Estimation of Hedonic Models Using a Multilevel Approach: An Application for the Swiss Rental Market

DRAGANA DJURDJEVIC[a], CHRISTINE EUGSTER[b] and RONNY HAASE[b]

JEL- Classification: C14, J16, R31
Keywords: hedonic model, rental market, multilevel modelling

1. Introduction

Hedonic regression models have traditionally been used to model housing prices. Typically, the price of the house is related to the implicit prices of the housing attributes. The hedonic model is often estimated by OLS, leading to regression coefficients which represent the implicit prices. Most hedonic models include characteristics of the location to measure the influence of neighbourhood on housing prices, thereby ignoring the inherent hierarchy in the housing decision. However, as pointed out by QUIGLEY (1985), a household begins its search process by choosing a town to live in, followed by the neighbourhood within the town and finally the house to inhabit within this neighbourhood and town. The hierarchical nature of the housing decision cannot be ignored and use of the traditional OLS estimation technique can lead to unreliable and biased estimates, especially where the latter is performed without consideration for the presence of segmented markets. With multilevel modelling, a hierarchical structure is recognised when it is present in the data. In addition, the parameters of the model are allowed to vary across space. Thus, both spatial correlation and heterogeneity are taken into account. More detailed information on multilevel modelling can be found in the books by GOLDSTEIN (1995) and BRYK and RAUDENBUSH (2002).

a Wuest & Partner, Rue de la Rôtisserie 1, CH-1204 Geneva, Switzerland; phone: ++41 (0) 22 319 30 00; fax: ++41 (0) 22 319 30 01; e-mail: djurdjevic@wuestundpartner.com; web page: www.wuestundpartner.com

b Wuest & Partner, Gotthardstrasse 6, CH-8002 Zurich, Switzerland; phone: ++41 (0) 44 289 90 00; fax: ++41 (0) 44 289 90 01; e-mail: eugster@wuestundpartner.com and haase@wuestundpartner.com; web page: www.wuestundpartner.com
Acknowledgement: We are grateful to the participants of the Swiss hedonic day, 27 June 2007, Geneva. We also thank Philippe Thalmann and Andrea Baranzini for their helpful comments. All remaining errors are ours.

Hierarchical data have been widely used in many applied studies in social, medical and biological sciences. A hierarchy consists of units grouped at different levels. For example, students (level 1 units) are clustered in classes (level 2 units). Individuals within the same group are more similar than individuals across other groups. Thus, groups tend to have differentiated behaviours and this differentiation implies that the group and its individuals both influence and are influenced by the group membership. In our education example, pupils from the same class tend to be more similar in their achievement than pupils chosen at random from the total population of all classes because they are taught together or they share the same environment. Ignoring this grouping may result in overlooking the importance of group effects in pupils' achievement.

In the relevant international literature, the hedonic theory is associated with Rosen's (1974) seminal paper. Rosen's approach has been applied to different fields, including housing (for instance, Bajari and Kahn, 2005), public and environmental economics, labour markets and to a lesser extent, marketing and industrial organisation (see for instance, Berry, Levinsohn and Pakes, 1995). Similarly, issues addressed through hedonic price modelling are very diverse and range from estimating the effect of airport noise or the value of urban air quality (see for instance, Graves, Murdoch, Thayer and Waldman, 1988) to building price indexes (for instance, Griliches, 1971; Hoesli, Favarger and Giacotto, 1997). The issue of market segmentation has also been addressed (see for instance, Bajic, 1985; Basu and Thibodeau, 1998).

The Swiss housing and rental markets are also segmented. In addition, the quality of data is very variable. As a consequence, most existing Swiss studies tend to focus on specific regions such as the Canton of Geneva (see for instance Bender, Gacem and Hoesli, 1994) or the Canton of Zurich (Salvi, Schellenbauer and Schmidt, 2004). In his study, Fahrländer (2006) develops a nationwide model for the private property markets using nonparametric techniques that allow for a more flexible specification for the transaction prices. These statistical adjustments are aimed at relaxing the traditional assumptions about the (log)-linearity between prices and variables describing size and age of the properties, thus allowing the modelling of prices for thinner markets. As regards the Swiss rental housing market, various studies have been published since the study by Thalmann (1987) on the rental market in the city of Lausanne. Rieder (2005) uses data from the Swiss Federal Statistical Office on the Swiss rental market from 2003 to explain rental differences at regional level. The study by Baranzini, Ramirez, Schaerer and Thalmann (2008) investigates the effect of environmental amenities on the rental markets of the Zurich and Geneva urban areas.

In this paper, we are concerned with a hedonic model that explains rents at nationwide level and allows for the geographical specifics of the Swiss market. Our focus is mainly a descriptive one: we establish the main features of a mixed effects model and stress the importance of the hierarchical nature of data in determining rents. This paper does not attempt to provide a complete model of rental prices. We therefore consider only a two-level model: the level 1 units are apartments which are clustered within municipalities that represent the level 2 units. The dwelling characteristics enter the fixed part of the model whereas the random part includes only one variable at municipality level. Of course, we can add other variables at the second level and thereby develop a more complete model.

The paper is organised as follows. Section 2 presents the data and the variables used for the empirical application and some descriptive statistics are reported. Section 3 describes the statistical model. The results are discussed in Section 4. Section 4 also provides a comparison between different models: the multilevel hedonic model with two random effects (for the intercept and the slope for dwelling area) and the hedonic model with cantonal dummies as well as cantonal interaction terms for dwelling area (which are estimated with OLS). The conclusions are presented in Section 5.

2. Data and Descriptive Statistics

2.1 Database

For the empirical analysis, we use data on rent rolls, which provide information on rental prices, as well as on apartment attributes relating to dwelling type, floor area, number of rooms and location of each apartment in terms of its postcode. The database originates from the Wüest & Partner evaluation of real estate properties owned by various institutional investors such as real estate and pension funds. Nearly 5 000 properties were evaluated, for example, in 2007. This represents a sample of about 20 000 apartments.[1] For this application, we want to describe the most up-to-date situation prevailing on the rental market for 2007. That is why the selected observations are limited to cases involving a new rental contract concluded since the first quarter of 2006. This results in a final sample

1 Observations with incomplete or unreliable information on key variables such as dwelling area had already been eliminated at an earlier stage.

of 11 913 apartments spread over 327 municipalities.[2] By way of comparison, the existing stock of rental apartments in Switzerland was 2 million in 2000 (Census, 2000). About 300 000 to 350 000 changes of tenancy are observed each year. Our sample thus represents about 3% of the new tenant market.

2.2 Representativity

We check the representativity of our data by comparing the distribution of some variables in our sample with that in the data from the Census 2000.[3] Table 1 reports the geographic differences between the two datasets.

Table 1: Geographical Representativity

Geographical region	Census (in %)	Sample (in %)
Zurich	22.0	34.6
Eastern	9.6	9.4
Central	8.8	8.1
North-western	15.5	22.9
Berne	13.0	6.2
South	7.2	1.6
Lake Geneva	15.6	12.2
Western	8.2	5.0
Total	100.0	100.0

Table 1 shows that the Zurich region is over-represented in our sample, while the regions of Berne and southern Switzerland are under-represented. In general, however, we can consider the data to be sufficiently representative of the Swiss rental market.

2 Initially, our sample contained 12 138 apartments. To allow specification of random effects at municipality level, only municipalities with at least 5 apartments were selected. This sample selection rule leads to a reduction of the sample size to 11 913 apartments. The analysis of the characteristics of the apartments between these two samples shows that there is no over-represented or under-represented category after the selection compared to the initial sample.

3 Actually, the sample should be compared with data on tenancy changes instead of data on the existing stock of apartments. Since no data on the new tenant market are available to us, we choose to compare our sample with data provided by the Census 2000.

Figure 1 shows that the distributions of living area between the two datasets are very similar. For example, the 95th percentile of dwelling area corresponds to 125 m^2 in our sample and 130 m^2 in the census data.

Figure 1: Comparison Between the Sample and the 2000 Population Census by Living Area

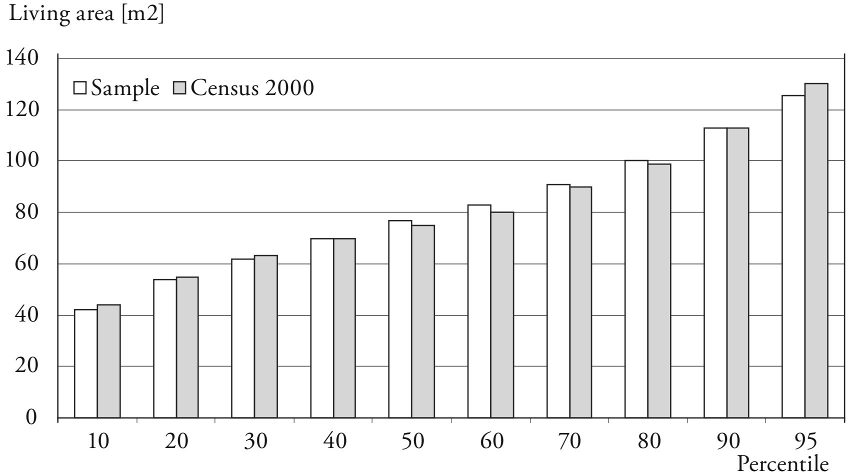
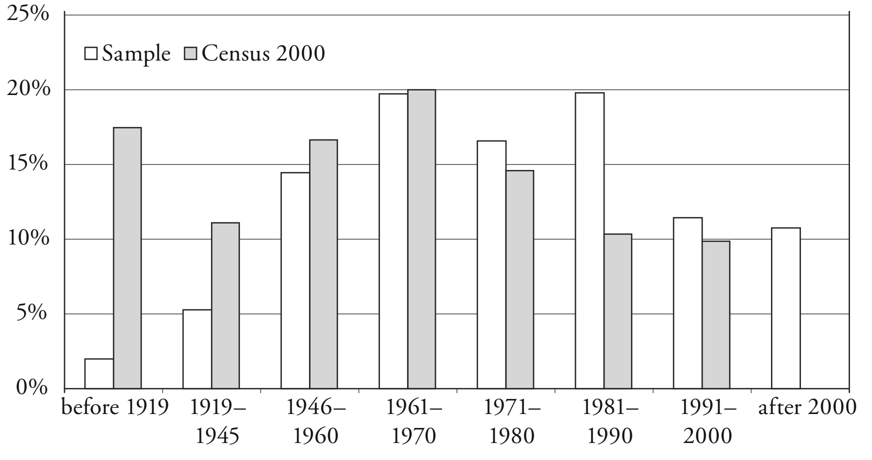


Figure 2 provides a further analysis by year of construction. It shows that the share of new apartments (built after 2000) amounts to 11% in our sample and that 42% of apartments are located in buildings constructed after 1980. Compared to the census data, construction years in the 70s and 80s are over-represented in our data. Furthermore, the share of old buildings (before the 50s) is higher in the census data. These differences are mainly attributable to different sample selection rules. First, the observation period in the census data is restricted to properties built before 2000. Second, our data are generated according to an inflow sampling scheme, whereas the census data samples all apartments at a particular point in time.

On the basis of the previous analyses, the claim that our sample is representative of the stock of apartments in Switzerland does not seem exaggeratedly strong.

Figure 2: Comparison Between the Sample and the 2000 Population Census by Year of Construction



2.3 Variables

The variables used to derive the rental model are divided into 2 groups. The group of structural variables includes the characteristics of the dwelling itself such as the year of construction, the dwelling area, the number of rooms and the floor level. We also use two variables for the assessment of condition and standard. These are discrete variables that are coded with a rating of 1 point for the lowest level to 5 for the highest level. For instance, a score of 1, applied to the condition, indicates a need for renovation, while 5 for the standard denotes a luxurious apartment. The definition of the ratings deserves some further comments. Wüest & Partner is mandated by institutional investors to value their real estate on a yearly (and semi-yearly) basis. For this purpose, Wüest & Partner uses a network of valuers, who apply a uniform procedure to appraise the real estate. To assess the quality factors (condition, standard and location within the municipality), different criteria are used within a guided procedure developed for the valuation purpose. For instance, the presence and the amount of refurbishment investment in the building elements (shell, roof, heating…) are considered in giving a rating to the condition variable. In assessing the standard, the fit-out standard of the kitchen (materials used and number of appliances) is analysed. In addition, the layout quality of the dwelling (in terms of the relationship between the number of rooms and the room sizes) and the sanitary accommodation are also rated. Finally, special features such as an indoor/outdoor swimming pool, an open chimney or a Swiss "Minergie" energy efficiency standard are used to distinguish between luxurious and the average apartments. In this guided procedure, each individual variable is graded and the overall rating is derived using a weighted average of all the individual variables.

The second group of variables relates to locational attributes. The first variable is called the "macro-location" and captures the differences on rental prices that are attributed to differences existing between municipalities. In hedonic models for the estimation of housing prices, the effect of locational attributes is typically specified by expanding the fixed part of the model (controlling for some socioeconomic factors such as accessibility, unemployment rate or the tax level). An important question arising with this approach is that the fixed effects can be overfitted in municipalities with a small number of observations. Moreover, we encounter the problem of proliferation of parameters in the hedonic equation, which results from the inclusion of many locational variables. Lastly, we want to predict a rental price for a particular apartment for each municipality in Switzerland. That is why the macro-location is generated in a separate spatial and temporal model, thus reducing the problem of n dimensions to a problem of one dimension.[4] Based on data in newspaper advertisements and Internet listings, this model computes virtual prices for each municipality. Formally, the listed m^2 price for one municipality depends on the prices of the neighbouring municipalities and on municipality-specific characteristics relevant in assessing the attractiveness of one municipality with respect to another. The variables at municipal level represent different socioeconomic factors such as accessibility, infrastructure (motorway, rail access, shopping facilities), demographic structure and trends, commuter flows, employment situation, tax burden, vacancy rates, building activity and the structure of the current building stock. Travelling times (average between private and public transport) rather than geographical distances are used as a weighting matrix in the spatial model given Switzerland's special feature of having natural barriers. The differences in the virtual prices

4 For Switzerland, different models using a wide range of socioeconomic characteristics exist for estimation of the macro-location (the model by Salvi, Schellenbauer and Schmidt, 2004 for the Canton of Zurich, the models by Scognamiglio, 2000 and Fahrländer, 2006 estimate the macro-location for each Swiss municipality). Having a separate model for the macro-location is helpful in simulating rents for a particular apartment for each municipality of Switzerland. The rental differences can be thus represented in a map of Switzerland (see, for instance, Figure 5).

reflect differences between municipalities.[5] Wüest & Partner produces a series of virtual prices for each segment of the real estate market: the retail market, office market and housing market, in which a distinction is drawn between single-family houses, owner-occupied apartments and rental apartments. This implies that a municipality may have a different degree of risk from the investor's point of view according to the type of investment which is made. The determination of these virtual prices uses a separate model and its discussion deserves a separate paper. Since our study is mainly descriptive, we will leave these questions for future research.

Finally, to assess environmental effects, we use a variable called the "micro-location". This variable comes from direct site observation and represents the quality of the building's location within the municipality or district in terms of desirability (prestige value of the residential location, attractiveness of the neighbouring residential area), noise, view within the municipality and accessibility (public transport links and frequency). The assessment of noise levels may reflect the existence of a road with heavy traffic or a railway line within 100m of the building and the presence of special features such as shooting range or industrial facilities that would produce high noise levels.[6] Variables used in the conditional analysis are briefly described in Table 2.

Most variables enter the model linearly or log-linearly. We use a quadratic term for the age of a building to allow more flexible specification. Extensions of the model are still possible by using additional nonlinear or interaction terms to capture rent differences for thinner markets or to model rents for non-standard properties such as luxurious apartments at prime locations, or markets of bigger or smaller apartments. As we mentioned in the introduction, this paper does not aim to construct a complete rental model. It deals with multilevel modelling as a means of taking account of clustering in the data.

5 These virtual prices are modelled on a quarterly basis. For the 5 major Swiss cities of Zurich, Geneva, Basel, Berne and Lausanne, different submarkets defined by urban districts are allowed. For instance, the city of Zurich is composed of 12 geographical areas such that 12 average m^2 prices for the "macro-location" are considered.

6 Assessing the impact of noise in hedonic models is difficult since the question concerns the reliability of scientific measures of noise as instruments for subjective perceptions by individuals. For Switzerland, the study by Baranzini, Ramirez, Schaerer, and Thalmann (2006) provides a comparison of perceived and scientific noise measures for the region of Geneva. The results show that the scientific measures of noise approximate well to the perceived measures.

Table 2: Model Variables

Variables	Description
Dependent variable	
Rent	Yearly rent in CHF (in log)
Structural variables	
Area	Dwelling area (in log)
Age	Age of the building (linear and squared effects)
Condition	Condition of the building (scale from 1 to 5)
Standard	Standard of the apartment (scale from 1 to 5)
Rooms	Number of rooms in the apartment
Number of apartments	Number of apartments in the building
Building type	Dummy for a mixed commercial and residential building
Floor level	Position of the apartment in the building
Locational variables	
Macro-location	Average asked price by m^2 (in log)*
Micro-location	Situation within the municipality (scale from 1 to 5)
Time variables	
dhj_1 (reference)	Dummy if the rental contract begins between the 1st and the 2nd quarter 2006.
dhj_2	Dummy if the rental contract begins between the 3rd and the 4th quarter 2006.
dhj_3	Dummy if the rental contract begins after 2007

Notes: to facilitate interpretation of the coefficients, the continuous independent variables have been deviated around their means such that the model is estimated with regard to the average dwelling; a rating of 1 (or 5) corresponds to the minimum (or maximum) mark for a property of the lowest (or highest) quality. * This corresponds to the virtual prices (by m^2) which are computed in a separate model mentioned on page 7.

3. The Multilevel Model

In this paper, we are concerned with the estimation of a hedonic model to derive rents for the whole Switzerland that relaxes the very simplistic assumption of a unitary housing market. The specification of segmented markets can be accomplished in several ways. One solution is to use the traditional hedonic model in which the fixed part is expanded through a series of dummy variables for the submarkets (cantons for example) or to estimate a separate equation for each submarket. Expanding the fixed part with dummies and/or interaction terms is difficult because the number of parameters tends to increase especially if the basic model already includes many variables. Second, market-specific parameter estimates are often based on a very small number of observations. In this case, it is difficult to make statistically reliable inferences.

Multilevel modelling is helpful in this latter situation: multilevel models can use not only the individual data in a particular submarket as in the fixed-effects models with interaction terms but also information in the pooled data of all submarkets. For markets with sparse data and/or a wide rental variability within this submarket, the estimation coefficients will rely on the pooled data: the estimates will shift towards the mean effect (or the pooled mean). This pooling allows statistical inferences to be made in a situation where no inference would have been possible using traditional methods. Multilevel modelling to a certain extent follows the principle of "borrowing the strength" of pooling to allow statistical inference in submarkets with sparse data. For submarkets with a large number of observations, the multilevel and ordinary regression models (with cantonal dummies for example) will produce similar estimates.

To take account of the fact that our data are clustered at municipality level, we consider two levels of hierarchy: level 1 corresponds to the apartment and level 2 to the municipality. In this hierarchical framework, we treat the municipalities as a random sample taken from a population of municipalities. We want to make inferences about the variations between municipalities in general. We are not interested in a specific municipality or a specific apartment.

3.1 Model with a Random Intercept Only

If we consider the simplest model with no explanatory variable except the intercept, the rent of apartment i in the municipality j is given by a function of the market-wide rent β_0 plus a differential u_{0j} for the municipality j. The coefficient β_0 is fixed and the terms u_{0j}, which are referred to as random effects in the model, represent the departure from the overall mean. The rent y_{ij} of apartment i located in the municipality j is defined as follows:

$$y_{ij} = \beta_{0j} + e_{ij}; \quad i = 1 \ldots N_j \text{ apartments, } j = 1 \ldots N \text{ municipalities} \tag{1}$$

$$\beta_{0j} = \beta_0 + u_{0j} \tag{2}$$

Equation (1) is based on individual data and thus represents a micro-model (also referred to in literature as the within-group equation). Equation (2) is based on aggregated data and represents a macro-model (also called the between-group equation), in which the fixed coefficient is allowed to vary between submarkets, i.e. the intercept of the micro-model is the dependent variable in the macro-model. Combining both models leads to the final model $y_{ij} = \beta_0 + u_{0j} + e_{ij}$ in which the two error terms u_{0j} and e_{ij} are assumed to be independent of each other (see GOLDSTEIN, 1987). They are referred to as random parameters since they are allowed to vary according to a distribution. The assumption of a normality distribution in this simple model leads to the presence of two variances: σ_e^2 the variance at level 1 and σ_{u0}^2 the variance at level 2.

3.2 Model with a Random Slope

Fully random multilevel models consider that implicit prices of housing attributes can vary between submarkets. The attribute parameters are allowed to vary according to a higher-level distribution. This is accomplished by specifying an additional macro-model:

$$\beta_{1j} = \beta_1 + u_{1j} \tag{3}$$

The implicit price of the housing attribute is equal to the average market-wide price plus a differential for each municipality. The complete model is thus given by the following expression:

$$y_{ij} = \beta_0 + \beta_1' X_{ij} + (u_{0j} + u_{1j} X_{ij} + e_{ij})$$

The error terms are again assumed to be normally distributed:

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u01} & \sigma_{u1}^2 \end{bmatrix}$$

The model now has four random terms: σ_e^2 the variance at level 1; and σ_{u0}^2, σ_{u1}^2 and σ_{u01}^2 the variances and the covariance term at level 2. The intercept and implicit prices are not assumed to be fixed, separate and independent but as coming from a higher-level distribution having a mean and a covariance structure. The mean of this distribution is the usual intercept and implicit price representing the average market-wide relationship. The variances and covariance capture the parameter drift. If the variances are zero, there is no need for macro-models and thus no significant submarket effects.

Conversely, the presence of segmented markets implies spatial dependence and/or autocorrelation in the error terms of the hedonic rent equation. Given the similarities in rents within a submarket, errors are more likely to be correlated within submarkets than across submarkets.

The model assumes random grouping, i.e. that municipalities in the sample are a random sample of all municipalities in the population.[7] It is estimated using an iterative procedure based on the maximisation of the likelihood function (for details, see GOLDSTEIN, 1995).[8] For this algorithm, a normal distribution is assumed for the parameters of the random effects. The normality assumption implies that values for municipalities that are far from the mean effect are rather improbable. This means that the municipality-specific parameters are more or less close to the "general situation". The model assumes a parametric structure for the covariance matrix of the error term. As an alternative, a Bayesian estimation or a method of moments for a nonparametric estimation of the covariance matrix can be used.

4. Estimation Results

4.1 Random Intercept Model

The simplest model provides useful preliminary information about the variance of rents at municipality level (see Table 3). We can decompose the total variance into two components corresponding to each level of the hierarchy: the variance between municipalities and the variance between apartments within the same municipality. The intra-class correlation, which is interpreted as the proportion of total variance due to differences between municipalities, is equal to 29%. This means that 29% of rental differences are explained by municipalities.

7 Actually, our data suffer from a sample selection problem which arises from the fact that only apartments owned by institutional investors are considered in the analysis. This can lead to some problems in the regional representativity of municipalities sampled in our data. That is why we make the restrictive assumption that random grouping is conditional on the previous selection rule.

8 The different models have been estimated using MlwiN 2.0, a program written and developed by Rasbash, Steele, Browne and Prosser from the Center for Multilevel Modeling, University of Bristol (and University of Nottingham).

Table 3: Estimated Parameters in the Random Intercept Model

Parameter	Coefficient	(std. err.)
Fixed		
Intercept	**9.555**	(0.013)
Random		
Level 2 (municipality level) σ^2_{u0}	**0.047**	(0.004)
Level 1 (apartment level) σ^2_{e}	**0.117**	(0.001)
Intra-class correlation	0.29	

Note : **bold** significant at 5% level.

After controlling for housing attributes, the intra-class correlation coefficient decreases to 24%. This small reduction strongly suggests that rents vary across municipalities and provides evidence for the existence of segmented markets.

4.2 Random Slope Model

In this section, the coefficient for dwelling area is treated as a random effect. The paper focuses on a simple hedonic model that can illustrate the value of the multilevel approach in allowing for segmented markets. To this end, we make the restrictive assumption that only the rate of increase of rents associated with an additional square footage of the dwelling area can vary across municipalities. The other explanatory variables are assumed to have a fixed effect. This latter assumption can be relaxed and the question of the specification of a complete model to explain rents is left for future research.

Table 4 shows that, after controlling for the explained variables, there is a significant unexplained variance in rents that remains at municipality level. Upon interpretation of the other effects, we find the expected relationships. For instance, a positive effect is found for an additional m^2 dwelling area, an additional room or a better quality score, whereas an apartment located in an older building has a lower rent. When looking at the effect of the dummy for mixed-use and residential buildings, there is no preconception about the sign of this effect. Here, a positive effect is obtained meaning that apartments located in mixed-use buildings have higher rents than apartments located in residential buildings.

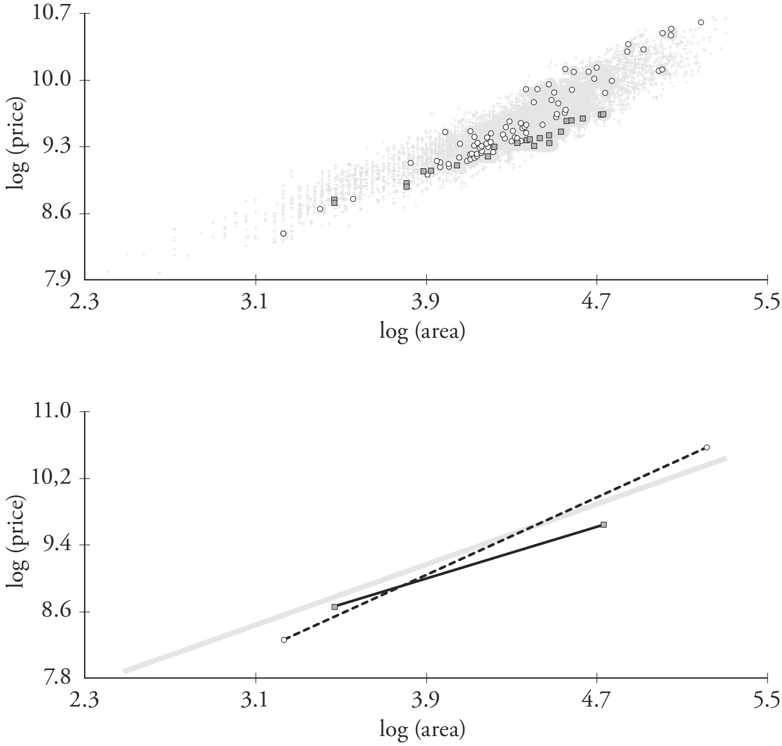
Table 4: Estimated Parameters in the Random Coefficient Model

Parameters	Coefficient	(std. err.)
Fixed effects		
Intercept (overall mean effect)	**9.549**	(0.019)
Dwelling area (overall mean effect)	**0.603**	(0.017)
Macro-location (in log)	**0.789**	(0.024)
Age in years	**–0.007**	(0.000)
Age squared	**0.006**	(0.000)
Number of rooms	**0.052**	(0.005)
Number of apartments	**–0.0002**	(0.000)
Floor level	**0.009**	(0.001)
Condition	**0.065**	(0.003)
Standard	**0.024**	(0.004)
Micro-location	**0.009**	(0.003)
Mixed building	**0.038**	(0.005)
Dhj_2	**0.007**	(0.003)
Dhj_3	**0.036**	(0.004)
Random effects		
Level 2 (municipality level)		
σ^2_{u0} (intercept)	**0.022**	(0.004)
σ_{u01} (covariance)	**0.020**	(0.004)
σ^2_{u1} (dwelling area)	**0.036**	(0.006)
Level 1 (apartment level)		
σ^2_{e}	**0.017**	(0.000)

Note : **bold** significant at 5% level.

In the paper, the rate of increase in rents for additional square footage is allowed to vary across municipalities, between municipalities with higher average rents and between municipalities with lower average rents. A correlation term that is not significantly different from zero implies that the rate of increase in rental prices for additional square footage is similar across municipalities regardless of the municipality's mean rent. This correlation coefficient can be calculated

Figure 3: Relationship between Dwelling Area and Rents (in log)



as $\sigma_{u01}/\sigma_{u0}\sigma_{u1}$ and is equal to 0.71 in our random coefficient model. This has some interesting implications. It suggests that there is a relatively strong association between average municipality rents and the municipality effects of square footage – indicating that municipalities with higher average rents tend to have a greater rate of increase in rental prices associated with increases in dwelling area.

Figure 3 provides an illustration of the variability of the effect of dwelling area at municipality level. The upper part of Figure 3 shows how the observations are distributed according to dwelling area. Two municipalities have been represented: one with circles and another with squares. We can draw a black line for the average relationship between rents and dwelling area. The circles are significantly above the average line whereas the squares are significantly below it. As a consequence, Figure 3 shows that different intercepts have to be considered. The graph also suggests a more complicated situation. At higher levels of dwelling area, the circles appear to be consistently above the squares. By contrast, at the other end of the scale, there does not seem to be much difference between these municipalities. The lower part of Figure 3 indeed suggests that the two municipalities have different slopes. The municipality represented by circles appears to be an exceptional municipality with a higher intercept and a much steeper slope than the other municipalities.

4.3 Comparison

This section presents the estimation results of two classical hedonic models in which the fixed part contains cantonal dummies and interaction terms to take account of segmented markets. These models are estimated using the standard OLS estimation method (OLS 1 for specification with cantonal dummies only and OLS 2 for specification with cantonal interaction terms for dwelling area). By way of comparison, estimates of the multilevel hedonic model (ML) are also presented.

In the traditional models, the estimate of the yearly rental price for the average dwelling is about CHF 13 500 with an increase of about 4% in relation to the reference canton for apartments located in the Cantons of Zurich, Berne, Solothurn and St. Gallen.[9] By comparison, the typical dwelling in the Cantons of Basel-Stadt and Baselland has a rental price which is on average 8% higher. The positive interaction term for the Canton of Basel-Stadt in the second specification indicates that the effect of dwelling area is greater than in the rest of Switzerland. For the other cantons (except for the Cantons of Baselland and Argau having no significant interaction terms), the effect of an additional square foot is lower than the overall mean effect. In these cantons, rental prices are on average higher than in other Swiss regions (positive coefficients for the dummies). Yet, there is no additional premium for a higher dwelling area. This is not the case for the Canton of Basel-Stadt, for which a higher intercept and a higher slope are observed. Although the other estimated fixed effects are very similar in specification 1 and specification 2, the LR test statistic rejects the first specification against the second specification.[10]

9 exp(9.51) = CHF 13 494. Since the continuous independent variables have been deviated around their means , 13 494 CHF is the yearly rent for the average dwelling. This latter corresponds to an apartment with 74 m^2, 3 rooms, located on the 2^{nd} floor of the building, being 30 years old and having an average condition, standard and micro-location. The discrete variables are equal to zero.

10 $\chi^2_5 = 58.65$ with a p-value of 0. Using LR tests, the OLS specifications are rejected against the ML model ($\chi^2_5 = 1241.82$ for OLS 1 and $\chi^2_{10} = 1183$ for OLS 2 with a p-value of 0).

Table 5: Comparison between the Different Models

Parameters	OLS 1		OLS 2		ML	
	Coeff.	(std. err.)	Coeff.	(std. err.)	Coeff.	(std. err.)
Fixed effects						
Intercept (overall mean effect)	**9.512**	(0.037)	**9.513**	(0.037)	**9.549**	(0.019)
Dwelling area (overall mean effect)	**0.614**	(0.019)	**0.630**	(0.011)	**0.603**	(0.017)
Macro-location (in log)	**0.766**	(0.025)	**0.761**	(0.022)	**0.789**	(0.024)
Age in years	**–0.007**	(0.000)	**–0.007**	(0.000)	**–0.007**	(0.000)
Age squared	**0.005**	(0.000)	**0.005**	(0.000)	**0.006**	(0.000)
Number of rooms	**0.053**	(0.005)	**0.054**	(0.005)	**0.052**	(0.005)
Number of apartments	**–0.0001**	(0.000)	**–0.0001**	(0.000)	**–0.0002**	(0.000)
Floor level	**0.006**	(0.001)	**0.007**	(0.001)	**0.009**	(0.001)
Condition	**0.063**	(0.003)	**0.063**	(0.003)	**0.065**	(0.003)
Standard	**0.018**	(0.004)	**0.019**	(0.004)	**0.024**	(0.004)
Micro-location	**0.007**	(0.003)	**0.007**	(0.003)	**0.009**	(0.003)
Mixed building	**0.033**	(0.005)	**0.032**	(0.005)	**0.038**	(0.005)
Dhj_2	0.006	(0.003)	0.005	(0.003)	**0.007**	(0.003)
Dhj_3	**0.037**	(0.004)	**0.037**	(0.004)	**0.036**	(0.004)
Cantonal dummies						
For canton of Zurich	**0.047**	(0.004)	**0.048**	(0.004)		
For canton of Bern	**0.052**	(0.006)	**0.052**	(0.006)		
For canton of Zug	**0.029**	(0.010)	**0.034**	(0.010)		
For canton of Solothurn	**0.040**	(0.010)	**0.038**	(0.010)		
For canton of Basel City	**0.085**	(0.008)	**0.092**	(0.007)		
For canton of Basel Land	**0.078**	(0.007)	**0.083**	(0.007)		
For canton of St.Gallen	**0.036**	(0.007)	**0.037**	(0.007)		
For canton of Aargau	**0.056**	(0.005)	**0.054**	(0.005)		
Cantonal interaction terms						
For canton of Zurich			**–0.048**	(0.009)		
For canton of Bern			*–0.030*	(0.016)		
For canton of Zug			**–0.053**	(0.025)		
For canton of Solothurn			**–0.063**	(0.024)		
For canton of Basel City			**0.036**	(0.015)		
For canton of Basel Land			0.031	(0.025)		
For canton of St.Gallen			**–0.049**	(0.018)		
For canton of Aargau			–0.006	(0.671)		
R^2	0.858		0.859		0.877	
Log-likelihood	5493.04		5512.95		6277.26	
SSE	245.85		244.98		185.61	

Notes: the reference canton is composed of other cantons; **bold** significant at 5% level; *italics* at 10% level; 11 913 observations.

As regards the fixed effects, the different regressions yield similar results. It is common to many hedonic pricing models that not all variables exhibit the proper sign when estimated with OLS. This is not the case in this application: there is no variable that is treated differently according to the type of estimation. However, the estimators differ in their estimates of the magnitude of the fixed effects. For example, the ML estimate for the number of apartments in the building is 100% greater (in absolute term) than the corresponding OLS estimates. This is also the case for the effect of floor level (+ 50%), of standard (+ 33%) and to a lesser extent for the effect of micro-location (+ 20%). A particular attention should be devoted to the effect of the macro-location, which is closer to one in the ML specification. In the hedonic model, we expect the coefficient for the macro-situation to be close to one since this variable reflects the relative attractiveness of municipality A with respect to municipality B. If the generalised m^2 asked prices in municipality A are 1% higher than in municipality B, then the yearly rental price of a specific apartment should be 1% higher in municipality A than in municipality B. However, as advertisements offer insufficient information on the characteristics of apartments, the m^2 asked price for one municipality is computed without correcting for the structural attributes (except for floor area). As a consequence, the contextual effects of the macro-location in the hedonic model are potentially confounded with the rent impact of apartment attributes. That is why the coefficient for the macro-location in the hedonic model is not equal to one.

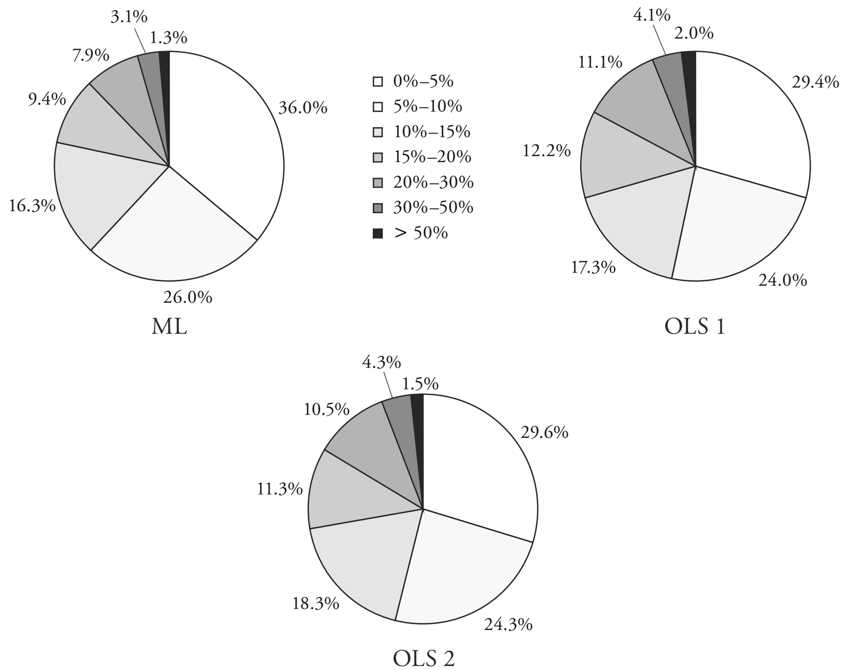
The locational attributes (cantonal dummies and interaction terms) included in specifications OLS 1 and OLS 2 measure large-scale effects. However, the very local influences may prove to be more important in the prediction of rental prices. If we use the sample R^2 and the sum of squared errors SSE to measure goodness of fit, the estimated SSE falls by 33% relative to the OLS results whereas the R^2 rises from 0.858 to 0.877, which is a small, though significant amount. Use of a likelihood ratio statistic to check whether the different specifications are significantly different yields the result that the OLS specifications are rejected against the ML specification.

This section also provides some information on the precision of models in terms of rental prediction. The following figure represents the absolute differences between actual and predicted rents (as % of actual rents).[11] The class 0%–5%

11 If we consider the simplest version of the multilevel model (model with a random intercept only), the true values of u_{0j} (level 2 residuals which represent departures from the overall mean β_0) are unknown, but we can obtain estimates of them given the observed data and the estimated parameters of the model. These estimated or predicted residuals are often referred to

means that the predicted rents lie between ± 5% of actual rents. Figure 4 shows that a higher precision is achieved in the multilevel hedonic model. As an illustration, about 62% of predictions lie within ± 10% of the actual rents. This share amounts to 53% for OLS 1 and to 54% for OLS 2. [12]

Figure 4: Comparison between ML and OLS Estimation Methods



---

as posterior residual estimates in Bayesian terminology. If we define the raw residual for the apartment i in the jth municipality by $r_{ij} = y_{ij} - \hat{y}_{ij}$, the raw residual for the jth municipality is obtained by averaging the individual residuals r_{ij} for all apartments in that municipality. Then, the level 2 residuals for this municipality are estimated using

$$\hat{u}_{0j} = \frac{n_j \hat{\sigma}^2_{u0}}{n_j \sigma^2_{u0} + \hat{\sigma}^2_e} \sum_{i=1}^{n_t} r_{ij}$$

where n_j represents the number of apartments in municipality j (see Goldstein, 1995).

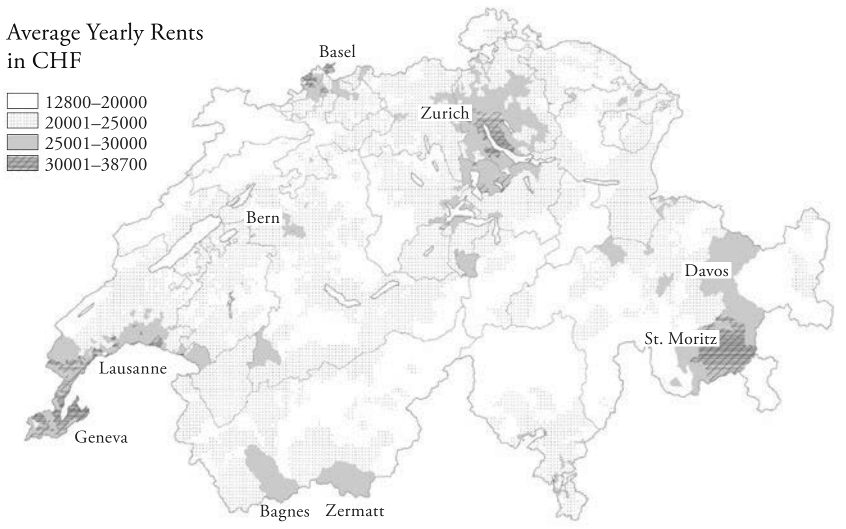
12 OLS actually corresponds to an estimation technique and not a model. For the sake of simplicity, however, the designation "OLS 1" (or 2) refers to specification 1 (or 2).

4.4 Simulation of Rents at Municipality Level

The estimated coefficients of the model allow us to simulate the yearly rental price that a particular apartment can achieve in each municipality of Switzerland. To this end, yearly average rental prices were calculated for an apartment with 4 rooms, 110 m^2 living area, located in a new building with 10 apartments. The apartment is of average standard and the location within the municipality is considered good. In municipalities for which we have no observations at all, because they are not in our sample, our best estimate of the rent is the overall mean (i.e. values given to all fixed and random coefficients correspond to the overall mean effects). This implies that rent differentials appearing in Figure 5 are attributable to differentials in the macro-location variable.

Figure 5 illustrates the municipality differences that can exist in rents for this particular apartment. For instance, rents are highest in Switzerland's five major cities as well as in the tourist regions (such as St. Moritz or Davos) and in the regions around Lake Zurich and Lake Geneva. On the other hand, rents are much lower in the rest of Switzerland, e.g. in the Canton of Jura.

Figure 5: Comparison of Rents for a Simulated Apartment at Municipality Level



Note: Geostat/Swiss topo

5. Conclusion

The hedonic approach has rapidly established itself as one of the most widely used analytical techniques for estimating the implicit prices of housing attributes. The popularity of hedonic housing price models stems from the multiple regression analysis, which represents a powerful statistical instrument. Multiple regression analysis allows the sorting out of crossed influences that can affect prices. Although this method offers many advantages, somewhat restrictive assumptions are required (in particular regarding the independence and the homoskedasticity of the residuals of the model). The model's failure to accommodate such restrictions due to the presence of excessive multicollinearity, heteroskedasticity and spatial autocorrelation may lead to estimates of implicit prices for housing attributes that are no longer reliable. Many solutions to most of the problems encountered with multiple regression analysis have been discussed in economic literature. An appropriate market segmentation and an adequate model specification in describing rental price variation at different spatial levels have been proposed.

This study presents an approach to the determination of rents for the whole of Switzerland that provides for the existence of different submarkets. We apply a suitable procedure that takes into account the hierarchical nature of the data. In the empirical part, we consider two levels of hierarchy: the micro-level (apartment) and the macro-level (municipality). With this multilevel structure, we find that a significant part of rental differences are attributable to municipality differences. Even after controlling for housing attributes, this unexplained variance remains at municipality level.

We compare the predictive ability of two specifications of the simpler hedonic model based on submarkets for cantons with the multilevel hedonic model. We find that – for our data – the multilevel hedonic model provides more accurate predictions that the traditional model. The simple model, including a series of cantonal dummies and interaction terms, performs better than the simple model with cantonal dummies only. Our results therefore strongly suggest that the determination of rents requires the definition of different submarkets. The advantage of multilevel modelling over the traditional hedonic model is that it allows statistical inferences in submarkets with sparse data. For markets with a small number of apartments, we can, for example, obtain individualised estimates for dwelling area. Yet we know that these effects are likely to be estimated imprecisely. Rather than accepting these inaccuracies, we may prefer to exploit the fact that the submarkets belong to a population of submarkets whose parameters have been estimated using a weighted combination of the estimated micro- and macro-effects. Unreliable implicit prices are thus shrunk towards the overall mean effect. These features of the multilevel models yield benefits in the estimation of implicit prices and can thus help accurate prediction. Moreover, if we assume that submarkets are considered as functionally related rather than separate and autonomous, then expanding the random part of the model responds more closely to the definition of submarkets than expanding the fixed part of the model.

Finally, in our application, we consider the particular example of implicit prices for dwelling area that can vary across municipalities. The purpose of this paper is to show how a simple hedonic model can be used to take account of the inherent hierarchy in determining rents. The model can, of course, be extended to consider more random effects.

References

Bajari, P., and M. Kahn (2005), "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities", *Journal of Business and Economic Statistics*, vol. 23(1), pp. 20–33.

Bajic, V. (1985), "Housing Market Segmentation and Demand for Housing Attributes: Some Empirical Findings", *AREUEA Journal*, vol. 13(1), pp. 58–75.

Baranzini, A., J. Ramirez, C. Schaerer and P. Thalmann (2008), "Using the Hedonic Approach to Value Natural and Land Uses in an Urban Area: An Application to Geneva and Zurich", *Economie Publique/ Public Economics*, vol. 20, pp. 1–21.

Baranzini, A., J. Ramirez, C. Schaerer and P. Thalmann (2006), "Feel it or Measure it – Perceived vs. Measured Noise in Hedonic Models", Cahier de recherche HES-SO/HEG-GE/C-06/7/1-CH.

Basu, S., and T. Thibodeau (1998), "Analysis of Spatial Autocorrelation in House Prices", *Journal of Real Estate Finance and Economics*, vol. 17(1), pp. 61–86.

Bender, A., B. Gacem and M. Hoesli (1994), «Construction d'indices immobiliers selon l'approche hédoniste», *Finanzmarkt und Portfoliomanagement*, vol. 8 (4), pp. 522–533.

Berry, S., J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium", *Econometrica*, vol. 63(4), pp. 841–889.

Bryk, S., and S. Raudenbush (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, Second Edition, Sage Publications, Inc., Thousand Oaks, CA.

Fahrländer, S. (2006), "Semiparametric Construction of Spatial Generalized Hedonic Models for Private Properties", *Swiss Journal of Economics and Statistics*, vol. 127 (4), pp. 501–528.

Goldstein, H. (1987), “Nonlinear Multilevel Models, with an Application to Discrete Response Data”, *Biometrika*, vol. 78 (1), pp. 45–51.

Goldstein, H. (1995), “Multilevel Statistical Models”, Third Edition, *Oxford University Press*.

Graves, P., J. Murdoch, M. Thayer and D. Waldman (1988), “The Robustness of Hedonic Price Estimation: Urban Air Quality”, *Land Economics*, vol. 64(3), pp. 220–233.

Griliches, Z. (1971), “Hedonic Price Indexes Revisited”, in *Price indexes and quality change*, Federal Reserve Board, Massachusetts, pp. 3–15.

Hoesli, M., P. Favarger and C. Giacotto (1997), “Real Estate Price Indices and Performance”, *Swiss Journal of Economics and Statistics*, vol. 133, pp. 29–48.

Quigley, J. (1985), “Consumer Choice of Dwelling, Neighborhood and Public Services”, *Regional Science and Urban Economics*, vol. 15, pp. 41–63.

Rosen, S. (1974), “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”, *Journal of Political Economy*, vol. 82, pp. 34–55.

Salvi, M., P. Schellenbauer and H. Schmidt (2004), *Preise, Mieten und Renditen: Der Immobilienmarkt transparent gemacht*, Ed.: Zürcher Kantonalbank, Zürich.

Rieder, T. (2005), „Die Bestimmung der Mieten in der Schweiz – ein hedonisches Modell unter Berücksichtigung der räumlichen Struktur zur Analyse der Mietpreise in der Schweiz“, Lizenzarbeit Universität Bern.

Thalmann, P. (1987), «Explication empirique des loyers lausannois», *Swiss Journal of Economics and Statistics*, vol. 123, pp. 47–70.

SUMMARY

This paper is an empirical application of a hedonic model to determine rents in Switzerland. Unlike traditional hedonic models, we take account of the fact that data are clustered at different levels. The resulting spatial autocorrelation and heterogeneity permits a consideration of different submarkets within Switzerland. For the empirical analysis, we consider the special case of the clustering of apartments at municipality level. Our results show that even after controlling for housing and locational attributes, a significant part of rental differences are still attributable to municipality differences.