Regional Science and Urban Economics 60 (2016) 260–275

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Urban house price surfaces near a World Heritage Site: Modeling conditional price and spatial heterogeneity

Markus Fritsch[a], Harry Haupt[a], Pin T. Ng[b,c,*]

[a] University of Passau, Department of Statistics, Passau 94030, Germany
[b] Northern Arizona University, W. A. Franke College of Business, USA
[c] Anhui University, School of Economics, Hefei, China

ARTICLE INFO

Article history:
Received 17 March 2015
Received in revised form 14 July 2016
Accepted 28 July 2016
Available online 2 August 2016

JEL classification:
R12
R23
R3
C21
C51

Keywords:
Hedonic pricing
Quantile regression
Spatial association
Spline smoothing

ABSTRACT

In housing price regression, a large bundle of non-separable structural and location characteristics, potentially affecting prices nonlinearly, constitute the relevant set of predictors. Spatial subcenters and complex spatial association structures may, therefore, exist or, stated differently, horizontal market segmentation might be prevalent. Moreover, it is not unlikely for the housing price generating market mechanisms to vary across different parts of the conditional price distribution. This can ultimately cause disparate price segments to exhibit varying functional relationships through different subsets of characteristics and lead to vertical market segmentation. In order to take nonlinearity, horizontal and vertical market segmentation into account within the scope of housing price regressions, we propose incorporating a semiparametric approach into the quantile regression framework. In our empirical application, we investigate rental data from the German city of Regensburg, which contains an Old Town on the World Heritage List. Focusing on location effects exerted by the World Heritage Site, we illustrate how statements about horizontal and vertical market segmentation can be derived from a semiparametric quantile regression model based on empirical evidence and economic reasoning.

1. Introduction

Housing price analysis constitutes a complex multivariate problem, as housing is a heterogeneous (or differentiated) good consisting of a variety of attributes, which in turn differ with respect to numerous characteristics. Considerable research interest lies in models capable of explaining the observed market outcomes as much as possible, generating accurate forecasts of prices (of unobserved houses), and foremost providing insights into the complex underlying price generating mechanisms. The latter, captured by the hedonic housing price surface (Rosen, 1974), is usually analyzed via the marginal and aggregated effects that physical and location characteristics exert on prices, and can be addressed by using regression methods. However, regression analysis of marginal and aggregated hedonic effects is a particularly complex problem for several reasons.[1]

First, one of the attributes of a house is its geographic location; hence, housing markets are intrinsically spatial, leading to cluster or segmentation effects (see e.g., Goodman, 1981; Dale-Johnson, 1982, for early accounts on this problem) and spatial association (see e.g., Dubin, 1988, 1992; Cheshire and Sheppard, 1995, 1998; Basu and Thibodeau, 1998; Clapp et al., 2002; McMillen and Redfearn, 2010). The statistical consequence is that housing price data may exhibit spatial breaks, spatial correlation and spatial heterogeneity. Generally, the focus of approaches based on spatial econometric methods is to control for spatial association (e.g., Basu and Thibodeau, 1998; Pace et al., 1998b; Cohen and Coughlin, 2008; Holly et al., 2011; McMillen, 2015, and the literature cited therein), while approaches

* Corresponding author at: Northern Arizona University, W. A. Franke College of Business, Flagstaff, AZ86011-5066, USA.
E-mail addresses: markus.fritsch@uni-passau.de (M. Fritsch), harry.haupt@uni-passau.de (H. Haupt), pin.ng@nau.edu (P. Ng).

[1] Green and Malpezzi (2003, p. 32f) survey methods for the analysis of house prices. The issue of regression based estimation of price indices is beyond the scope of this paper (see Diewert, 2003; Silver and Heravi, 2007, or McMillen, 2008, for recent contributions).

http://dx.doi.org/10.1016/j.regsciurbeco.2016.07.011
0166-0462/

relying on methods from spatial statistics explicitly try to estimate the spatial structure (e.g., Clapp et al., 2002; Gelfand et al., 2003; Banerjee et al., 2004; Majumdar et al., 2006). Cressie (1993) discusses models of both worlds while Pace et al. (1998a) and more recently Kauermann et al. (2012) elaborate on inherent differences.

Second, housing prices are generically nonlinear with respect to the characteristics of the house. Hence, simple (local) linearization strategies are either incompatible with economic theory (e.g., Epple, 1987; Sheppard, 1999; Ekeland et al., 2004) or do not produce satisfying empirical results (e.g., Goodman, 1978; Halvorsen and Pollakowski, 1981), or both. Simulations of Cropper et al. (1988) reveal that parametric hedonic pricing models do not perform well. McMillen and Redfearn (2010) extend these arguments and provide Monte Carlo and empirical evidence in favor of nonparametric methods. These findings are in line with the recent applied econometric literature on hedonic pricing (e.g., Anglin and Gencay, 1996; Parmeter et al., 2007; Haupt et al., 2010).

We attempt to carefully address each of these issues and their delicate interplay. To begin with, we try to account for nonlinearities in both the structural and the spatial component of the hedonic regression. For this, we employ a flexible semiparametric specification incorporating additive structures by using univariate splines and triograms for modeling nonlinearities and spatial association. Such a strategy imposes only weak assumptions on the structure of the inherent data generating processes while avoiding the curse of dimensionality of fully nonparametric approaches. Appropriately specified, the resulting price surface permits the existence of irregularly distributed spatial subcenters – horizontal segmentation of a housing market.

We then incorporate the semiparametric approach for modeling nonlinearities and spatial association into the quantile regression framework (Koenker and Bassett, 1978; Koenker 2005). Quantile regression is a type of regression analysis that estimates the conditional median or quantiles of the dependent variable. While the popular least squares regression provides an estimate of the conditional mean given certain values of the covariates, quantile regression directly estimates the various quantiles of the dependent variables given the same set of values of the covariates. In addition to its inherent robustness against outliers in the dependent variable measurements, the main attraction of quantile regression lies in its ability to provide estimates for different aspects of measures of central tendency and dispersion. This provides a more complete picture of the conditional distribution of the dependent variable. The virtue of quantile regression in this context is that it may reveal different market mechanisms being prevalent across different portions of the conditional distribution of the housing price. Stated more explicitly, quantile regression considers possible varying functional relationships defining the influence of different subsets of characteristics on housing price across different segments in the distribution of the housing price. These potential disparities in functional relationships can be directly estimated via different conditional quantiles, thus, allowing vertical segmentation of a housing market.

In a nonlinear and spatially structured regression, however, it is not straightforward to assess the complexity of the spatial structure, the degree of nonlinearity, and the relevant subset of characteristics that affect housing price. This problem is further complicated by the fact that these properties of the regression may differ horizontally (for housing submarkets) or vertically (across the conditional distribution of the housing price). Thus, identification of the functional impact of relevant physical and location specific characteristics on hedonic housing price regressions constitutes a particularly challenging problem (see e.g., Parsons, 1990; Arguea and Hsiao, 1993, and Cheshire and Sheppard, 1995). Calculating the effective dimension of a hedonic price surface allows comparison to ad hoc parametric approaches but may be difficult in nonlinear models (see the discussion in Sections 4 and 5 of McMillen and Redfearn, 2010). This highlights another benefit of embedding the semiparametric approach into the quantile regression framework: The so-called exact-fit-property of quantile regressions allows us to estimate the effective dimension of a hedonic price surface – even in highly nonlinear contexts (see e.g., Koenker et al., 1994; Koenker and Mizera, 2004) – and constitutes an essential piece of information that enables us to determine the functional form and interpret the corresponding hedonic shadow prices and conditional predictions. To sum up, incorporating a flexible semiparametric specification in the quantile regression framework facilitates estimation of submarket specific hedonic price surfaces (and hedonic shadow prices) which subsequently enables us to simultaneously account for horizontal and vertical market segmentation.

Using a representative sample of the rental housing market in Regensburg, Germany in 2001, we empirically illustrate the proposed method for modeling of horizontal and vertical segmentation of the hedonic housing price surface. We employ flexible nonparametric univariate and bivariate smoothers. The resulting spatial structure under an anisotropic assumption enables us to capture the potential heterogeneity in the spatial distribution of housing prices in Regensburg. The nonparametric approach is purely data-driven while avoiding the necessity of complex ad hoc parametrizations of traditional approaches. We detail specification, estimation and model selection for different methods allowing for horizontal and vertical variation in the hedonic housing price surfaces across different quantiles.

In our empirical analysis, we focus on one of the key amenities of Regensburg, the distance to the boundary of the Old Town (historic city center). This area is on the UNESCO World Heritage List.[2] After controlling for relevant housing characteristics in a flexible nonlinear fashion, maps of the estimated hedonic price surfaces reveal substantial spatial inequalities between neighborhoods to the east and to the west of the historic city center besides a strong historic city center effect. Though this pattern holds generally, variations emerge across different quantiles. In addition, the effects of housing-quality inducing variables differ across quantiles. Economically, these inequalities are founded by urban development outside the historic city center taking quite different routes in the eastern neighborhoods compared to the western neighborhoods, leading to the city districts in the east and west diverging in terms of their structural amenities and disamenities.

The remainder of the paper is organized as follows: Section 2 introduces spatially additive hedonic price quantile regression models and embeds our approach in the relevant literature. In Section 3, we apply the proposed method to investigate and economically interpret the complexity and variability in hedonic shadow prices and price predictions using spatially structured urban housing market data. Finally, Section 4 summarizes our findings.

2. Smooth modeling of spatial heterogeneity

To attempt to address the issues listed in Section 1, we define a hedonic price equation to be of the form $p = h(\mu, \eta, \epsilon)$, with a scalar dependent price variable p, structural regression component μ, spatial component η and latent (error) component ϵ. It has been known since the 1950s that any multivariate function can be represented by finite composition (addition) of continuous functions of

[2] "Located on the Danube River, the Old Town of Regensburg with Stadtamhof is an exceptional example of a central-European medieval trading center, which illustrates an interchange of cultural and architectural influences. The property encompasses the city center on the south side of the river, two long islands in the Danube [...], and the area of the former charity hospital [...]. A navigable canal, part of the European waterway of the Rhine-Main-Danube canal, forms the northern boundary of Stadtamhof." (Quote from UNESCO World Heritage Centre (2016) at http://whc.unesco.org/en/list/1155).