investments, there is a noticeable lack of reports or studies that analyze these costs with state-of-the-art statistical techniques. Specifically, past studies usually assume linear relationships between costs and building attributes, they do not control for cluster-specific or longitudinal effects and do not account for the simultaneous structure of cost categories. Therefore, in this study we provide a novel approach to real estate cost benchmarking: We analyze the effects of building attributes on electricity, heating and maintenance costs for office buildings in Germany in a multivariate structured additive regression (STAR) model simultaneously, modeling potentially nonlinear effects as P(enalized)-Splines and controlling for cluster-specific and individual heterogeneity in a three-way random effects structure. This way, we gain insights into how building attributes influence costs, and how cost levels vary across cities, companies and buildings. We furthermore derive quality-adjusted time indices for the two major German submarkets, the former German Democratic Republic and the old West German states. The results obtained can be used to derive portfolio allocation strategies and for planning, constructing, operating and redeveloping real estate.

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