Hedonic Analysis of Housing Sales Prices with Semiparametric Methods

Vincenzo Del Giudice, University of Naples Federico II, Naples, Italy
Benedetto Manganelli, University of Basilicata, Potenza, Italy
Pierfrancesco De Paola, University of Naples Federico II, Naples, Italy

ABSTRACT

This study estimates a hedonic price function using a semiparametric regression based on Penalized Spline Smoothing, and compares the price prediction performance with conventional parametric models. The excellent results obtained show that the semiparametric models allow to obtain a significant improvement in the prediction of housing sales prices.

KEYWORDS

Additive Models, Penalized Spline, Property Prices, Real Estate Appraisal, Semiparametric Regression

INTRODUCTION

Describe the general perspective of the chapter. Toward the end, specifically state the objectives of the chapter.

The evolution of real estate markets is influenced by quantitative and qualitative characteristics, as well as by differentiation and the change in the mode of appreciation of the real estate goods. These aspects suggest the development of new and advanced models for hedonic analysis of property prices, able to recognize the different forms of appreciation, based on survey and statistical analysis of market data (Morano et Al., 2014; Manganelli et Al., 2014).

In this paper, it’s proposed a semi-parametric statistical model that can to improve the performance of usual predictive multiparametric models.

In international literature, many studies have applied some special non-parametric or semiparametric additive regressions for the formulation of hedonic price models for the analysis of housing market. Mainly, these studies make use to Generalized Additive Models, among the most common non-parametric multivariate regression techniques, and the “backfitting algorithm” (Friedman & Stuetzle, 1981) that represents main method for resolution of additive models in base to available statistics data.

About Generalized Additive Models, they are based on the sum of q non-parametric functions, relating to q variables (T), plus a constant term (α); also providing for the use of a link function, note and parametric [G(·)], such as to connect the different functions that bind the dependent variable (Y) for each predictor (Ruppert et Al., 2003):

\[ E[Y|T] = G[\alpha + \sum_{j=1}^{q} f_j(T_j)] \]
where $T$ is a generic vector of numerical explanatory variable: $T = (T_1, ..., T_q)^T$.

On the other hand, the backfitting algorithm represents a very flexible tool, which allows to build models with very articulate components defined by an iterative procedure which suggests the algorithm that can be used for the estimation of individual functions in an additive model, removing the effects of all other functions (other of $i$-th considered) towards the dependent variable. This iteration of the algorithm is then performed until the individual function of the model does not suffer more variations between one iteration and the other (Hastie & Tibshirani, 1990).

An alternative approach with limited computational difficulties in estimating the individual functions that define an additive model, consists to place and match to each of these functions some specific smoothing spline function.

Currently the use of smoothing spline functions interest many scientific fields, like chemistry, natural and physical sciences, medicine, economy (limitedly to production costs only).

Early implementations of smoothing spline functions are recognized in some applications of physical sciences (Whittaker, 1923), with few significant studies in the economic (Engle et Al., 1986; Koenker, Ng & Portnoy, 1994; Craig & Ng, 2001), albeit very limited are currently the applications in real estate field.

It should be also reported that studies relating to the property market and to urban economy, focused on the use of smoothing spline functions, are mainly due to Anderson (1982, 1985), which estimate forecast models of urban residential density; Sunderman et al. (1990) formulate models to check the iniquity of the U.S. taxation systems on the real estate; Speyrer and Ragas (1991) highlight the impact of the risk of flooding on property values; Zheng (1991) applies some density functions to examine the spatial structure of Tokyo’s metropolitan area; Coulson (1992) formulates a model for estimating hedonic price based on smoothing spline for a properties sample in Pennsylvania; Bao and Wan (2004) analyze the residential property market in Hong Kong based on a sample of about 170,000 transactions relating to real estate units.

Recently, international scientific research has also further implemented the semiparametric techniques regressive employing functions smoothing spline, combining them with mixed models characterized by a partial systematic component. This has given rise to applications of semi-parametric regression models with mixed effects that use specific penalized spline functions, greatly limiting the random effects (Parise et Al., 2001; Coull, Ruppert & Wand, 2001; Coull, Schwartz & Wand, 2001; Zheng & Little, 2004; Opsomer et Al., 2008; Montanari & Ranalli, 2006).

Semi-parametric models applied to real estate appraisals are currently subject of specialized literature and, particularly, it concerns choice and processing of property prices and real estate features (Bin, 2004; Clapp, 2004; Gencay & Yang, 1996; Pace, 1998).

**MODEL SPECIFICATION**

Provide broad definitions and discussions of the topic and incorporate views of others (literature review) into the discussion to support, refute or demonstrate your position on the topic.

The relationship between selling price and explanatory variables is examined with a semi-parametric additive model, characterized by the combination of a generalized additive model, which expresses the relationship between the non-linear response and the explanatory variables, and a linear mixed effects model, which expresses the spatial correlation of observed values (Del Giudice & De Paola, 2014):

$$prezzo = \beta_0 + \beta_{serv} + \beta_{affacci} + \beta_{man} + \beta_{liv} + f_1(\text{epoca}) + f_2(\text{cvend}) + f_3(\text{sup int}) + f_4(\text{sup balc}) + f_5(\text{sup coll}) + \epsilon_1$$

(1)
Sources of Soil Pollution
[www.igi-global.com/chapter/sources-of-soil-pollution/206531?camid=4v1a](www.igi-global.com/chapter/sources-of-soil-pollution/206531?camid=4v1a)

The Effects of Cloud Approach in Short Chain Administration
[www.igi-global.com/article/the-effects-of-cloud-approach-in-short-chain-administration/120470?camid=4v1a](www.igi-global.com/article/the-effects-of-cloud-approach-in-short-chain-administration/120470?camid=4v1a)