UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Finance

The Pursuit of Preferences: A Hedonic Real Estate Appraisal

Sebastian Backlund and Per Svennerholm

Supervisor: Evert Carlsson
Master Degree Project No. 2016:119
Graduate School

SCHOOL OF BUSINESS, ECONOMICS AND LAW AT UNIVERSITY OF GOTHENBURG

Abstract

Graduate School

Master of Science in Finance

The Pursuit of Preferences:
A Hedonic Real Estate Appraisal

by Sebastian Backlund and Per Svennerholm

This study modifies the hedonic Builder's Model through assigning specific variable distributions, in an attempt to appraise individual real estates. Using data from six different submarkets surrounding Gothenburg, we first calibrated a pooled constrained F-test to determine that the submarkets should be estimated individually. Marginal prices were then estimated to confirm the model's economic significance as they reflected the submarkets' descriptive statistics. After economic confidence was established, we estimated Tobin's Q and saw that four submarkets were trading at a premium. However, fundamental economic factors such as low interest rates and scarcity of land motivates the premium and neglects the suggestion of an overvaluation. Lastly, an out-of-sample *post hoc* forecast was calibrated with a log-log model used as a reference model. Our model does not statistically outperform the reference model albeit the lower absolute residuals and volatility. When analyzing the forecast results, we found that misleading input data and the lack of time to fully calibrate the model were the most important factors interfering with our results.

Keywords: *Preferences, Hedonic Regression, The Builder's Model, The Modified Builder's Model, Pooled Constrained Regression, Marginal Prices, Tobin's Q, Individual Real Estate Appraisal*

Acknowledgements

Firstly, both of us would like to thank each other for the countless hours spent on this study. Without such great friendship, the long nights and days would never have been this enjoyable. Hopefully our journeys will cross again when moving on to our professional lives. Privately, our families have always supported us. Never had we been where we are today if it had not been for all them advises, reprimands, and driving to various practices and tournaments. We could therefore not be more grateful to have them all in our lives.

Since the foundation of all knowledge is repetition, we feel obliged to all our *Alma mater*, with a special acknowledgement to Ph.D. Mattias Sundén for his guidance through our econometric obstacles. Finally, Ph.D. Evert Carlsson, the man with superior knowledge and outstanding charisma. The long and interesting conversations we have had together have always enlightened us with much more than financial knowledge. His courses have always possessed with the greatest challenge. However, the piles of work he has put us through has not only spurred our financial interest, but also prepared us for our future careers. For all this, we are pleased to have had guidance from *"the guru"*, whom we will remember, admire, and be grateful to long after this study has been finalized.

Contents

List of Figures

List of Tables

List of Abbreviations

TP = Transaction price in MSEK
BA = Building area measured in square meters (m^2)
LS = Lot size of the real estate in units of square meters (m^2)
H = The height above reference point expressed in meters
TTC = Travel time in minutes to downtown Gothenburg (Kungsportsplatsen)
DTW = Distance to water measured in meters to closest shoreline
Q = The house's taxation point [0, 60] represent the quality of the building
SQ = School quality per submarket expressed in percentile rank
AGE = Age of building measured in years at transaction date
K = Dummy variable (= 1 if a terrace house; 0 otherwise)
R = Dummy variable (= 1 if a town house; 0 otherwise)
HOX = Aggregated villa price movement in Gothenburg

1 Introduction

The purpose of this study was to develop a model that appraises individual real estates as the sum of house and lot value. This is important because a real estate is associated with the overall economic health and wealth of a nation as it constitutes a large portion of a household's balance sheet (European Central Bank, 2016; Chin and Chau, 2002). Previous Swedish literature has primarily focused on the macro factors driving the real estate prices, *e.g.* Englund (2011). Our method follows the hedonic approach where we focus on the bundles of characteristics associated with the real estate. We used data between 2008 and 2015 from six different submarkets surrounding Gothenburg, when modifying the empirically tested Builder's Model through assigning specific variable distributions. The variables were the characteristics associated with the real estate, *e.g.* travel time to city and quality of house. Before testing the model on out-of-sample data, we calibrated a pooled constrained F-test to test an appropriate market size. This is important because previous research has concluded that a smaller submarket size will yield statistical gains, which our findings supported was (Goodman and Thibodeau, 2003).

We are interested in economic significance of our model because the parameters then reflects the "real world", *i.e.* economic significance is defined as the magnitude and sign of the estimated parameters. Besides interpreting our parameters, we also investigated the economic significance of the model by estimating the marginal price of characteristics. Previous research has focused on marginal prices as they revealed the key value drivers of a good in which can be used to estimate the demand curve (Palmquist, 1984; Rosen, 1974). Since we calibrated our own variable distributions, the marginal prices will display the model's economic significance if they reflect our set distributions and the submarkets' descriptive statistic. Conclusively, we found that the model was economically significant because the characteristics located exclusively in a submarket were observed as key value drivers, and they correspond to the descriptive statistics.

After economic significance was confirmed, we investigated what consumers have paid for a square meter (m^2) real estate in relation to the production cost of one m^2 through Tobin's Q. The increase in real estate prices as of late has spurred the word "housing bubble" to appear in Swedish newspapers more than occasionally. We therefore wanted to investigate the validity of such statement since a "housing bubble" is a threat against the Swedish economy. We found that four out of six submarkets were trading at a Tobin's Q above one. This implies that the submarkets were overvalued since consumers have paid more for an existing m^2 real estate than it would have cost to purchase a new one. Consequently, an economic rationale for construction firms exists in these submarkets, where construction firms should exploit the market conditions and to build more houses. However, we observed that the reason behind the Tobin's Q above one was due to the decreasing domestic interest rates and scarcity of land. Henceforth, the high Tobin's Q was the result of fundamental economic factors and does not necessarily imply the submarkets were overvalued.

Lastly, we conducted an one-step ahead *post hoc* forecast using out-of-sample data. Since our objective was to develop the foundation of a model that appraises real estate prices as the sum of a house and lot value, the first step was to outperform a reference model. We therefore calibrated a log-log model as our reference model. Our conclusion from the forecast was that there was no statistical difference between the two models albeit the lower absolute residuals and volatility seen in our Modified Builder's Model. When analyzing the forecast and its residuals, we found that misleading input data and our time limitations might have harmed the prediction accuracy of our Modified Builder's Model.

The discussions of characteristics as value drivers can be traced back to Lutz (1910), whom concluded real estate prices reflect the characteristics of its community (Coulson, 2008b). In the 1920s, regression analysis was added to the discussion of characteristics being value drivers. For example, Haas (1922) tried to develop a model that appraises farmlands in Blue Earth County, Minnesota. He made the classical economic assumption of people being rational, *i.e.* they wanted to maximize their own utility, when regressing historical transaction prices of farmlands. The explanatory variables he included in the regression were the depreciation rate of buildings per acre, land classification index, productivity soil index, and distance to market. He utilized the parameter estimates from the regression to make a *post hoc* forecast of farmlands in 1918 and 1919. He found that his model overestimated the farmlands by 6.63 USD in 1918 and 2.96 USD in 1919.

Waugh (1928) also used a linear regression, but focused his attention to the characteristics consumers paid for when purchasing vegetables on the Boston market in 1927. His interest was to understand which quality factors determined the price the buyer and seller agreed upon. For example, when he analyzed 200 individual lots of asparagus, the color of the asparagus was observed as the most important characteristic. For tomatoes, he noted a lower demand on weekends. This resulted in lower tomato prices on the weekends and decreased the importance of the vegetable's color. Thus, he concluded that the quality characteristics of a good was reflected in the transaction price.

The name "hedonic analysis" was not founded until Court's automobile study in 1939 (Coulson, 2008a). Through decomposing the price of an automobile via its characteristics, *i.e.* the horsepower, braking capacity, and window area, he formed an automobile price index. Since Court named the type of analysis as hedonic, he is referred to as "the founding father of hedonic analysis" (Coulson, 2008a).

Most hedonic models used today are derived from the work of Lancaster (1966) and Rosen (1974). They both argued that a good consists of a bundle of characteristics in which the consumer receives utility when consuming the good. A buyer is therefore paying for the characteristics of a good and not for the good *per se*. Lancaster (1966) assumed there is a linear relationship between the price of a good and its characteristics, whereas Rosen (1974) assumed a nonlinear relationship if the characteristics can be separated and repackaged by the consumer. Diewert (2003) explained the separable assumption as a consumption trade-off between characteristics that maximizes the utility. Henceforth, a consumer will consume a quantity of characteristic in accordance with his or her utility curve. Furthermore, researchers typically assume Rosen's (1974) nonlinear relationship between the price and the characteristics. For this assumption to be valid, the market must be in equilibrium (Colson and Zabel, 2012).

Our model was derived from the existing Builder's Model. According to Diewert and Shimizu (2015), the Builder's Model is derived from Rosen's assumptions. We therefore impose the same assumption in our model. Furthermore, the Builder's Model was developed to create separated quality indices for house and lot values to help government agencies when estimating the assets of the household sector (Diewert and Shimizu, 2015; Eurostat, 2013). Several studies have used the Builder's Model on both real estates and condominiums when estimating price indices in the Netherlands and Japan, respectively (*cf.* Diewert and Shimizu (2015) and Diewert *et al.* (2011)). Considering, the Builder's Model was designed to create real estate indices, we attempt to make it applicable on individual real estates by assigning specific variable distributions.

The remaining sections of this thesis are organized as follows. In section two, the econometric framework of the models is presented. Section three discusses the data gathering process. In the fourth section, we calibrate the reference model and the Modified Builder's Model before presenting the empirical results of the pooled constrained regressions, the marginal prices, and Tobin's Q. In section five, we calibrate the forecast and demonstrate the empirical result before the concluding remarks are made in the final section.

2 The Models

2.1 The Reference Model

Previous research suggests that there is no standard functional form used in hedonic regressions. However, Malpezzi *et al.* (1980) recommend the semi-log functional form (*cf.* Equation 2.1) for its simplicity of interpreting the coefficients, increased flexibility, and reduced variance when implementing non linearity[1]. The log-log functional form (*cf.* Equation 2.2) shares many of the advantages of the semi-log model, but the interpretation of coefficients becomes more straight-forward as they are now interpreted as elasticities (Coulson, 2008b). The functional form of a log-log model utilizes the log of all non-zero variables. Dummies and other variables, which can take a negative or zero value, will be entered linearly (Coulson, 2008b). Freeman (1993) argues that the functional form should be based upon a goodness-of-fit criteria such as the R^2.

$$log(Y_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_i X_N + \varepsilon_i \tag{2.1}$$

$$log(Y_i) = \beta_0 + \beta_1 log(X_1) + \beta_2 X_2 + ... + \beta_i log(X_N) + \varepsilon_i \tag{2.2}$$

2.2 Pooled Constrained Regression

Goodman and Thibodeau (2006) demonstrate that determining the number of submarkets in a model is likely to increase the prediction accuracy of the forecast. They suggest an F-test when the submarkets are nested and

[1] The semi-log functional form is achieved by taking the natural logarithm of the transaction price.

conditional on the number and composition of them (Goodman and Thibodeau, 2003). The pooled F-test they conducted, compares an unrestricted model, where all submarkets were estimated independently (*cf.* Equation 2.4), to a restricted model, where all submarkets' parameters were constrained to equality (*cf.* Equation 2.3, where $j \equiv submarket, i \equiv variable$and $t \equiv observation$).

$$log(Y_{j,t}) = \sum_{i=1}^{N} \beta_i \, log(X_{i,j,t}) + \varepsilon_{j,t} \tag{2.3}$$

$$log(Y_{j,t}) = \sum_{i=1}^{N} \beta_{i,j} \, log(X_{i,j,t}) + \varepsilon_{j,t} \tag{2.4}$$

The F-test for statistical significance of spatial disaggregation was given by:

$$F_{d,\sum_{i=1}^{N}(n_i - v_i)} = \frac{\frac{SSE_r}{d}}{\frac{SSE_u}{\sum(n_i - v_i)}} \tag{2.5}$$

Where SSE_r and SSE_u was the sum of squared error in the *restricted* model and *unrestricted* model, respectively. Furthermore, *d* was the number of restrictions, n_i was the number of observations in submarket i, and v_i was the number of estimated parameters in the specific submarket. The test statistics follows an F-distribution with *d* and n_i-v_i degrees of freedom. The aggregated market was deemed appropriate if failure to reject the null hypothesis was concluded (Goodman and Thibodeau, 2003).

2.3 The Builder's Model

2.3.1 The Foundation

The Builder's Model was developed to address the problem of creating a house and lot index separately. According to Eurostat (2013), the difficulties of creating a separate house and lot index primarily arise from:

- Houses are *heterogeneous* by nature since two identical real estate's can never occupy the same location simultaneously.
- Depreciation, renovations, and remodeling interferes with the assumption of a constant quality index.

In a parallel universe, the first issue would not impose any problems, and hedonic analysis would be less needed (Coulson, 2008b). Furthermore, as a response to the second issue, Diewert *et al.* (2011) argues that it is essential to decompose the real estate into two components; lot and house, since a lot does not depreciate in value over time in contrast to a house. This resulted in the development of the simple Builder's Model. From this model, three versions have emerged: (1) *the vacant land method*; (2) *the construction cost method*; and (3) *the hedonic regression method*. The first two methods use a fixed value for the lot and house, respectively, in order to find starting values for the nonlinear regression. For example, the construction cost method typically uses the reported constant production cost of a m^2 listed by government agencies (Diewert *et al.*, 2011). This ensures a constant quality price index for the house.

2.3.2 The Model

Consider an equation with exogenous transaction prices ($TP_{j,t}$) explained by house ($BA_{j,t}$) and lot size ($LS_{j,t}$) and their corresponding parameters, λ_j and θ_j (*cf.* Equation 2.6):

$$TP_{j,t} = \theta_j LS_{j,t} + \lambda_j BA_{j,t} + \varepsilon_{j,t} \tag{2.6}$$

This form is regarded as the *simple* Builder's Model, which is normally used for newly built houses. Therefore, adjustments are necessary for existing and resold houses (Diewert *et al.*, 2011). Considering, an old house is worth less than a new as the house structure depreciates in value over time, *ceteris paribus*. Therefore, ($AGE_{j,t}$) and a depreciation rate (γ) is implemented to capture the *net depreciation rate*.[2] As a result of adding the depreciation rate, the simple model (*cf.* Equation 2.6) converts into a nonlinear model defined as (Eurostat, 2013):

$$TP_{j,t} = \theta_j LS_{j,t} + \lambda_j BA_{j,t}(1-\gamma)^{AGE_{j,t}} + \varepsilon_{j,t} \tag{2.7}$$

Previous literature estimates the *net depreciation rate* to range in-between 0.5 and 2 percent, but adds that it can be on the downside of the "true" depreciation because renovations and reconstructions are left out (Eurostat, 2013). Furthermore, the model presented (*cf.* Equation 2.7) has not yet taken into account the effect of additional characteristics that could explain variations in the transaction price (Eurostat, 2013). For example, a lot's value could diverge depending on characteristics such as the distance to water and travel time to city. Whereas a house's value might differ depending on

[2] Diewert *et al.* (2011) uses $(1-\gamma)AGE_{j,t}$ whereas Diewert and Shimizu (2015) uses: $(1-\gamma)^{AGE_{j,t}}$.

characteristics such as the quality of house or type of dwelling (villa, row house, town house). Lot characteristics are expressed as ($\sum_{i=1}^{N} X_{i,j,t}\eta_{i,j}$) and house characteristics as ($\sum_{i=1}^{N} Z_{i,j,t}\rho_{i,j}$) (*cf.* Equation 2.8). Consequently, the functional form with the additional characteristics is expressed as:

$$TP_{j,t} = \left[\theta_j LS_{j,t}\left[1+\sum_{i=1}^{N} X_{i,j,t}\eta_{i,j}\right] + \lambda_j BA_{j,t}\left[1+\sum_{i=1}^{N} Z_{i,j,t}\rho_{i,j}\right](1-\gamma)^{AGE_{j,t}}\right] + \varepsilon_{j,t} \tag{2.8}$$

In order to capture the marginal utility of increasing the number of m^2 of a lot, Diewert *et al.* (2011) add linear splines to the model. They implement splines because the marginal utility is expected to be increasing at a decreasing rate. This means that at a certain threshold, an additional unit of m^2 will not add as much utility to the consumer. Henceforth, the spline coefficients' ($\theta_{k,j}$) expresses the marginal price to pay for an additional m^2 dependent on the current size, *ceteris paribus*. Also, imposing splines will capture the cost curve of producing a real estate, *i.e.* economies of scale (Industrial Systems Research, 2013). The f and g (*cf.* Equation 2.9 - 2.10) are the number of m^2 set for each segment, where $k \equiv k^{th} spline$.[3]

$$Splinelot_{k,j,t} = [\theta_{k,j} min(LS_{j,t}, g) + \theta_{k,j} max(LS_{j,t} - g, 0)]\left[1+\sum_{t=1}^{T} X_{i,j,t}\eta_{i,j}\right] + \varepsilon_{j,t} \tag{2.9}$$

$$Housevalue_{k,j,t} = \lambda_{k,j} BA_{j,t}\left[1+\sum_{i=1}^{N} Z_{i,j,t}\rho_{i,j}\right](1-\gamma)^{AGE_{j,t}} + \varepsilon_{j,t} \tag{2.10}$$

2.4 The Marginal Price of a Characteristic

The partial derivative of a hedonic equation with respect to the characteristic ($\frac{\partial TP_{j,t}}{\partial X_{i,j,t}}$) reflects the marginal price, *i.e.* the consumer's willingness to pay for one additional unit. Lancaster's (1966) linear assumption states the marginal price will be constant, *i.e.* not dependent on the quantity consumed. Rosen's nonlinear assumption argues that the marginal price will depend on the quantity consumed (Rosen, 1974; Palmquist, 1984). The marginal price of the nonlinear Builder's Model, which is derived from Rosen's theory, depends on the other characteristics of the real estate as the model consists of multiple interaction terms (*cf.* Equation 2.8).

2.5 *Q*-Theory

Tobin's *Q* is the ratio of the market value of an asset (firm or house) divided by their replacement cost: $Q = \frac{Market\,value\,of\,installed\,capital}{Replacement\,cost\,of\,capital}$ (Foote, 2010;

[3] The general function of a linear spline: $f\min(X, 100) + f\max(X - 100, 0)$.

Sims, 2011). A Tobin's Q above one indicates the market values the house more than it would cost to build a new house. This gives construction firms an incentive to invest as they will make a profit from building and selling new houses. For example, "When a house is worth more than the wood, nails, and labor it takes to build one, builders are going to build a lot of homes" (Foote, 2010). Q theory makes the assumption of an efficient market in which future interest rates, for example, will be incorporated in the transaction price (Stevens, 2005). A decrease in interest rates is therefore associated with an increase in Tobin's Q as the discount factor of the expected future earnings will increase. Henceforth, the cash flow the house is expected to generate in the future is discounted by a larger factor, which will make the numerator (market value of installed capital) in the Tobin's Q equation lower, *ceteris paribus*. Consequently, it will be more profitable for construction firms to build new houses, which results in an increase in housing investments (Foote, 2010).

2.6 Real Estate Appraisal

Inserting the characteristics from out-of-sample transactions into the estimated model constitutes a forecast for prices of real estates. When comparing two or several models, Goodman and Thibodeau (2003, 2006) use different descriptive statistics, such as average residuals, proportional errors $(\frac{Residuals}{Price})$, absolute average residuals, absolute average proportional errors, and volatility. Another frequent measure is Theil's coefficient (*cf.* Equation 2.11), which represents the R^2 of the forecast and demonstrates how much of the variation in the transaction prices the model captures. However, to evaluate if the difference between two models is statistically different, the Morgan-Granger-Newbold test can be applied. The MGN test (*cf.* Equation 2.12) determines if the difference in forecast errors of the models, the quadratic loss function, is zero using a t-distribution (Clapp and Giaccotto, 2000)[4]. Henceforth, the two-tailed t-test's null hypothesis is that the forecast error is equal to zero, while the alternative hypothesis is that there is a difference.[5]

$$U_t^2 = 1.0 - \frac{\frac{1}{n_t}\sum_{i=1}^{n_t}(y_{i,t} - \hat{y}_{i,t})^2}{\frac{1}{n_t}\sum_{i=1}^{n_t}(y_{i,t} - \bar{y}_{i,t})^2} \tag{2.11}$$

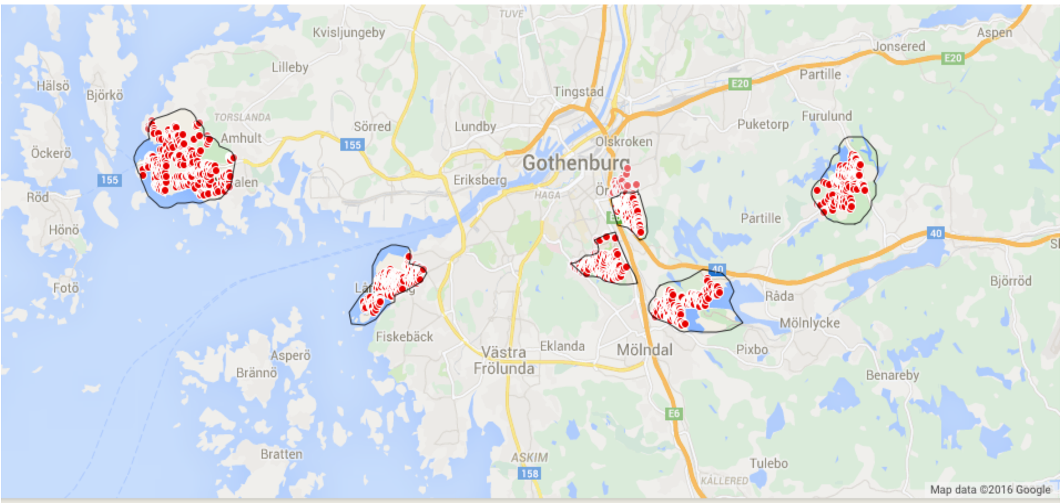
$$MGN = \frac{r}{\left[\frac{(1-r^2)}{(T-1)}\right]^{(1/2)}}, \quad \text{where} \quad r = \frac{x'z}{[(x'x)(z'z)]^{(1/2)}} \tag{2.12}$$

[4] The mean squared error is referred to as the quadratic loss function.
[5] It is a two-tailed test because it is a quadratic loss function.

3 Data

We gathered transaction data for six different submarkets, which was delineated by Booli (*cf.* Figure 2.1).[1] The submarkets have been chosen for their environmental similarities as they pair-wise have similar landscape, social class, and proximity to downtown Gothenburg.[2] Furthermore, our data set was a pooled cross-section time-series sample between 2008 and 2015. A house sold in 2008 will not have the same price if it was sold in 2015 due to real changes in the economy, *e.g.* the decrease in domestic interest rate after the financial crisis of 2008 (Riksbanken, 2016). Therefore, we used the Valueguard's Villa Housing index in Gothenburg to account for aggregate price movements (Wilhelmsson, 2000). The housing statistics were collected from Booli, Hemnet, and Skatteverket's databases, whereas the land characteristics were obtained from Hitta.se, Västtrafik.se, and Google Maps. The characteristics we used have all been tested in previous literature and are expected to yield a significant impact. Motivations for each variable, their expected impact, submarket descriptive statistics, and their respective gathering source are found in Appendix A. Due to lack of data, we have excluded variables found in previous research such as air pollution, traffic noise, and crime rate.

FIGURE 3.1: Submarkets



The six submarkets listed from left to right: Hjuvik (W), Långedrag (SW), Krokslätt (S), Örgryte (E), Stensjön (SE), and Öjersjö (E)

[1] The submarket Örgryte was expanded based on our local knowledge of the submarket. Therefore, transactions found at Booli from neighboring areas that people refer to as Örgryte were included: Skår, Jakobsdal, Överås, Orangerigatan, Bö, Bäckeliden, Danska vägen, and Santessonsgatan.

[2] Långedrag *versus* Hjuvik, Stensjön *versus* Öjersjö, and Örgryte *versus* Krokslätt.

Initially, there were 1,894 observations, however, the sample was reduced to 1,647 units because houses were sold past our time period, duplicates (identical observations), missing characteristics, and extreme outliers (*cf.* Table 3.1). Henceforth, including these observations in the regression would cause the parameters to suffer from omitted variable bias and multicollinearity (Eurostat, 2013). Regarding outliers; houses sold below their taxation value were excluded. This was decided because the taxation value should reflect 75 percent of the market value since the last address declaration. We therefore believe these houses have been sold within the family at a discount or something similar (Skatteverket, 2016). We also deleted lots, new houses, and real estates sold for twice or half the average price per m^2 in the specific submarket. The lots were excluded due to the lack of house structure, new houses because of no depreciation, and the remainders for their extremes (Diewert *et al.*, 2011).[3] Additional transactions for the first quarter of 2016 were gathered, but to be used in the forecast (*cf.* Table 3.2). Applying the same methodology when assembling the out-of-sample data used for prediction, a total of 58 transactions were gathered.

TABLE 3.1: In-sample Descriptive Statistics

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	1647	5.38	2.19	1.32	25	MSEK
Building area	1647	152	54	25	510	m^2
Quality of House	1647	30	5	13	54	Taxation Points
Age of house	1647	45	25	1	86	Years
Terrace house	1647	0.01	0.30	0	1	$\int_0^1$
Town house	1647	0.11	0.31	0	1	$\int_0^1$
Lot size	1647	826	441	60	4488	m^2
Height	1647	21	15	0	87	Meters
Travel time to city	1647	35	13	10	79	Minutes
Distance to water	1647	984	839	10	3520	Meters
School quality	1647	77	9	54	94	Percentile

Quantitatively describing the characteristics in the sample.

[3] The lots were saved for analysis when assigning distributions and estimating average m^2 prices.

TABLE 3.2: Out-of-sample Descriptive Statistics

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	58	7.13	2.23	3.53	13.2	MSEK
Building area	58	153	47	80	295	m^2
House quality	58	31	4	23	45	Points
Age of house	58	46	27	2	87	Years
Terrace house	58	0.17	0.33	0	1	$\int_0^1$
Town house	58	0.07	0.25	0	1	$\int_0^1$
Lot size	58	802	366	163	1729	m^2
Height	58	19	13	2	69	Meters
Travel time to city	58	34	13	13	62	Minutes
Distance to water	58	886	778	41	2910	Meters

The observed characteristics used to forecast transaction prices in 2016.

4 Calibration and Empirical Results

4.1 The Reference Model

We applied Freeman's (1993) approach when determining the functional form for the reference model, *i.e.* the higher goodness-of-fit. We began by adding all characteristics: distance to water ($DTW_{j,t}$), height above reference point ($H_{j,t}$), travel time to city ($TTC_{j,t}$), school quality ($SQ_{j,t}$), quality of house ($Q_{j,t}$), row house ($R_{j,t}$), town house ($K_{j,t}$), age ($AGE_{j,t}$), and HOX (HOX_j) to capture the aggregated price movement in submarket *j*. We tested two different functional forms, the semi-log and the log-log. We noted that the log-log model experienced a higher adjusted R^2 of 0.8391 compared to the semi-log's 0.8192. The functional form of the reference model was therefore determined to be in log-log (*cf.* Equation 4.1).

$$\begin{aligned} log(TP_{j,t}) =& \beta_{0,j} + \beta_{1,j} log(BA_{j,t}) + \beta_{2,j} log(LS_{j,t}) + \beta_{3,j} log(DTW_{j,t}) + \\ & \beta_{4,j} log(Q_{j,t}) + \beta_{5,j} log(TTC_{j,t}) + \beta_{6,j} log(AGE_{j,t}) + \\ & \beta_{7,j}(K_{j,t}) + \beta_{8,j}(R_{j,t}) + \beta_{9,j}(H_{j,t}) + \beta_{10,j}(HOX_j) + \varepsilon_{j,t} \end{aligned} \tag{4.1}$$

To control for multicollinearity, we conducted a Variance Inflationary test (VIF). The test implied that there was no presence of multicollinearity among the explanatory variables, based on the test's rule thumb of (10) (*cf.* Appendix C) (South Florida, 2016). Also, when conducting the Breush-Pagan and White-test, we observed signs of heteroskedasticity (*cf.* Appendix C), which was corrected by using robust standard errors (Alem, 2014).

4.2 The Pooled Constrained Regression

4.2.1 Aggregated, Pair-wise, or Individually

When deciding an appropriate market size, we followed the methodology proposed by Goodman and Thibodeau (2003, 2006). The six submarkets were therefore considered to be estimated aggregated, pair-wise based on environmental similarities, and individually. For example, in the aggregated market size all submarkets have the same parameter estimates. To fulfill the assumption of equal functional form in all submarkets, we utilized the reference model for the pooled constrained regressions.

The test statistics implied that all submarkets should be estimated individually at 0.1 significance level (*cf.* Table 4.1). Pair-wise outperformed the aggregated consolidation of submarkets, whereas the individual outperformed the pair-wise. Henceforth, albeit submarkets' similar nature, spatial discrepancies still exists. We observed such a discrepancy in the descriptive statistics (*cf.* Appendix A), where the average distance to water ($DTW_{j,t}$) in Långedrag and Hjuvik were 457 meters and 534 meters, respectively. Furthermore, due to lack of variation in school quality ($SQ_{j,t}$) within each submarket, *i.e.* all real estates within a submarket had the same level of school quality, the variable was excluded in the following sections.[1]

TABLE 4.1: F Statistics – Pooled Submarkets

Statistics	Aggregated	Pair-wise	Individual
Number of submarkets	1	3	6
Sum of Squared Errors	62.96	42.30	33.97
F Statistic	252.55	56.56	

F-test statistics for pooled submarket regressions.

We also tested for equal distribution of $R_{j,t}$, $K_{j,t}$, $AGE_{j,t}$, $TTC_{j,t}$, and $Q_{j,t}$ for all submarkets and $DTW_{j,t}$ for Örgryte and Krokslätt. We believe these variables should have the same impact independently on locality. For instance, $Q_{j,t}$ had a similar mean and volatility in all submarkets (*cf.* Appendix

[1] Resulting in a reduction in number of unknown parameters by six in both models.

A), which suggest similar parameters estimates. However, all F-tests were rejected at 0.001 significance level (*cf.* Table 4.2). Therefore, no distributional constraints were imposed on the reference model.

TABLE 4.2: F Statistics – Pooled Distributions

Statistic	K	R	AGE	Q	DTW	TTC	Individually
Number of Restrictions	5	5	5	5	1	5	0
Sum of Squared Errors	34.43	34.43	34.76	34.30	34.25	34.78	34.18
F Statistic	329.58	329.58	332.70	328.32	1639.27	332.89	

F-test statistics for pooled regression on individual coefficients.

4.3 The Modified Builder's Model

We will develop a modified version of the Builder's Model that originates from *the hedonic regression method*. This was decided because the model will be evaluated based on economic significance in which motivates all parameters to be estimated simultaneously.

4.3.1 The Simple Model

At first, we set up the Simple Builder's Model (*cf.* Equation 2.6) in order to decompose the transaction price ($TP_{j,t}$) into two components: house (λ_j) and lot value (θ_j). This resulted in an adjusted R^2 of 0.5931. Considering the relatively low R^2 and that our sample only consists of old houses, we added a *net depreciation rate* to capture the aging effect on the house structure. We deviate from Diewert and Shimizu (2015) by calibrating the *net depreciation rate* (γ_j) using a natural exponential function $\left(e^{-\gamma AGEj,t}\right)$ considering the rate is continuously compounding.[2] This resulted in an increase in the adjusted R^2 to 0.5948.

Furthermore, Diewert *et al.* (2011) removed all houses above the age of 50 years since these houses typically need abnormal renovations expenditures. We choose another direction by setting a condition; if a house was older than 25 years, the house was classified as 25 years old. This upper bound was partly decided because the observed average age of houses in Örgryte was close to 70 years. Also, the fundamentals of a house such as power lines, drains, and heating can all be utilized for long time periods if maintained properly (Dinbyggare, 2016). The life-span of such characteristics typically ranges between 25 and 50 years before they need a replacement (Villaägarna, 2016). Thus, these factors should already be reflected in

[2] $e^{-\gamma AGE_{j,t}} \approx (1-\gamma)^{AGE_{j,t}}$.

the price. With this restriction (*cf.* Equation 4.2) the adjusted R^2 increased to 0.6035.

$$TP_{j,t} = \theta_j LS_{j,t} + \lambda_j BA_{j,t} e^{-\gamma_j (25 \geq AGE_{j,t})} + \varepsilon_{j,t} \tag{4.2}$$

4.3.2 Adding Explanatory Variables

To capture more variations in the $TP_{j,t}$, we implemented additional explanatory variables. For the lot we added the characteristics: $DTW_{j,t}$, $H_{j,t}$, and $TTC_{j,t}$.[3] For the house we incorporated: $Q_{j,t}$, $R_{j,t}$, and $K_{j,t}$. The parameter for the aggregated price movement was raised to the power of β_j to reflect the submarkets' sensitivity to the index. If β_j was equal to one, then submarket *j's* price movement was equal to the aggregated. When adding these additional explanatory variables, the adjusted R^2 of our model increased from 0.6035 to 0.7192.

$$\begin{aligned} TP_{j,t} = HOX^{\beta_j} \, [\theta_j LS_{j,t} \, [1 + \beta_{1,j} X_{DTW,j,t} + \beta_{2,j} X_{TTC,j,t} + \beta_{3,j} X_{H,j,t}] \\ + \lambda_j BA_{j,t} \, [1 + \beta_{4,j} X_{Q,j,t} + \beta_{5,j} X_{K,j,t} + \beta_{6,j} X_{R,j,t}] \\ * \, e^{-\gamma_j (25 \geq AGE_{j,t})}] + \varepsilon_{j,t} \end{aligned} \tag{4.3}$$

4.3.3 Adding Splines

To capture the nonlinear marginal utility of increasing m^2 and the economies of scale seen in real estate production, we calibrated linear splines similar to Diewert *et al.* (2011). After analyzing our data set, we observed that consumers have different preferences for different m^2 segments on both the house and lot. For example, small real estates were sold at a higher m^2 price than the average m^2 price in all submarkets [4]. We therefore deviate from Diewert *et al.* (2011) when implementing linear splines on both the house and lot. Furthermore, our intuition was that a house should be priced the same independent on locality, *i.e.* consumer preferences for the same house should be equal in our submarkets. However, we observed small average m^2 price differences for various $BA_{j,t}$ segments in all submarkets. We therefore set different linear splines in all submarkets (*cf.* Appendix C), and expect minor differences in a house's value dependent on locality. Also, in some submarkets the m^2 price appeared to rise after an upper threshold was surpassed. In these submarkets, we calibrated a third spline ($\theta_{3,j}$), which ensured the full economies of scales "curve" was captured. Furthermore, the linear splines on lots were set after analyzing each submarkets' lot sizes in

[3] Recall from the pooled constrained regression that the school quality variable ($SQ_{j,t}$) was removed since all submarkets should be estimated individually.

[4] Smaller real estate $\equiv$ lower number of m^2 of house and lot.

relation to the $TP_{j,t}$. Large differences in the lot sizes between submarkets were noted (*cf.* Appendix A). We noted that consumers tend to pay more per m^2 for smaller lots. To capture this we set the first spline below the average lot size in each submarket. For example, approximately 22 percent of the lots were below 225 m^2 in Örgryte. Based on our findings we found it reasonable to assume that consumers in Örgryte value the first 225 m^2 more than the exceeding m^2 (*cf.* Equation 4.4). Finally, using Örgryte as an example, the living area was divided into two categories; less than 180 and above 180 m^2 (*cf.* Equation 4.5). Consequently, when implementing the splines in our model, the adjusted R^2 increased to 0.7300.

$$\begin{aligned} Splinelot_{k,Org,t} = {} & HOX^{\beta_{Org}} \left[\theta_{1,Org} min(LS_{Org,t}, 225) \right. \\ & \left. + \theta_{2,Org} max(LS_{Org,t} - 225, 0)\right] \\ & * \left[1 + \sum_{i=1}^{N} X_{i,Org,t}\eta_{i,Org}\right] + \varepsilon_{Org,t} \end{aligned} \tag{4.4}$$

$$\begin{aligned} Splinehouse_{k,Org,t} = {} & HOX^{\beta_{Org}} \left[\lambda_{1,Org}\, min(BA_{Org,t}, 180) \right. \\ & \left. + \lambda_{2,Org}\, max(BA_{Org,t} - 180, 0)\right] \\ & * \left[1 + \sum_{i=1}^{N} Z_{i,Org,t}\rho_{i,Org}\right] e^{-\gamma_j(25 \geq AGE_{Org,t})} + \varepsilon_{Org,t} \end{aligned} \tag{4.5}$$

4.3.4 Assigning Distributions

The Builder's Model was developed to estimate a separate house and lot index at an aggregate level using postcode dummies for locality (Haan and Gong, 2015). Consequently, we modified the functional form to fit the purpose of our study. We believe that by implementing specific variable distributions, the model will capture the price impact of characteristics on *individual* real estates.

When analyzing the removed lots from our sample, consumers' preferences for lots were observed. For instance, in Långedrag and Hjuvik, $DTW_{j,t}$ was noted to explain large variations in the $TP_{j,t}$. A lot located inside 50 meters from the ocean in Långedrag was sold for approximately 11 MSEK, whereas a similarly sized lot located approximately 1000 meters from the ocean was sold for approximately 2 MSEK. To capture the approximately 600 percent price difference, we calibrated a variable distribution for $DTW_{Langedrag,t}$.

When we analyzed the removed lots, we observed several "saddle points" for the impact of the characteristics (*cf.* Appendix C). Continuing with Långedrag and $DTW_{j,t}$ as an example, if the middle of the house was closer than 25 meters from the shoreline, the relatively high price impact seemed to reflect the probability of beach access. Between 25 and 50 meters from the shoreline, we observed another "saddle point" that might reflect the probability of seeing the edge of the water, and so forth. To capture these "saddle points", we calibrated logistic equations (*cf.* Equation 4.6) to assign a specific factor. This factor represents the percentage increase in the price per m^2 on the lot as it approaches the shoreline (*cf.* Figure 4.1).[5] For each segment, *e.g.* 0-25 meters from the shoreline, a logistic equation was calibrated. We calibrated *four* different logistic equations in Långedrag to capture the price impact of $DTW_{j,t}$. The parameters of the logistic equations ($c_{dtw}, a_{dtw}, b_{dtw}$) were calibrated in Excel to fit a curve reflecting the price impact of $DTW_{j,t}$. These parameters were therefore fixed values, which were not estimated in the regression (*cf.* Appendix C).

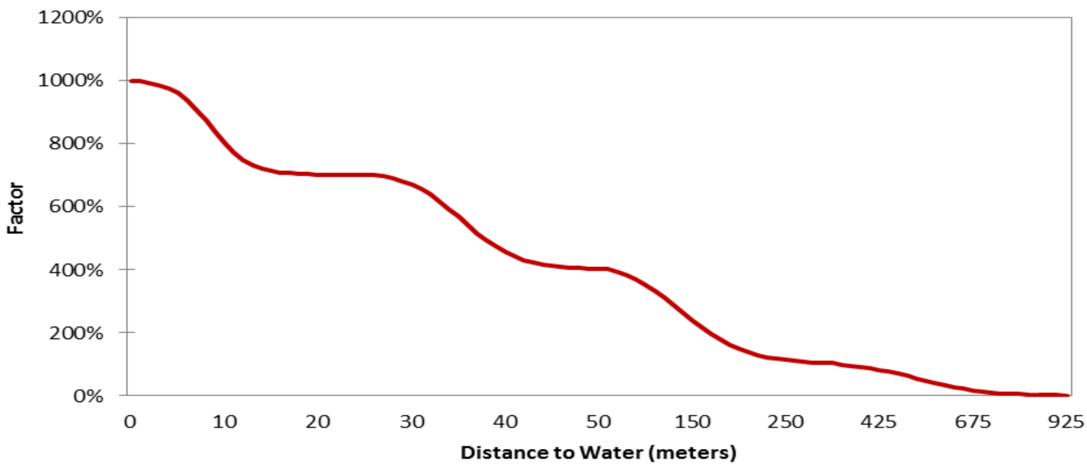
Furthermore, the least valuable lot with respect to $DTW_{Langedrag,t}$ was when the factor was equal to zero, *i.e.* $DTW_{Langedrag,t}$ approaches 1000 meter from the shoreline. [6] The highest factor the m^2 price can receive was 1000 percent (*cf.* Figure 4.1), *i.e.* 1000 percent more expensive than the least valuable lot. However, $DTW_{j,t}$ was measured from the middle of the house, which implies that half of the house is located in the water to receive the 1000 percent factor. The highest factor in our data set was approximately 700 percent in Långedrag. This was in accordance with the example stated earlier, since this observation was located closer to the shoreline, *i.e.* closer than 50 meters.

$$\mathit{factor} = \frac{c_{dtw}}{1 + a_{dtw}\, e^{b_{dtw} X_{dtw}}} \tag{4.6}$$

[5] In Appendix C - Assign Distributions, a deeper explanation of motivations, calibrations, and factors for all characteristics are presented.

[6] The $DTW_{j,t}$ factor will be zero outside of 1000 meters in all submarkets.

FIGURE 4.1: Distance to water distribution in Långedrag



The lot price factor in relation to distance to water

For a house ($BA_{j,t}$) the case was different because the $\lambda_{k,j}$ estimates were based on a standard quality of house. Henceforth, the quality of house ($Q_{j,t}$) distribution was calibrated so that 31 taxation points had no impact on the price (*cf.* Appendix C). According to Villaägarna, 31 taxation points reflect a standard $Q_{j,t}$ both on the exterior and interior of the house. Luxury items, a garage, and a newly renovated roof are all examples of characteristics included in $Q_{j,t}$ (Skatteverket, 2014). We calibrated the $Q_{j,t}$ distribution to decrease the m^2 price, $\lambda_{k,j}$, if it was below the standard quality of 31 taxation points and an increase if it was above. However, when analyzing our data set and out-of-sample data on apartments, small impacts were observed when $Q_{j,t}$ ranged between 15 and 45 taxation points (*cf.* Appendix C). With the distributional assumptions, the adjusted R^2 of our model increased to 0.8151.

In our model, we were more interested in economic significance rather than statistical significance, although a combination was preferred.[7] Economic significance demonstrates that our model is reasonable as the estimates reflected the "real world". To improve the economic significance of the model, we imposed the restriction of equal distribution between all submarkets for $AGE_{j,t}$, $TTC_{j,t}$, $Q_{j,t}$, $K_{j,t}$, and $R_{j,t}$ and equal distribution for $DTW_{Krokslatt,t}$ and $DTW_{Orgryte,t}$. Firstly, $DTW_{Orgryte,t}$ and $DTW_{Krokslatt,t}$ were assumed to have minimal explanatory power on $TP_{j,t}$ as all observations were located 1000 meters from the closest shoreline. Secondly, our

[7] With economic significance, we refer to the sign and magnitude of the estimated parameters. For example, the estimated price for a m^2 should not be negative.

intuition for estimating γ_t equally, was that houses should depreciate at the same rate as all submarkets are located close to each other. Similarly, for $TTC_{j,t}$ we believe consumer preferences are equal in all submarkets as the data represents walking and using the public transportation system to get to Kungsportsplatsen for all real estates, irrespective of their locality. For $Q_{j,t}$ we believe all consumers have the same preferences based on analyzing our sample and out-of-sample data on apartments in Gothenburg (*cf.* Appendix C). Lastly, when analyzing transaction prices for new houses in Tölö Trädgårdar in Kungsbacka, we observed that a detached real estate was sold for approximately 5 percent more per m^2. When imposing an equality constraint on the parameters for $K_{j,t}$ and $R_{j,t}$, both parameter estimates (β_K and β_R) were negative. Before the restrictions, the parameters held both negative and positive signs, which contradicts our findings in Tölö Trädgårdar.

When applying these restrictions, the number of parameters decreased by 15 and the adjusted R^2 was reduced by 0.0017 to 0.8129. The number of insignificant parameters reduced from 17 to 3 and all coefficients had the expected sign. For example, $\theta_{2,Orgryte}$ was negative before, which indicate a buyer would pay less for 225+ m^2 than 225 m^2. The functional form we have calibrated (*cf.* Equation 4.7) is referred to as the *Modified Builder's Model* (MBM), and includes 36 unknown parameters to be estimated: $\theta_{k,j}$, $\lambda_{k,j}$, γ, β_R, β_K, and β_j.

$$Value\,of\,Property\,(TP_{j,t}) \;=\; Lotvalue_{j,t} + Housevalue_{j,t} + \varepsilon_{j,t} \tag{4.7}$$

$$Lotvalue_{j,t} = HOX^{\beta_j}\left[[1 + lotfactor]\,[Splinelot_{k,j,t}]\right] \tag{4.8}$$

$$Housevalue_{j,t} = HOX^{\beta_j}\left[[1 + housefactor]\,[Splinehouse_{k,j,t}]\,e^{-\gamma(25\geq AGEj,t)}\right] \tag{4.9}$$

$$Splinelot_{k,j,t}(g) = \theta_{k,j}min(LS_{j,t}, g) + \theta_{k,j}max(LS_{j,t} - g, 0)$$

$$Splinehouse_{k,j,t}(f) = \lambda_{k,j}min(BA_{j,t}, f) + \lambda_{k,j}max(BA_{j,t} - f, 0)$$

$$housefactor = \beta_R R_{j,t} + \beta_K K_{j,t} + \frac{c_Q}{1 + a_Q e^{b_Q X_{Q,t}}}$$

$$lotfactor = \frac{c_{DTW}}{1 + a_{DTW} e^{b_{DTW} X_{DTW,j,t}}} + \frac{c_{TTC}}{1 + a_{TTC} e^{b_{TTC} X_{TTC,t}}} + \frac{c_H}{1 + a_H e^{b_H X_{H,j,t}}}$$

To control for multicollinearity in the MBM, the Condition Indices test on the coefficients was used. Since two variance proportions was never above 0.5 in conjunction with a condition index above 30, we rejected the presence of multicollinearity (*c.f.* Appendix C) (South Florida, 2016).

4.4 Coefficients

All estimated parameters in the MBM displayed their correct signs (*cf.* Table 4.3).[8] The distributions between θ and λ values were disproportional to our prior beliefs, *i.e.* low θ and high λ estimates. For example, $\theta_{2\,Krokslatt}$ suggests that if a lot exceeds 300 m^2 (the threshold of the first spline, $\theta_{1,Krokslatt}$), the marginal price of an additional m^2 was 203 SEK, *ceteris paribus*.[9] Furthermore, γ of 1.3 percent validated the authors' expectations as it was within the range proposed by Diewert and Shizmu (2015). Although, $\lambda_{3,Stensjon}$, β_k, and $\theta_{2,Orgryte}$ were insignificant, they still held the correct sign.[10] The generated standard errors of parameters, except for a few $\theta_{k,j}$, were comparable to other studies considering they typically ranged between 20 to 40 percent (Malpezzi, Ozanne, and Thibodeau, 1980). The adjusted R^2 of 0.8129, was in accordance with previous studies that typically range between 0.55 and 0.90. The coefficients have maintained the relation in transaction prices between submarkets (*cf.* Appendix A), and thereby demonstrating economic significance [11].

Almost all parameters in the reference model were significant at a 5 percent level. In Örgryte, Krokslätt and, Öjersjö $AGE_{j,t}$ was insignificant, and held the wrong sign in Örgyte. Several other parameters showed unreliable signs. For example, the estimated parameters in Hjuvik and Krokslätt implied that a row house and a terrace house were more expensive than a detached real estate. Another example was found in Långedrag where $TTC_{Langedrag,t}$ showed a positive sign. This implies that a real estate's value increases as the distance to downtown Gothenburg is increasing. We observed that Stensjön and Öjersjö have the highest intercept, besides Hjuvik, although they have the lowest average $TP_{j,t}$ (*cf.* Appendix A). Finally, the adjusted R^2 value for the reference model was 0.8384, which was higher than the MBM's of 0.8129.

[8] All coefficients presented in Table 4.3 are expressed in 2008 prices. HOX^{β_j} adjusts for time

[9] Similar estimates were found in all submarkets.

[10] In general, a terrace house should not be more expensive than a detached dwelling, (*cf.* Appendix A for variable motivations).

[11] For example, lot values in Långedrag were more expensive than in Öjersjö on average.

TABLE 4.3: The MBM Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr > I t I
γ	0.013045	0.00124	10.52	<.0001
$\beta_{Langedrag}$	0.780037	0.0682	11.44	<.0001
$\lambda_{1,Langedrag}$	33475.72	2088.2	16.03	<.0001
$\lambda_{2,Langedrag}$	22258.53	1657.9	13.43	<.0001
$\lambda_{3,Langedrag}$	46215.76	3950.8	11.7	<.0001
$\theta_{1,Langedrag}$	1313.604	72.1074	18.22	<.0001
$\theta_{2,Langedrag}$	330.219	49.351	6.69	<.0001
β_{Hjuvik}	0.598836	0.0795	7.53	<.0001
$\lambda_{1,Hjuvik}$	22747.07	2137.7	10.64	<.0001
$\lambda_{2,Hjuvik}$	21192.64	1666.9	12.71	<.0001
$\lambda_{3,Hjuvik}$	26831.21	3895	6.89	<.0001
$\theta_{1,Hjuvik}$	852.6362	79.953	10.66	<.0001
$\theta_{2,Hjuvik}$	240.3608	56.4521	4.26	<.0001
$\beta_{Stensjon}$	0.76258	0.1188	6.42	<.0001
$\lambda_{1,Stensjon}$	38496.1	3325.5	11.58	<.0001
$\lambda_{2,Stensjon}$	9323.466	2791	3.34	0.0009
$\lambda_{3,Stensjon}$	16228.64	9073	1.79	0.0739
$\theta_{1,Stensjon}$	797.8536	358.7	2.22	0.0263
$\theta_{2,Stensjon}$	207.1607	76.5229	2.71	0.0069
$\beta_{Ojersjo}$	0.832065	0.1177	7.07	<.0001
$\lambda_{1,Ojersjo}$	26005.82	2382.1	10.92	<.0001
$\lambda_{2,Ojersjo}$	13547.88	2438.3	5.56	<.0001
$\theta_{1,Ojersjo}$	1394.94	432.5	3.23	0.0013
$\theta_{2,Ojersjo}$	172.001	48.4861	3.55	0.0004
$\beta_{Orgryte}$	1.115428	0.0691	16.13	<.0001
$\lambda_{1,Orgryte}$	37497.63	1880.4	19.94	<.0001
$\lambda_{2,Orgryte}$	10579.5	1982.3	5.34	<.0001
$\theta_{1,Orgryte}$	1984.287	261.8	7.58	<.0001
$\theta_{2,Orgryte}$	7.630001	63.5493	0.12	0.9044
$\beta_{Krokslatt}$	0.980765	0.1175	8.34	<.0001
$\lambda_{1,Krokslatt}$	35290.18	2962	11.91	<.0001
$\lambda_{1,Krokslatt}$	15346	2266	6.77	<.0001
$\theta_{1,Krokslatt}$	1084.336	545.8	1.99	0.0471
$\theta_{2,Krokslatt}$	202.839	102.9	1.97	0.0489
β_R	-0.04661	0.0225	-2.07	0.0382
β_K	-0.00569	0.0243	-0.23	0.8151

The estimated parameters used to estimate marginal prices, average prices, and to forecast.

TABLE 4.4: The Reference Model Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr > | t |
$Intercept_{Langedrag}$	11.3606	0.4851	23.42	<.0001
$Intercept_{Hjuvik}$	11.77144	0.44	26.75	<.0001
$Intercept_{Stensjon}$	12.20988	0.4611	26.48	<.0001
$Intercept_{Ojersjo}$	11.53101	0.5016	22.99	<.0001
$Intercept_{Orgryte}$	11.05363	0.9945	11.11	<.0001
$Intercept_{Krokslatt}$	11.51868	0.8388	13.73	<.0001
$\beta_{TTC,Langedrag}$	0.284087	0.115	2.47	0.0136
$\beta_{TTC,Hjuvik}$	0.03547	0.0631	0.56	0.5742
$\beta_{TTC,Stensjon}$	0.017338	0.0837	0.21	0.8359
$\beta_{TTC,Ojersjo}$	-0.10005	0.1079	-0.93	0.354
$\beta_{TTC,Orgryte}$	-0.20454	0.0813	-2.52	0.012
$\beta_{TTC,Krokslatt}$	-0.29527	0.0894	-3.3	0.001
$\beta_{AGE,Langedrag}$	-0.03808	0.016	-2.38	0.0176
$\beta_{AGE,Hjuvik}$	-0.08757	0.0139	-6.31	<.0001
$\beta_{AGE,Stensjon}$	-0.09919	0.0137	-7.22	<.0001
$\beta_{AGE,Ojersjo}$	-0.01264	0.0133	-0.95	0.3417
$\beta_{AGE,Orgryte}$	0.001688	0.0354	0.05	0.962
$\beta_{AGE,Krokslatt}$	-0.03318	0.0245	-1.36	0.1756
$\beta_{K,Langedrag}$	-0.05579	0.0277	-2.02	0.0439
$\beta_{K,Hjuvik}$	-0.06025	0.0448	-1.34	0.1789
$\beta_{K,Stensjon}$	-0.04627	0.0317	-1.46	0.1443
$\beta_{K,Ojersjo}$	-0.02363	0.0374	-0.63	0.5274
$\beta_{K,Orgryte}$	0.004167	0.0293	0.14	0.887
$\beta_{K,Krokslatt}$	0.154576	0.0713	2.17	0.0302
$\beta_{R,Langedrag}$	-0.01581	0.0509	-0.31	0.756
$\beta_{R,Hjuvik}$	0.094345	0.0389	2.42	0.0155
$\beta_{R,Stensjon}$	-0.01665	0.0348	-0.48	0.6323
$\beta_{R,Ojersjo}$	-0.05944	0.0445	-1.33	0.1821
$\beta_{R,Orgryte}$	-0.10959	0.0285	-3.84	0.0001
$\beta_{R,Krokslatt}$	0.054413	0.0551	0.99	0.3232

The estimated parameters used to forecast.

4.5 Key Value Drivers

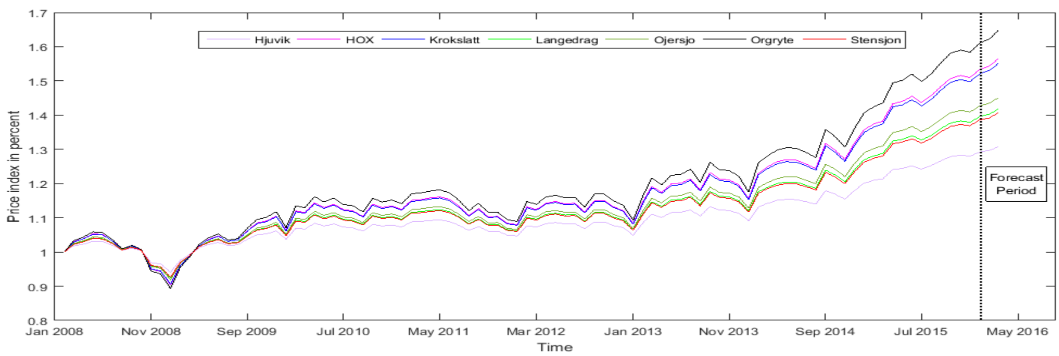
Marginal prices were estimated through the partial derivative of the $TP_{j,t}$ with respect to the continuous characteristics (*cf.* Equation 4.10). For the binary variables, $K_{j,t}$ and $R_{j,t}$, the marginal price was not estimated. This was because they were calculated as the difference in $TP_{j,t}$ when the structure was a detached dwelling and when it was a town ($K_{j,t}$) or row house ($R_{j,t}$) (Coulson, 2008b). Their marginal price therefore represents the interaction effect of their parameter estimate (β_K and β_R). Lastly, the submarkets' sensitivity to the aggregated price movement (HOX^{β_j}) demonstrates their sensitivity to real changes in the economy, *i.e.* macro events such as interest rate changes (*cf.* Figure 4.2). All marginal prices were converted into December 2015 year's value to simplify the comparison between submarkets.

$$Marginal\ Price_{i,j,t} = \frac{\partial TP_{j,t}}{\partial X_{i,j,t}} \tag{4.10}$$

4.5.1 Aggregated Price Movement

Figure 4.2 displays the submarkets' sensitivity to the aggregated price movement (HOX^{β_j}) in Gothenburg between 2008 and 2015. Örgryte and Krokslätt exhibited the highest price increase by approximately 60 and 50 percent, respectively. Since Örgryte had a price movement greater than the aggregated, $\beta_{Orgryte}$ was greater than one (*cf.* Table 4.3). This implies that Örgryte was the most sensitive submarket to real changes in the economy. For example, the sharp decline in domestic interest rates in late 2014 caused a price jump, where Örgryte's $TP_{j,t}$ increased the most. In contrast, Hjuvik was the least exposed submarket to macro economic changes and experienced a price increase that was approximately 30 percent less than Örgryte.

FIGURE 4.2: Aggregated Price Movement



The submarkets' sensitivity to the price movement in Gothenburg.

4.5.2 Marginal Prices

TABLE 4.5: Marginal Prices by the Ocean

Marginal Prices		Långedrag			Hjuvik			Improvement
Type	Characteristic	80^{th} Percentile	20^{th} Percentile	Average	80^{th} percentile	20^{th} Percentile	Average	Units
LOT	LS	3,499	926	2,005	1,253	628	1,160	1 m^2
	DTW	7,867	794	6,203	3,990	98	2,724	1 Meter
	H	72,795	12,714	41,618	68,790	2,722	30,632	1 Meter
	TTC	217	12	151	6	0	5	1 Minute
House	BA	25,331	22,601	25,239	23,604	19,992	21,771	1 m^2
	Q	41,465	1,560	34,929	21,619	9,493	24,443	1 Taxation Point

Marginal Prices (SEK) in Långedrag and Hjuvik.

The marginal price of characteristics were higher in Långedrag than in Hjuvik (*cf.* Table 4.5). This confirmed the economic significance of the MBM since the average $TP_{j,t}$ was higher in Långedrag. The economic significance was also validated by the same value drivers in both submarkets. Henceforth, the marginal prices that affects the $TP_{j,t}$ the most were the same in both submarkets, *i.e.* reflecting their similar nature. For example, in both submarkets $DTW_{j,t}$ had a higher marginal price than an additional m^2 of lot. The same were seen for $Q_{j,t}$ and $BA_{j,t}$. We also noted that $H_{j,t}$ had a larger impact than $DTW_{j,t}$, but the observed range for $H_{j,t}$ was lower than $DTW_{j,t}$ (*cf.* Appendix A). Henceforth, $DTW_{j,t}$ drives the value of the lot more albeit its lower marginal price. Furthermore, Hjuvik generally exhibit lower spread between the average and the percentiles compared to Långedrag. The large discrepancies between the percentiles in Långedrag were justified by the descriptive statistics. Henceforth, several real estates were sold above 20 MSEK in Långedrag, whereas the most expensive house in Hjuvik was sold for 15 MSEK (*cf.* Appendix A).

TABLE 4.6: Marginal Prices by the Lake

Marginal Prices		Stensjön			Öjersjö			Improvement
Type	Characteristic	80^{th} Percentile	20^{th} Percentile	Average	80^{th} percentile	20^{th} Percentile	Average	Units
LOT	LS	1,243	479	827	757	425	725	1 m^2
	DTW	1,348	2	732	1,785	645	1,145	1 Meter
	H	20,581	5,215	12,739	38,500	3,400	20,053	1 Meter
	TTC	1,260	92	924	131	14	84	1 Minute
House	BA	36,904	9,364	16,028	28,932	14,883	22,130	1 m^2
	Q	24,588	5,302	16,858	23,709	4,692	16,123	1 Point

Marginal Prices (SEK) in Stensjön and Öjersjö.

For submarkets located by the lake, Stensjön and Öjersjö, less discrepancies in marginal prices were observed (*cf.* Table 4.6). Continuing, $Q_{j,t}$, $BA_{j,t}$ and $H_{j,t}$ have higher marginal prices than the remaining characteristics in both submarkets. The marginal price of $H_{j,t}$ was higher in Öjersjö. According to the descriptive statistics (*cf.* Appendix A), both the range and maximum values of $H_{j,t}$ were greater in Stensjön. The higher marginal price in Öjersjö was thus motivated, because the probability of a lot located higher than their neighbors increases more with a unit improvement than it does in Stensjön. Furthermore, $TTC_{j,t}$ was higher in Stensjön, which was motivated by the set distribution of the characteristic, *i.e.* the impact on price should be close to zero in Öjersjö as no real estates where located within 30 minutes from Kungsportsavenyn. Lastly, since the characteristics driving both submarkets were the same, with small differences in magnitudes, economic significance of the MBM was validated. Henceforth, the marginal prices reflected their similar nature and average $TP_{j,t}$.

TABLE 4.7: Marginal Prices by the City

Marginal Prices		Örgryte			Krokslätt			Improvement
Type	Characteristic	80^{th} Percentile	20^{th} Percentile	Average	80^{th} percentile	20^{th} Percentile	Average	Units
LOT	LS	5,758	33	2,187	2,140	452	839	1 m^2
	DTW	0	0	0	0	0	0	1 Meter
	H	66,930	11,031	36,192	16,952	4,137	9,921	1 Meter
	TTC	221,860	96,780	160,047	68,553	3,915	41,046	1 Minute
House	BA	43,795	12,412	29,395	38,843	16,996	23,288	1 m^2
	Q	47,406	9,493	45,706	30,919	6,745	25,089	1 Point

Marginal Prices (SEK) in Örgryte and Krokslätt.

All marginal prices except $DTW_{j,t}$ were higher in Örgryte than in Krokslätt (*cf.* Table 4.7). From Table 4.7 we observed that both submarkets were driven by the same characteristics. For example, the marginal price was equal to zero for $DTW_{j,t}$, which confirms the authors' prior belief since the observations were located outside 1000 meters from the closest shoreline. Furthermore, $TTC_{j,t}$ was valued relatively high in both submarkets, especially when observing the upper percentiles. This motivates the characteristic as a key value driver in both submarkets. Since $TTC_{j,t}$ was expressed in minutes, a greater distance is expressed compared to the characteristics measured in meters, *e.g.* $H_{j,t}$. Furthermore, the average marginal price for $Q_{Orgryte,t}$ was roughly 82 percent higher than $Q_{Krokslatt,t}$. Since both submarkets' average $Q_{j,t}$ were equal to 30 taxation points (*cf.* Appendix A), consumers in Örgryte were willing to pay more for the characteristic, *ceteris paribus*.

By observing Table 4.5 - 4.7, the economic significance of the MBM was confirmed. Henceforth, the MBM captured the key value drivers, set by our distributions, and the marginal prices between submarkets reflect their descriptive statistics. The former was for example observed in Örgryte and Krokslätt where $TTC_{j,t}$ was observed as one of their key value drivers. In Långedrag and Hjuvik $DTW_{j,t}$ and $H_{j,t}$ were noted to be key value drivers, whereas Stensjön and Öjersjö had $H_{j,t}$ as one of their key value drivers. Henceforth, in accordance with our prior beliefs we observed that consumers value the submarket's specific characteristics when purchasing a real estate. Lastly, we noted that the marginal prices reflected the $TP_{j,t}$, *e.g.* Långedrag was more expensive on average than Hjuvik.

4.6 Average Prices

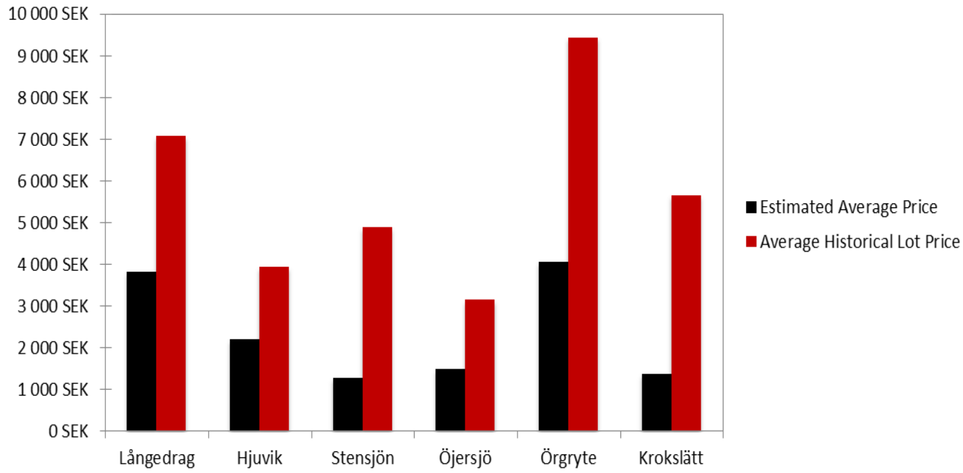
To investigate if the real estate market was overvalued we analyzed average prices through Tobin's Q. The average price was calculated by dividing the estimated price for the lot and house with their corresponding house ($BA_{j,t}$) and lot size ($LS_{j,t}$).[12] As a result, the units were expressed in SEK/m^2. By adding the estimates of SEK/m^2 for the house and lot, we were able to compare what consumers have paid for an existing real estate to what it cost to build a new one, *i.e.* buy a new m^2 of lot and build a m^2 of house. Furthermore, to estimate the cost of a new m^2 house we gathered production cost data (SEK/m^2) from Statistics Sweden between 2008 and 2014 in the Gothenburg area. By estimating the cumulative average growth rate (CAGR) between these years, we calibrated an estimation of production cost in 2015. Thereafter, we used the average price of previously sold lots to estimate the SEK/m^2 of a lot.[13] After adding the SEK/m^2 production cost to the estimated SEK/m^2 of lot, we had the numerator in Tobin's Q (*cf.* Equation 4.11). To compare "apples to apples" we converted all the estimated SEK/m^2 of house and lot into 2015's values, by using the HOX index. All prices in Figure 4.3 - 4.5 are therefore presented in December 2015 years value.

$$Tobin's\,Q = \frac{Market value\,of\,installed\,cap.}{Replacement\,cost\,of\,cap.} = \frac{Estimated\,avg.\,real\,estate\,price}{Total\,avg.\,production\,cost} \tag{4.11}$$

The estimated lot values were undervalued on average in all submarkets when compared to the average new SEK/m^2 (*cf.* Figure 4.3). For example, in Krokslätt the average lot price was estimated to approximately 1370 SEK/m^2, which was lower than the average market value of roughly 5659 SEK/m^2. [14]

[12] When estimating the parameters, we received an estimated lot and house value.
[13] The removed lots from our data set.
[14] One can easily be conceived by Figure 4.3 that a lot was more expensive in Örgryte than in Långedrag. However, taking into account that the average lot was roughly 339 m^2 smaller in Örgryte, the total lot value still favors Långedrag.

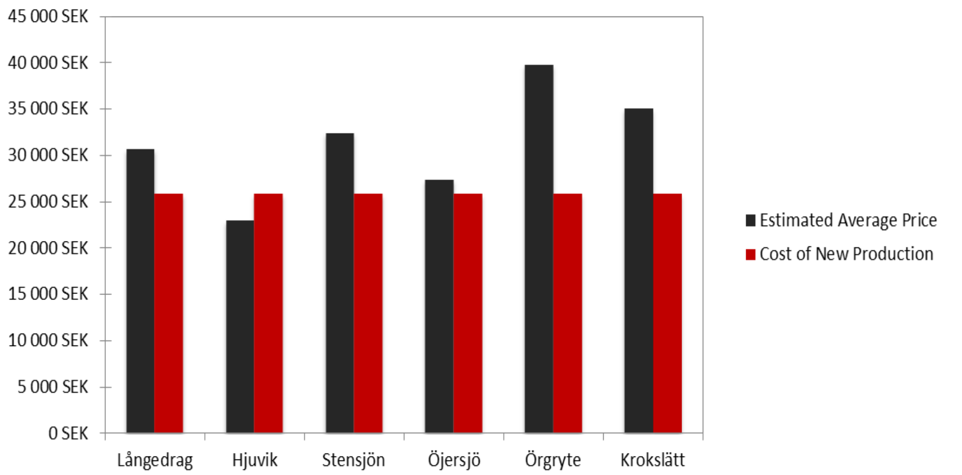
FIGURE 4.3: The Price of a m^2 - Lot



The estimated lot (SEK/m^2) in each submarket

In Figure 4.4 we noted that all submarkets except Hjuvik have a higher average SEK/m^2 than the production cost of building a new m^2. This implies that the replacement cost of a m^2 was lower than what consumers paid for it on average, *i.e.* a Tobin's Q above one for the house. As a parallel to corporate finance, the ratio states that a firm is trading at higher levels than their asset values, *i.e.* market value is higher than book value. Furthermore, this implies it was economically rational for a construction firm to build houses, since consumers paid more for the bricks and wood than it costs the firm to install them. This applies to all submarkets except in Hjuvik, where a construction firm would have made a loss.

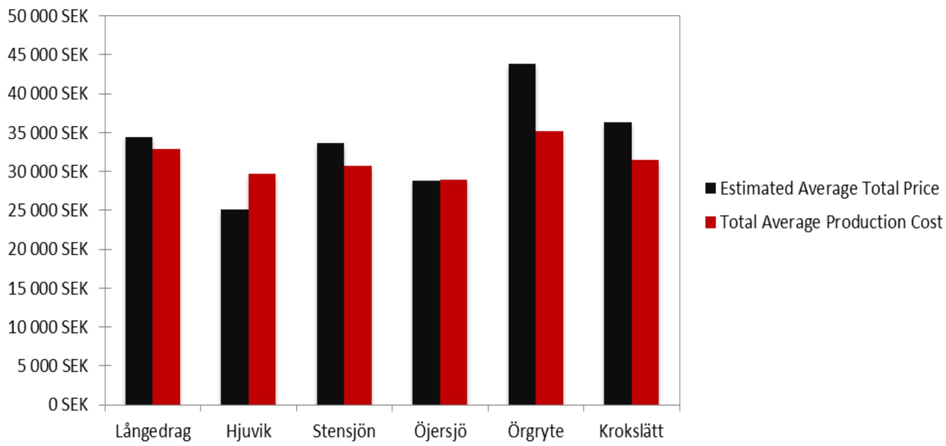
FIGURE 4.4: The Price of a m^2 - House



The estimated house (SEK/m^2) in each submarket

In order to build a new house, one must first acquire a lot where the structure can be established. We therefore looked at the sum of purchasing a m^2 of house and lot. From Figure 4.5, we observed that all submarkets except Hjuvik and Öjersjö, experienced a Tobin's Q above one for the real estate.[15] The result implies that it was economically rational for construction firms to purchase new lots and build houses in all submarket besides Hjuvik and Öjersjö. Henceforth, these two submarkets were trading at a discount.

FIGURE 4.5: The Price of a m^2 - Real Estate



The sum of the estimates of house and lot (SEK/m^2).

5 The Forecast

We calibrated the forecast by using the estimated coefficients and applying them to the out-of-sample data, including an updated HOX index to account for any aggregated price movements.

5.1 Peer Evaluation

The MBM's forecast accuracy was higher on an aggregated level, which was seen when analyzing the residuals and volatility (*cf.* Table 5.1). The average absolute residual was 23,874 SEK, *i.e.* 0.43 percent lower than the reference model on average. Since the volatility was lower, 48,656 SEK or 1.02 percent, the confidence level for the margin of error was narrowed. Furthermore, both models display a high goodness of fit in Theil's coefficient. The

[15] A real estate SEK/m^2 is equal to the sum of a house and lot SEK/m^2.

higher Theil's coefficient of MBM resulted in a higher rank than the reference model. However, the MGN test displayed that such statement cannot be made with statistical assurance. Henceforth, although the descriptive statistics favored the MBM on an aggregated level, the difference between the two models was not statistically significant.

TABLE 5.1: Forecast Comparison

		MBM	Reference Model	Difference
Test	Morgan-Granger-Newbold	-1.4087	-1.4087	-
	Theil's Coefficient	0.9979	0.9976	0.0003
Residuals	Absolute Average	634,162 SEK	658,036 SEK	23,874 SEK
	Volatility	451,803 SEK	500,459 SEK	48,656 SEK
Proportional Error	Absolute Average	9.19%	9.62%	0.43%
	Volatility	6.28%	7.30%	1.02%

Comparing the forecast accuracy of the two models.

5.2 Individual Real Estate Appraisals

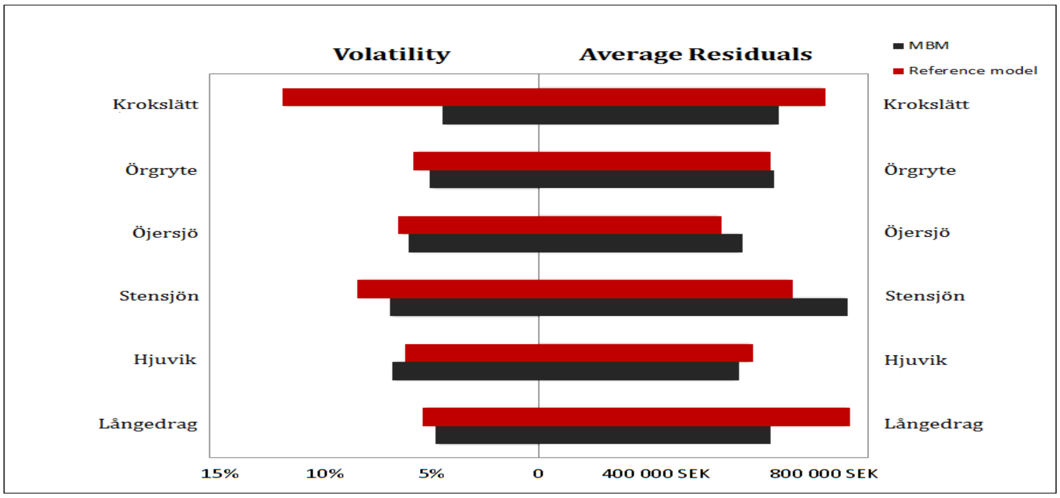
When analyzing the forecast in each submarket, the MBM had lower volatility in all submarkets besides Hjuvik (*cf.* Table 5.1).[1] The reference model had lower absolute residuals in Stensjön and Öjersjö, but higher in all other submarkets. The pair-wise submarkets were therefore reflected in the models' prediction accuracy, *i.e.* the MBM had lower absolute average residuals in Långedrag and Hjuvik (the submarkets by the ocean), Örgryte and Krokslätt (the submarkets by the city), whereas the reference model was better when a lake and forest were present (Stensjön and Öjersjö). Figure 5.1 also indicates that the two best forecasted submarkets were Långedrag and Örgryte.

Through analyzing specific transactions, explanations for model discrepancies and the existence of residuals were provided. For example, in Långedrag one transaction that was located 51 meters from the shoreline, was wrongly estimated by approximately one percent in the MBM. Whereas the reference model was off by approximately nine percent (*cf.* Appendix D). Henceforth, the MBM's factor for $DTW_{j,t}$ (*cf.* Figure 4.1) increased the lot value, which improved the forecast accuracy. Another example was found in Örgryte, where the $TTC_{j,t}$ was regarded as one of the key value drivers. The MBM had an absolute average proportional error of 5.23 percent for all houses within 15 minutes from the city. The absolute average proportional errors for the reference model on the same properties was equal to

[1] *cf.* Appendix D for complete descriptive statistics of the forecast in each submarket.

9.36 percent. The MBM model therefore seizures more of the effect from the key value drivers. However, transactions with the "average characteristics" demonstrates a tendency of favoring the reference model. For example, the reference model outperforms the MBM in Öjersjö and Stensjön. In these sub-markets we observed lower deviations in the out-of-sample data compared to the in-sample data (*cf.* Appendix A & B).

FIGURE 5.1: Forecast Statistics



Describing the forecast accuracy of the two models

Furthermore, several reasons of residuals were observed when analyzing the some of the MBM's worst appraisals. In Örgryte, the real estate with the largest absolute proportional error was underestimated by 13.9 percent.[2] When observing real estates sold on the same street in the in-sample data set, we noted that they were sold for more than the average price per m^2. Moreover, a house in Stensjön that was included in the forecast was also found in the in-sample data. The $TP_{j,t}$ increased by 198 percent without any sign of variations in characteristics besides time. In Långedrag, a transaction far away from the shoreline was underestimated by approximately 12 percent. The estimated lot value was economically rational compared to recently sold lots in close proximity. A closer look at the house reveals two unquantifiable characteristics on an aggregated level: no insight from people passing by on the street and an old classical structure of the villa. Finally, a transaction in Krokslätt demonstrates the outcome of an auction where the real estate rose 45.8 percent above the asking price.

[2] *cf.* Figure 4.3, 4.4, 4.5, Tables in Appendix B, and Tables in Appendix D for the motivations of all the examples and reasons mentioned in the paragraph.

6 Conclusion

6.1 Highlights

The purpose of this study was to develop a starting point of a refined model that predicts individual real estate prices as the sum of the house and lot value. A real estate's large portion of the household's balance sheet reveals the importance of the subject as both financial institutions and government agencies depend on a proper asset pricing model. Continuing, the MBM had an adjusted R^2 of 0.8129, which was within the upper range compared to previous studies that used other models. The MBM demonstrates economic significance when analyzing marginal prices and observing the relationship in key value drivers in-between submarkets as they reflected the descriptive statistic in $TP_{j,t}$. We therefore believe the undervaluation of lots and the overvaluation in houses, can be explained by misleading input data and missing characteristics. An example of misleading input data was $Q_{j,t}$ because it might not reflect the true quality of the house since the address declaration is both optional and subjective (*cf.* Appendix A). An example of missing characteristics were the "mingle factors" that we observed, *e.g.* Örgryte, but excluded due to time constraints [1].

From the pooled regression we observed that statistical gains were made when minimizing the market size, *i.e. "smaller is better"*. Although the submarkets are similar in nature and share the same key value drivers, they should be estimated individually. This follows our prior beliefs as the average $TP_{j,t}$ in each submarket was different, *i.e.* the magnitude of the coefficients should reflect the submarket's $TP_{j,t}$. Conclusively, estimating submarkets individually increased the adjusted R^2 of the reference model, which imposed a tougher benchmark for the MBM to outperform.

When identifying the submarkets' key value drivers, we observed the economic significance of our model. We noted that the $TP_{j,t}$ was reflected in the marginal price and that the submarket's specific characteristics were key value drivers, *e.g.* $TTC_{j,t}$ in Örgryte and Krokslätt. This implies that households were looking for specific characteristics when deciding to purchase a real estate. Then the choice of submarket will reflect their willingness to

[1] We define "mingle factors" as exclusive characteristics only found at a specific street or in a small group of houses in close proximity *e.g. beau monde*.

pay for these characteristics. For example, if a family is moving to Långedrag, living close to the ocean is deemed to be most valuable whereas the second "best" would be to have an ocean view. The same conclusion was derived from Hjuvik's marginal prices, but at a lower price. From Figure 4.2, we noted that the submarkets in close proximity to Gothenburg have experienced the highest price increase. We believe this reflects a change in preferences, where consumers were more interested in living in urban submarkets than on the country side.

From Tobin's Q we noted that it is economically rational for construction firms to build houses in all submarkets except Hjuvik. However, to build a new house, one must own a lot. For that reason and because our disproportional distribution of $\lambda_{j,k}$ and $\theta_{j,k}$, the results in Figure 4.5 become the most interesting when analyzing Tobin's Q. We observed that consumers have paid more money for an old m^2 real estate (house and lot) than it would have cost them to purchase a new m^2 in four out of six submarkets. Consumers living in the submarkets trading at a premium should sell their real estate and move to their pair-wise submarket that was trading at a discount, *i.e.* Hjuvik or Öjersjö. Since both Örgryte and Krokslätt were trading at a premium, the substitution method does not apply to these consumers.

Trivially, construction firms should focus their attention on the submarkets with a Tobin's Q above one. However, this might not be as easy as suggested due to scarcity of land. For example, we only observed a few sold lots in Örgryte. Thus, construction firms will not be able to take advantage of the opportunity. The opposite analogy was observed in Hjuvik and Öjersjö. Most transactions of real estates were recorded in Hjuvik, whereas Öjersjö had most number of sold lots out of all submarkets. Based on these findings, we believe that construction firms seem to be well aware of the market conditions and exploit them in the submarkets where it is feasible. Our conclusion was strengthened when gathering data for 2016 in Stensjön where we noted multiple lots were sold during the spring. Lastly, from Q Theory we know that Tobin's Q increases with decreasing interest rates. Since interest rates have decreased during our sample period, we would expect a high Tobin's Q. Adding the scarcity of land and the interest rate findings together, we feel that the high Tobin's Q in four submarkets was not a sign of overvaluation, but simply the result of fundamental economic factors. Therefore, we think it is a bold statement to say that the housing market is overvalued, at least when analyzing our six submarkets.

On an aggregated level the MBM outperformed the reference model based on their descriptive statistics. When analyzing each submarket, the MBM was better in four out of six submarkets in terms of residuals, and in five out of six submarkets in terms of volatility. The reference model outperformed the MBM in two submarkets, Öjersjö and Stensjön. Potential explanations were found in the high intercept of the reference model and the "average" characteristics. The former suggests that some factors have been left out in the MBM, which the reference model's intercept was able to capture. The later explanation suggests that the MBM captures more variation around the edges but at the expense of the average. This was not a surprise, considering a "linear" model (OLS) is by definition best on average. Furthermore, Långedrag and Örgryte were the best appraised submarkets. Considering we have focused most of our attention on lot analysis in Långedrag, due to time constraints, this was not a surprise. In Örgryte, we have great *ex ante* knowledge as one of the authors have been raised in the submarket. This was out of great importance considering only a few lots have been sold in Örgryte during the past eight years, *i.e.* very few observations were available to conduct a lot analysis. Therefore, the characteristics consumers value in Örgryte were partly based on "inside information". Henceforth, the common denominator for these two submarkets was the identification of key value drivers, since they served as the foundation of appropriately set distributions. Conclusively, one with enough time on their hands will be able to analyze any submarket of interest and identify its key value drivers.

The residual analysis revealed several reasons to why the MBM occasionally experienced high residuals. Firstly, in Örgryte an example of "mingle factors" were evident, which means the submarket should have been divided up even further. This can be solved by the implementation of dummy variables for locality. However, due to time constraints we could not implement it into MBM. Secondly, in Långedrag an example of existing unquantifiable characteristics was seen. The quality of house nor type of house takes into account that an old classical structure with high quality were typically sold at a higher price than a newer house with the same quality in our submarkets, *ceteris paribus*. Thirdly, in Stensjön we saw a large increase in the $TP_{j,t}$ of one particular house sold twice, where our data did not show any differences. We believe this was due to both misleading input data on the house and the psychological state of the buyer. Henceforth, the first time the house was sold was in the aftermath of the financial crisis, whereas we currently have all-time low domestic interest rates and upcoming amortization laws. Lastly, we observed the outcome of an auction in Krokslätt, where the $TP_{j,t}$ was 45 percent higher than the asking price. Together, all these reasons recognize the complexity of creating an asset pricing model. However, the first step towards perfection is always problem identification.

6.2 Further Research

The MBM provides a way of making the traditional Builder's Model applicable to appraise individual real estates. A good example of success was found in Långedrag where the authors put most of their attention when analyzing lots before assigning distributions. More detailed analysis of all submarkets should therefore result in a better appraisal. Also, a more flexible model that optimize the curvature of consumers' preferences is to prefer, *i.e.* without the touch of a human. Consumer's preferences could then be updated recursively and constantly analyzed. An example would be to use Rosen's (1974) two-step approach and estimate the demand curve. Thereafter one can implement the "mingle factors" by the usage of a dummy variable for locality. Furthermore, instead of using an index to account for variations in the macro environment, stylized financial statements of the household is preferred. However, one must gather individual data and solve the problem of endogeneity, reverse causality, that arises from cost of debt and equity. Thereafter, income and costs such as electricity, water, taxation, insurance could be added to the financial statements. Lastly, we tried to implement other characteristics such as the number of bidders, number of houses for sale, and consumer's future views, but they were excluded due to the lack of individuality [2].

[2] By consumer's future view, we mean that if one does not think they will be unemployed tomorrow, their willingness to spend might not be affected by their current employment situation.

Bibliography

Alem, Yonas (2014). *Graduate Econometrics Lecture 4: Heteroskedasticity*.

Booli Search Technologies AB (2016). *Booli Slutpriser*. URL: www.booli.se.

Chin, Tung-Leong and K.W Chau (2002). "A Critical Review of Literature on the Hedonic Price Model". In: *The International Journal for Housing Science and Its Applications*. URL: http://goo.gl/7DceeK.

Clapp, John M. and Carmelo Giaccotto (2000). "EVALUATING HOUSE PRICE FORECASTS". In: *Working Paper*. URL: http://pages.jh.edu/jrer/papers/pdf/past/vol24n01/01.1_26.pdf.

Colson, N. Edward and Jeffrey Zabel (2012). "What Can We Learn From Hedonic Models Where Markets Are Dominated By Foreclosures?" In: *Pennsylvania State University*. URL: http://goo.gl/qCp7fS.

Coulson, Edward (2008a). "Hedonic Methods and Housing Markets - Chapter 1: An Introduction and Discussion of Origins". In: *Department of Economics, Penn State University*. URL: http://goo.gl/4oiN1G.

— (2008b). "Hedonic Methods and Housing Markets - Chapter 2: A Brief Survey and Interpretation of Hedonic Parameters". In: *Department of Economics, Penn State University*. URL: http://goo.gl/WtYwuV.

Court, A.T. (1939). *Hedonic Price Indexes With Automotive Examples*. URL: http://goo.gl/E3FFlL.

Diewert, Erwin (2003). "Hedonic Regressions. A Consumer Theory Approach". In: *The National Bureau of Economic Research*. URL: http://goo.gl/i5EMKN.

Diewert, W. Erwin and Chihiro Shimizu (2015). "Hedonic Regression Models for Tokyo Condominium Sales". In: *Discussion Paper 15-07*.

Diewert *et al.*, W. Erwin (2011). "Hedonic Regressions and the Decomposition of a House Price Index into Land and Structure Components". In: *Working Paper*.

Dinbyggare (2016). *Stambyte – fräscha stammar viktiga för husets hållbarhet*. URL: `http://goo.gl/7Wv7kn`.

Englund, Peter (2011). "Svenska huspriser i ett internationellt perspektiv". In: *Sveriges riksbank, Riksbankens utredning om risker på den svenska bostadsmarknaden*. URL: `http://goo.gl/385ozb`.

European Central Bank (2016). *Sectoral Breakdown of MFI loans and deposits vis-a-vis other domestic residents*. URL: `http://sdw.ecb.europa.eu/reports.do?node=1000003262`.

Eurostat (2013). *Handbook on Residential Property Prices Indices (RPPIs)*. Luxembourg: Publications Office of the European Union, 2013. Chap. Decomposing an RPPI into Land and Structures Components, pp. 81–101.

Eurostat (2013). *Handbook on Residential Property Prices Indices (RPPIs)*. Luxembourg: Publications Office of the European Union, 2013. Chap. Hedonic Regression Methods, pp. 49–65.

Foote, Chris (2010). "Chapter 18: Investment and the Housing Market". In: *Harvard University, Department of Economics*. URL: `http://goo.gl/cB1AQB`.

Freeman, Albert M. (1993). "The measurement of environmental and resource values : theory and methods". In:

Goodman, Allen C. and Thomas G. Thibodeau (2003). "Housing market segmentation and hedonic prediction accuracy". In: *Journal of Housing Economics* 12.3, pp. 181–201. URL: `http://www.sciencedirect.com/science/article/pii/S1051137703000317`.

Goodman, Allen C. and Thomas G. Thibodeau (2006). "The Spatial Proximity of Metropolitan Area Housing Submarkets". In: *Forthcoming in Real Estate Economics*. URL: `http://goo.gl/VYjbWY`.

Google Maps (2016). *The Google Maps Elevation API*. URL: `https://goo.gl/T7lGch`.

Haan, Jan de and Yunlong Gong (2015). *Accounting for Spatial Variation of Land Prices in Hedonic Imputation House Price Indexes: A Semi-Parametric Approach*.

Haas, G. C. (1922). "Sales prices as a basis for farm land appraisal". In: *Technical Bulletin* 9.

Hittapunktse AB (2016). *Tomter*. URL: `www.hitta.se`.

Illustrerad vetenskap (2008). *Varför vill vi bo vid vattnet?* URL: `http://illvet.se/manniskan/psykologi/varfor-vill-vi-bo-vid-vattnet`.

Industrial Systems Research (2013). *Housebuilding and the New Homes Market: a Survey*. Industrial Systems Research. Chap. Size of Firms, Economies of Scales Diversification, pp. 51–52.

Li, Mingche M. and James H. Brown (1980). "Micro-Neighborhood Externalities and Hedonic Housing Prices". In: *Land Economics* 56, pp. 125–141. URL: `http://econpapers.repec.org/article/uwplandec/v_3a56_3ay_3a1980_3ai_3a2_3ap_3a125-141.htm`.

Macpherson *et al.* (2005). "The composition of hedonic pricing models". In: *Journal of Real Estate Literature* 13, pp. 1–44.

Malpezzi, Stephen, Larry Ozanne, and Thomas Thibodeau (1980). "CHARACTERISTIC PRICES OF HOUSING IN FIFTY-NINE METROPOLITAN AREAS". In: *A report on research supported by Department of Housing and Urban Development Office of Policy Development and Research*, pp. 24–26. URL: `http://www.huduser.gov/portal/Publications/pdf/HUD-50814.pdf`.

Orrberg (2015). *ALLT FLER DRÖMMER OM ATT BO VID VATTNET*. URL: `https://www.jm.se/bostader/inspiration/artiklar/2015/oktober/bo-vid-vatten/`.

Palmquist, Raymond B. (1984). "Estimating the Demand for the Characteristics of Housing". In: *The Review of Econimcs and Statistics* 66.3, pp. 394–404. URL: http://web.mit.edu/14.33/www/attach/palmquist.pdf.

Riksbanken (2016). *Sök räntor & valutakurser*. URL: http://goo.gl/1U273U.

Rosen, Sherwin (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition". In: *Journal of Political Economy* 82, pp. 34–55.

SCB (2015). *Antal och andel personer och hushåll efter boendeform den 31 december 2014*. URL: http://goo.gl/YhN3MT.

— (2016). *Kostnader per kvm för nybyggda ordinära gruppbyggda småhus*. URL: http://goo.gl/nN3LyX.

Service HNS AB Hemnet (2016). *Antalet villor till salu*. URL: www.hemnet.se.

Sims, Eric (2011). "Graduate Macro Theory II: Notes on Investment". In: *University of Notre Dame*. URL: http://goo.gl/2ZeovC.

Skatteverket (2014). "Fastighetstaxering Småhus, 2015 - 2017". In: *SKV389 utgåva 9*.

Skatteverket (2016). *Fastighetstaxering*. URL: http://goo.gl/4d6hE7.

Skatteverket (2016). "Typregister Småhus". In: *Oracle Fusion Middleware Forms Services*.

Skolverket (2016). *Slutbetyg årskurs 9*. URL: http://goo.gl/oh72v0.

Sköld, Lovisa (2014). "Villa vanligaste boendeformen". In: *Statistiska Central Byrån* 2014:4.

South Florida, University of (2016). *Collinearity*. URL: http://goo.gl/QNpdI1.

Stevens, Jim (2005). "The Q Theory of Investment". In: *The American University of Bulgaria*. URL: http://goo.gl/cB1AQB.

Valueguard AB (2016). *Nasdaq OMX Valueguard-KTH Housing Index (HOX) Göteborg villa*. URL: www.valueguard.se/goteborgvilla.

Villaägarna (2016). *Del 11: Dags att byta VVS-rör och elldedningar?* URL: http://goo.gl/UKy0CE.

Västtrafik (2016). *Reseplaneraren*. URL: www.vasttrafik.se.

Wabe, J. Stuart (1971). "A Study of House Prices as a means of Establishing the Value of Journey Time, the Rate of Time Preference and the Valuation of some Aspects of Environment in the London Metropolitan Region". In: *Applied Economics* 3:4, pp. 247–255. DOI: http://dx.doi.org/10.1080/00036847100000012.

Wilhelmsson, Mats (2000). "The Impact of Traffic Noise on the Values of Single-family Houses". In: *Journal of Environmental Planning and Management* 43.6, pp. 799–815.

A In-sample Data

Transaction Price

All closing prices for villas, terrace houses, and town houses originate from Booli's search engine (Booli Search Technologies AB, 2016). All apartments were excluded due to the owner structure of the building.

Descriptive Statistics - Transaction Price

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	6.92	2.89	3.2	25	MSEK
Hjuvik	366	4.88	1.77	1.8	15	MSEK
Stensjön	237	4.12	1.02	1.58	8.2	MSEK
Öjersjö	219	4.04	1.28	1.32	7.8	MSEK
Örgryte	300	6.87	1.87	3	12.7	MSEK
Krokslätt	224	4.79	1.11	2.30	9.0	MSEK

Describing the transaction details in each submarket.

Building area

From Booli's website, the size of each house was retrieved. The area reported was the so-called living area in which gross floor area was excluded. For more recently sold houses, gross floor area data was listed on the website (Booli Search Technologies AB, 2016). However, this characteristic was omitted in order to have a large enough sample size to conduct our study. Lastly, in accordance with previous studies, the authors believe that building area will have a positive impact on the transaction price (Macpherson *et al.*, 2005).

Descriptive statistics - Building area

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	172	58	52	510	Meters
Hjuvik	366	144	53	30	411	Meters
Stensjön	237	127	38	42	280	Meters
Öjersjö	219	133	44	25	310	Meters
Örgryte	300	182	56	88	365	Meters
Krokslätt	224	145	40	56	360	Meters

Quantitatively describing the building area in each submarket.

Quality of House

The quality of house was estimated using the reported taxation points found at Skatteverket's public database. Every third year an owner of a house gets to fill out the form that determines the total taxation value. However, the declaration is optional and one might choose to stay with Skatteverket's estimated quality. Furthermore, one of the factors determining the total real estate taxation value is the quality of house, taxation points, which is estimated irrespectively of the real estate's locality.[1] These taxation points reflects several characteristics of the house, such as condition of the kitchen which is divided into three categories: simple-, normal- and high quality standard. Other examples it also takes into account are the house facing (brick, wood or concrete) and if there is an existing garage/-carport or not (Skatteverket, 2014). Moreover, the quality of house is subjectively determined by the house owner as he or she fills out the declaration by themselves. Furthermore, since a house owner only gets to update the information in the address declaration in-between every third year when they feel the reported taxation points are invalid. Therefore, a gap between the reported years exist in which real estate transactions were made. For approximately twenty percent of all observations, a change in taxation points between two declaration periods was observed in conjunction with a recorded transaction of the real estate. To account for such discrepancies, we calibrated a linear interpolation on a monthly basis.

Descriptive statistics - Quality of House

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	32	5	19	50	Taxation Points
Hjuvik	366	30	5	13	47	Taxation Points
Stensjön	237	30	4	19	44	Taxation Points
Öjersjö	219	31	5	13	54	Taxation Points
Örgryte	300	30	5	19	49	Taxation Points
Krokslätt	224	30	4	19	42	Taxation Points

Describing the quality of the houses in each submarket.

Previous research have included the quality of house in their models, but with different proxies. Li and Brown (1980) used age and Diewert *et al.* (2011) utilized the number of rooms. Thereby using the Swedish tax agency standardized scoring system seemed reasonable in comparison to

[1] Ten points for a house located in the south of Sweden is equal to ten points for a house located elsewhere.

their methods. Conclusively, we believe that a higher quality will yield a higher transaction price.

Age of Building

The age of a building was another characteritic obtained from the Swedish tax agency that influence the taxation value of a house. (Skatteverket, 2016). It is measured as the building year of the structure. A house with 100 m^2 of building area that had its first 50 m^2 built ten years ago and the other 50 m^2 built five years ago will have a building age of seven and a half years. (Skatteverket, 2014).[2]

Due to lack of historical data, the tax agency had limited the building year to 1929 *i.e.* houses built before 1929 were assigned the same age (Skatteverket, 2014). Previous studies have also constrained this variable, *i.e.* if the building year in the sample was older than year 1900, it was equal to 1900 (Wabe, 1971). Continuing, the age of the building was estimated by subtracting the reported building year from the sales date. This ensures that the information known to the buyer was present in the model. One might also assume that the age is correlated with quality due to depreciation. However, the correlation between the two variables was estimated to -0.0350. Finally, age has been used in several previous hedonic studies in which its significance and negative impact have been demonstrated (Malpezzi, Ozanne, and Thibodeau, 1980; Goodman and Thibodeau, 2003). The same result is therefore expected in this study.

Descriptive statistics – Age of Structure

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	53	23	2	86	Years
Hjuvik	366	28	17	1	86	Years
Stensjön	237	46	19	1	85	Years
Öjersjö	219	22	12	1	74	Years
Örgryte	300	70	15	16	86	Years
Krokslätt	224	52	23	2	86	Years

Describing the age of the structure in each submarket.

[2] 7.5 Years= 0.5*10 years+0.5*5 years

Type of Building

The last structure characteristic included in the estimation of taxation value is the type of building. The three types of structures was therefore gathered from Skatteverket's database: Villa, Terrace house (K), and Town house (R). Malpezzi *et al.* (1980) have used building type in their study, *e.g.* single-family attached or detached. However, they argued that the sign of the coefficients indicate that they capture locational effects rather than structural. Diewert *et al.* (2011) used a similar method but also included the type of construction *i.e.* brick, wood or concrete *etcetera*. Such attributes was not included in this study due to lack of data. Finally, since submarkets such as Örgryte have a wide-range of different house types, the inclusion was motivated. All else equal, the authors believe that a villa should be worth more than its adversaries, considering it is a detached property.

Descriptive statistics - Terrace house

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	0.08	0.27	0	1	Binary
Hjuvik	366	0.04	0.20	0	1	Binary
Stensjön	237	0.12	0.33	0	1	Binary
Öjersjö	219	0.25	0.43	0	1	Binary
Örgryte	300	0.12	0.33	0	1	Binary
Krokslätt	224	0.04	0.20	0	1	Binary

Describing terrace houses with regards to villas for each submarket.

Descriptive statistics - Town house

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	0.03	0.18	0	1	Binary
Hjuvik	366	0.002	0.05	0	1	Binary
Stensjön	237	0.10	0.30	0	1	Binary
Öjersjö	219	0.06	0.24	0	1	Binary
Örgryte	300	0.32	0.47	0	1	Binary
Krokslätt	224	0.13	0.34	0	1	Binary

Describing town houses with regards to villas for each submarket

Lot Size

After collecting the addresses from Booli, the lot size data was gathered through Hitta.se's real estate search application. As previous research suggest, a greater lot size is expected to result in a higher transaction price (Macpherson *et al.*, 2005).

Descriptive statistics - Lot size

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	859	366	199	3519	m^2
Hjuvik	366	1097	341	219	2352	m^2
Stensjön	237	722	445	145	4488	m^2
Öjersjö	219	926	555	152	4076	m^2
Örgryte	300	520	324	60	1865	m^2
Krokslätt	224	761	375	132	2242	m^2

Quantitatively describing the lot sizes for each submarket.

Height

If a house is not in close reach of water nor the edge of the shoreline can be seen, the third best alternative is a great view (Diewert and Shimizu, 2015; Skatteverket, 2014).[3] Using Google Map's API, the height above the sea was estimated for specific coordinates. The coordinates were estimated from the middle of the house and was retrieved from Hitta.se. Thereafter by subtracting the lowest estimation in each neighborhood, a low point equal to zero for each submarket was calculated. The height was thereby in relative terms to its submarket and should reflect the view of the real estate. Therefore, the higher altitude of the real estate, the higher transaction price is to be expected (Diewert and Shimizu, 2015).

[3] An if statement was calibrated to eliminate the effect of height if the property is located within 50 meters from the shoreline.

Descriptive statistics - Height

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	16	8	0.2	33	Meters
Hjuvik	366	15	11	0	50	Meters
Stensjön	237	29	10	0	59	Meters
Öjersjö	219	14	9	0	39	Meters
Örgryte	300	15	11	0	55	Meters
Krokslätt	224	44	19	2	87	Meters

Quantitatively describing the submarkets' height above reference point.

Travel Time to City

As of 2014, 42.9 percent of the Swedish household and approximately 60 percent of all Swedish teenagers live in an owned villa (SCB, 2015; Sköld, 2014). Consequently, the travel time to a workplace is important. Due to lack of individual data, travel time to city will serve as a proxy and has been used in previous research (Wabe, 1971; Diewert and Shimizu, 2015). Data has been gathered using Västtrafik's online travel time application. Henceforth, by inserting each address and choosing the alternative that reaches Kungsportsplatsen the fastest on a weekday between 7 and 8 am, including both walking and public transportation.

Descriptive statistics - Travel Time to City

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	39	4	32	50	Minutes
Hjuvik	366	51	8	39	79	Minutes
Stensjön	237	33	3	25	46	Minutes
Öjersjö	219	38	4	33	50	Minutes
Örgryte	300	16	4	10	33	Minutes
Krokslätt	224	24	4	15	35	Minutes

Quantitatively describing each house's travel time to Kungsportsplatsen. It includes both the time to walk and to use the public transportation system.

Distance To Water

Using the coordinates from Hitta.se, Google Maps' distance tool was utilized to estimate the proximity to the shoreline. As previously stated, the intuition behind the variable was that close proximity to the shoreline was the best alternative, whereas the second best was to see the edge. The magazine Illustrerad Vetenskap, argues that consumers' willingness to pay for the attribute is derived from a psychological perspective. For example, a view of the far horizon, could express a sign of freedom, whereas water and its sound provides a calming effect on the overall psych (Illustrerad vetenskap, 2008). Thus, people are willing to pay more to leave their stressful environment at work (Orrberg, 2015).

Water Specifics

Submarket	Water name (1)	Water name (2)	Type of water
Långedrag	Kattegatt	N/A	Ocean
Hjuvik	Kattegatt	N/A	Ocean
Stensjön	Rådasjön	Norra Långevattnet	Lake
Öjersjö	Stora Kåsjön	Stora Hålasjön	Lake
Örgryte	Delsjön	N/A	Lake
Krokslätt	Delsjön	Norra Långevattnet	Lake

The different lakes and ocean used when estimating the distance to water.

Descriptive statistics - Distance to Water

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	457	271	14	1011	Meters
Hjuvik	366	534	367	10	1700	Meters
Stensjön	237	405	262	20	1040	Meters
Öjersjö	219	439	256	13	1120	Meters
Örgryte	300	1863	303	1160	2550	Meters
Krokslätt	224	2394	473	1560	3520	Meters

Quantitatively describing the distance to closest shoreline.

School Quality

Descriptive statistics on all ninth graders in Gothenburg back to 2007 and each submarkets' school grades were retrieved from the Swedish National School Agency for Education (Skolverket, 2016). The data included in the regression for school quality was the percentile of the average school grade for each specific submarket in relation to Gothenburg. If a family was to move to Gothenburg, their comparison was on relative terms. Consequently, since a higher education typically results in a higher salary, on average, the higher percentile was expected to yield a higher transaction price. Lastly, only what was known to the market was included in the regression. This implies that the percentiles will change each July as the grades for the school year ending in June will be publicly available.

Schools

Submarket	School name
Långedrag	IES Gothenburg
Långedrag	Montessoriskolan Skäret
Långedrag	Nya Påvelundsskolan 1
Långedrag	Nya Påvelundsskolan 2
Hjuvik	Nordlyckeskolan 7-9
Hjuvik	Torslandaskolan 6-9
Stensjön	Kvarnbyskolan
Öjersjö	Öjersjö Brunn Skola
Öjersjö	Furulunds skola 6-9
Örgryte	Böskolan
Örgryte	Montessoriskolan Casa
Krokslätt	Kunskapsskolan Krokslätt
Krokslätt	Internationella Engelska Skolan
Krokslätt	Montessoriskolan Kvarnhjulet

The schools used when estimating the submarkets' school quality.

Descriptive statistics - School Quality

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Långedrag	301	85	2	82	91	Percentile
Hjuvik	366	72	6	54	83	Percentile
Stensjön	237	76	3	64	79	Percentile
Öjersjö	219	65	6	54	75	Percentile
Örgryte	300	85	6	73	94	Percentile
Krokslätt	224	73	8	57	82	Percentile

Quantitatively describing the submarkets' school quality in relation to Gothenburg.

B Out-of-sample Data

Descriptive Statistics Långedrag

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	8	8.41	1.93	5.3	12.5	MSEK
Building area	8	166	45	80	218	m^2
Quality of House	8	34	5	27	43	Taxation Points
Age of structure	8	66	25	11	87	Years
Terrace house	8	0	0	0	0	$\int_0^1$
Town house	8	0	0	0	1	$\int_0^1$
Lot size	8	918	260	456	1429	m^2
Height	8	19	10	3	33	Meters
Travel time	8	37	2	33	39	Minutes
Distance to water	8	585	295	51	965	Meters

Representing the data used to forecast 2016 in Långedrag.

Descriptive Statistics Hjuvik

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	18	6.90	1.67	3.53	10.05	MSEK
Building area	18	165	41	97	245	m^2
Quality of House	18	30	3	27	36	Taxation Points
Age of structure	18	24	14	4	59	Years
Terrace house	18	0.06	0.23	0	1	$\int_0^1$
Town house	18	0	0	0	0	$\int_0^1$
Lot size	18	1057	300	344	1729	m^2
Height	18	11	7	2	26	Meters
Travel time	18	20	6	41	62	Minutes
Distance to water	18	410	385	41	1440	Meters

Representing the data used to forecast 2016 in Hjuvik.

Descriptive Statistics Stensjön

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	9	5.98	1.47	4.16	8.35	MSEK
Building area	9	134	24	105	173	m^2
Quality of House	9	32	4	28	39	Taxation Points
Age of structure	9	43	19	11	67	Years
Terrace house	9	0.11	0.31	0	1	$\int_0^1$
Town house	9	0	0	0	0	$\int_0^1$
Lot size	9	772	326	239	1189	m^2
Height	9	29	7	20	41	Meters
Travel time	9	31	4	27	41	Minutes
Distance to water	9	308	212	62	640	Meters

Representing the data used to forecast 2016 in Stensjön.

Descriptive Statistics Öjersjö

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	6	4.99	1.15	3.58	6.6	MSEK
Building area	6	123	31	90	176	m^2
Quality of House	6	29	2	25	31	Taxation Points
Age of structure	6	18	8	2	24	Years
Terrace house	6	0.33	0.47	0	1	$\int_0^1$
Town house	6	0.17	0.37	0	1	$\int_0^1$
Lot size	6	654	451	163	1535	m^2
Height	6	11	8	2	21	Meters
Travel time	6	33	2	29	37	Minutes
Distance to water	6	495	195	158	697	Meters

Representing the data used to forecast 2016 in Öjersjö.

Descriptive Statistics Örgryte

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	11	9.31	2.39	6.6	13.2	MSEK
Building area	11	156	51	94	236	m^2
Quality of House	11	30	5	26	45	Taxation Points
Age of structure	11	76	6	70	87	Years
Terrace house	11	0.46	0.50	0	1	$\int_0^1$
Town house	11	0.27	0.45	0	1	$\int_0^1$
Lot size	11	461	189	182	774	m^2
Height	11	19	12	3	35	Meters
Travel time	11	17	3	13	23	Minutes
Distance to water	11	1929	315	1240	2300	Meters

Representing the data used to forecast 2016 in Örgryte.

Descriptive Statistics Krokslätt

Variable	No. of Obs.	Mean	Std. Dev.	Min.	Max.	Units
Transaction price	6	6.02	0.97	4.81	7.85	MSEK
Building area	6	152	70	81	295	m^2
Quality of House	6	31	4	23	36	Taxation Points
Age of structure	6	60	16	29	77	Years
Terrace house	6	0.17	0.37	0	1	$\int_0^1$
Town house	6	0	0	0	0	$\int_0^1$
Lot size	6	697	214	461	1006	m^2
Height	6	32	20	2	69	Meters
Travel time	6	21	2	18	24	Minutes
Distance to water	6	2057	435	1550	2910	Meters

Representing the data used to forecast 2016 in Krokslätt.

C Calibration

Linear Splines

Splines - (g and f-values)

Type	Building Area			Lot size		Units
Spline	1	2	3	1	2	
Långedrag	BA$\leq$ 100	100<BA$\leq$ 275	BA>275	LS$\leq$ 600	LS>600	m^2
Hjuvik	BA$\leq$ 100	100<BA$\leq$ 275	BA>275	LS$\leq$ 800	LS>800	m^2
Stensjön	BA$\leq$ 100	100<BA$\leq$ 200	BA>200	LS$\leq$ 400	LS>400	m^2
Öjersjö	BA$\leq$ 125	BA >125		LS$\leq$ 225	LS>225	m^2
Örgryte	BA$\leq$ 180	BA >180		LS$\leq$ 225	LS>225	m^2
Krokslätt	BA$\leq$ 120	BA >120		LS$\leq$ 300	LS>300	m^2

The splines used in MBM.

Assign Distributions

Factor Calibration

To capture the nonlinear relationship between the characteristics (both house and lot) and the transaction price, a logistic function was used to generate a specific *factor* in each submarket. Thus, the factors were set to reflect consumers' willingness to pay for a characteristic such as distance to water, height above reference point, taxation point *etcetera*. Thereafter, these factors were multiplied with the splines for the house and lot, respectively.

None of the parameters included in the logistic equation (*cf. Equation 4.6*) were estimated in the regression, but were inserted as fixed values. Furthermore, $c_{i,j}$ reflects the bearing capacity of the function, *i.e.* the maximal value for the factor. Whereas, $b_{i,j}$ is the steepness of the curve and $a_{i,j}$ is used to determine the midpoint of the S-curve[1]. The same methodology was applied for all characteristics. Henceforth, when the first order derivative is positive, the b_i, j parameter shifts to a negative sign, *i.e.* $e^{-b_{H,j} * X_{H,j}}$. Lastly, to solve the logistic functions we used the Excel add-in *Problem Solver*. After setting targets for each $X_{i,j}$, we minimized the sum of squared errors between our target and the generated output by changing the parameters

[1] The midpoint of the logistic curve is equal to $\frac{log(a_{i,j})}{b_{i,j}}$

of the logistic function. As a result, we had optimized the parameter estimates of $a_{i,j}$, $b_{i,j}$ and $c_{i,j}$ to fit the consumer preferences we observed when analyzing lots.

Distributional Motivations

Height

The assigned factor to the variable differs among submarkets, due to the rationale of the lot analysis and prior knowledge of the submarkets. For example, in Örgryte height will add more value to the lot compared to the other submarkets. This was motivated by the the observed transaction prices in relation to their altitude as well as local knowledge. Thus, the carrying capacity (max value of the factor) of the logistic function was set to be equal to 200 percent whereas it was equal to 100 percent in the remaining submarkets. Also, regardless of locality, the point of inflection of the logistic s-curve was approximately equal to the average height in each submarket. Meaning that the lot price increase by half the max factor in comparison to the sunk lot, *ceteris paribus*.

Travel Time to City

The travel time to city factor was calibrated the same for all submarkets. It was decided that if travel time exceeded 30 minutes to Kungsportsavenyn, the factor should not add any additional value to the real estate. For example, in Långedrag and Hjuvik, we argue that the factor should not add any value to the lot considering there is a trade-off between living in these submarkets with close proximity to the water. Consequently, the factor has a greater impact on real estates located in the urban submarkets, Örgryte and Krokslätt. Therefore, the carrying capacity target was set to be equal to 300 percent, representing zero minutes. While the midpoint value of one and a half, reflects the effect of being 15 minutes away from downtown.

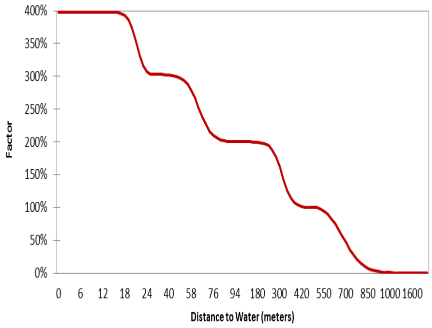
Quality of House

To calibrate the impact of quality, both apartments and villas where analyzed. Starting off by concluding that the average quality in all submarkets was equal approximately to 31, which was what a standard house has according to Skatteverket. Thus, a house in close range of 31 taxation points should therefore have little impact on the price of a m^2. Thereafter, when analyzing the relationship between quality and transaction price, no visible deviations was found for houses between 15-45 taxation points. However, outside of this spectra, larger deviations were observed. Thereby given cause for the factor of 100 and negative 100 percent factor. Hence, as quality approaches 0, the negative 100 percent factor, indicate that the m^2 essentially becomes worthless. Indicating one was paying for the land, *i.e.* a new house must be built.

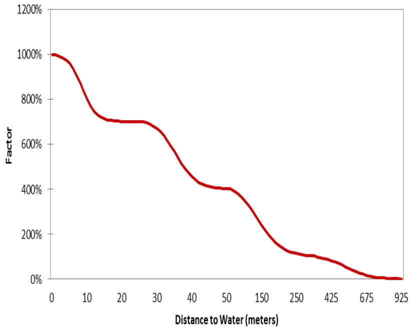
To confirm these findings, an analysis of apartments on sale at Hemnet's website was conducted. Apartments have higher liquidity in which indicates observations with equal locality effects was observed. Furthermore, through ocular econometrics we found that a non-standard home was either more or less expensive depending on the apartments' current condition. The same conclusion was also confirmed when talking to family and friends about their willingness to pay for higher quality. In which therefore support the previous findings.

Calibrated Distribuitions

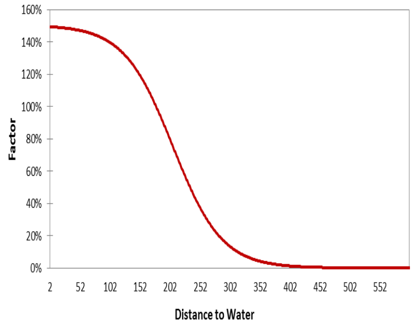
(A) Hjuvik



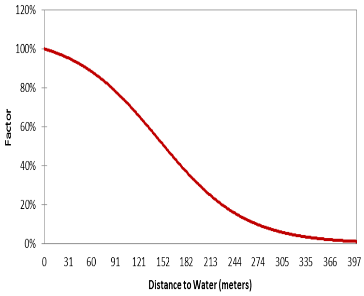
(B) Långedrag



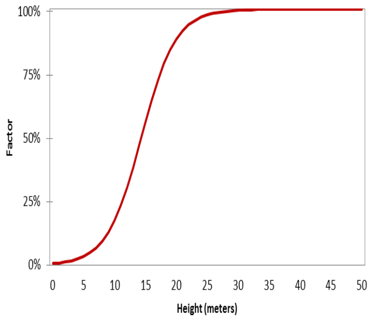
(C) Öjersjö



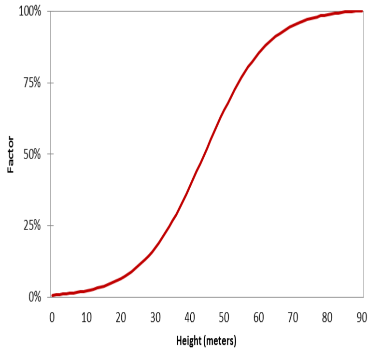
(D) Stensjön, Örgryte , Krokslätt



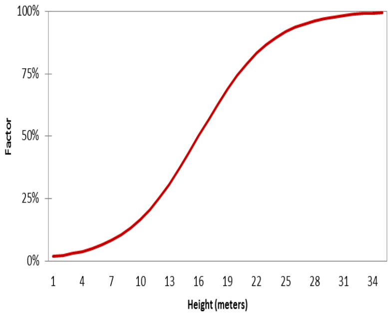
(E) Hjuvik



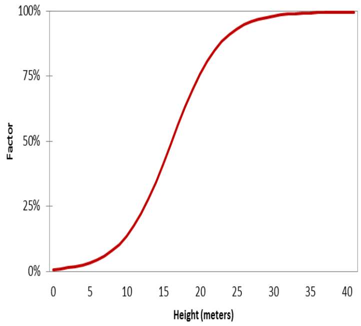
(F) Krokslätt



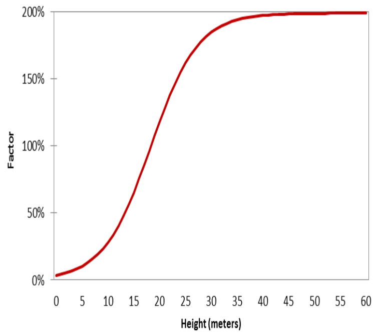
(G) Långedrag



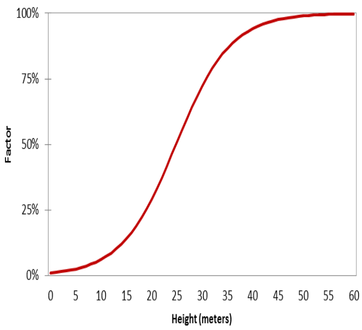
(H) Öjersjö



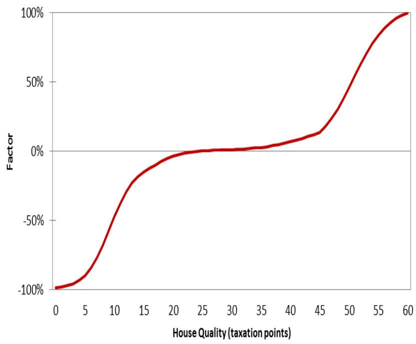
(I) Örgryte



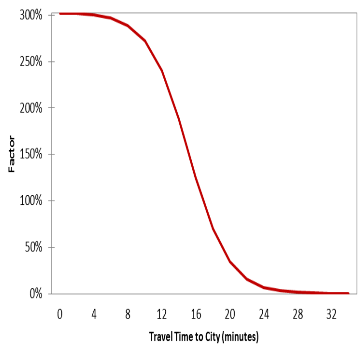
(J) Stensjön



(K) All



(L) All



Demonstrating the distributions graphically assigned to each variable for all submarkets

Reference Model Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
$Intercept_{Langedrag}$	11.3606	0.4851	23.42	<.0001
$Intercept_{Hjuvik}$	11.77144	0.44	26.75	<.0001
$Intercept_{Stensjon}$	12.20988	0.4611	26.48	<.0001
$Intercept_{Ojersjo}$	11.53101	0.5016	22.99	<.0001
$Intercept_{Orgryte}$	11.05363	0.9945	11.11	<.0001
$Intercept_{Krokslatt}$	11.51868	0.8388	13.73	<.0001
$\beta_{DTW,Langedrag}$	-0.20057	0.0168	-11.97	<.0001
$\beta_{DTW,Hjuvik}$	-0.16937	0.0141	-12	<.0001
$\beta_{DTW,Stensjon}$	-0.05181	0.0136	-3.8	0.0001
$\beta_{DTW,Ojersjo}$	-0.11397	0.0134	-8.51	<.0001
$\beta_{DTW,Orgryte}$	0.165783	0.0969	1.71	0.0871
$\beta_{DTW,Krokslatt}$	0.05494	0.0903	0.61	0.5429
$\beta_{BA,Langedrag}$	0.474014	0.0345	13.75	<.0001
$\beta_{BA,Hjuvik}$	0.450211	0.0367	12.26	<.0001
$\beta_{BA,Stensjon}$	0.30329	0.0373	8.13	<.0001
$\beta_{BA,Ojersjo}$	0.590198	0.0377	15.64	<.0001
$\beta_{BA,Orgryte}$	0.436663	0.0381	11.46	<.0001
$\beta_{BA,Krokslatt}$	0.380644	0.0477	7.98	<.0001
$\beta_{LS,Langedrag}$	0.132688	0.0314	4.23	<.0001
$\beta_{LS,Hjuvik}$	0.150336	0.0317	4.74	<.0001
$\beta_{LS,Stensjon}$	0.099235	0.0221	4.5	<.0001
$\beta_{LS,Ojersjo}$	0.073295	0.0266	2.75	0.006
$\beta_{LS,Orgryte}$	-0.00915	0.021	-0.44	0.6625
$\beta_{LS,Krokslatt}$	0.1071	0.0303	3.53	0.0004
$\beta_{H,Langedrag}$	0.006954	0.00185	3.75	0.0002
$\beta_{H,Hjuvik}$	0.002794	0.00102	2.73	0.0064
$\beta_{H,Stensjon}$	0.001706	0.000889	1.92	0.0551
$\beta_{H,Ojersjo}$	0.005843	0.00122	4.78	<.0001
$\beta_{H,Orgryte}$	0.006829	0.00103	6.66	<.0001
$\beta_{H,Krokslatt}$	0.001142	0.00124	0.92	0.357
$\beta_{TTC,Langedrag}$	0.284087	0.115	2.47	0.0136
$\beta_{TTC,Hjuvik}$	0.03547	0.0631	0.56	0.5742
$\beta_{TTC,Stensjon}$	0.017338	0.0837	0.21	0.8359
$\beta_{TTC,Ojersjo}$	-0.10005	0.1079	-0.93	0.354
$\beta_{TTC,Orgryte}$	-0.20454	0.0813	-2.52	0.012
$\beta_{TTC,Krokslatt}$	-0.29527	0.0894	-3.3	0.001
$\beta_{Q,Langedrag}$	0.098232	0.0618	1.59	0.1123
$\beta_{Q,Hjuvik}$	0.20412	0.0804	2.54	0.0113
$\beta_{Q,Stensjon}$	0.193001	0.0733	2.63	0.0085
$\beta_{Q,Ojersjo}$	0.156826	0.0801	1.96	0.0503
$\beta_{Q,Orgryte}$	0.157489	0.0617	2.55	0.0108
$\beta_{Q,Krokslatt}$	0.247883	0.0779	3.18	0.0015
$\beta_{AGE,Langedrag}$	-0.03808	0.016	-2.38	0.0176
$\beta_{AGE,Hjuvik}$	-0.08757	0.0139	-6.31	<.0001
$\beta_{AGE,Stensjon}$	-0.09919	0.0137	-7.22	<.0001
$\beta_{AGE,Ojersjo}$	-0.01264	0.0133	-0.95	0.3417
$\beta_{AGE,Orgryte}$	0.001688	0.0354	0.05	0.962
$\beta_{AGE,Krokslatt}$	-0.03318	0.0245	-1.36	0.1756
$\beta_{K,Langedrag}$	-0.05579	0.0277	-2.02	0.0439
$\beta_{K,Hjuvik}$	-0.06025	0.0448	-1.34	0.1789
$\beta_{K,Stensjon}$	-0.04627	0.0317	-1.46	0.1443
$\beta_{K,Ojersjo}$	-0.02363	0.0374	-0.63	0.5274
$\beta_{K,Orgryte}$	0.004167	0.0293	0.14	0.887
$\beta_{K,Krokslatt}$	0.154576	0.0713	2.17	0.0302
$\beta_{R,Langedrag}$	-0.01581	0.0509	-0.31	0.756
$\beta_{R,Hjuvik}$	0.094345	0.0389	2.42	0.0155
$\beta_{R,Stensjon}$	-0.01665	0.0348	-0.48	0.6323
$\beta_{R,Ojersjo}$	-0.05944	0.0445	-1.33	0.1821
$\beta_{R,Orgryte}$	-0.10959	0.0285	-3.84	0.0001
$\beta_{R,Krokslatt}$	0.054413	0.0551	0.99	0.3232
$\beta_{HOX,Langedrag}$	0.728136	0.0692	10.52	<.0001
$\beta_{HOX,Hjuvik}$	0.608894	0.0516	11.8	<.0001
$\beta_{HOX,Stensjon}$	0.687975	0.0534	12.88	<.0001
$\beta_{HOX,Ojersjo}$	0.637916	0.0547	11.67	<.0001
$\beta_{HOX,Orgryte}$	0.992394	0.0691	14.37	<.0001
$\beta_{HOX,Krokslatt}$	0.835029	0.0822	10.16	<.0001

The estimated coefficients used to forecast.

Econometric Validation

Heteroskedasticity

Heteroskedasticity Test - Reference Model

Dependent Variable	Test	Statistic	DF	Pr >ChiSq	Variables
ltp	White's Test	572.4	219	<0.0001	*Cross all variables*
ltp	Breusch-Pagan	77.02	5	<0.0001	*1, ldtw, lba, lls, lttc, lq, SQ, K, R, HOX*

Test statistics for reference model.

Multicollinearity

Multicollinearity Test Reference Model

Type	1	2	3	4	5	6	7	8	9	10	11	12
Variable	Intercept	log(BA)	log(LS)	H	log(TTC)	log(DTW)	log(Q)	SQ	log(AGE)	HOX	K	R
Variance Inflationary Factor	0	1.51	2.98	1.30	2.83	2.11	1.26	1.43	1.72	1.02	1.59	2.19

Test statistics for reference model.

Collinearity Test Statistics - MBM

.	Eigenvalue	Condition No.	γ	br	bk	θ_1^{lan}	θ_2^{lan}	λ_1^{lan}	λ_2^{lan}	λ_3^{lan}	β_{lan}	θ_1^{hju}	θ_2^{hju}	λ_1^{hju}	λ_2^{hju}	β_{hju}	θ_1^{ste}	θ_2^{ste}	λ_1^{ste}	λ_2^{ste}	λ_3^{ste}	β_{ste}	θ_1^{oje}	θ_2^{oje}	λ_1^{oje}	λ_2^{oje}	β_{oje}	λ_1^{org}	λ_2^{org}	θ_1^{org}	θ_2^{org}	β_{org}	θ_1^{kro}	θ_2^{kro}	λ_1^{kro}	λ_2^{kro}
			Proportion of Variation																																	
1	5.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	4.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	4.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	3.9	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	3.9	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	3.7	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	1.1	2.1	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	1.0	2.2	0.0	0.3	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	1.0	2.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0	1.0	2.3	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.6	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.6	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.5	3.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.5	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.7
5	0.5	3.3	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.1
6	0.4	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	0.4	3.6	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.4	3.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.3	0.0	0.1
9	0.3	3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.4	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0
0	0.3	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.2	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.3	4.4	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.3	0.5	0.0	0.0	0.0	0.0
2	0.2	4.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.2	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.1	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.2	4.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.2	5.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.1	0.0	0.4	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0
6	0.2	5.3	0.0	0.0	0.0	0.3	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	0.2	5.6	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.4	0.3	0.0	0.0	0.0	0.0
8	0.2	5.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.1	0.1
9	0.1	7.7	0.0	0.0	0.0	0.6	0.1	0.4	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0	0.1	9.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.1	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	10.3	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.7	0.1	0.0	0.0	0.0	0.0	0.0
2	0.0	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.1	0.8	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	12.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.2	0.7	0.0
4	0.0	13.4	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.2	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.1	0.0	0.1	0.0
5	0.0	15.9	0.7	0.0	0.0	0.0	0.0	0.3	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0

D The Forecast

Forecast Comparison

		Residual		Proportional Error	
		Absolute Average	*Volatility*	*Absolute Average*	*Volatility*
Långedrag	*MBM*	631,508 SEK	402,696 SEK	7.46 %	4.68 %
	Reference Model	847,806 SEK	466,130 SEK	9.99 %	5.26 %
Hjuvik	*MBM*	546,107 SEK	442,271 SEK	7.75%	6.68 %
	Reference Model	582,953 SEK	446,112 SEK	8.29 %	6.08 %
Stensjön	*MBM*	842,545 SEK	568,287 SEK	13.35 %	6.78 %
	Reference Model	692,062 SEK	610,389 SEK	11.34 %	8.28 %
Öjersjö	*MBM*	556,897 SEK	222,755 SEK	12.34 %	5.94 %
	Reference Model	498,960 SEK	265,062 SEK	11.19 %	6.39 %
Örgryte	*MBM*	641,503 SEK	560,878 SEK	6.89 %	4.96 %
	Reference Model	633,087 SEK	404,975 SEK	7.6 %	5.7 %
Krokslätt	*MBM*	653,094 SEK	281,841 SEK	10.74 %	4.39 %
	Reference Model	783,225 SEK	704,139 SEK	12.64 %	11.64 %

TABLE D.0: *Comparing the two forecast models in each submarket.*

Forecast Estimates

Submarket	Transaction Price	Reference Model	MBM			Units
	Actual Value	Total Estimated Value	Total Estimated Value	Estimated House Value	Estimated Lot Value	
Långedrag	5.30	5.92	5.34	2.82	2.52	MSEK
Långedrag	7.98	7.80	7.27	4.68	2.59	MSEK
Långedrag	12.50	13.63	12.78	6.30	6.48	MSEK
Långedrag	9.50	8.39	8.36	6.00	2.36	MSEK
Långedrag	8.65	9.97	9.58	6.13	3.45	MSEK
Långedrag	8.10	9.54	8.99	6.34	2.65	MSEK
Långedrag	8.15	7.30	7.18	5.32	1.86	MSEK
Långedrag	7.10	7.23	7.19	4.42	2.78	MSEK
Hjuvik	6.85	7.00	6.73	3.44	3.29	MSEK
Hjuvik	7.30	7.22	6.02	3.16	2.86	MSEK
Hjuvik	9.60	8.52	8.97	6.05	2.91	MSEK
Hjuvik	3.53	4.00	3.55	2.73	0.82	MSEK
Hjuvik	5.83	5.69	5.24	3.51	1.73	MSEK
Hjuvik	6.70	6.42	6.47	3.37	3.11	MSEK
Hjuvik	7.23	8.04	6.95	3.37	3.58	MSEK
Hjuvik	7.96	6.82	6.84	4.03	2.81	MSEK
Hjuvik	10.05	9.26	9.43	5.46	3.96	MSEK
Hjuvik	8.20	8.57	8.48	5.06	3.42	MSEK
Hjuvik	7.05	7.91	7.61	3.88	3.73	MSEK
Hjuvik	8.05	8.79	8.66	4.62	4.04	MSEK
Hjuvik	7.85	8.07	6.71	3.75	2.96	MSEK
Hjuvik	5.50	4.55	4.12	3.12	1.00	MSEK
Hjuvik	4.25	4.30	4.30	2.14	2.16	MSEK
Hjuvik	4.62	4.50	4.44	2.59	1.85	MSEK
Hjuvik	7.40	9.11	7.38	3.98	3.40	MSEK
Hjuvik	6.20	5.67	5.48	3.45	2.04	MSEK
Stensjön	6.30	6.06	5.60	4.59	1.01	MSEK
Stensjön	8.35	8.00	7.30	5.72	1.57	MSEK
Stensjön	8.07	5.93	5.95	4.60	1.36	MSEK
Stensjön	5.51	5.65	5.14	4.42	0.72	MSEK
Stensjön	4.63	5.59	5.39	4.16	1.23	MSEK
Stensjön	7.15	6.01	5.83	4.71	1.11	MSEK
Stensjön	4.70	5.21	5.27	4.62	0.65	MSEK
Stensjön	4.16	4.77	4.80	4.03	0.77	MSEK
Stensjön	4.98	5.11	4.93	4.10	0.83	MSEK
Öjersjö	5.00	4.43	4.27	3.27	1.00	MSEK
Öjersjö	6.60	6.37	6.36	4.87	1.49	MSEK
Öjersjö	4.88	3.93	4.02	3.34	0.68	MSEK
Öjersjö	3.65	4.16	4.15	2.62	1.53	MSEK
Öjersjö	3.58	2.99	2.88	2.48	0.40	MSEK
Öjersjö	6.21	6.07	6.53	5.09	1.44	MSEK
Örgryte	11.40	12.09	11.09	8.56	2.53	MSEK
Örgryte	8.32	9.68	9.35	6.96	2.39	MSEK
Örgryte	13.20	12.51	11.37	9.10	2.27	MSEK
Örgryte	12.55	12.21	12.43	9.93	2.51	MSEK
Örgryte	7.25	7.82	7.77	6.16	1.61	MSEK
Örgryte	7.00	8.05	6.71	4.63	2.09	MSEK
Örgryte	7.81	7.35	6.98	4.85	2.13	MSEK
Örgryte	7.03	8.27	7.05	4.62	2.43	MSEK
Örgryte	9.30	9.58	9.11	6.54	2.58	MSEK
Örgryte	12.00	12.08	10.91	8.44	2.47	MSEK
Örgryte	6.60	6.39	5.77	4.17	1.61	MSEK
Krokslätt	6.50	7.96	7.22	5.71	1.51	MSEK
Krokslätt	4.81	4.69	3.98	3.13	0.85	MSEK
Krokslätt	5.85	7.83	6.35	5.30	1.05	MSEK
Krokslätt	5.87	5.75	5.22	4.32	0.90	MSEK
Krokslätt	5.25	5.57	5.40	4.38	1.02	MSEK
Krokslätt	7.85	8.55	8.92	8.08	0.84	MSEK

The actual and estimated transaction prices for both models on every