The role of spatial and temporal structure for residential rent predictions

Roland Füss^{a,b,*}, Jan A. Koller^a

^a Swiss Institute of Banking and Finance (s/bf), University of St.Gallen, Rosenbergstrasse 52, 9000 St. Gallen, Switzerland

^b Research Associate at the Centre for European Economic Research (ZEW), Mannheim, Germany

ARTICLE INFO

Keywords:

Classification and regression tree (CART) technique
Forecast evaluation
Hedonic pricing model
Rental prices
Spatiotemporal autoregressive (STAR) model

ABSTRACT

This paper examines the predictive power of five linear hedonic pricing models for the residential market with varying levels of complexity in their spatial and temporal structures. Unlike similar studies, we extend the out-of-sample forecast evaluation to one-day-ahead predictions with a rolling estimation window, which is a reasonable setting for many practical applications. We show that the in-sample fit and cross-validation prediction accuracy improve significantly when we account for spatial heterogeneity. In particular, for one-day-ahead forecasts, the spatiotemporal autoregressive (STAR) model demonstrates its superiority over model specifications with alternating spatial and temporal heterogeneity and dependence structures. In addition, sub-market fixed effects, constructed on the basis of statistical TREE methods, improve the results of predefined local rental markets further.

© 2016 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

1. Introduction

The accurate prediction of rental prices is essential for many participants in the real estate market, such as investors, regulators, and policy-makers. For instance, rent forecasts play a crucial role in the property valuation adopted in discounted cash-flow models. In addition, when imputed market rents for owner-occupied dwellings are subject to income tax, as they are in Luxembourg, the Netherlands, or Switzerland, for example, local rent predictions serve as a tax base. Furthermore, the imputed rent constitutes a major component of the gross domestic income measure. A profound knowledge of the structure and development of rents in the local and national housing markets is also of importance for public housing policy.

Hedonic pricing models are the approach that is applied by far the most extensively to the prediction of rental and residential sales prices. In their most basic form, they model real estate prices as a linear function of the properties' attributes. Recent approaches have focused on the temporal structure, but have also paid particular attention to the spatial structure. In this paper, we compare the predictive power of hedonic models with varying levels of complexity of spatial and temporal structures. We focus mainly on comparisons of in-sample model fits, as well as the cross-validation prediction and one-day-ahead forecasting accuracies of the different models. We stress the fact that, while one-day-ahead forecasting, without ex ante information, is the very nature of many practical applications, it has still attracted little attention in the literature.

We restrict the analysis to fully parametric and linear models in order to ensure that the results remain interpretable for practical purposes. Starting with a simple baseline model that includes only physical attributes of the dwellings, we incrementally add elements to account

* Corresponding author at: Swiss Institute of Banking and Finance (s/bf), University of St.Gallen, Rosenbergstrasse 52, 9000 St. Gallen, Switzerland. Fax: +41 0 71 224 7088.

E-mail addresses: roland.fuess@unisg.ch (R. Füss), jan.koller@c-alm.ch (J.A. Koller).

<http://dx.doi.org/10.1016/j.ijforecast.2016.06.001>

0169-2070/© 2016 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

for spatial and temporal effects. We show that the consideration of a few spatial and temporal components can increase the forecast accuracy substantially. Likewise, significant spatial and temporal effects underpin the inconsistency of hedonic coefficients in the absence of these components, since the assumption of *i.i.d.* errors of the *baseline* model is violated. We demonstrate the empirical bias resulting from spatial heterogeneity, which in general prohibits causal inference. Thus, we aim to improve the prediction accuracy for rental prices by accounting for both the space and time dimensions in our model specifications.

Our paper contributes to the literature in at least three ways. First, we analyze the importance of the temporal and spatial structure of a model for residential markets. Most notably, the models include a spatiotemporal autoregressive (STAR) specification, as suggested by [Pace, Barry, Clapp, and Rodriguez \(1998\)](#).¹ Only a few other studies on rental price modeling exist, such as those by [Sirmans and Benjamin \(1991\)](#) and [Valente, Wu, Gelfand, and Sirmans \(2005\)](#), with the latter predicting apartment rents based on spatial econometric techniques. However, unlike our approach, they model an explicit spatial process, whereas our STAR model merely makes strong structural assumptions about the spatiotemporal dependencies.

Second, our choice of model specifications allows us to determine the *marginal effects* of increasing the complexity of the spatial and temporal structure. Importantly, we pay particular attention to the different prediction approaches, a point that has often been underrated in similar studies. In the forecasting application with a rolling estimation window, we account for temporal heterogeneity to the same degree in all models. We then improve the *baseline* model incrementally by gradually accounting for spatial heterogeneity, as well as temporal and spatial dependence. Finally, we account for spatial heterogeneity by constructing rental sub-markets using a classification and regression tree (CART) method. To our knowledge, only [Clapp and Wang \(2006\)](#) have used the CART methodology for sub-market construction in the context of housing price sales data.

Third, we use an exclusive rental apartment data set for the canton of Zurich, Switzerland. In countries in which the majority of households are tenants, the higher turnover in the rental market reduces the estimation error due to the rich data availability. Thus, the Swiss housing market, with its low homeownership rate, serves as an ideal testing ground for evaluating the predictive power of different hedonic rental price models.² Moreover, the canton of Zurich constitutes a residential market within a homogeneous regional economy, which makes it an ideal laboratory for our analyses.

Our empirical results show that including sub-market dummy variables based on ZIP codes improves the prediction accuracy significantly. However, while this approach

systematically ignores the spatial structure, the construction of sub-markets based on a regression tree technique displays a superior estimation strategy. The forecast evaluation shows that augmenting the hedonic model with spatially-lagged variables is particularly successful for one-day-ahead forecasting. These results highlight the STAR model's superiority over other specifications, and emphasize the importance of local dynamics in one-day-ahead forecasting.³

The rest of the paper is organized as follows: Section 2 provides an overview of the theoretical and empirical approaches to modeling the temporal and spatial structure in hedonic pricing models. Section 3 introduces the research design and the models that are compared for prediction purposes. The empirical predictive power of these models is evaluated in Section 4. Section 5 summarizes our results and provides concluding remarks.

2. Space–time modeling

Hedonic pricing models have been the workhorse of the housing literature for decades. Since the seminal work of [Rosen \(1974\)](#), the capitalization of dwelling amenities in house prices, based mainly on linear hedonic functions, has been studied extensively in the literature.⁴ One of the most important distinguishing features of hedonic housing price models is the underlying assumption about the spatial and temporal structure. Moreover, hedonic pricing models take a wide range of functional forms. In this section, we demonstrate the range of spatial and temporal effects in empirical housing applications, which serve as individual components in our model comparison and evaluation. Concerning the functional form, the potential model complexity has been influenced largely by advances in computer technology. In particular, semi- and non-parametric approaches have attracted a considerable amount of attention. For an overview of these models, see [Anglin and Gencay \(1996\)](#) and [McCluskey, McCord, Davis, Haran, and McIlhatton \(2013\)](#), as well as [Martins-Filho and Bin \(2005\)](#) for the application of artificial intelligence methods to the field of real estate.

Despite the successful application of non-linear and non-parametric methods, we restrict our analysis to linear and parametric methods for two reasons. First,

¹ See also [Pace, Barry, Gilley, and Sirmans \(2000\)](#) for an application of the STAR model to housing prices in Baton Rouge, Louisiana.

² Switzerland has one of the lowest homeownership rates internationally, at approximately 44%, according to [Eurostat \(2015\)](#).

³ Note that one-day-ahead forecasting has the conceptual feature of using only information that is available at the time of the prediction. The goal of this study is not to predict prices far into the future, since the temporal market dynamics in the model only account for short-term dynamics, because they are captured by temporally and spatially lagged variables. However, many applications require medium- or long-term forecasts of rental prices, e.g., for investment decisions. For these kinds of applications, the findings in this study can serve as a qualitative basis, since essentially the same out-of-sample prediction could be conducted with a longer time horizon. In contrast, short-term time horizons are more important in regions in which only a few rental price transactions are observable, the local housing market is very dynamic, or a precise valuation at a specific date is required, such as for imputed rents.

⁴ Instead of using house prices, the present study is one of the few to apply hedonic pricing techniques to rental price data. For reviews of the earlier literature, see [Bourassa, Hoesli, and Peng \(2003\)](#), [Malpezzi \(2003\)](#) and [Sirmans, Macpherson, and Zietz \(2005\)](#).

the focus of this study is on the marginal effects of the spatial and temporal components, and comparisons between models with different functional forms would unnecessarily dilute this intention. Second, semi- and non-parametric approaches often yield results that are difficult to interpret, and therefore, when considering price prediction, especially in practical applications, it is questionable whether the parsimony of time- and space-discrete models should be dismissed in favor of more complex approaches. By restricting our focus throughout this study to linear models, we use the following general pricing model:

$$P = \alpha(t, s) + X\beta + X^T\phi_T + X^S\phi_S + X^{TS}\phi_{TS} + X^{ST}\phi_{ST} + P^T\psi_T + P^S\psi_S + P^{TS}\psi_{TS} + P^{ST}\psi_{ST} + \varepsilon, \quad (1)$$

where t and s refer to time and space, respectively. The rental price is a process $P(t, s) \equiv \{p(t, s), (t, s) \in (\mathbb{R}, \mathbb{R}^2)\}$, i.e., a process in continuous time and location. The rental price vector P has the dimensions $N \times 1$ and the matrix X has dimensions $N \times k$, which comprise the housing-specific physical characteristics. The corresponding $(k \times 1)$ -vector of coefficients captures the hedonic prices of physical attributes. Note that the intercept $\alpha(t, s)$ is a function of time and location, and thus is potentially time- and location-specific, i.e., it accounts for *heterogeneity* along these dimensions. The coefficients ϕ and ψ measure the effects of temporal and spatial *dependencies*. In particular, these are the coefficients of the spatial and temporal lags of X and P , respectively, indicated by superscripts T and S . While the ϕ coefficients are $(k \times 1)$ -vectors, the ψ -coefficients are scalars. The error term is normally distributed, i.e., $\varepsilon \sim N(0, \sigma)$. The way of modeling the temporal and spatial effects is determined largely by the underlying data structure, and therefore the next section presents the most widely applied approaches, as well as combinations of them.⁵

2.1. Aggregation of space and time

Housing naturally has a specific location, while observations of price offers or transactions naturally happen at specific points in time. In general, therefore, the resulting observations are continuous in time and space. For this reason, real estate prices should not be treated as cross-sectional, time series or panel data, as Dube, Legros, and Thanos (2014) point out. Aggregation along either of these dimensions is problematic, and a subject of current research.⁶ In addition, the ways in which space and time are aggregated (or not) largely determine the set of feasible models and estimation methods. The models in this paper exhibit both *continuous* and *discrete* elements. In the *continuous* case, the location s is a pair of values containing latitude and longitude coordinates:

$$s \in \{\text{lat}, \text{lon}\} \in \mathbb{R}^2, \quad (2)$$

where *lat* denotes the latitudinal and *lon* the longitudinal values.

In the *discrete* case, the Euclidean space is partitioned into H aggregate sub-markets that are defined implicitly by:

$$h : \mathbb{1}(s \in R_h), \quad (3)$$

where R_h , with $h = 1, \dots, H$, are distinct partitions of the location space, and $\mathbb{1}(\cdot)$ is the indicator function. Similarly, the time dimension can be either divided into discrete partitions (e.g., in years) or treated as a continuous value. The left panel of Fig. 1 illustrates the aggregation of time and space into discrete partitions by including simple time and location dummies to account for potential *heterogeneity*. On the other hand, the right panel depicts the spatiotemporal specification, which models spatial and temporal *dependency*. In particular, the current observation (dark shaded dot) may depend on temporally lagged ($t1, t2$, and $t3$) or spatially lagged ($s1, s2, s3$, and $s4$) observations.

2.2. Spatial dimension

It is widely accepted that real estate transaction prices depend not only on physical attributes, but also on the local “market conditions”. There are two ways to interpret local market conditions. The first is to regard them as an unobserved effect that constitutes a locally homogeneous environment by differentiating the local from the global hedonic pricing coefficients. This is referred to as *spatial heterogeneity*. The second interpretation of local market conditions is traced back to a contagion effect between transactions. In particular, the occurrence of market transactions affects the price of (spatially lagged) observations. This is referred to as *spatial dependence*. Can (1990) relates spatial heterogeneity to neighborhood effects and spatial dependence to spillover effects.⁷ Although the theoretical distinction between these two effects has been accepted widely, the way in which these effects should be incorporated into econometric models is a subject of ongoing discussion. Since spatial dependence leads to empirical spatial heterogeneity, it is hard to identify the true data-generating process (DGP). Therefore, to a certain extent, the choice of the model structure is arbitrary. In most applications, the choice is driven simply by data availability.⁸

2.2.1. Spatial heterogeneity

Similarly to panel data, the simplest and most obvious way to account for potential cross-sectional dependence is to allow for fixed-effects. In spatial data, cross-sectional dependence often exhibits a spatial structure. If this is indeed due to the spatial heterogeneous hedonic pricing function, the inclusion of location-specific dummy variables in the case of a discrete space is a promising way

⁵ For instance, Liu (2013) used time and regional dummies combined with a spatiotemporal autoregressive (STAR) model.

⁶ See Dube et al. (2014) for a detailed discussion of these issues.

⁷ See also Can (1992) for a discussion of this distinction in the spatial dimension.

⁸ Anselin (2010), who provides an extensive review of the evolution of spatial econometrics over the last three decades, illustrates the trade-offs that emerge from these issues.

Fig. 1. Illustration of the time and space data structure. The figure shows the aggregation of discrete time and space in the data structure. The left panel illustrates the aggregation of time and space into discrete partitions. This structure can be modeled by including time and location dummies in order to account for potential *heterogeneity*. The right panel illustrates the modeling of spatial and temporal *dependency*. The dark shaded circle indicates the current observation, which depends on temporally lagged ($t1$, $t2$, and $t3$) or spatially lagged ($s1$, $s2$, $s3$, and $s4$) observations.

to control for this type of heterogeneity. Several studies have found evidence of local differences in the marginal effects of property characteristics (see e.g. Goodman & Thibodeau, 1998). Pace et al. (1998) argued that the use of indicator variables is the simplest way to control for spatial and temporal dependencies. However, they also argue that this strategy is only feasible with limited numbers of dependencies, since each indicator is related to a coefficient, which leads to an incidental parameter problem for large numbers of distinct locations and dependencies.

In this context, the definition of sub-markets is paid particular attention, and has been addressed by several studies (see e.g. Bourassa, Hamelink, Hoesli, & MacGregor, 1999; MacLennan, 1977; Michaels & Smith, 1990). The problem with administrative boundaries, e.g., based on MSA or ZIP-code definitions, is that such sub-markets do not follow the spatial heterogeneity structure. Typically, housing data come with information on the spatial positions of the observations (e.g., ZIP codes or school districts). These predefined sub-markets have been used widely as dummy variables in hedonic regression equations. Over the last two decades, several approaches for housing market segmentation based on statistical methods have been examined. Bourassa et al. (2003) used principal component analysis to identify orthogonal factors of the dwelling properties, then utilize cluster analysis to identify homogeneous sub-markets on the basis of these factors. Goodman and Thibodeau (2003) applied a hierarchical method for sub-market construction.

Goodman (1981) identified three requirements that a concise sub-market definition must fulfill: homogeneity, parsimony, and contiguity.⁹ Indeed, finding contiguous homogeneous sub-markets is supposed to solve the spatial

heterogeneity problem. Comparisons in terms of prediction accuracies show that constructed sub-markets are not superior to *a priori* sub-market definitions. In an analysis of residential property transactions in Auckland, Bourassa et al. (2003) found sub-markets as defined by professional appraisers to perform better. They compared the predictive powers of models with sub-markets based on a principal component analysis with a predefined set of sub-markets. The main result of their analysis was the finding that sub-market definitions that disregard spatial contiguity are worse for practical purposes. Goodman and Thibodeau (2003) derived housing sub-markets by utilizing a data set for metropolitan Dallas that contains 28,561 transactions of single-family houses. Using a segmentation approach based on a hierarchical model, they found predictions derived for smaller sub-markets to outperform pooled estimates. However, the predictive performance of ZIP-code predefined sub-markets is comparatively good.

The literature on the construction and analysis of segments in the rental market is sparse. As one of the few contributors, Des Rosiers and Theriault (1996) analyzed rental property data in the Quebec area, and identified five significantly different market segments.¹⁰ In this paper, we follow the approach suggested by Clapp and Wang (2006) and use the regression tree approach to identify sub-markets. The reason for this choice is that the regression tree is a method that produces geometrically simple (rectangular) sub-markets based on an optimality criterion, which aims to improve the prediction results

⁹ See also Clapp and Wang (2006) for theoretical requirements for sub-market construction.

¹⁰ Concerning heterogeneity in a continuous sense, *soft boundary* methods, originally developed in the field of geostatistics, are available. In this study, though, we restrict our focus to *hard boundaries*. See Bourassa, Cantoni, and Hoesli (2007) for a more detailed discussion of geostatistical models with continuous spatial domains and lattice models.

directly. In addition, the method is easy to implement and fulfills all of the requirements suggested by Goodman (1981).

2.2.2. Spatial dependence

Most recent studies have focused on modeling the spatial dependence of house prices through the application of new techniques from the field of spatial statistics to housing data (see e.g. Basu & Thibodeau, 1998; Can, 1992; Dubin, 1998; Pace & Gilley, 1997). Similarly, standard lattice models, such as spatial autoregressive (SAR) and conditional autoregressive (CAR) models, have been applied widely in the real estate literature.¹¹ A comparison and summaries of the two different spatial dependence specifications are given by Dubin (1998). A taxonomy of recent applications that have included spatial dependencies is provided by Bourassa, Cantoni, and Hoesli (2010), who differentiate between models with location dummies, lattice models of spatial dependence, and geostatistical models. Geostatistical models are characterized by a continuous spatial dimension. Using a data set of US house price transactions, the authors found geostatistical models to perform well. While their methodological approach is closely related to our study, we also test models with temporal dependence.

2.3. Temporal dimension

While there has been a significant amount of research on spatial effects over the last three decades, the temporal dimension has attracted very little attention. A few studies have used aggregate panel data to determine the fundamental drivers of housing prices (see e.g. Adams & Füss, 2010; Harter-Dreiman, 2004; Hort, 1998). In many hedonic pricing applications, the time dimension is neglected entirely, since transactions take place within a relatively short time window (i.e., within one year). This special case of temporal aggregation shrinks the time dimension to a single point (see Section 2.1).¹² The following subsection discusses these remaining issues, i.e., temporal heterogeneity and dependence, which are conceptually similar to the spatial dimension.¹³

2.3.1. Temporal heterogeneity

Temporal heterogeneity is the variation in hedonic parameters over time. In applications with relatively long time periods, temporal heterogeneity seems almost

natural. Indeed, the time dimension is even the major focal point for the purpose of hedonic price index construction.¹⁴ Two estimation methods for index construction are applied widely: incorporating time dummies in the regression, and estimating the hedonic equation separately for each period. In the first case, the parameters in β are assumed to be constant over time, and only the intercept parameter $\alpha(t, s)$ is allowed to vary over time. In the second case, the assumption of time-invariant hedonic slope parameters β is relaxed. Note that the rolling-window approach in our out-of-sample predictions allows us to estimate β on a day-to-day basis. Thus, the one-day-ahead forecasting approach accounts for temporal heterogeneity to a certain degree in all models, similarly to the work of Liu (2013) and Munneke and Slade (2001). A crucial distinction between the cross-validation and one-day-ahead forecasting results is the incremental improvement from accounting for temporal heterogeneity.

2.3.2. Temporal dependence

For discrete-time data, dynamic panel models have been the first choice in the literature on housing markets. However, these models are used only if the data come in an aggregated form, which is not the case for individual housing transactions or offer price data. In the case of non-aggregated data, the continuous-time information naturally often comes with continuous-space data and is captured by spatiotemporal models, which typically include autoregressive components. It is only recently that the temporal dimension has been addressed by a few studies. In particular, Pace et al. (1998) formulated a model that incorporated both the spatial and temporal dependencies in a concise methodological framework.¹⁵ As in our study, Liu (2013) applied the STAR model of Pace et al. (1998) to a Dutch data set of housing transactions. The main reason for the increased application of spatiotemporal models can be traced back to the more structured way of modeling spatial patterns.

In addition, the temporal dependence structure shows potential for forecasting applications. In particular, many practical applications only have historical information available for pricing a dwelling. In this case, time dummy methods are limited. This favorable aspect of the STAR model has not been stressed adequately in the literature to date, and thus is a main focus in the one-day-ahead forecasting application of this study.¹⁶ In real forecasting applications, the time dimension is indeed of particular importance, since the identification of temporal dependencies is the main purpose. These dependencies are measured by including spatial and temporal lags,

¹¹ For an overview, see e.g. Anselin (2010).

¹² For instance, Bourassa et al. (2010) treat the data as cross-sectional. By neglecting time effects, the implicit assumption is made that the observations occur simultaneously, which generally is not a realistic assumption. Even if the time window of the presumed cross-sectional data is short, planned but not yet realized transactions or price offers may still have an effect on previous transactions.

¹³ One important exception is that spatial econometrics allow for feedback loops, due to the connectivity induced by the weighting matrix. The size of the spillover and feedback effects depends on the estimated spatial lag, as well as on the strength of the spatial weights.

¹⁴ See Diewert, de Haan, and Hendriks (2015) for a general discussion of residential real estate index construction and Diewert, Saeed, and Silver (2009) for a specific discussion of the time dimension in hedonic indexes. For an overview of residential real estate index construction in Switzerland, see Fahrlander (2008).

¹⁵ A review of the literature on spatiotemporal modeling is provided by Liu (2013).

¹⁶ The reason for this limited interest in real forecasting might be that traditional applications have been ex post analyses, such as housing index construction.

while other informational inefficiencies are deliberately not exploited, to keep the methods simple. Hence, only two of the five models incorporate temporal dependencies in our model comparison on in-sample estimation and forecast evaluation. A summary of the different hedonic model specifications according to their temporal and spatial structure (baseline, ZIP, TREE, TAR, and STAR model) is given in Table 1.

3. Methodology

3.1. Prediction approaches

We compare the predictive power of the models in Table 1 using three different prediction approaches: *in-sample*, *cross-validation*, and *one-day-ahead forecasting*. As in other studies, the in-sample prediction is an overall model fit, while the cross-validation technique serves as an out-of-sample robustness test. In the in-sample prediction, the estimation is performed over the whole sample period. In contrast, the cross-validation technique is based on splitting the data set into training and test samples using a random re-sampling procedure. Finally, unlike in most hedonic prediction analyses, we also test the forecasting accuracy of the models.¹⁷ Many studies on housing prediction do not account for the temporal dimension, and therefore make out-of-sample analyses in the form of cross-validation.¹⁸ In contrast, our data come with specific dates, which allows us to model the temporal dependence. Therefore, the forecasting application is a real one-day-ahead forecast. For that purpose, we define a rolling estimation window of 500 days (Section 4.5 demonstrates that varying the window size does not change the results substantially). Based on the estimation window, we fit the models and forecast all rental prices observed on the subsequent day. Note that the rolling estimation window does account for temporal parameter heterogeneity to some extent.

In all three prediction approaches, we use several accuracy measures based on the residual $\hat{\varepsilon}$ from a rental price regression in natural logarithms. An important and widely applied error variation measure is the mean squared error, defined as $MSE = n^{-1} \sum_{i=1}^n \hat{\varepsilon}_i^2$. In order to get an impression of the size of the prediction, we report the square root of the MSE , i.e., the root mean squared error $RMSE = \sqrt{MSE}$. As a similar measure, but with less weight on the tails, we also present the mean absolute error $MAE = n^{-1} \sum_{i=1}^n \text{abs}(\hat{\varepsilon}_i)$.

3.2. Model specifications

The subsequent sections introduce the models that we consider in our comparative study of in-sample estimation and forecast evaluation. The specifications of these models

follows the incremental enhancement from a simple baseline model up to a spatiotemporal autoregressive (STAR) model (see Table 1). All of the models can be embedded in the general framework formulated in Eq. (1). The *baseline* model is a simple benchmark model in terms of a hedonic pricing function, where the rental price is a linear combination of physical apartment characteristics, stacked in the vector X . In addition, yearly time dummy variables absorb the temporal fixed-effects in the price variations, i.e.,

$$\text{Baseline: } P = \alpha(t, s) + X\beta + \varepsilon,$$

$$\text{where } \alpha(t, s) = \alpha_t, \quad (4)$$

with $\alpha_t = \text{const}$ for $t = 1, \dots, \Gamma$.

A more advanced specification is the *ZIP* model, which accounts for spatial heteroscedasticity in a simple way. Specifically, predefined sub-markets, constituted by ZIP codes, are included in the form of dummy variables:

$$\text{ZIP: } P = \alpha(t, s) + X\beta + \varepsilon,$$

$$\text{where } \alpha(t, s) = \alpha_t + \alpha_s, \quad (5)$$

with $\alpha_t = \text{const}$ for $t = 1, \dots, \Gamma$ and $\alpha_s = \mathbb{1}(s \in R_h)$ for $h = 1, \dots, H$ (ZIP codes).

The *TREE* model has the same functional form as the *ZIP* model:

$$\text{TREE: } P = \alpha(t, s) + X\beta + \varepsilon,$$

$$\text{where } \alpha(t, s) = \alpha_t + \alpha_s, \quad (6)$$

but with $\alpha_t = \text{const}$ for $t = 1, \dots, \Gamma$ and $\alpha_s = f(\text{lat}, \text{lon}) = \sum_{h=1}^H c_h \mathbb{1}((\text{lat}_i, \text{lon}_i) \in R_h)$ for $h = 1, \dots, H$ (regions), where the prediction in each region is a constant c_h .

Hence, the only difference between the *ZIP* and *TREE* models is the definition of the sub-markets. As was outlined in Section 2.2.1, the use of predefined boundaries may be problematic, because such sub-markets do not follow the spatial heterogeneity structure. Since we have information about the precise locations of apartments, we can optimize the sub-market homogeneity through statistical methods. For this purpose, we use the classification and regression tree (CART) approach.¹⁹ The basic principle of a tree is the partitioning of the predictor space, such that the predictions are kept constant at each partition. For our sub-market selection, we first estimate the *baseline* model, and use it to derive the residuals $\hat{\varepsilon}$. In the second step, the idea is to predict this residual surface by partitioning the coordinate space into H partitions. The prediction is a (local) constant on each partition. This second task is solved using a regression tree (see Section 3.3 for a detailed description of the *TREE* regression approach). As an additional restriction, price predictions, based on a model with regional indicator variables, require a sufficient number of data points in each region. Hence, sub-markets must be specified so that

¹⁷ Note that Liu (2013) makes one-step-ahead forecasts based on an annual frequency.

¹⁸ For instance, Bourassa et al. (2010) do not account for the temporal dimension.

¹⁹ Since Breiman, Friedman, Olshen, and Stone (1984), regression trees have gained popularity, and have been applied extensively in computational statistics. The CART approach was also used for sub-market identification by Clapp and Wang (2006).

Table 1

Overview of the model specifications.

Model	Temporal structure		Spatial structure	
	Heterogeneity	Dependence	Heterogeneity	Dependence
Baseline	Time dummies	No	No	No
ZIP	Time dummies	No	ZIP dummies	No
TREE	Time dummies	No	TREE dummies	No
TAR	Time dummies	Temporal lags	TREE dummies	No
STAR	Time dummies	Temporal lags	TREE dummies	Spatial lags

The table lists the different hedonic model specifications, with varying levels of complexity of the temporal and spatial structures. We compare the models' predictive power using three different prediction approaches: *in-sample*, *cross-validation*, and *one-day-ahead forecasting*. For all model specifications, we derive in-sample predictions as an overall model fit by including all observations over the whole sample period. We further conduct an in-sample cross-validation, which is based on splitting the data set into training and test samples using a random re-sampling procedure. We also test the forecasting accuracies of the models based on real one-day-ahead forecasts. For that purpose, we define a rolling estimation window of 500 days.

there are enough data points available in each partition for prediction purposes. In order to avoid overfitting, and to make the results comparable to those of the baseline model, we restrict the number of sub-markets to the number of ZIP-code regions; i.e., the *ZIP* and *TREE* models have the same numbers of parameters. This allows us to evaluate the marginal improvement derived from the two different sub-market definitions.

3.3. Classification and regression tree (CART)

The following subsection demonstrates the use of regression trees for sub-market construction. As was mentioned above, the residuals for this purpose are estimated from a regression of the rental price on the physical apartment attributes. Then, a regression tree is used to predict these residuals using the coordinates as the only predictor variables. The concepts and notation in this second step are based largely on the work of [Hastie, Tibshirani, and Friedman \(2009\)](#).

We begin by assuming that the price of an apartment depends on physical characteristics X and the price of the local amenities A . Thus, the rental price of dwelling i can be written as:

$$P_i = \alpha + X_i\beta + A_i + \varepsilon_i, \quad (7)$$

where ε_i is the error term. We further assume that the local amenities are unobserved and depend on the latitude and longitude coordinates, $A_i = A(\text{lat}_i, \text{lon}_i)$. The goal of the construction of the sub-markets is to identify H homogeneous partitions of the whole market, such that

$$A(\text{lat}_i, \text{lon}_i) = A_h, \quad (8)$$

for all $\{\text{lat}_i, \text{lon}_i\} \in R_h$, $h = 1, \dots, H$,

where R_h , with $h = 1, \dots, H$, define partitions of the Euclidean space. In order to identify these sub-markets, we run a regression of the rental price P on the physical characteristics of the dwelling. In the next step, we predict the residuals of the first regression using the regression tree approach, based on only the two predictors $(\text{lat}_i, \text{lon}_i)$, for $i = 1, \dots, N$ apartments.

The idea of the regression tree is to partition the Euclidean space into H regions, where the prediction in each region is a constant c_h :

$$f(\text{lat}, \text{lon}) = \sum_{h=1}^H c_h \mathbb{1}(\{\text{lat}_i, \text{lon}_i\} \in R_h). \quad (9)$$

The goal is to identify both the splitting variable (*lat* or *lon*) and the splitting points, such that the prediction error is minimized. This problem is solved by a recursive algorithm.²⁰ The illustration in [Fig. 2](#) shows 60 tree-based sub-markets for the canton of Zurich that result from this procedure. The left panel shows the tree structure, which is pruned to only seven splits for illustrative purposes. The right panel shows the resulting sub-markets, where each leaf of the tree corresponds to a partition on the map, which represents one sub-market.

3.4. A spatiotemporal autoregressive (STAR) model

We derive the spatiotemporal autoregressive (*STAR*) model presented by [Pace et al. \(1998\)](#) by starting with a simple hedonic model without spatial effects:

$$P = \alpha + X\beta + u, \quad (10)$$

where P is the price, X is a set of exogenous factors, and β contains a set of parameters. As the time subscripts indicate, we assume heterogeneity over time. Note that this simple hedonic model does not account for potential dependence in the error term u . We resolve this shortcoming by assuming a structural dependence in the error term of the form

$$u = Wu + \varepsilon, \quad (11)$$

where ε is a white noise process and W is the weighting matrix. Following [Liu \(2013\)](#), the *STAR* model can be written in compact form as:

$$(I - W)P = (I - W)X\beta + \varepsilon, \quad (12)$$

where I is the identity matrix.

In a standard spatial model, W would be the spatial weighting matrix. However, the *STAR* model generalizes the weighting matrix to a set of spatial and temporal lags, as well as combinations of these lags. In their most general specification, [Pace et al. \(1998\)](#) suggest that W be of the form:

$$W = \phi_T T + \phi_S S + \phi_{TS} TS + \phi_{ST} ST, \quad (13)$$

with S and T being the spatial and temporal lag operators, respectively. By applying these operators sequentially, additional lags can be constructed. For instance, ST is the

²⁰ For details, we refer the reader to [Hastie et al. \(2009\)](#).

Fig. 2. Map of regression-tree-based sub-markets in the canton of Zurich. The figure illustrates the sub-market identification according to the regression tree approach based on residual location values. In this approach, the splitting variable (lat or lon) and the splitting points are determined in such a way as to minimize the prediction error. The illustration shows 60 tree-based sub-markets resulting from this procedure for the canton of Zurich. The left panel shows the reduced tree structure for seven splits. The right panel shows the resulting sub-markets, with each partition on the map corresponding to a leaf of the tree, representing a sub-market.

spatial lag of the temporal lag and TS the temporal lag of the spatial lag. Substituting Eq. (13) into Eq. (12) leads to the general regression equation:

$$P = \alpha + X\beta + TX\phi_T + SX\phi_S + TSX\phi_{TS} + STX\phi_{ST} + TP\psi_T + SP\psi_S + TSP\psi_{TS} + STP\psi_{ST} + \varepsilon. \quad (14)$$

This is the unrestricted form of a STAR model, which includes all spatial and temporal lags, as well as the corresponding interactions. Note that X is a vector of k explanatory variables. For instance, the term ϕ_{ST} is a vector of k coefficients that corresponds to the spatial lag of the temporal lag of the explanatory variables. In order to simplify the notation, and for the sake of consistency, we rewrite the temporal and spatial suffixes. For instance, SX is rewritten as X^S . Finally, the spatiotemporal autoregressive (STAR) model includes all combinations of linear spatial and temporal lags:

$$\text{STAR: } P = \alpha(t, s) + X\beta + X^T\phi_T + X^S\phi_S + X^{TS}\phi_{TS} + X^{ST}\phi_{ST} + P^T\psi_T + P^S\psi_S + P^{TS}\psi_{TS} + P^{ST}\psi_{ST} + \varepsilon. \quad (15)$$

In general, a spatial weighting matrix is an $N \times N$ matrix where each element represents a bilateral weight, which is defined as a proximity measure. However, observations with non-zero (positive) weights are restricted, for two reasons. First, observations in the future are not expected to have an effect, and therefore are excluded. Second, the spatial dependence diminishes along the space. As a simplification and for computational reasons, only a predefined number of observations enter the weighting matrix, namely the closest previous m_s neighbors. Using

this restriction, the spatial weighting matrix reduces to a sparse matrix with only $N \times m_s$ non-zero elements. Similarly, the temporal weighting matrix defines the weights given to past observations in order to compute the temporal lag. The number of prior observations considered is restricted to m_T .

As the final model, we introduce the temporal autoregressive (TAR) model, which is a truncated version of the STAR model, and therefore its components correspond to those of the STAR model. It is an enhancement relative to the TREE model, since it includes a temporal moving average of rental prices.²¹ This model accounts for global time trends in the rental market in terms of autocorrelation:

$$\text{TAR: } P = \alpha(t, s) + X\beta + X^T\phi_T + P^T\psi_T + \varepsilon. \quad (16)$$

4. Empirical results

4.1. Data

We test our models using an exclusive data set that consists of 28,728 rental apartments offered in the canton of Zurich, Switzerland, between 2002 and 2014. The data are taken from the Swiss online rental market platform. The information available about the rental objects and the descriptive statistics are shown in Table 2. Note that the rental price and living area (*living area*) are in log

²¹ In this context, the moving average refers to a *rolling average*, and should not be confused with a *moving-average (MA) process*.

Table 2
Descriptive statistics.

Panel A: Descriptive statistics of continuous variables					
Variable	Mean	Std. Dev.	Q 5%	Median	Q 95%
log(Rental price)	7.218	0.755	5.908	7.222	8.503
log(Living area)	3.505	1.781	1.000	3.500	5.500
Number of rooms	4.403	0.432	3.5	4.5	5.0
Panel B: Descriptive statistics of binary variables					
Variable	Percentage of observations		Total		
	Yes	No			
Elevator	43.80	56.20	100.00		
Parking	30.92	69.08	100.00		
Garage	53.58	46.42	100.00		
Panel C: Descriptive statistics of apartment type					
Apartment type	Percentage of obs.	Age	Percentage of obs.		
Standard	82.22	Less than 4 years	18.29		
Duplex	6.58	4 to 14 years	18.46		
Attic flat	3.15	15 to 29 years	18.45		
Roof flat	4.31	30 to 43 years	17.29		
Studio	0.35	44 to 70 years	17.10		
Single room	0.33	More than 70 years	10.41		
Furnished flat	1.67	Total	100.00		
Terrace flat	0.41				
Single flat	0.23				
Loft	0.54				
Other	0.21				
Total	100.00				
Panel D: Observations by year					
Year	Observations	Proportion of the sample			
2002	637	2.2%			
2003	1 102	3.8%			
2004	1 557	5.4%			
2005	2 085	7.3%			
2006	2 281	7.9%			
2007	2 445	8.5%			
2008	2 691	9.4%			
2009	3 074	10.7%			
2010	2 975	10.4%			
2011	3 092	10.8%			
2012	3 053	10.6%			
2013	2 878	10.0%			
2014	858	3.0%			
Total	28 728	100.0%			

The table reports the descriptive statistics of rental prices for the canton of Zurich. The data are sourced from the Swiss online rental market platform. Note that the rental price and living area (*living area*) in Panel A are in log values. Panels B and C show the proportions of binary housing attributes and apartment types covered in the sample. Panel D reports the numbers of observations over the entire sample period from 2002 to 2014. Note that fewer observations are available for 2014 because only the first three months of the year are included.

terms. Panel D in Table 2 shows that the number of objects containing all relevant information increased between 2002 and 2009, and has been relatively stable since. Far fewer observations are available for the last year because the sample ends in the first half of 2014.

In general, the rent offered in housing and apartment markets is not the same as the contract price. Specifically, if the price offered for a dwelling is too high, the contract may not be finalized. We do not observe directly whether the contract associated with an advertisement is concluded or not. However, filtering and eliminating objects that have been re-advertised within a short time period acts as a good proxy for the market price.²² After eliminating the re-advertisements, we have more than 28,000 observations

at our disposal that can be identified as market contract prices.²³ Fig. 3 shows a map of the canton of Zurich, with the observations indicated by gray dots. The administrative borders shown are defined by ZIP codes.

With regard to our comparative study of prediction models, the rental market in the canton of Zurich has several advantageous features. First, the density of observations is relatively high, because the area is small (91 square kilometers) and the number of observations is large. This is because Switzerland has a

last price of the (series of) re-advertisements is considered as the contract price, with the previous advertisements being ignored in the forecasting application.

²³ For extensive discussions of listing prices, contract prices, and market prices, refer to Allen, Rutherford, and Thomson (2009), Anglin, Rutherford, and Springer (2003), and Knight (2002).

²² In particular, if the same apartment appears on the platform again within 30 days, it is identified as a re-advertisement. In that case, only the

Fig. 3. Map of study area. The figure shows a map of the canton of Zurich, with the distribution of observations indicated by gray dots (approx. 28,000 observations). It also lists the six towns with the highest populations. The canton of Zurich accommodates 171 municipalities and covers an area of 91 square kilometers. The administrative borders of these municipalities, shown on the map, are defined by ZIP codes.

high percentage of renters. In the canton of Zurich, in particular, the homeownership rate in 2005 was as low as 7% (see [Statistik Stadt Zürich, 2005](#)). The resulting high density of observations allows us to identify the temporal and spatial structure better.

4.2. Preliminary analysis: spatial dependence

The use of methods that include spatial components is appropriate only if rental prices demonstrate spatial dependence. The magnitude of the spatial dependence in the rental price is measured by the lags in the *TAR* and *STAR* models. The presence of spatial dependence can also be captured by a variogram. In particular, we would expect spatial dependence in the residuals of the *baseline* model, which contains no spatial elements. Therefore, we estimate the empirical variogram for the residuals that result from the regression of this benchmark model and compare it with those for the residuals of the alternative models. For this purpose, we briefly introduce the variogram function, which represents the variance between observations conditional on a specific distance. Consider the set of all location pairs (s_i, s_j) with a distance g , denoted by $N(g)$. The semivariogram is then defined by the function $\gamma(g)$, with

$$\gamma(g) = \frac{1}{2|N(g)|} \text{Var}(u(s_i) - u(s_j)), \quad (17)$$

where $|N(g)|$ is the number of distinct pairs in $N(g)$ and $u(s)$ are the residuals from the hedonic regression. There is evidence of spatial dependence if the correlation between the residuals is non-constant for different distances. In terms of the variogram function $\gamma(g)$, this means that the functional value is low (high) for small (large) distances. The semivariogram for the residuals of the *baseline* model in Eq. (10) is shown in Fig. 4. Since the semivariogram increases with distance, the residuals do indeed exhibit spatial dependence, and thus, the consideration of spatial elements might be appropriate. Nevertheless, the question of whether location dummies or spatial lags are more suitable for accounting for spatial heterogeneity is an empirical one. However, the inclusion of spatial components in the regression should make the semivariogram flatter. The next subsection delineates whether and how the spatial dependence will decrease with an increasing complexity of the spatial specification.

4.3. In-sample estimation

This subsection discusses the economic and statistical significance of the coefficients derived from in-sample estimations. The prediction accuracy results are presented in the next subsection. For the in-sample prediction, we run an OLS regression of rental prices on a set of apartment attributes, as well as on the spatial and temporal lags. Table 3 shows the regression results of the five models. The

Table 3
In-sample estimation.

Variables	Baseline		ZIP		TREE		TAR		STAR	
	coeff.	p-value	coeff.	p-value	coeff.	p-value	coeff.	p-value	coeff.	p-value
Panel A: Coefficients of physical attributes										
log(Living area)	0.628	0.000	0.566	0.000	0.537	0.000	0.537	0.000	0.536	0.000
Rooms	0.032	0.000	0.059	0.000	0.065	0.000	0.065	0.000	0.069	0.000
Elevator	0.068	0.000	0.033	0.000	0.020	0.000	0.020	0.000	0.034	0.000
Parking	−0.031	0.000	0.002	0.523	0.005	0.024	0.005	0.025	0.007	0.001
Garage	−0.001	0.825	0.013	0.000	0.019	0.000	0.019	0.000	0.027	0.000
Apartment	0.178	0.000	0.170	0.000	0.162	0.000	0.162	0.000	0.138	0.000
Duplex	0.204	0.000	0.205	0.000	0.199	0.000	0.199	0.000	0.164	0.000
Attic flat	0.372	0.000	0.352	0.000	0.334	0.000	0.334	0.000	0.286	0.000
Roof flat	0.205	0.000	0.223	0.000	0.213	0.000	0.214	0.000	0.178	0.000
Studio	0.192	0.000	0.146	0.000	0.103	0.000	0.103	0.000	0.073	0.001
Single room	0.111	0.002	0.041	0.177	−0.022	0.000	−0.022	0.416	−0.049	0.040
Furnished flat	0.484	0.000	0.416	0.000	0.355	0.000	0.356	0.000	0.282	0.000
Terrace flat	0.327	0.000	0.316	0.000	0.301	0.000	0.302	0.000	0.253	0.000
Single flat	0.229	0.000	0.201	0.000	0.195	0.414	0.195	0.000	0.178	0.000
Loft	0.234	0.000	0.270	0.000	0.272	0.000	0.273	0.000	0.228	0.000
Age 5–14	−0.078	0.000	−0.076	0.000	−0.074	0.000	−0.074	0.000	−0.060	0.000
Age 15–29	−0.120	0.000	−0.126	0.000	−0.139	0.000	−0.138	0.000	−0.104	0.000
Age 30–43	−0.154	0.000	−0.177	0.000	−0.191	0.000	−0.191	0.000	−0.133	0.000
Age 44–70	−0.099	0.000	−0.168	0.000	−0.189	0.000	−0.189	0.000	−0.142	0.000
Age 70+	0.096	0.000	−0.014	0.002	−0.102	0.000	−0.102	0.000	−0.095	0.000
Q2	0.000	0.915	0.007	0.024	0.007	0.018	0.007	0.021	0.004	0.158
Q3	−0.002	0.689	0.005	0.122	0.004	0.177	0.003	0.270	0.002	0.440
Q4	0.000	0.958	0.010	0.001	0.010	0.000	0.010	0.002	0.007	0.015
Panel B: Coefficients of spatial and temporal lags										
lag(T) log(Living area)							−0.066	0.333	−0.067	0.392
lag(T) Rooms							0.016	0.434	0.011	0.622
lag(T) Elevator							−0.017	0.499	−0.040	0.246
lag(T) Parking							0.003	0.924	0.001	0.959
lag(T) Garage							0.007	0.793	−0.037	0.215
lag(S) log(Living area)									−0.504	0.000
lag(S) Rooms									−0.048	0.000
lag(S) Elevator									−0.060	0.000
lag(S) Parking									0.005	0.256
lag(S) Garage									−0.033	0.000
lag(ST) log(Living area)									0.415	0.003
lag(ST) Rooms									−0.022	0.624
lag(ST) Elevator									−0.074	0.226
lag(ST) Parking									−0.093	0.137
lag(ST) Garage									0.097	0.075
lag(TS) log(Living area)									−0.037	0.785
lag(TS) Rooms									0.015	0.704
lag(TS) Elevator									0.069	0.186
lag(TS) Parking									−0.008	0.874
lag(TS) Garage									0.095	0.077
lag(T) log(Rent)							0.080	0.097	0.099	0.174
lag(S) log(Rent)									0.816	0.000
lag(ST) log(Rent)									−0.702	0.000
lag(TS) log(Rent)									0.026	0.791
Time dummies	yes		yes		yes		yes		yes	
Location dummies	no		ZIP		TREE		TREE		TREE	
R ²	0.722		0.810		0.853		0.853		0.883	
No. of parameters	35		94		94		102		120	

Panel A of the table reports the coefficients of the *baseline*, *ZIP*, *TREE*, *TAR*, and *STAR* regressions. Panel B shows the coefficients of the spatial and temporal lags. The spatial and temporal lags are indicated by *lag(S)* and *lag(T)*, respectively. Cross-effects are indicated by two lag indicators: the spatial lag of the temporal lag is denoted by *lag(ST)* and the temporal lag of the spatial lag by *lag(TS)*. The variable *log(Rent)* refers to lags of the rental price as the endogenous variable.

model complexity and the number of parameters increase from left to right, i.e., from the *baseline* to the *STAR* model.

Panel A of Table 3 shows that the most important physical attributes of the apartments have highly significant effects and the expected signs. The effect with the highest

significance is that for the apartment's living area. Since both the rental price and living area are denoted in natural logarithms (*log*), the coefficients can be interpreted as elasticities. Thus, a 1% increase in apartment size is associated with a rental price increase of 0.54% to 0.63%, depending

Fig. 4. Empirical semivariogram of the baseline model. The figure shows the semivariogram estimated from the residuals of the *baseline* model. The increasing pattern, which reflects spatial dependence as a function of distance (in km), indicates that the residuals still show spatial autocorrelation. The dashed line illustrates the limiting value (for large distances). A flattening curve means that the dependence decreases with an increasing distance. Once the limiting value is reached, the spatial dependence has diminished.

on the model. The size of this effect is comparable to those found in other studies on dwelling prices. The number of rooms also has an increasing effect on the rental price, i.e., dividing the same apartment area into more rooms increases the rent. However, the effect is small, ranging from 3.2% to 6.9% per room, depending on the model specification.

An interesting finding that reflects the problem of spatial heterogeneity involves the coefficients for *Parking* and *Garage*. The former variable indicates that outside parking is available, while the latter indicates that the apartment has a parking garage. The coefficients for these variables are significantly positive in all models, except for the baseline model. This result is due to the fact that prices for locations tend to be higher in central regions where parking spaces are rare. This is a good example of one way in which spatial heterogeneity can lead to biased estimates. The coefficients for the age of an apartment are closely related to this finding. Indeed, the *baseline* model would suggest that old apartments (70 years old or older) are more expensive than new buildings. Again, this can be attributed to the fact that, on average, buildings in central locations are older, particularly those in historic city centers. Finally, the type of apartment has a considerable effect on the rental price, presumably due to households' specific tastes.

In regard to spatial and temporal dependence, the regression coefficients of the corresponding lags are shown in Panel B of Table 3. First, and most notably, the coefficient for the spatial lag of the explained variables, denoted by $[lag(S) \log(Rent)]$, is highly significant. The magnitude of the effect is 0.816, indicating an almost one-to-one spillover effect of spatially lagged objects. This finding is largely in line with those of Liu (2013), who estimates a coefficient of between 0.87 and 0.93 (depending on the year) for the spatial dependence in house prices.

Second, the spillover effect of spatially lagged objects is also present for the apartments' attributes. The corresponding coefficients are those starting with $lag(S)$. All effects (except for the *Parking* variable) are significantly

negative. This finding is in line with the local competition concept, which states that an increased availability of objects with favorable characteristics (i.e., large area, many rooms, and an elevator) has negative effects on rental prices. Therefore, we can conclude that the spatial effects estimated in the STAR model reflect reasonable market dynamics. In particular, competition takes place through both lower prices and better attributes of objects close to the apartment at hand.

In contrast to the highly significant spatial lags, the temporal lags $[lag(T)]$ are almost negligible. Similarly, the cross-effects $[lag(TS) \text{ and } lag(ST)]$ are of minor importance, with the exception of the temporal lag of the spatial lag. The highly significant negative sign indicates that price spillovers are only temporary, i.e., the effect on the price is mean-reverting. In particular, the temporal lag is calculated on the basis of the last 180 observations ($m_T = 180$), which corresponds to an average time of approximately 31.3 days.

We gain further insight into the remaining spatial dependence of the *ZIP*, *TREE*, *TAR*, and *STAR* models by estimating the semivariogram of the regression residuals. Fig. 5 shows the corresponding semivariograms. A comparison to the *baseline* model (shown in Fig. 4) reveals that the spatial dependence decreases incrementally with increasing spatial elements. The exceptions are the *TREE* and *TAR* models, which, by construction, include only temporal components. These findings are largely in line with our expectations. In addition, they demonstrate that the *TREE* model has already captured a large portion of the spatial dependence. This result shows that no clear distinction between spatial dependence and spatial heterogeneity can be determined empirically.

4.4. Forecasting performance

The in-sample analysis suggests the presence of spatial heterogeneity. On the one hand, this is an important assumption for the consistency of the coefficients, as the effect of the variable *Parking availability* has illustrated. On the other hand, accounting for spatial heterogeneity can improve the prediction accuracy. In order to prevent model overfitting, we also test the out-of-sample predictive power. We use two different approaches: cross-validation and one-day-ahead forecasting. In the cross-validation approach, we create a random test sample, which consists of 15% of the observations. The rest of the data (85%) are used as the estimation sample. In order to maintain the feasibility of models with time and location dummies, the random sampling mechanism must ensure that the estimation sample includes at least one observation with each time-region combination. In order to meet this requirement, the random sampling process is applied to each time-region combination, i.e., a sample of 85% of the observations within each time-region combination is selected at random. This specification allows a quasi-out-of-sample prediction, while spatial and temporal heterogeneity can still be incorporated in the form of time and region dummy variables. Panels A and B of Table 4 show the in-sample and cross-validation prediction results.

Fig. 5. Empirical semivariogram of alternative models. The figure shows the semivariograms estimated from the regression residuals of the alternative models ZIP, TREE, TAR, and STAR. For the ZIP and TREE specifications, the spatial dependence decreases with the increasing number of spatial elements. The TREE and TAR models fail to capture the spatial dependence adequately, as they only consider temporal components.

For the one-day-ahead forecasts, we use a rolling estimation window of 500 days. The reason for employing this approach is to ensure that strictly historical information is used for the estimation of the models, which is particularly important in practical prediction applications. The results of the out-of-sample predictions are shown in Panel C of Table 4.

The overall fits of the models (Panel A) in terms of R^2 values are similar for the three prediction approaches. The *baseline* model with only physical attributes fits the data quite well, with an R^2 of 0.72. The inclusion of location dummy variables in the ZIP model raises this number to 0.81, an increase of approximately nine percentage points. As expected, the ZIP-code classification already captures a substantial part of the spatial discrepancies in rental prices. This finding is in line with the results of Bourassa et al. (1999), who found sub-market choice using sophisticated statistical methods to provide only a small improvement in the prediction accuracy. However, the use of tree-based

dummy variables increases the R^2 to about 0.85, while the TAR model is no better than the TREE model in terms of R^2 values. This suggests that the past general rental price level, as captured by a moving average of the rents, does not contain much information for current rental prices.²⁴ In contrast, the inclusion of spatial lags in the STAR model raises the R^2 to 0.88, an increase of three percentage points. These results of the overall fit of the models are comparable to those found in other similar studies.²⁵

²⁴ Note again that the moving average in this context means *rolling average* and should not be confused with a *moving-average (MA) process*.

²⁵ For instance, the R^2 values found by Liu (2013) were 0.78 for the baseline model and 0.88 for the STAR model (for the year 2007). We also used an F -test to determine whether the increases in the R^2 values as the model complexity increases are significant. The test results confirm that the increase in explanatory power with an increasing model complexity (from left to right in Table 3 or in Panel A of Table 4) is significant at the 1% level. The only exception to this result is the change from the TREE model to the TAR model, where the minor increase in the R^2 value is insignificant.

Table 4
Forecast accuracy.

	Baseline	ZIP	TREE	TAR	STAR
Panel A: In-sample prediction (global estimation)					
Root mean squared error (RMSE)	0.214	0.176	0.155	0.155	0.138
Mean absolute error (MAE)	0.157	0.127	0.113	0.113	0.097
R^2	0.722	0.810	0.853	0.853	0.883
Prediction error less than...					
...20%	0.726	0.807	0.850	0.850	0.888
...15%	0.604	0.697	0.743	0.744	0.802
...10%	0.438	0.523	0.568	0.568	0.644
...5%	0.230	0.284	0.310	0.310	0.375
Panel B: Out-of-sample prediction (random sampling)					
Root mean squared error (RMSE)	0.214	0.177	0.155	0.155	0.138
Mean absolute error (MAE)	0.157	0.128	0.113	0.113	0.097
R^2	0.720	0.809	0.852	0.852	0.883
Prediction error less than...					
...20%	0.724	0.807	0.850	0.850	0.888
...15%	0.602	0.696	0.743	0.744	0.802
...10%	0.437	0.523	0.567	0.567	0.645
...5%	0.230	0.284	0.309	0.309	0.375
Panel C: One-day-ahead forecast (rolling window)					
Root mean squared error (RMSE)	0.204	0.169	0.152	0.152	0.133
Mean absolute error (MAE)	0.149	0.123	0.111	0.112	0.095
R^2	0.714	0.804	0.842	0.841	0.879
Prediction error less than...					
...20%	0.746	0.817	0.852	0.850	0.893
...15%	0.630	0.706	0.745	0.744	0.809
...10%	0.464	0.538	0.570	0.567	0.648
...5%	0.242	0.296	0.314	0.311	0.374

The table reports the prediction accuracy results for the *baseline*, *ZIP*, *TREE*, *TAR*, and *STAR* models. Panel A shows the in-sample prediction results, Panel B the cross-validation prediction results, and Panel C the one-day-ahead forecasting results. For the one-day-ahead forecasting, the reported R^2 is the pseudo R^2 (calculated as the squared correlation between the forecasted and actual values).

A similar pattern to that of the R^2 values can be identified for the prediction measures *RMSE* and *MAE*. We focus on these measures for comparing different prediction approaches. Most notably, the cross-validation results (Panel B) are almost the same as those for the in-sample fit, which confirms the robustness of the in-sample results. While the prediction errors in the *baseline* model are large (21.4% *RMSE* and 15.7% *MAE*), these figures decrease substantially for the *STAR* model, to 13.8% and 9.7%.

In the one-day-ahead forecasting application (Panel C), the results of all models improve slightly, due to the rolling-window approach, which accounts in part for the temporal heterogeneity of all coefficients in all models. However, the magnitude of the improvement is small. For instance, the *baseline* model improves by only one percentage point in terms of *RMSE*. Similar improvements are obtained for the other models. The largest relative improvement, though still small, is achieved for the *STAR* model, which fits with our expectations. In particular, note that the spatial lags are based on past but local observations. Therefore, capturing the local market trend seems to be most important when forecasting out-of-sample, i.e., when time-fixed effects are not feasible.

4.5. Robustness tests

The forecast evaluations in the previous subsection are derived on the basis of specific choices of model parameters. These parameters refer mainly to the size

of the rolling window, the number of neighbors used in the determination of the spatial and temporal lags, and the regression tree parameter for sub-market definitions. This section varies these parameters in order to determine the sensitivity of the prediction accuracy for the one-day-ahead forecasts. The *RMSEs* of the robustness tests for different parameter values are summarized in Table 5.

First, we vary the length of the estimation window for the one-day-ahead forecasts in Panel A of Table 5. In the case of a high temporal heterogeneity, we would expect a higher accuracy for shorter estimation windows. Surprisingly, the *RMSEs* for all models are slightly lower for a time window of 600 days than 500 days; however, the differences are small. In addition, increasing the estimation window further means that we need more observations at the end point. Hence, we refrain from increasing the estimation window further. In contrast, when the rolling estimation window is reduced to 400 days, the *RMSEs* are marginally lower for all models.

Second, we address the arbitrary choice of the parameters in the *STAR* model. As was mentioned in Section 3.2, the number of observations used to calculate the temporal and spatial lags in the *STAR* model is chosen following Pace et al. (1998). Panel B of Table 5 presents the *RMSEs* resulting from a variety of different parameter combinations of m_T and m_S .²⁶ Deviations of the parameters from

²⁶ Note again that the parameter m_T denotes the number of observations used to calculate the time lag, and m_S the number of observations used to calculate the spatial lag.

Table 5

Robustness tests.

Panel A: RMSEs for varying lengths of the rolling estimation window					
Estimation window (days)	Baseline	ZIP	TREE	TAR	STAR
400	0.204	0.170	0.152	0.153	0.135
500	0.204	0.169	0.152	0.152	0.133
600	0.203	0.169	0.151	0.151	0.132
Panel B: RMSEs of the STAR model with varying m_S and m_T					
	m_T				
	100	140	180	220	260
m_S	10	0.1320	0.1326	0.1327	0.1326
	20	0.1322	0.1328	0.1329	0.1327
	30	0.1322	0.1328	0.1329	0.1327
	40	0.1322	0.1328	0.1329	0.1327
	50	0.1322	0.1328	0.1329	0.1327
Panel C: RMSEs of the TREE model with a varying tree complexity parameter					
Complexity parameter	Number of sub-markets		RMSE		
20.00	5		0.1705		
5.00	10		0.1522		
1.20	30		0.1517		
0.55	60		0.1512		

The table shows the results of the robustness tests for varying model parameters. All tests are performed in the one-day-ahead forecasting application. Panel A reports the *root mean squared error* (RMSE) for varying lengths (in days) of the rolling estimation window. Panel B shows the RMSEs when the numbers of neighbors (m_S and m_T) used for determining the spatial and temporal lags are chosen differently, where m_S and m_T denote the numbers of observations used to calculate the spatial and time lags, respectively. Panel C shows the prediction accuracy with varying regression tree parameters and the corresponding numbers of tree-defined sub-markets. The values in bold are those used in the forecasting application in Section 4 (standard values).

their initial values ($m_T = 180$ and $m_S = 30$) do not change the outcomes notably. Although the effect of the change in the spatial lag parameter m_S on the prediction results is negligible, the parameter m_T changes the RMSEs slightly. Nevertheless, the prediction results are highly robust with respect to the choice of these parameters. In particular, the RMSEs vary by less than 0.06 percentage points.

A third arbitrary parameter choice is the number of sub-markets defined in the TREE model.²⁷ Panel C of Table 5 shows the RMSEs for different choices of the level of complexity of the regression tree, which results in different numbers of sub-markets. As the resulting values indicate, the RMSEs do not change substantially even with very few sub-markets. This indicates that the macro location is most important in explaining differences in rental prices.

5. Conclusion

We test the forecast performances of five different hedonic model specifications with varying levels of complexity of temporal and spatial structures for the residential rental market. The baseline model with no spatial and temporal elements served primarily as a benchmark. By fitting the models to rental apartment data from the canton of Zurich, Switzerland, and evaluating the forecast accuracy on daily rental price data, we reveal several important findings. With regard to the hedonic prices, we demonstrate the importance of accounting for spatial heterogeneity. In particular, if only physical attributes are included in the model, the high dependence

of certain variables (e.g., available parking) on the location can cause large biases in the estimates. Similarly, the prediction results are substantially worse in the absence of location dummy variables. Indeed, we show that including sub-market dummy variables based on ZIP codes improves the prediction results considerably. Since sub-markets defined by ZIP codes do not account systematically for the specific spatial structure, we test a more sophisticated sub-market definition by accounting for empirical spatial heterogeneity. We construct sub-markets based on the residual values of the *baseline* model using a regression tree. The construction of more sophisticated sub-markets improves the prediction accuracy substantially.

Including temporal lags (in the STAR model) reveals two further interesting findings. First, we find empirical evidence of a functioning local market competition. In particular, there is a positive dependence on the rental prices of neighboring objects, but a negative dependence on the favorable attributes of neighboring objects. This result suggests that the presence of other apartment offerings in the neighborhood with low rental prices and favorable (high) attributes has a negative effect on the rental price of an object. Second, and more importantly for the purposes of this paper, including spatial dependence increases the prediction accuracy considerably. In particular, when comparing the STAR model with the TREE model, the RMSE decreased from 0.155 to 0.138 for the in-sample prediction and cross-validation. However, the greatest improvement is found in the one-day-ahead forecast, where the RMSE decreases from 0.152 to 0.133. This finding is in line with the hypothesized expectation that local dynamics are important for one-day-ahead forecasting. Thus, the STAR model is the best candidate for empirical applications where no ex ante information is available.

²⁷ Note that our previous estimations set the number of sub-markets equal to the number of regions in the ZIP model.

Another important outcome of the one-day-ahead forecasting application is the finding that the prediction accuracy is better than those in the in-sample and cross-validation predictions. Hence, it follows that rental prices exhibit some level of temporal heterogeneity, which is accounted for by the rolling estimation window. For real world applications, e.g., for the purpose of price setting by real estate developers or valuation by appraisers, all models except for the baseline model are suitable in terms of prediction accuracy. In particular, the prediction errors are less than 18% in terms of the root mean squared error (RMSE), and the mean absolute error (MAE) is less than 13%. The best prediction is achieved by the STAR model for one-day-ahead forecasting, with an RMSE of 13.3% and a MAE of 9.7%. These results are highly robust to variations in the model parameters. For instance, changing the rolling estimation window length has almost no effect on the prediction results. Most notably, however, reducing the number of sub-markets in the TREE model from 60 to 10 increases the prediction error by only 0.1 percentage points.

The temporal structure, and the temporal autocorrelation in particular, has only minor effects. Therefore, the TAR model does not provide (marginal) improvements. The temporal dimension merits greater attention in future research for out-of-sample forecasting applications, by testing autoregressive dynamics in hedonic prices for instance (under temporal heterogeneity). In addition, it may be promising to use more sophisticated forecasting models rather than simple temporal lags in terms of moving averages. However, the empirical evidence in this study shows that the dynamics in the residential market are *local*, and therefore it is questionable whether the consideration of *global* trends could indeed improve the prediction accuracy.

Acknowledgments

The authors are grateful to Dick van Dijk (the editor), the associate editor, and the anonymous referee for valuable suggestions, which have significantly improved the article.

References

- Adams, Z., & Füss, R. (2010). Macroeconomic determinants of international housing markets. *Journal of Housing Economics*, 19(1), 38–50.
- Allen, M., Rutherford, R., & Thomson, T. (2009). Residential asking rents and time on the market. *Journal of Real Estate Finance and Economics*, 38(4), 351–365.
- Anglin, P. M., & Gencay, R. (1996). Semiparametric estimation of a Hedonic price function. *Journal of Applied Econometrics*, 11(6), 633–648.
- Anglin, P. M., Rutherford, R., & Springer, T. M. (2003). The trade-off between the selling price of residential properties and time-on-the-market: The impact of price setting. *Journal of Real Estate Finance and Economics*, 26(1), 95–111.
- Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1), 3–25.
- Basu, S., & Thibodeau, T. G. (1998). Analysis of spatial autocorrelation in house prices. *Journal of Real Estate Finance and Economics*, 17(1), 61–85.
- Bourassa, S., Cantoni, E., & Hoesli, M. (2007). Spatial dependence, housing submarkets, and house price prediction. *Journal of Real Estate Finance and Economics*, 35(2), 143–160.
- Bourassa, S. C., Cantoni, E., & Hoesli, M. (2010). Predicting house prices with spatial dependence: A comparison of alternative methods. *Journal of Real Estate Research*, 32(2), 139–160.
- Bourassa, S. C., Hamelink, F., Hoesli, M., & MacGregor, B. D. (1999). Defining housing submarkets. *Journal of Housing Economics*, 8(2), 160–183.
- Bourassa, S. C., Hoesli, M., & Peng, V. S. (2003). Do housing submarkets really matter? *Journal of Housing Economics*, 12(1), 12–28.
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and regression trees*. New York: Chapman & Hall.
- Can, A. (1990). The measurement of neighborhood dynamics in Urban house prices. *Economic Geography*, 66(3), 254–272.
- Can, A. (1992). Specification and estimation of Hedonic housing price models. *Regional Science and Urban Economics*, 22(3), 453–474.
- Clapp, J. M., & Wang, Y. (2006). Defining neighborhood boundaries: Are census tracts obsolete? *Journal of Urban Economics*, 59(2), 259–284.
- Des Rosiers, F., & Theriault, M. (1996). Rental amenities and the stability of Hedonic prices: A comparative analysis of five market segments. *Journal of Real Estate Research*, 12(1), 17–36.
- Diewert, W. E., de Haan, J., & Hendriks, R. (2015). Hedonic regressions and the decomposition of a house price index into land and structure components. *Econometric Reviews*, 34(1–2), 106–126.
- Diewert, E., Saeed, H., & Silver, M. (2009). Hedonic imputation versus time dummy Hedonic indexes. In W. E. Diewert, J. S. Greenlees, & C. R. Hulten (Eds.), *Price index concepts and measurement* (pp. 161–196). University of Chicago Press.
- Dube, J., Legros, D., & Thanos, S. (2014). Putting time into space: Expectations and causality issues in the price determination process. In *Conference Paper, conference of the international association for applied econometrics*.
- Dubin, R. A. (1998). Predicting house prices using multiple listings data. *Journal of Real Estate Finance and Economics*, 17(1), 35–59.
- (2015). *Data Explorer - Distribution of Population by Tenure Status, Type of Household and Income Group*. Technical Report. Eurostat.
- Fahrlander, S. S. (2008). Indirect construction of Hedonic price indexes for private properties. *Swiss Journal of Economics and Statistics (SJES)*, 144(IV), 607–630.
- Goodman, A. C. (1981). Housing submarkets within Urban areas: Definitions and evidence. *Journal of Regional Science*, 21(2), 175–185.
- Goodman, A. C., & Thibodeau, T. G. (1998). Housing market segmentation. *Journal of Housing Economics*, 7(2), 121–143.
- Goodman, A. C., & Thibodeau, T. G. (2003). Housing market segmentation and Hedonic prediction accuracy. *Journal of Housing Economics*, 12(3), 181–201.
- Harter-Dreiman, M. (2004). Drawing inferences about housing supply elasticity from house price responses to income shocks. *Journal of Urban Economics*, 55(2), 316–337.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer-Verlag.
- Hort, K. (1998). The determinants of urban house price fluctuations in Sweden 1968–1994. *Journal of Housing Economics*, 7(2), 93–120.
- Knight, J. R. (2002). Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Economics*, 30(2), 213–237.
- Liu, X. (2013). Spatial and temporal dependence in house price prediction. *Journal of Real Estate Finance and Economics*, 47(2), 341–369.
- MacLennan, D. (1977). Some thoughts on the nature and purpose of house price studies. *Urban Studies*, 14(1), 59–71.
- Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In T. O'Sullivan, & K. Gibb (Eds.), *Housing economics and public policy: Essays in honor of Duncan MacLennan* (pp. 67–85). Oxford: Blackwell Science.
- Martins-Filho, C., & Bin, O. (2005). Estimation of Hedonic price functions via additive nonparametric regression. *Empirical Economics*, 30(1), 93–114.
- McCluskey, W., McCord, M., Davis, P., Haran, M., & McIlhatton, D. (2013). Prediction accuracy in mass appraisal: A comparison of modern approaches. *Journal of Property Research*, 30(4), 239–265.
- Michaels, R. G., & Smith, V. K. (1990). Market segmentation and valuing amenities with Hedonic models: The case of hazardous waste sites. *Journal of Urban Economics*, 28(2), 223–242.
- Munneke, H. J., & Slade, B. A. (2001). A metropolitan transaction-based commercial price index: A time-varying parameter approach. *Real Estate Economics*, 29(1), 55–84.
- Pace, R. K., Barry, R., Clapp, J. M., & Rodriguez, M. (1998). Spatiotemporal autoregressive models of neighborhood effects. *Journal of Real Estate Finance and Economics*, 17(1), 15–33.

- Pace, R. K., Barry, R., Gilley, O. W., & Sirmans, C. F. (2000). A method for spatial-temporal forecasting with an application to real estate prices. *International Journal of Forecasting*, 16(2), 229–246.
- Pace, R. K., & Gilley, O. W. (1997). Using the spatial configuration of the data to improve estimation. *Journal of Real Estate Finance and Economics*, 14(3), 333–340.
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55.
- Sirmans, G. S., & Benjamin, J. D. (1991). Determinants of market rent. *Journal of Real Estate Research*, 6(3), 357–380.
- Sirmans, S. G., Macpherson, D. A., & Zietz, E. N. (2005). The composition of Hedonic pricing models. *Journal of Real Estate Literature*, 13(1), 1–44.
- Statistik Stadt Zürich, 2005. Veränderung der Wohnverhältnisse in der Stadt Zürich zwischen den Eidg. Volkszählungen von 1990 und 2000, Publikationsreihe Analysen.
- Valente, J., Wu, S., Gelfand, A., & Sirmans, C. (2005). Apartment rent prediction using spatial modeling. *Journal of Real Estate Research*, 27(1), 105–136.