UNIVERSITÄT
BERN

Faculty of Economics and
Social Sciences

Department of Economics

Semiparametric Construction of Spatial Generalized Hedonic Models for Private Properties

Stefan Sebastian Fahrlaender

05-07

June 2005

DISCUSSION PAPERS

Gesellschaftsstrasse 49
CH-3012 Bern, Switzerland
http://www.vwi.unibe.ch

SEMIPARAMETRIC CONSTRUCTION OF SPATIAL GENERALIZED HEDONIC MODELS FOR PRIVATE PROPERTIES

STEFAN SEBASTIAN FAHRLÄNDER*

Version: 22nd June 2005

Abstract

This paper analyzes the empirical hedonic prices for non-standard condominiums and single family houses using nonparametric estimates as well as a generalized additive model for the spatial generalization of the attractiveness of all Swiss communities. We find that the assumption of log-linearity for continuous variables does not hold but can be replaced by partwise log-linear or quadratic terms.

Due to the topographical segmentation of Switzerland, driving times seem to be more adequate than geographical distances to explain the price level of a village with the price level of its neighbours. We show, that using metric multidimensional scaling, the driving times between the villages can be converted into artificial coordinates with three principal axis to serve as a basis for the prediction of the macro-locations of the Swiss villages.

JEL Classification Codes: C52, R31.

Keywords: Hedonic prices, private property, Switzerland, robust regression, splines, multidimensional scaling.

---

* Universitätstrasse 118, 8006 Zürich, Switzerland, +41 43 233 85 90; fahrlaender@gmx.net.

This paper benefited from helpful comments by Klaus Neusser and Andreas Ruckstuhl. The usual disclaimer applies. A special thanks goes to the data providing Swiss banks and insurance companies as well as to Wüest & Partner.

First draft: April 2005. This draft has been presented at the Brown Bag Seminar of the Department of Economics at the University of Bern as well as to specialists of UBS, Wüest & Partner and Zürcher Kantonalbank in June 2005.

1 INTRODUCTION

In recent years probably hundreds of articles discussing the hedonic estimation of prices and their dynamics for private properties have been published (For recent reviews and meta-analysis see for example MAURER, PITZER and SEBASTIAN, 2000; MALPEZZI, 2002; NELSON, 2003 and SIRMANS, MACPHERSON and ZIETZ, 2005).[1] The greatest part of international literature discusses the parametric specification of hedonic models and indexes and is mostly focussed on specific effects like airport noise, infrastructure projects etc. in single metropolitan areas or submarkets within metropolitan boundary.

Various studies on direct index construction for private properties have been published for Switzerland and Swiss regions.[2] While BENDER, GACEM and HOESLI (1994) and HOESLI, FAVARGER and GIACOTTO (1997) focus on the Swiss Canton of Geneva, the study of BIGNASCA ET AL. (1996) and SALVI, SCHELLENBAUER and SCHMIDT (2004) discuss the estimation and the price path for private properties in the Canton of Zurich. SCOGNAMIGLIO (2000) calculates nation-wide direct indexes. In 2000 Fahrländer and Hausmann[3] developed indirectly constructed and regionally differentiated indexes for Switzerland.

For several years, hedonic models for single family houses and condominiums have been widely used by Swiss banks, insurance companies and others as instruments to estimate market values of private properties in day-to-day business. Whereas models initially were used for standard properties, the application over time went more and more towards the appraisal of non-standard objects and more complex models become necessary.

With the availability of more – especially regionally dispersed – data, as well as newest statistical methods, the possibilities to measure hedonic prices increase and the models can be improved, especially for non-average properties and thin markets.

---

1 For the theoretical foundations see ROSEN (1974). NELSON (2003) published a meta-analysis comparing the findings of 23 studies concerning airport noise and hedonic property values.

2 On the discussion of direct and indirect index construction see for example MURRAY and SARANTIS (1999) or MAURER, PITZER and SEBASTIAN (2000).

3 See FAHRLÄNDER (2001a and b) and FAHRLÄNDER and HAUSMANN (2001). The indexes have been presented to specialists of several Swiss banks including UBS and Zürcher Kantonalbank in winter 2000 and spring 2001.

This article is focussed on two main questions:

Firstly: Starting from a rather simple parametric log-linear hedonic model (section 3.1), nonparametric methods as well as interaction terms are introduced to measure the empirical hedonics and to improve the quality of fit for properties with non-standard specification throughout all Swiss regions.[4] These questions go along with the ongoing discussion of hedonic models in the international literature (see for example WALLACE, 1996; EKELAND, HECKMAN and NESHEIM, 2003 or CLAPP, 2004). In this step, fixed effects are used to get an estimate for the attractiveness of the location (e.g. the macro-location of 2'910 Swiss villages) for the 4th quarter 2004 (section 3.2).[5]

Secondly: Due to thin markets in certain regions fixed effects for the macro-location can only be estimated for approximately one third of the Swiss villages. In addition, certain estimated effects are based only on a small number of observations and therefore might be overfitted and not represent the true mean. Therefore a model to review the estimated effects is necessary as well as predictions for the villages, where no or only little observations are available (section 3.3).

In section 3.4 the models then are combined and in-sample as well as out-of-sample analysis are conducted.

We find that improvements for the fit of non-standard objects are possible using nonparametric or partwise log-linear and quadratic terms for continuous variables, for discrete variables, factors as well as interaction terms are introduced (sections 3.2 and 3.4). Since Switzerland, with its numerous mountains and lakes, is geographically extremely segmented, geographical coordinates do not seem to be adequate to predict the macro-location of the Swiss villages. Using metric multidimensional scaling (MDS) it can be shown, that a matrix of the driving times between all the Swiss villages can be transformed into a system of three principal coordinates.[6] Using thin plate regression splines for these artificial coordinates as

---

[4] Since the hedonic demand and supply curves are not observed, it is assumed that supply is (completely) elastic and therefore the observed hedonic prices reflect rather the demand than the supply side (see for example ROSEN, 1974 or KINOSHITA, 1987). The question whether these prices reflect a long-term market equilibrium or not, is not in the focus of this article.

[5] The macro-location of the properties is doubtlessly the most important and the only exogenous variable for the measurement of market values of private properties.

[6] The driving times are a weighted average of individual and public transport, under load, i.e. under consideration of average speed, traffic jams, access, waiting and changing times for public transport. See VRTIC ET AL. (2005).

well as the level of communal income tax, typologies of the communities and some special effects for the "Golden Shores", we find, that the macro-location of the Swiss villages can be predicted accurately (sections 3.3 and 3.4).

2 DATA

The analysis is based on arm's length transactions of single family houses and condominiums all over Switzerland, which took place in 2004. The data was compiled by Swiss banks and insurance companies in day-to-day business and covers approximately 50 percent of the yearly arm's length transaction market (data pool).[7] Due to the wide range of participating companies, which possibly cover different customer segments – i.e. banks covering retail and private banking, insurance companies and the "Alternative Bank" – and due to the size of the samples, it is believed, that the available transactions widely represent the Swiss market and most of its regions in 2004.

The data pool can be divided into objects with their own identifiable site area (single family houses) and objects without their own site area (condominiums). The later include also terraced apartments and other special cases. Only sale prices from arm's length transactions are included in the samples (for the variables see Appendix A). Special cases like ground leases, objects with depreciating easements or big development reserves are excluded.[8] Since only the political community and postal codes are available, but no geo-coordinates, the models are based on an estimated indicator for the attractiveness of the village (macro-location).[9] In addition, a rather rough assessment of the location within the community (micro-location) is determined as well.

For the description of the property itself, details concerning size, construction year, assessments of the standard and the condition as well as some other information are available.

---

7 Alternative Bank Schweiz, Banque Cantonale Vaudoise, Helvetia Patria Versicherungsgesellschaft, Luzerner Kantonalbank, Thurgauer Kantonalbank, UBS, Zürcher Kantonalbank, Zurich Financial Services.

8 Since the prices of garages are included in the total prices, the sales prices are adjusted by regional differentiated prices for garages. A consideration of the number of garages as a percentual increase of the sale price does not seem to be adequate since the value of a garage would for example raise and fall with the size of the property.

9 In the recent years many political communities have merged and since in the big cities and tourist areas further differentiation is necessary, a variable village is generated where possible out of the code for the political community and the post code. Out of 2'780 political communities (as of December 2004) 2'910 macro-locations can be identified mainly in cities like Bern, Basel, Geneva, Lausanne, Lugano, Luzern, St. Gallen, Winterthur and Zürich as well as some important villages and tourist resorts that do not represent a political community like Verbier or Crans Montana.

Out of the comparison of the theoretical condition, based on the ageing process and the assessment of the condition, a factor refurbishment level is generated.[10]

Due to non-divisionibility, site areas in Switzerland often are larger than necessary for a certain cubic content of single family houses. Using assumptions for regional differentiated planning and construction laws, and out of the site areas, an excessive site area can be approximated.[11]

Since the data is collected on a day-to-day basis, when using commercial valuation models, the quality of the seized variables is generally high except for sale prices, which may differ by a factor of 10 or 100. In addition, the assessments of the condition, the standard and the micro-location are rather rough and partially subjective. It can also be shown, that continuous variables often are rounded to $5m^2$ or $10m^2$, which equals up to 10 percent for small condominiums. Therefore we believe that a large part of the variance of econometric models can be explained by the inaccuracies of the raw data as well as the outcome of the bargaining process.

For the main analysis 9'510 transactions of condominium and 8'756 transactions of single family houses of the period January to December 2004 are included. For out-of-sample analysis additional 1'005 (condominiums) resp. 1'054 (single family houses) observations, about half from transactions during the 1st quarter 2005 and half from transactions during the year 2004, are available.[12]

---

10 5 levels with 0=no refurbishment and 4=completely refurbished.

11 In Switzerland planning and construction laws and even measurement rules may differ from community to community or at least from Canton to Canton.

12 Data compiled by Banque Cantonale Vaudoise as well as from Zürcher Kantonalbank were not available for this analysis.

3 MODELS AND RESULTS

Starting from simple log-linear models for both condominiums and single family houses (section 3.1) implicit and explicit assumptions are challenged and proposals are tested to improve the quality of fit for non-average objects (section 3.2). The models in this first and second step are based on fixed effects for villages where a reasonable number of observations are available. In a third step, models are discussed to fit the fixed effects and to predict levels of the macro-location for villages with no or only very little observations (section 3.3). In section 3.4 the models are re-estimated using predicted macro-locations and in-sample as well as out-of-sample analysis are conducted.

3.1 SIMPLE MODELS AS A STARTING POINT

In the beginning the ln of the adjusted prices (aprice) is fitted in a fixed effects model of the form

$$\ln(aprice) = \alpha_j(village) + \beta_1 \cdot \ln(nwf) + \beta_2 \cdot \ln(bauj) + \beta_3 \cdot \text{cond} + \beta_4 \cdot \text{stand} + \beta_5 \cdot \text{micro} + \beta_6 \cdot \text{yearqu} + \varepsilon \qquad (1)$$

for condominiums, without intercept but fixed effects α_j for each village j.[13]

The estimations are based on robust regression using M-estimators with Huber's ψ-function to control the influence of outlayers (see Appendix B).[14]

The use of robust methods is crucial for the estimation of the fixed effects for villages, since one single outlayer may result in a completely biased estimation for the level of the village.

These rather simple hedonic models already show very good results and the coefficients correspond with the theoretical expectations and are all highly significant (see Tab. 1). Since a large part of the two samples consist of standard and quasi-standard objects, possible improvements may not result in a dramatic change of general reference values like residual standard error or R^2. However, the employed setup includes several assumptions, which might

---

[13] The prices are regionally adjusted by modelled prices for excessive numbers of garages. For single family houses, basically the same model is fitted but instead of the dwelling, the natural logarithms of the cubic content and the site area are used for the size of the objects.

[14] All statistical analysis is computed with R (see R DEVELOPMENT CORE TEAM, 2004). For an overview on robust statistics, especially M-estimators and MM-estimators see for example HAMPEL ET AL. (1986), ROUSSEEUW and LEROY (1987), MONTGOMERY, PECK and VINING (2001) or RUCKSTUHL (2004).

be very strong and should be tested. In addition, it should be noted, that due to the use of fixed effects for the macro-location an overfit likely occurs in villages with few observations.

Table 1: Results of the starting models (robust estimation)[15]

	Condominiums			Single family houses		
	Coefficient	Standard error	Sign. level	Coefficient	Standard error	Sign. Level
ln(*nwf*)	0.917	0.004	***			
ln(*volsia03*)				0.586	0.006	***
ln(*land*)				0.106	0.003	***
ln(*bauj*)	5.837	0.246	***	6.407	0.188	***
cond	0.091	0.003	***	0.051	0.003	***
stand	0.131	0.003	***	0.110	0.003	***
micro	0.119	0.003	***	0.123	0.003	***
att				-0.050	0.004	***
2nd quarter	0.019	0.005	***	0.005	0.005	n.s.
3rd quarter	0.026	0.004	***	0.018	0.005	***
4th quarter	0.027	0.004	***	0.018	0.004	***
S.E. residuals	0.123			0.128		
D.F. residuals	8'802			7'586		
R^2	0.972			0.951		

Note: *** representing a 1‰ significance level.

3.2 MODEL EXPANSION AND NONPARAMETRIC TESTS OF THE ASSUMPTIONS

For all variables the models in section 3.1 are based on equal coefficients for all Swiss regions, which might be a very strong assumption especially in the light of different cultures, urban and rural areas, as well as tourist regions, where markets are strongly driven by foreign demand. Therefore it is tested, if regional significant differences in the coefficients exist or not.[16]

In both models constant agios for the levels of the factors regarding condition, standard and micro-location are used. These assumptions are strong and the agios should rather be

15 For practical reasons, the levels of the factor for the macro-location are not shown.

16 For results see Appendix D.

estimated for each level. In addition interactions between micro-location and standard exist especially for the luxurious market segments.[17]

It is obvious that there is a strong correlation between the condition of the objects and the year of construction, especially in the light of the fact, that the major part of each sample is mainly newly constructed objects and in the second-hand market, properties are usually sold prior to refurbishment. Therefore, besides the year of construction a factor for refurbishment levels, as a difference between the theoretical and the assessed condition, is used instead of the condition itself.

It is also tested, if there is an agio to be paid for holiday homes. A positive result – an agio of approx. 8.5 percent – is found.[18]

Although only a short period is viewed in this sample, it is possible that a factor per yearly transaction quarter for Switzerland is not sufficient and that regional factors should be introduced. It can be shown, that most of Switzerland does not differ significantly in this short period except for the Geneva-Lausanne region as well as for the tourist areas, where regional factors per transaction quarter partly show significant differences from the rest of the country.

Besides these smaller improvements, the log-linearity between the selling price and the continuous variables is tested. Under the assumption of decreasing marginal utility of the size of the objects, log-linearity for dwelling, cubic content and site area is a possible model but still a strong assumption. The non-parametric estimation for the dwelling of condominiums (see Fig. 1) is based on a cubic regression spline with 25 knots and shows that the assumed log-linearity is a good model but that improvements are still possible.[19] Since the coefficient can be interpreted as elasticity, already a rather small change in the slope may influence the predicted price to a significant extent. For smaller objects, the utility of additional floor space appears to be more important than for bigger objects. For the extended model a partwise log-

---

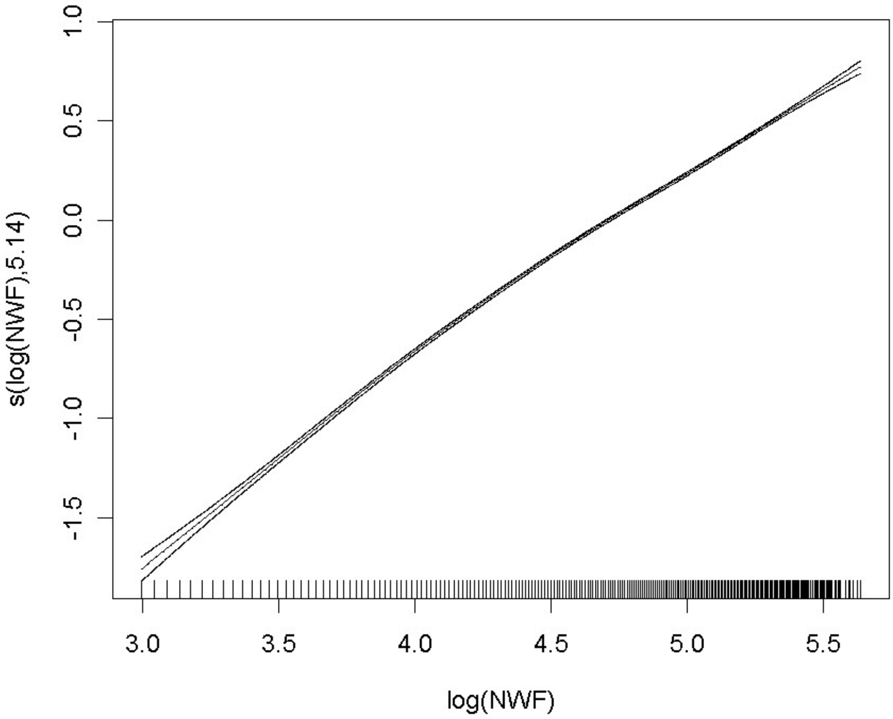
[17] Regional differences as well as differences for big and small properties have also been tested but no significant results are found.

[18] This result is consistent for both condominiums and single family houses. The reasons for this agio can be differences in the standard that are not assessed properly as well as a proximity to funiculars and other winter sport facilities but could also be found in lacking knowledge of the buyers of the local market. The market can also be driven by the quota system for the foreign demand.

[19] Splines are one possibility of nonparametric estimation. Cubic regression splines estimate a cubic regression for each segment of the spline and paste them together while setting the 2nd derivative to zero at the knots, e.g. the ends of the segments. See for example HASTIE and TIBSHIRANI (1990), GREENE (1997), HASTIE, TIBSHIRANI and FRIEDMAN (2001) or MÄCHLER (2003).

linear term with five segments is fitted for all Swiss regions with an additional term for large apartments in tourist regions (for the results see Appendix D).

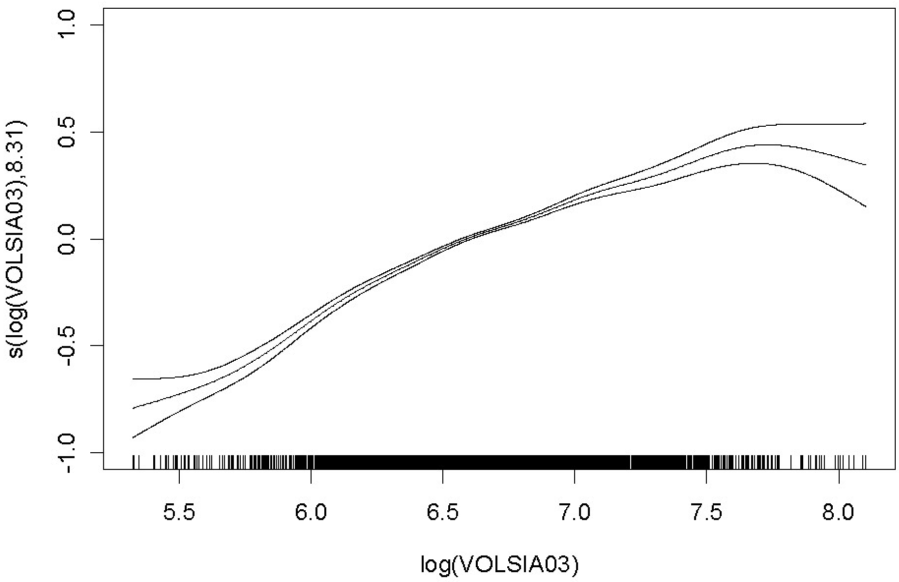
Figure 1: Non-parametric estimation of the coefficient for ln(*nwf*)



For single family houses the assumption of log-linearity between price and size does not hold. Although both ends of the distribution of the cubic content the available data is rather weak and unsure, log-linearity must be rejected but can be replaced by a partwise log-linear term.[20]

[20] For very large objects it is not clear if an observation represents a very large single family house or a refurbished old farmhouse with additional barns and sheds.

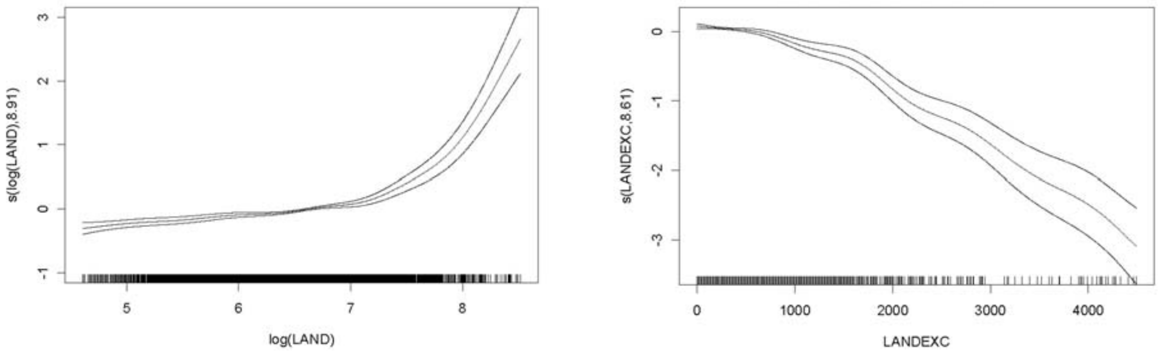
Figure 2: Non-parametric estimation of the coefficient for ln(*volsia03*)



According to construction laws a standard sized house can usually be constructed on a lot of approximately 400 m^2 by which many of the seized properties are based on too large a site area. Due to colinearity of the cubic content and the average lot size, in the simple model, problems may occur when predicting prices of objects with a site area much larger than average according to construction laws. Using approximations the non-necessary "excessive" site area is separated from the site area and both variables are used for the fit.[21] Up to approx. 1'000 m^2 log-linearity holds quite well for the site area but for larger sites this assumption in combination with a variable for the excessive site area does not hold anymore. For the estimated excessive site area, the demonstrated semi-logarithmic fit shows non-linearity which can partly be adjusted with a log-linear model but as well as for the site area itself, for the excessive site area a partwise log-linear term allows a more accurate fit.

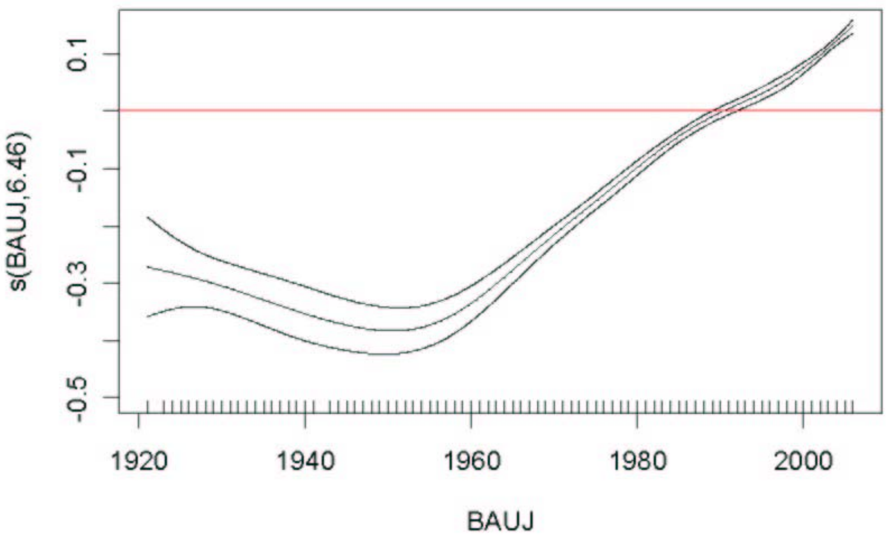
[21] These approximations do not allow an estimation of the necessary site area per squaremeter but should allow a reasonable approximation which may correspond with the market reality equivalent to an exact calculation.

Figure 3: Non-parametric estimation of the coefficient for ln(*land*) and *landexc*



Using refurbishment levels instead of the factor condition the construction year must be interpreted twofold. Firstly, the construction year still reflects the condition of the objects, e.g. a five-year-old condominium would still be assessed with condition 5 but is obviously already occupied and therefore not precisely new anymore. Secondly, the construction year also reflects the building quality and the style of the architecture. Prices generally decrease with the age of a condominium until an age of approximately 40 years (see Fig. 4).[22] The disagio for post-world-war-two properties then again decreases, since during this period objects with other construction quality, architecture, room heights etc. were built. Therefore instead of a log-linear fit like proposed in the starting model, a quadratic term for the construction year fits better especially for objects with construction year before the 1960ies.

Figure 4: Non-parametric estimation of the coefficient for the construction year



Note: Cubic regression spline for condominiums.

---

22 For single family houses comparable results are found.

Using again fixed effects for the estimation of the macro-location, the discussed expansions allow a better fit for non-standard objects for both condominiums and single family houses (see Appendix D and E for detailed results). These improvements can be shown in analysis regressing the residuals on the variables rather than with reference values like standard error or R^2 because a big part of the samples are standard and quasi-standard objects and therefore little affected by the discussed adjustments.

3.3 SPATIAL GENERALIZATION

Since the parameters for the attractiveness of the villages – the macro-location – are so far estimated using fixed effects, two important questions arise:

Firstly, for some of the villages, due to a small number of observations, the fixed effects can be overfitted, since with a small number of observations, the estimated effect does not necessarily correspond with the true mean for this village.[23] Therefore the estimated parameters for the macro-location have to be verified.

Secondly, only of about 730 (condominiums) respectively 1'000 (single family houses) fixed effects can be estimated, because for the other villages, none or too few observations are available. Thus a model for the estimation of the macro-location of these villages becomes necessary.[24]

In literature, little can be found concerning the questions mentioned above, since most of the studies cover a limited area or region and mostly fixed effects are used for the estimation of the macro-location. But for Switzerland, already several models exist:

- The Swiss consultancy Wüest & Partner uses widely modelled and generalized m^2-prices for both single family houses and condominiums covering all Swiss communities.[25]
- In the original models of the Zürcher Kantonalbank for the Canton of Zurich, the macro-location is modelled using the Cantonal planning regions as well as the commuter time by car to the City of Zurich (see BIGNASCA ET AL. 1996, p. 23ff.). In a later study, the income

23 See for example STAHEL (2002), p. 152ff..

24 Out of the 2'910 villages, these numbers equal only about one quarter respectively one third of the possible macro-locations. On the other hand, the villages with observations are doubtlessly those, where a market occurs at all. These regions also show the largest proportion of the housing stock as well as of the population.

25 They are modelled quarterly, based on advertisements in newspapers and verified using transaction data where available.

tax burden, the average taxable per capita income and, for communities near the City of Winterthur, the driving time to Winterthur are added to the model for the macro-location (see SALVI, SCHELLENBAUER and SCHMIDT 2004, p. 38ff.).

- SCOGNAMIGLIO (2000) generally distinguishes general, geographic, macroeconomic and socioeconomic variables are necessary to model the macro-location. In total this author uses 50 indicators to model the macro-location for each Swiss political community. These indicators are not published but three examples such as weighted distances to the next centres, number of overnight stays in hotels and the average taxable income per capita are declared.

A simple and logical model is employed to explain the unknown level of the prices of a community with no observations using the estimated levels of its next geographical neighbours, e.g. using distances. The idea is to assume a spatial correlation between the prices of private properties, since communities compete with each other.

But for two reasons, in Switzerland geographical distances do not provide a sound measure to determine the next neighbours:

Firstly, geographical coordinates do not consider the topography of Switzerland. For that reason, in a model with distances prices for example south of the main ridge of the Swiss Alps, would influence the prices on the northern side and vice versa. The same applies for lakes and other natural barriers.

Secondly, and this is partly influenced by the topographical structure as well, the spatial correlation is from a theoretical point of view rather based on attainability i.e. driving times than on Euclidian distances. The often curved roads and steep ascents may slow down the average speed on one road and therefore sometimes shorter distances result in longer driving times than on longer routes. In addition geographical coordinates neglect the course of roads and major motorways.

The solution is a system of artificial coordinates based on driving times between the 2'910 Swiss villages but the problem is that this system cannot be fitted in a two-dimensional system of coordinates.[26] For three villages, two-dimensional coordinates can be calculated by

---

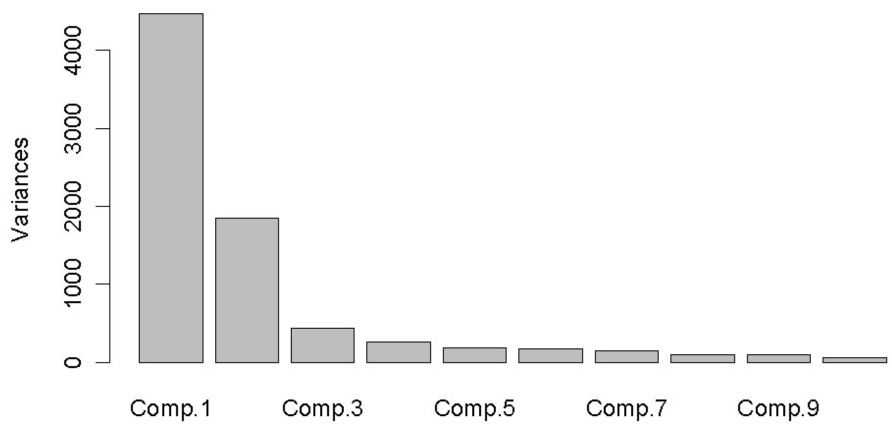
[26] The driving times are a weighted average of individual and public transport, under load, e.g. under consideration of average speeds, traffic jams, walking times to bus stops and train stations, waiting and changing times. The original matrixes of the driving times have been provided by the ETH Zurich. See VRTIC ET. AL. (2005) and FRÖHLICH and AXHAUSEN (2004).

using the driving times between the villages. Adding a fourth village, usually a third axis becomes necessary and for n = 2’910 villages the extreme case is n-1 = 2’909 axis.

In praxi, such a big number of dimensions would not be applicable and the question arises how many dimensions p would allow a sufficient representation of the driving times. Using classical, metric multidimensional scaling (MDS), such a system of coordinates can be modelled (see Appendix C).[27]

As the Screeplot of the principal component analysis of the principal coordinates of the MDS shows (Fig. 5), the driving times between the Swiss villages could be represented in a system of artificial coordinates with p=2, since the other principal coordinates explain only a small proportion of the variance.

Figure 5: Screeplot of the first ten principal components of the MDS

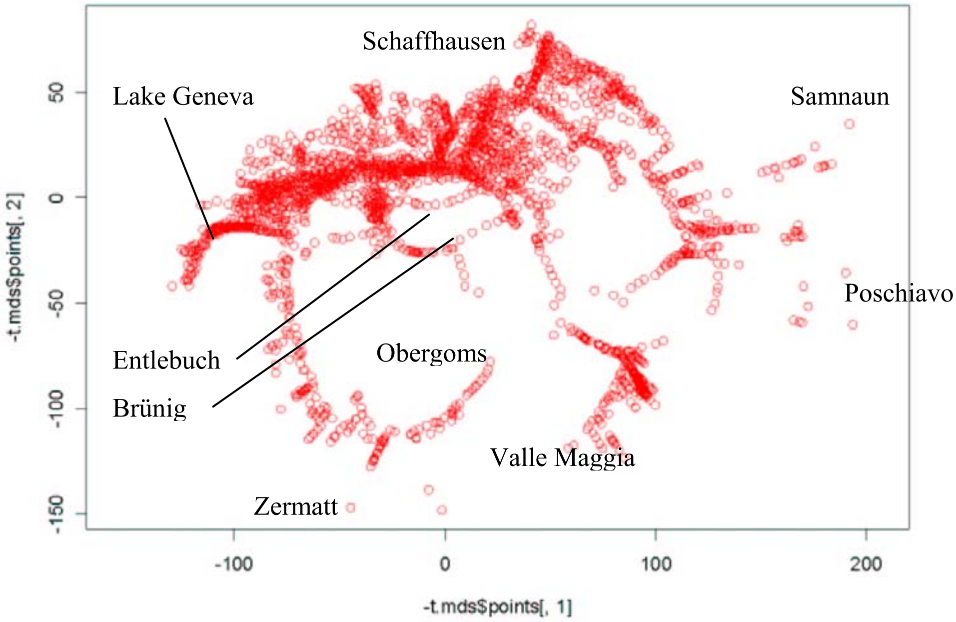


The plot of the villages in the 1st and 2nd principal coordinates of the MDS does not show a completely different but unusual picture of Switzerland (see Fig. 6). The lower mainland north of the Alps with a dense network of highways can clearly be distinguished from the rather remote mountain valleys and the urban area in Southern Switzerland. Over all the 1st and 2nd principal coordinates distinguish between the urban and rather rural regions of the country.

---

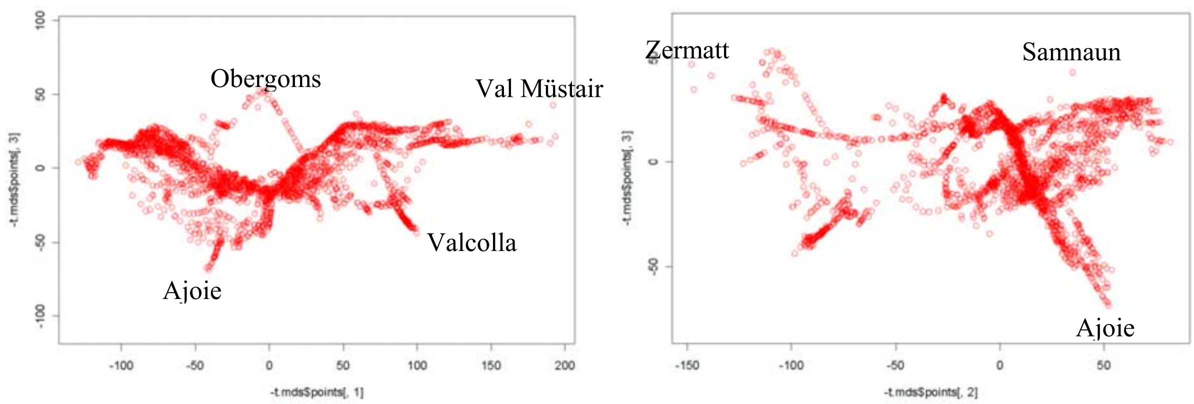
27 See for example COX and COX (1994) and MÄCHLER (2004).

Figure 6: Plot of the 1st and 2nd principal coordinates of the MDS



Using the 3rd principal coordinate, there is not much change for the mayor part of the villages in the lower mainland and along the main traffic axis but there is an additional differentiation of the rural and mountaineous regions (Fig. 7).

Figure 7: 1st and 3rd as well as 2nd and 3rd principal coordinates of the MDS



Using three principal coordinates, a three-dimensional system of driving-times based artificial coordinates is constructed, explaining approximately 90 percent of the variance. This system of artificial coordinates is used to model the macro-locations of the Swiss villages.

Since the available macro-locations from the models with fixed effects are generally trusted, a three dimensional thin plate regression spline allowing high variability is used in a generalized additive model as well as other indicators of the attractiveness of the villages such as the income tax burden[28] as well as special dummies for cities and the two "Golden Shores" at the Lake Geneva and the right side of the Lake of Zürich.[29]

Since the attainability of the city centres and the next neighbours are from a theoretical point of view the main predictors for the level of prices, the model is mainly based on the regression spline for the driving-times based artificial coordinates. In the rather urban regions the degree of urbanity plays a role especially due to the reachability of the working places, better public transport etc.. Although the effect of tax burden on real estate prices itself is not clear, the tax burden can serve as an indicator of the purchasing power of the population and, assuming a positive correlation between purchasing power and real estate prices, therefore also as regressor for the macro-location. Since for habitation soft factors like view, sun, image etc. play an important role, different spatial indicators are tested. It can be shown, that especially the villages along Lake Geneva and along the sunny side of Lake Zürich attract people willing to pay more for a condominium or single family houses than in neighbouring villages, what probably can be understood as an interaction between low taxes, great view and proximity to the Cities of Geneva, Lausanne and Zurich.

While the thin plate regression spline itself explains already about 80% of the variability of the macro-locations the additional regressors add another 3% of explained variability to this still rather simple model for the macro-location of condominiums. The model for single family houses is slightly weaker but also good (see Tab. 2). Comparing the predicted macro-locations and the macro-locations from the models with fixed effects, it can be shown, that the prediction is accurate for the mayor part of the villages except for some outlayers especially in the model for single family houses.

---

28 The tax burden is in accordance with „Tukey's first aid transformations" arcus sinus transformed. See for example STAHEL (2002, p. 278). The tax burden on the communal level has been calculated by Tribut and been provided for this study by Credit Suisse. See CREDIT SUISSE AND TRIBUT (2005).

29 For Generalized Additive Models and thin plate regression splines see for example HASTIE and TIBSHIRANI (1990) or HASTIE, TIBSHIRANI and FRIEDMAN (2001). The name thin plate spline refers to a physical analogy involving the bending of a thin sheet of metal and is the two-dimensional analog of the cubic spline in one dimension.

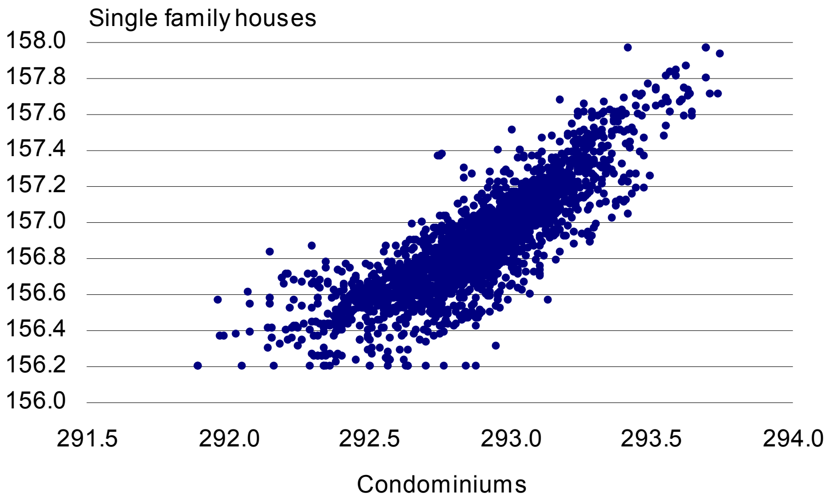
Table 2: Semi-parametric fit of the estimated macro-locations

	Condominiums			Single family houses		
	Coefficient	Standard error	Sign. level	Coefficient	Standard error	Sign. level
Intercept	5.684	0.000	***	5.063	0.001	***
TAX	-0.011	0.001	***	-0.018	0.003	***
CITY	0.191	0.045	***	0.338	0.104	***
ZSEERE	0.726	0.133	***	1.274	0.333	***
GESEE	0.327	0.084	***	0.520	0.192	**
Spline (χ^2)	2'320		***	2'329		***
E.D.F.	234			202		
S.E. residuals	0.000			0.000		
R^2_{korr}	0.832			0.774		

Note: ** represents a 1% and *** a 1‰ significance level. The indicator variables CITY, ZSEERE and GESEE have been divided by 1'000 to increase the readability of the coefficients and standard errors.

Using these models, complete macro-locations are predicted for all the 2'910 villages for both condominiums and single family houses. For villages with many observations, the final macro-location is taken from the model with fixed effects, for those with some observations a weighted average of the level from the model with fixed effects and the prediction is used and for villages with no or little observations only the predicted macro-location is used. The comparison of the macro-locations of condominiums and single family houses shows, that there is, like theoretically expected, a strong Pearson-correlation (0.85) between the submarkets (see Fig. 8).

Figure 8: Scatterplot of the macro-locations for condominiums and single family houses



3.4 RE-ESTIMATION AND OUT-OF-SAMPLE ANALYSIS

The modelling of the macro-locations allows re-estimations of the models using the predicted macro-locations instead of fixed effects. Since in the models with fixed effects, levels are only estimated for villages with a reasonable number of observations, in the re-estimation for condominiums with the predicted macro-locations the sample size is increased by 705 observations.

The coefficients in both models are comparable and do not change significantly (see Appendix D for detailed results). In the re-estimation using predicted macro-locations, the coefficient for the macro-location now is not exactly 1 but not significantly different from 1 with a small standard error. Apparently the predicted macro-locations are quite accurate with few macro-locations being biased. Overall, the prediction of the macro-locations and the re-estimation result in a slight increase of the standard error of the residuals respectively in a slightly lower R^2. The analysis of the residuals shows no systematic regional biases but higher variances of the residuals in regions with thin markets.

Over the total sample of 9'510 observations two thirds of the condominiums can be explained by a difference of ±10% of the transaction price and 95% of the objects with a maximal difference of ±20%. In the subsample of the additional observations with purely predicted macro-locations the respective proportions are 59% and 93%.[30]

In addition to these 9'510 observations, some 1'005 new observations mainly of transactions of the 1st quarter 2005 became available for out-of-sample analysis. Predictions of the market values reveal a robust standard error of 0.135 which is, considering the small sample size and possible changes of the general market price level in the 1st quarter 2005, about as good as the in-sample-results (see Tab. 3).[31]

Table 3: Sample sizes and robust standard error of the samples (condominium)

	n	Standard error
Starting model	8'811	0.123
Observations used for fixed effects and modelling macro-locations (in-sample)	8'811	0.120
Total in-sample	9'510	0.124
Out-of-sample	1'005	0.135

30 Outlayers are excluded.

31 For the prediction of the prices, the transactions of the 1st quarter 2005 are considered as transactions of the 4th quarter 2004.

The re-estimation of the model for single family houses using the predicted macro-locations with 1’159 additional observations also shows comparable results like the model with fixed effects (see Appendix E). In total the re-estimation confirms the results of the fixed effects model but results in an increase of the total variance (0.138 compared with 0.125 in the fixed effects model), especially in the little villages with thin markets. 64% of the single family houses can be explained with an error of at most ±10% and 95% with ±20%. The corresponding proportions of the subsample are 58% and 92%.[32]

For single family houses as well, some 1’054 new observations of transactions mainly of the 1st quarter 2005 became available for out-of-sample analysis. The true transactions prices can be explained with a robust estimated standard error of 0.158 compared with 0.135 in the in-sample-analysis (see Tab. 4).

Table 4: Sample sizes and robust standard error of the samples (single family houses)

	n	Standard error
Starting model	7’598	0.128
Observations used for fixed effects and modelling macro-locations (in-sample)	7’598	0.125
Total in-sample	8’756	0.135
Out-of-sample	1’054	0.158

In the out-of-sample analysis an overestimation of approx. 1% results for both types of properties but it can be shown, that in many cases the compiled transaction price is most likely wrong by the factor 10 or 100. After removing such extreme observations an under-estimation on the portefeuille level of 0.18% (condominiums) respectively an over-estimation of 0.21% (single family houses) result. The analysis of the residuals reveals no structures in particular not concerning the Swiss regions.

32 Outlayers are excluded.

4 CONCLUDING REMARKS

In this article we show, that the assumptions of log-linearity between the transaction prices of private properties and the continuous variables describing size and age of the objects do not hold for non-standard objects. The predicting power of hedonic models can be increased for such properties using non-parametric methods or partwise linear and quadratic terms. In addition, in a nationwide model, regional differences in the elasticities should be considered as well as interaction terms and regional differences of the development of the prices over time. In addition the condition of non-refurbished properties and their age shows a strong colinearity and we therefore propose to consider refurbishment levels instead of the condition itself. Considering that these proposed model expansions only affect only the fit for non-standard objects, the resulting decreases of the standard errors show an improved predicting power for such objects.

Since the available data does not allow the estimation of fixed effects for all 2'910 Swiss villages, a model to predict a level for the macro-location for the remaining villages as well as to avoid overfitting of the fixed effects is necessary. We show that using MDS and principal component analysis a matrix of the driving times between the Swiss villages – instead of geographical coordinates – can be transformed into a system of artificial coordinates with three principal coordinates. Using thin plate splines for these coordinates as well as local tax burden and regional effects, a good prediction of the macro-locations for all Swiss villages is possible for both condominiums and single family houses.

However, if more detailed and especially geo-coded information for the transactions became available the possibilities for nationwide hedonic pricing models would increase again.

APPENDIX

A: VARIABLES

Compiled variables – overview

	Condominiums (Median)	Single family houses (Median)
Date of acquisition: *yearqu*	2nd quarter 2004	2nd quarter 2004
Total sale price (in CHF, arm's length transactions)	493'000	670'000
Cubic content (in m^3 SIA 416)[33]: *volsia03*	n.a.	748
Site area (in m^2): *land*	n.a.	599
Dwelling area (in m^2 SIA 416): *nwf*	110	n.a.
Construction year: *bauj*	1997	1984
Condition (factor with 5 levels): *cond*[34]	5	5
Standard (factor with 5 levels): *stand*	3	3
Detached / attached (indicator variable): *att*	n.a.	detached
Micro-location (factor with 5 levels): *micro*	4	4
Number of underground parking spaces	1	1
Main residence / holiday home (indicator): *holi*	main	main
Other information (elimination criteria)	-	-

Source: DATAPOOL. For details see WÜEST & PARTNER (2002).

Generated variables – overview

	Condominiums (Median)	Single family houses (Median)
Adjusted sale price (in CHF): *aprice*	488'000	662'000
Excessive site area (in m^2): *landexc*	n.a.	180
Refurbishment level (5 levels): *refurb*	0	0
Macro-location: *macro*	-	-

33 See SIA (2003).

34 The levels are 1=requiring refurbishment, 2=in disrepair, 3=sound condition, 4=good condition, 5=mint condition. For the other factors, 1 represents the lowest level and 5 the highest possible level. The Banque Cantonale Vaudoise as well as the Zürcher Kantonalbank use partly other indicators and therefore some of their input data have to be transformed.

B: ROBUST STATISTICS

In the ordinary least squares method (OLS) usually extreme outlayers are identified and removed from the sample what can already be considered as robust method (see Rousseeuw and Leroy (2003), p. viii). In robust regression, fitting is done by iterated re-weighted least squares using ψ-functions to control the influence of extreme observations.

The influence of an observation x on the estimator $\hat{\theta}$ in the univariate case is measured by the sensitivity curve

$$SC\left\langle x_1, x_2, \ldots x_{n-1}, \hat{\theta} \right\rangle = \frac{\hat{\theta}(x_1, x_2, \ldots x_{n-1}, x) - \hat{\theta}(x_1, x_2, \ldots x_{n-1})}{1/n} \tag{2}$$

which for large n corresponds to the influence function (IF). The sensitivity γ*,

$$\gamma^* := \sup_x IF\left(x; \hat{\theta}; F\right), \tag{3}$$

is the maximal influence of the observation x on the estimator $\hat{\theta}$ given a distribution *F*. For that reason the mean and the standard deviation are not robust estimators since for x → ∞, the IF as well as the γ^* are infinite. Hence the breakdown point φ_n^*,

$$\varphi_n^*\left(\hat{\theta}; x_1, x_2, \ldots, x_n\right) = \frac{m}{n}, \tag{4}$$

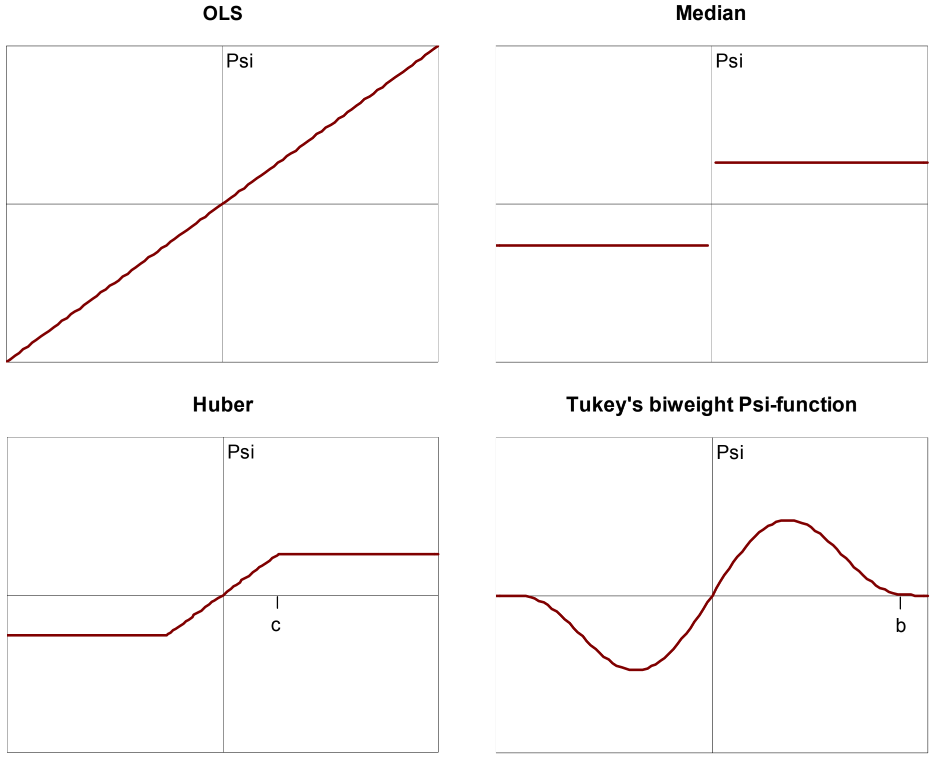
with m=maximal number of extreme observations is 0 for the mean but approx. 0.5 for example for the median. For the multivariate case, the median is not usable and therefore other methods are suggested.

Under certain conditions the p normal equations for the OLS can be written as

$$\sum_{i=1}^{n} \psi\left(\frac{y_i - \sum_{i=1}^{n} x_{ij}\beta_j}{\sigma}\right) x_{ij} = 0 \text{, with j = 1, 2, …,p} \tag{5}$$

and the result is still a maximum-likelihood estimation. Each ψ-function, fulfilling the above equation, is called M-estimator (see Fig. 9).

Figure 9: Psi-functions of typical M-estimators



Source: Ruckstuhl (2004b, p. 9).

M-estimators are asymtotically normal distributed with covariance matrix $\sigma^2\tau C^{-1}$ with $C=1/nX^TX$, what up to the correction term τ (>1) corresponds with the covariance matrix of the OLS. The covariance matrix $\hat{V}$ is estimated as

$$\hat{V} = (\hat{\sigma}^2/n)\hat{\tau}\hat{C}^{-1}, \tag{6}$$

with

$$\hat{C} = \frac{1/n\sum_{i=1}^{n}\omega_i \underline{x}_i \underline{x}_i^T}{1/n\sum_{i=1}^{n}\omega_i}, \text{ with weights } \omega_i . \tag{7}$$

Besides the choice of a ψ-function an estimator for the scale parameter σ is necessary. Since in outlayers are not removed from the sample but weighted, the s_{MAV}

$$s_{MAV} = median_i(|\varepsilon_i|)/0.6745 \tag{8}$$

is recommended as asymptotic scale parameter. The correction 1/0.6745 leads to a consistent estimation of σ for exactly normal distributed residuals (SEE RUCKSTUHL 2004, p. 9).

For the R^2 ROUSSEEUW AND LEROY (2003, p. 44ff) recommend

$$R^2 = 1 - \left(\frac{median|\varepsilon_i|}{\underset{i}{median}\left(\left| y_i - \underset{j}{median}(y_j) \right| \right)} \right)^2 . \tag{9}$$

In the multivariate case, M-estimators like Huber have a breakdown point of 1/p, e.g. with 7 explanatory variables about 14 percent. Another disadvantage of M-estimators is the fact, that they are usually able to identify outlayers in y-direction but usually fail with leverage points (outlayers in x-direction).

For this reason modified M-estimators (MM-estimators) using a redecending ψ-function like Tukey's biweight ψ-function of the form

$$\psi_{b_1}(u) := 1 - (1 - (\frac{u}{b_1})^2)^3 \text{, if } |u| < b_1 \text{, else 1,} \tag{10}$$

with $b_1 = 4.685$ were proposed (YOHAI, STAHEL and ZAMAR 1991, in RUCKSTUHL 2004b, p. 20, see Figure 9). The MM-estimator combines the advantages of other proposed estimators without having their disadvantages. In addition the MM-estimator has a breakdown point φ_n^* of 0.5. This estimator works with random resampling and does therefore not always produce exactly the same results. The result of a MM-estimation is again a M-estimation with the above described properties.

In this paper, generally M-estimators with Huber's ψ-function are used since in fixed effects models the resampling algorithm of the MM-estimator often leads to singularities. Where possible the MM-estimator is used to confirm the robustness of the Huber M-estimator.

C: METRIC MULTIDIMENSIONAL SCALING

The idea of metric MDS is to measure the dissimilarities out of the data and to plot them into a system of p coordinates (see for example COX and COX, 1994; BACKHAUS ET AL.,2000 or MÄCHLER 2004).

The values of the symmetric matrix of the driving times D with 2'910 rows and columns can be interpreted as approximate Euclidian distances

$$d_{ij} = \sqrt{\sum_{k=1}^{p} \left(x_{ik} - x_{jk}\right)^2} \tag{11}$$

between each pair of villages i j, where x_{ik} is the coordinate of the village i in the k^{th} principal coordinate.

With $a_{ij} = -\frac{1}{2} d_{ij}^2$ the matrix D is transformed into matrix A. This matrix A is transformed into a matrix B by centering with its means of rows and columns. Through Eigenvector- and Eigenvalue-decomposition of B

$$B = U \Lambda U = XX^T , \tag{12}$$

with Λ = vector of the Eigenvalues, U = centering matrix and X = matrix of the coordinates, at what $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ and $X = U\Lambda^{\frac{1}{2}}$, the coordinates X can be calculated. In the case of the geographical distances, the d_{ij} represent exact Euclidian distances and the two-dimensional system of coordinates can be found. In this case the first two Eigenvalues are positive and the other n-2 Eigenvalues equal 0.

Using driving times as distances d_{ij} this is not the case and more than two Eigenvalues are positive.

D: ESTIMATION OUTPUT FOR CONDOMINIUMS (robust estimation)

	Fixed effects			**With predicted macro-locations**		
	Coefficient	Standard error	Sign. level	Coefficient	Standard error	Sign. level
Intercept				13.530	18.912	n.s.
macro[35]				0.996	0.006	***
dnwf50[36]	-1.657	0.152	***	-1.576	0.153	***
lnwf50	1.170	0.033	***	1.149	0.033	***
dnwf5075	-0.855	0.161	***	-0.814	0.161	***
lnwf5075	0.961	0.032	***	0.951	0.032	***
dnwf7500	-0.463	0.188	**	-0.382	0.188	*
lnwf7500	0.874	0.037	***	0.855	0.037	***
lnwf0150	0.775	0.019	***	0.775	0.019	***
dnwf150p	-0.511	0.168	**	-0.411	0.167	*
lnwf150p	0.879	0.027	***	0.858	0.027	***
dnwftou3	-1.415	0.302	***	-1.596	0.309	***
lnwftou3	0.296	0.061	***	0.333	0.062	***
bauj	-0.295	0.019	***	-0.307	0.019	***
$bauj^2$	0.000	0.000	***	0.000	0.000	***
refurb: 1	0.052	0.006	***	0.056	0.006	***
refurb: 2	0.077	0.008	***	0.082	0.008	***
refurb: 3	0.146	0.011	***	0.153	0.011	***
refurb: 4	0.249	0.019	***	0.251	0.019	***
stand: 3	0.043	0.012	***	0.047	0.012	***
stand: 4	0.176	0.012	***	0.177	0.012	***
stand: 5	0.255	0.018	***	0.254	0.018	***
micro: 3	0.037	0.010	***	0.033	0.010	***
micro: 4	0.149	0.010	***	0.147	0.010	***
micro: 5	0.248	0.015	***	0.241	0.015	***
stand5&micro4	0.040	0.016	**	0.045	0.016	**
stand4&micro5	0.050	0.013	***	0.058	0.012	***
stand5&micro5	0.148	0.020	***	0.158	0.020	***
2^{nd} quarter	0.022	0.006	***	0.024	0.006	***
3^{rd} quarter	0.001	0.005	.	0.008	0.005	.
4^{th} quarter	0.019	0.004	***	0.020	0.004	***
2^{nd} quarter lem[37]	-0.011	0.001	n.s.	-0.011	0.001	n.s.
3^{rd} quarter lem	0.030	0.008	***	0.036	0.008	***
4^{th} quarter lem	0.021	0.007	***	0.020	0.007	***
2^{nd} quarter tou[38]	-0.003	0.017	n.s.	0.002	0.017	n.s.
3^{rd} quarter tou	0.075	0.012	***	0.076	0.013	***
4^{th} quarter tou	0.020	0.001	*	0.017	0.001	.
holi	0.085	0.007	***	0.084	0.007	***
S.E. residuals	0.120			0.124		
D.F. residuals	8'774			9'472		
R^2	0.973			0.971		

Note: "." represents a significance level of 10%, * a 5%, ** a 1% and *** a 1‰ significance level.

---

35 Since the macro-location is a generated variable, the calculated standard error should be adjusted to calculate the correct T-value.

36 Linear term for ln(nwf) with dnwf50 (intercept for dwellings below $50m^2$) and lnwf50 (slope)

37 Dummy for transactions in the metropolitan area Genève-Lausanne, 2^{nd} quarter 2004.

38 Dummy for transactions in tourist areas, 2^{nd} quarter 2004.

E: Estimation output for single family houses (robust estimation)

	Fixed effects			With predicted macro-locations		
	Coefficient	Standard error	Sign. level	Coefficient	Standard error	Sign. level
Intercept				-4.536	13.899	n.s.
macro				0.996	0.007	***
DVOL600	-1.274	0.140	***	-1.213	0.137	***
LVOL600	0.736	0.020	***	0.732	0.019	***
LVOL1300	0.536	0.012	***	0.541	0.012	***
DVOL2300	0.638	0.346	.	0.767	0.352	*
LVOL2300	0.453	0.047	***	0.439	0.048	***
DVOL23P	13.760	0.243	***	4.178	0.091	***
LVOL23P[39]	-1.196	0.243	***			
LLA2000	0.093	0.009	***	0.085	0.009	***
DLA2000P	-11.728	0.712	***	-8.645	0.697	***
LLA2000P	1.634	0.095	***	1.220	0.093	***
LLAE250	0.007	0.002	***	0.009	0.002	***
DLAE600	-0.096	0.078	n.s.	-0.101	0.080	n.s.
LLAE600	0.024	0.014	.	0.026	0.014	*
DLAE1800	0.389	0.116	***	0.373	0.118	***
LLAE1800	-0.049	0.018	**	-0.045	0.018	**
DLAE18P	11.831	0.649	***	9.464	0.648	***
LLAE18P	-1.605	0.087	***	-1.275	0.086	***
bauj	-0.156	0.014	***	-0.151	0.014	***
$bauj^2$	0.000	0.000	***	0.000	0.000	***
refurb: 1	0.037	0.006	***	0.038	0.006	***
refurb: 2	0.044	0.008	***	0.042	0.008	***
refurb: 3	0.106	0.009	***	0.110	0.009	***
refurb: 4	0.039	0.015	***	0.033	0.016	*
stand: 3	0.128	0.009	***	0.140	0.009	***
stand: 4	0.245	0.010	***	0.260	0.009	***
stand: 5	0.286	0.012	***	0.302	0.013	***
micro: 3	0.097	0.010	***	0.095	0.010	***
micro: 4	0.216	0.010	***	0.219	0.010	***
micro: 5	0.347	0.011	***	0.357	0.011	***
stand5&micro5	0.135	0.015	***	0.150	0.016	***
2nd quarter	0.009	0.005	.	0.012	0.006	*
3rd quarter	0.013	0.005	**	0.011	0.005	**
4th quarter	0.011	0.004	**	0.009	0.004	**
2nd quarter lem	-0.004	0.011	n.s.	-0.001	0.011	n.s.
3rd quarter lem	0.040	0.010	***	0.045	0.010	***
4th quarter lem	0.023	0.009	***	0.031	0.009	***
2nd quarter tou	-0.065	0.026	***	-0.105	0.026	***
3rd quarter tou	-0.005	0.022	n.s.	-0.034	0.021	n.s.
4th quarter tou	0.066	0.020	***	0.006	0.019	n.s.
att	-0.045	0.004	***	-0.043	0.004	***
holi	0.084	0.015	***	0.089	0.014	***
S.E. residuals	0.125			0.135		
D.F. residuals	7'555			8'714		
R^2	0.954			0.950		

Note: "." represents a significance level of 10%, * a 5%, ** a 1% and *** a 1‰ significance level.

[39] Since the segment of single family houses with a cubic content over 2'300m³ is based on a very few observations consisting of large villas as well as old farmhouses the slope for this segment is put to zero.

REFERENCES

BACKHAUS, K. ET AL. (2000), *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*, Berlin etc..

BENDER, A. R., B. GACEM and M. HOESLI (1994), "Construction d'indices immobiliers selon l'approche hédoniste", in: *Finanzmarkt und Portfoliomanagement*, vol. 8, no. 4, p. 522 – 533.

BIGNASCA, F. ET AL. (1996), *Immobilienmarkt Zürich: Immobilienpreise und Bauinvestitionen unter der Lupe*, Ed.: Zürcher Kantonalbank, Zürich.

CLAPP, J. M. (2004), "A Semiparametric Method for Estimating Local House Price Indices", in: Real Estate Economics, vol. 32, no. 1, p. 127 – 160.

COX, T. F. and M. A. A. COX (1994), *Multidimensional Scaling*, Monographs on Statistics and Applied Probability 59, London.

CREDIT SUISSE and TRIBUT (2005), *Steuerbelastung 2004 nach Gemeinden* (mimeo).

EKELAND, I., J. P. HECKMAN and L. P. NESHEIM (2003), *Identification and estimation of hedonic models*, NBER Working Paper 9910, Cambridge.

FAHRLÄNDER, S. (2001a), *Hedonische Handänderungspreisindizes: Konzeption und Methodik*, Ed.: Wüest& Partner, Zürich (mimeo).

FAHRLÄNDER, S. (2001b), "Performance von Wohneigentum", in: *Immo-Monitoring 2002, Band I, Wohnungsmarkt*, Ed.: Wüest & Partner, Zürich.

FAHRLÄNDER, S. and U. HAUSMANN (2001), "Ronco ist nicht Zug, Zug ist nicht Genf: Grosse Unterschiede in der regionalen Wertentwicklung von Wohneigentum", in: Neue Zürcher Zeitung, no. 151, Zürich.

FRÖHLICH, P. and K. W. AXHAUSEN (2004), "Sensitivity of accessibility measurements to the underlying transport network model", Arbeitsberichte Verkehrs- und Raumplanung der ETH Zürich, no. 245.

GREENE, W. H. (1997), *Econometric Analysis*, third edition, London etc..

HAMPEL, F. R. ET AL. (1986), *Robust Statistics: The approach Based on Influence Functions*, New York.

HASTIE, T. and R. TIBSHIRANI (1990), *Generalized Additive Models*, London.

HASTIE, T., R. TIBSHIRANI and J. FRIEDMAN (2001), *The Elements of Statistical Learning*, New York.

HOESLI, M., P. FAVARGER and C. GIACOTTO (1997), "Real Estate Price Indices and Performance", in: *Swiss Journal of Economics and Statistics*, vol. 133, p. 29 – 48.

KINOSHITA, T. (1987), "Working Hours and Hedonic Wages in the Market Equilibrium", in: *Journal of Political Economy*, vol. 95, no. 6, p. 1262 – 1277.

MÄCHLER, M. (2003), "Nichtparametrische Regression": Skript und Übungen zum Block RG-2e des Nachdiplomkurses in angewandter Statistik 7 der ETH Zürich, Ed.: Seminar für Statistik der ETH Zürich, Zürich.

MÄCHLER, M. (2004), "Multivariate Statistik II – Explorative Methoden": Skript und Übungen zum Block MU-2b des Nachdiplomkurses in angewandter Statistik 8 der ETH Zürich, Ed.: Seminar für Statistik der ETH Zürich, Zürich.

MALPEZZI S. (2002), "Hedonic Pricing Models: A Selective and Applied Review", prepared for *Housing Economics*: Essays in Honor of Duncan Maclennan, http://www.bus.wisc.edu/realestate.

MAURER, R., M. PITZER and S. SEBASTIAN (2000), "Konstruktion transaktionsbasierter Immobilienpreisindizes: Theoretische Grundlagen und empirische Umsetzung für den Wohnungsmarkt in Paris", http://www.finance.uni-frankfurt.de.

MONTGOMERY, D. C., E. A. PECK and G. G. VINING (2001), *Introduction to Linear Regression Analysis*, New York.

MURRAY, J. and N. SARANTIS (1999), "Price-Quality Relations and Hedonic Price Indexes for Cars in the United Kingdom", in: *International Journal of the Economics of Business*, vol. 6, no. 1, p. 5 – 27.

NELSON, J. (2003), "Meta-Analysis of Airport Noise and Hedonic Property Values: Problems and Prospects", Forthcoming in: *Journal of Transport Economics and Policy*, http://www.stat.lsa.umich.edu.

R DEVELOPMENT CORE TEAM (2004), "R: A language and environment for statistical computing". *R Foundation for Statistical Computing*, http://www.R-project.org.

ROSEN, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", in: *Journal of Political Economy*, vol. 82, p. 34 – 55.

ROUSSEEUW, P. J. and A. M. LEROY (2003), *Robust Regression & Outlier Detection*, New York.

RUCKSTUHL, A. (2004), "Einführung in die robusten Schätzmethoden": Skript und Übungen zum Block RG-2d des Nachdiplomkurses in angewandter Statistik 8 der ETH Zürich, Ed.: Seminar für Statistik der ETH Zürich, Zürich.

SALVI, M., P. SCHELLENBAUER and H. SCHMIDT (2004), *Preise, Mieten und Renditen: Der Immobilienmarkt transparent gemacht*, Ed.: Zürcher Kantonalbank, Zürich.

SCOGNAMIGLIO, D. (2000), *Methoden zur Immobilienbewertung im Vergleich: Eine empirische Untersuchung für die Schweiz*, Dissertation, Bern.

SIA (2003), *Schweizer Norm 504 416: Flächen und Volumen von Gebäuden*, Zürich.

SIRMANS, G. S., D. A. MACPHERSON and E. N. ZIETZ (2005), "The composition of Hedonic Pricing Models", in: *Journal of Real Estate Literature*, vol. 13, no. 1, p. 3 – 43.

STAHEL, W. A. (2002), *Statistische Datenanalyse: Eine Einführung für Naturwissenschaftler*, 4. Auflage, Braunschweig, Wiesbaden.

YOHAI, V., W. A. STAHEL and R. H. ZAMAR (1991), "A procedure for robust estimation and inference in linear regression", in: Directions in Robust Statistics and Diagnostics, Part II, vol. 34, Ed.: Stahel W. and S. Weisberg, *IMA Volums in Mathematics and Its Applications*, New York, p. 365 – 374.

VRTIC, M. ET AL. (2005), *Erzeugung neuer Quell-/Zielmatrizen im Personenverkehr*, Bericht an die Bundesämter für Raumentwicklung, für Strassen und Verkehr, IVT ETH Zürich, Emch und Berger and TU Dresden, Zürich.

WALLACE, N. E. (1996), "Hedonic-Based Price Indexes for Housing: Theory, Estimation, and Index Construction", in: FRBSF Economic Review, no. 3.

WÜEST & PARTNER (2002), *Bewertungsfunktionen für Wohneigentum: Hauptdomizil, Ferien- und Zweitwohnungen: Grundlagen/Variablendefinition/Glossar/Ergebnisse*, Zürich.