Johansson, Jörgen; Sandberg, Krister

Conference Paper

Hedonic Prices for Co-operative Flats in the City of Umeå estimated with Spatial Autoregressive GMM

41st Congress of the European Regional Science Association: "European Regional Development Issues in the New Millennium and their Impact on Economic Policy", 29 August - 1 September 2001, Zagreb, Croatia

Provided in Cooperation with:

European Regional Science Association (ERSA)

Estimation of Hedonic Prices for Co-operative Flats in the City of Umeå with Spatial Autoregressive GMM*

Krister Sandberg† and Jörgen Johansson‡
Department of Economics and CERUM
Umeå University
SE-901 87 Umeå, Sweden

Paper presented at the 41st Congress of the
European Regional Science Association in Zagreb, Croatia

June 2001

Abstract

In this paper we estimate hedonic prices for co-operative flats in the city of Umeå, Sweden, during 1998 and 1999. Structural and neighbourhood characteristics together with accessibility measures are used as attributes in the hedonic price function. Since there are indications of spatial dependence Ordinary Least Squares estimation is inappropriate. Instead Spatial Autoregressive GMM estimation is used. Two attractive nodes, although with different functions, are found in the city. Thus there are signs supporting the view that Umeå has developed into a multi-nodal structure for property values.

Keywords: SAR-GMM, Hedonic prices, Co-operative flats, Spatial dependence

JEL Classification: D46, D61, R20, R21

*The authors wish to thank HSB and Riksbyggen in Umeå for providing data and Associated Professor Lars Westin for valuable comments on the paper.

†Krister.Sandberg@econ.umu.se

‡Jorgen.Johansson@econ.umu.se

1. Introduction

The size and attributes of public and private real estate in a city develops in a dynamic process involving actors on both the supply and demand sides. While the city managers, at least in theory, may want to improve the supply in order to maximize the utility of its inhabitants, property owners and constructors may want to develop profitable buildings and sites.

However, if not the supply fits the utility driven demand for housing by households or the profit driven demand for nonresidential space by firms, the value of the property and the attractivity of the city is set under pressure by competition from other cities or other parts of the city. In the current movement of the economy towards a knowledge society this competition for movable labour by supply of "attractivity" has been even more emphasized. Hence, the supply side clearly has an interest in the valuation at the demand side of the attributes associated with its real estate.

The paper thus takes its starting point in the assumption that the attractiveness of a city, of areas within the city, and of individual residential and nonresidential units may be measured through the valuation of the attributes associated with the supply. The scope of the paper is although not to estimate the total value of the city but, as a first step and in order to develop appropriate statistical tools, to determine and quantify the valuation of housing attributes of co-operative flats in the city of Umeå, Sweden. This is made on the basis of hedonic price theory and through use of spatial autoregressive GMM estimation.

The basis for the theory of hedonic prices lies in the assumption that a good, in our case a co-operative flat, should be seen as a bundles of characteristics matching the household's utility function, as formulated by Lancaster (1966)[26]. The knowledge on which these characteristics are and how they are valued in the Swedish real estate market is still quite fragmentized. The Swedish Association of Municipal Housing Companies, SABO[31] in 1997 presented one of the first studies on the subject in a discussing format. In 1999 the Swedish government initiated a study resulting in SOU 2000:33[34] where the importance of distance to the CBD (Central Business District) on the monthly fee for multifamily houses was investigated for eight Swedish municipalities. The results were somewhat mixed, with e.g. a positive "monthly fee gradient" in some municipalities. On the other hand, various consultants offers estimates of attribute values for segments of the market to their clients although generally without a transparent methodological part.

By looking at the market for co-operative flats, this study moves into an area not previously studied in Sweden. It is assumed that the buyer implicitly reveals his preferences and his valuation of the attributes of the flat through the price he pays. Since each flat is purchased by the household with the highest bid it is sometimes then also assumed that the market prices give the outer envelope of the valuation of each attribute by all households in the market.

The theory is applied to the city of Umeå, situated along the Baltic coast in the northern part of Sweden. Umeå has, due to the university, experienced a rapid population growth in the last 50 years. It now has 104 000 inhabitants in the municipality and 137 000 in the city region[1], a medium size city with a rank among the twenty largest. Due to the growth, the city is presently standing at the crossroad between a peripheral expansion and density increasing investments in more central locations.

The largest accumulation of work places in the city region is in the area around the university[2]. The impact of this cluster of working places a bit away from the CBD on the prices for co-operative flats has not been investigated before. Umeå is suited for this kind of study since it is located 100 kilometers from nearest city of major importance, so direct influences from nearby cities ought not to influence the valuation of co-operative flats in some part of the city.

When studying the prices of co-operative flats it is important to distinguish these flats from flats with right of tenancy. A fundamental difference is the kind of ownership they represent. A co-operative flat is coupled with a membership and a share in a housing co-operative, which legally is the owner of the building. This gives the owner of the flat an indirect ownership of the building while a right of tenancy does not represent any ownership at all.

The share connected with the co-operative flat although has an economic value and may be sold on the market. A basic idea, that makes this possible, is that the share is eternal. Once sold, a flat should never be returned to the co-operative. Occasionally, this does anyhow happen when an owner must sell but cannot find a buyer. This is often the case for co-operatives with too high monthly fees in relation to their market attractivity. The effect is that the co-operative loses money and may eventually go bankrupt.

Owners of co-operative flats have more rights and responsibilities than someone who rents by tenure. In a co-operative there are for instance more possibilities

[1] SOU 2000:87 Regionalpolitiska utredningens slutbetänkande

[2] The university area consists not only of the university but also the university hospital and a science park. In all, this is the most important area for working in Umeå.

to change the standard of the flat and alter its appearance. The membership in the co-operative also gives a possibility to influence the maintenance of the building and the level of the monthly fee. This type of self-maintenance may have a positive value to some households and may also make the maintenance cheaper. In a flat rented by tenure, all these responsibilities lies on the owners of the housing company. The real estate tax is paid by the co-operative (thus indirect by the members). Since a co-operative flat is considered as comparable with other assets the holder must additionally pay wealth tax but may also make tax deduction for the interest on their housing mortgages.

When a co-operative is started, different strategies may be chosen to finance it. One alternative is that the co-operative first take all loans and then sells the flats with high monthly fees to its members. Another alternative is to let the members raise the money thereby directly finance the loans by the co-operative. The flats are then instead sold with a relatively high price while the monthly fee is kept at a lower level. The strategies distribute risks and capital costs differently between the co-operative and the individual owners of the flats. Since the co-operative has to be non-profit and may set the monthly fee freely, the cost of living in a co-operative usually also depends on the age of the co-operative. Old co-operatives may have lower costs since most of their loans are amortized already, which induce low fees. This in turn ceteris paribus should result in higher prices on the market. On the other hand, old houses may have lower standard, or need maintenance and renovation, which in turn should reduce the price.

The aspects mentioned above makes up the price paid on the market for the co-operative flats. In the remainder of the paper we identifies some of the aspects behind the price level and how influential they are.

In the next chapter the theory of hedonic prices and previous studies in this field of research are presented. This is followed by the econometric specification of our model. Chapter 3 describes the data for co-operative flats in Umeå and the explanatory variables. The empirical examination is outlined in chapter 4 and this is followed by our conclusions in the final chapter.

2. The theory of hedonic prices

The theory of hedonic prices was originally developed from the need to determine the impact of quality changes in products and consumers willingness to pay for such improvements. Hedonic prices are defined as implicit prices of attributes and are revealed from observed prices on differentiated goods and the specific amounts

of characteristics associated with them, see for example Lancaster (1966)[26]. To our knowledge, Haas[23] performed the first hedonic price study in his master thesis from 1922. He estimated how farm prices depend on the depreciated cost of buildings per acre, the land classification index, the soil productivity index, and the distance to the market in an early multivariable regression. Colwell & Dilmore [17] moreover argues that Haas "is the author of the first modern, empirical urban or regional economics paper". The first author to coin the term "hedonic prices" was Court (1939)[18] in a study of automobile price indices. The technique was "rediscovered" in the early 1960s by Adelman & Griliches (1961)[1], while Rosen (1974)[30] formalized the theory and extended the hedonic price theory to estimation of attribute supply-, and demand functions.

The empirical literature on hedonic prices for single-family houses is numerous and to a large extent, for example, Blomquist et al. (1988)[8] and Sivitanidou (1996)[33] based on American data. Among the non-American exceptions are studies by Wigren (1987)[37], Englund et al. (1998)[19], and Cheshire & Sheppard (1995)[13] using Swedish and British data respectively. A common feature among hedonic price studies is their interest to examine single specific characteristic and its influence on prices for single-family houses. To name but a few, the presence of lake view by Blomquist (1988)[7], the impact of nearby power lines as analyzed by Colwell (1990)[16], and the introduction of casinos as studied by Buck et al. (1991)[11].

Studies based on multifamily houses and co-operative flats are, compared with single-family house studies, few. Two Swedish exceptions are Eriksson (1997)[20] who studied how owners of multifamily houses value attributes of the houses and Werner (2000)[36] who analyzed the role of architectural thinking on co-operative flats and its influence on prices.

The impact on prices of distance to the CBD has been analyzed in a number of Swedish qualitative studies, e.g. Gavlefors & Roos (1992)[21], Lindgren & Rosberg (1992)[28], and Andersson (1998)[3]. The conclusions in these studies are that a high level of service and waterfront location influences the price in a positive way. The impact is however not ranked nor quantified. Others have studied the problem with a focus on quantification, e.g. Archer et al. (1996)[6] and Andersson (1997)[2]. Heikkila et al. (1989)[24] question the concepts of a monocentric impact on the price structure and argue that cities may have polycentric structures and uses Los Angeles as an example.

The concept of implicit or hedonic prices was as mentioned formalized in Rosen (1974)[30]. In order to derive the hedonic price model as it was developed by

Rosen we will here follow Sheppard (1999)[32]. The good considered, in our case co-operative flats, may be described by m characteristics. Each flat is then represented by the vector $\mathbf{z} = (z_1, ..., z_m)$. An element z_i measures the amount of the ith characteristic embedded in each flat. A price function based on this vector of characteristics is the hedonic price function $p(\mathbf{z}) = p(z_1, ..., z_m)$.

The preferences of the household may be represented by the utility function:

$$U = u(\mathbf{z}, \mathbf{y}, \boldsymbol{\alpha}) \tag{1}$$

Above, $\mathbf{z}$ is consumption of a co-operative flat, $\mathbf{y}$ is consumption of a composite good, and $\boldsymbol{\alpha}$ is a vector of parameters that characterize the household preferences. The price a household would be willing to pay for co-operative flats may be derived from the utility function as a function of the embodied characteristics, a given household income M, and an achieved utility level. This gives the household's bid rent function:

$$\gamma(\mathbf{z}, M, U, \boldsymbol{\alpha}) \tag{2}$$

and implicitly:

$$U = u(\mathbf{z}, M - \gamma, \boldsymbol{\alpha}) \tag{3}$$

The derivative of the bid rent function with respect to z_i, $\frac{\partial \gamma}{\partial z_i}$, gives the rate at which the household would be willing to change it's expenditure on a co-operative flat when characteristic i increases, while keeping other levels constant.

Problem (MAX CoF) *The household chooses a co-operative flat with characteristic* $\mathbf{z}$*, and it's consumption of the composite goods* $\mathbf{y}$ *by solving:*

$$\max_{\mathbf{z},\mathbf{y}} \quad u(\mathbf{z}, \mathbf{y}, \boldsymbol{\alpha}) \tag{4}$$
$$\text{s.t. } M \geq p(\mathbf{z}) + \mathbf{y}$$

The equilibirum price on the market, $p(\mathbf{z})$, reflects the market valuation of a flat with a set of attributes given i.e. amortization and interest schemes avilable, the annual fee to the co-operative, expected costs for repair and for improvements of the flat, as well as for the buildings owned by the co-operative etc. for the entire period the household intend to keep the co-operative flat.

The Lagrangian to (MAX CoF) with the Lagrangian parameter ϑ is:

$$L = u(\mathbf{z}, \mathbf{y}, \boldsymbol{\alpha}) + \vartheta \left[p(\mathbf{z}) + \mathbf{y} - M \right] \tag{5}$$

The first order conditions are:

$$\begin{aligned}
\frac{\partial L}{\partial z_i} &= \frac{\partial u}{\partial z_i} + \vartheta \frac{\partial p}{\partial z_i} = 0 \quad ;\forall i \\
\frac{\partial L}{\partial \mathbf{y}} &= \frac{\partial u}{\partial \mathbf{y}} + \vartheta = 0 \\
\frac{\partial L}{\partial \vartheta} &= p(\mathbf{z}) + \mathbf{y} - M = 0
\end{aligned}$$

Rearranging these conditions give:

$$\frac{u_i}{u_y} = p_i \qquad ;\forall i \tag{6}$$

where $u_i = \frac{\partial u}{\partial z_i}$, $u_y = \frac{\partial u}{\partial \mathbf{y}}$, and $p_i = \frac{\partial p}{\partial z_i}$ the hedonic price of characteristic i.

Combination of the first order conditions with the implicit differentiation of (3) yields that the household's optimal choice of a flat is characterized by equality between the slope of the bid rent and the hedonic price with respect to each characteristic:

$$\frac{\partial \gamma}{\partial z_i} = \frac{\partial p}{\partial z_i} = p_i \quad ;\forall i \tag{7}$$

Under the assumption of optimizing behavior, equation (7) indicates that if we are able to estimate the hedonic price for a characteristic, then this observation provides local information about the household's preferences or willingness to pay for the attribute in the vicinity of the observed choice. Hence, this justifies the use of the hedonic price approach in the analyses of the market for co-operative flats when the mix of attributes is developed not far away from the current market situation (and as long as new attributes, commodites etc not are introduced).

The vector $\mathbf{z}$ consists as mentioned of a set of characteristics which subjectively are determined by the household. However, generally the characteristics may be divided into three broader groups. Among hedonic price studies it is common to identify structural ($\mathbf{s}$), neighborhood ($\mathbf{n}$), and accessibility ($\mathbf{a}$) attributes with $\boldsymbol{\beta}$, $\boldsymbol{\eta}$, and $\boldsymbol{\psi}$ as the corresponding parameter vectors. Given this, the hedonic price function of a general regression model may be formulated as,

$$p(\mathbf{z}) = f(\mathbf{s}, \mathbf{n}, \mathbf{a}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\psi}) + \boldsymbol{\varepsilon} \tag{8}$$

Before we move to the estimation of this function, we may however observe that an overall impression with regard to earlier hedonic price studies, is the

lack of consideration on the possibility of spatial dependence, also termed spatial autocorrelation, in the material.

The Swedish statistician Bertil Matérn (1947)[29] acknowledged the existense of spatial variation, or as he called it typological variation, when considering how trees at different locations are associated with different growth rates. In other words, there might be some inherent systematic dependence between observations that cannot be explained by traditional variables. Two different types of spatial dependence are considered in this study. The first, is present when there is correlation in the dependent variable between observations in the space, cf. Anselin (1988)[4] and Can (1992)[12]. The second, arises when the error term of an observation is correlated with the error terms of observations located nearby i.e. lack of stochastic independence between observations. This was drawned to public attention by among others Cliff & Ord (1972)[14] and Bodson & Peeters (1975)[9]. See also Cliff & Ord (1973)[15] for a further discussion of the problem. In time series analysis this problem is well known, and it must be solved. If not, the model is simply not specified correctly and will violate the standard error assumptions under normality of the linear regression model. It gives inefficient estimates which also may be biased. Spatial dependence in the second type is formally incorporated in (8) via an autoregressive error term.

$$p(\mathbf{z}) = f(\mathbf{s}, \mathbf{n}, \mathbf{a}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\psi}) + \boldsymbol{\varepsilon} \tag{9}$$

$$\boldsymbol{\varepsilon} = \lambda \mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\xi}$$

In (9) $\mathbf{W}$, with elements w_{rs} corresponding to observation pair r and s, is the generalized weight matrix, $\mathbf{W}\boldsymbol{\varepsilon}$ is a spatial lag for the error term, λ is the autoregressive coefficient and $\boldsymbol{\xi}$ is a vector of well-behaved error terms $\boldsymbol{\xi} \backsim N(0, \delta^2 I)$. The autoregressive coefficient is usually not known and must therefore be estimated jointly with the regression coefficients. The covariance matrix for this error term has the following form:

$$E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] = \boldsymbol{\Omega} = \delta^2 \left[(\mathbf{I} - \lambda \mathbf{W})'(\mathbf{I} - \lambda \mathbf{W})\right]^{-1} \tag{10}$$

with $\mathbf{I}$ as the unit matrix,

The Moran's I test is commonly used in order to test for the presence of spatial autocorrelation. The test is here defined as

$$I = \frac{N}{S} \frac{\sum_r \sum_s w_{rs}(x_r - \mu)(x_s - \mu)}{\sum_r (x_r - \mu)^2} \tag{11}$$

where N is the number of observations while x_r and x_s are the observed prices for co-operative flats in locations r and s (with mean μ) in the data material. S is a scaling constant givewn by the sum of all weights

$$S = \sum_r \sum_s w_{rs} \tag{12}$$

When using row standardized weights, which is to prefer according tho Anselin (1995)[5], S equals N since the weights of each row adds to one. The test statistic is compared with its theoretical mean, $I = -1/(N-1)$. So, $I \to 0$ as $N \to \infty$. The null hypothesis $H_0 : I = -1/(N-1)$ is tested against the alternative hypothesis $H_a : I \neq -1/(N-1)$. If H_0 is rejected and $I > -1/(N-1)$, this indicates a positive spatial autocorrelation. That is, high values and low values are more spatially clustered than would be assumed purely by chance. For the other event, if H_0 is again rejected but $I < -1/(N-1)$, it indicates negative spatial autocorrelation. Hence observations with high and low prices are mixed together. Obviously the test is quite crude. One apparent drawback is that it to a large extent is determined by the a priori choice of the spatial weight matrix. However this is also a test for how well the weight matrix performs and what kind of relationship that exists.

In the following, to complement the results from the Moran's I tests we also present the G-statistic and G-I* statistic for spatial association[22]. For a chosen critical distance d, they are respectively defined as:

$$G(d) = \sum_r \sum_s w_{rs}(d) x_r x_s / \sum_r \sum_s x_r x_s \tag{13}$$

$$G_r^* = \sum_s w_{rs}(d) x_s / \sum_s x_s \tag{14}$$

We will return to this topic in chapter 4 and test the impact of different weight matrices.

3. Attributes of the realized sale of co-operative flats in Umeå 1998/99

The data set consists of 194 observations on realized sales of co-operative flats in the city of Umeå from late 1998 and 1999. The observations are divided over 11 co-operatives spread across the city. The co-operatives are presented in the city

map in Figure 1 below and indicated by stars. Two major housing firms in Umeå, HSB and Riksbyggen, have provided the data on the individual flats and their location at street level.

Figure 1: The Umeå City Map 1999

The independent variables used in this paper are, as was mentioned above, grouped into three categories based on their structural-, neighborhood- and accessibility characteristics. The low level of aggregation in the sample (individual flats and block location) makes access to data limited and restrict the choice of independent variables. The structural variables available for each flat are:

· The monthly fee

· The floor size

· The number of rooms

Since price, fee, size, and rooms all are scale dependent, the use of price per square meter, p/m^2, instead of the price p, for each flat as the dependent variable

gives a more scale neutral measure. It is also more common in the real estate market to relate objects to their prices per square meter. Price per square meter will also make the Moran's I test presented later in the paper more appropriate since the average size of the flats not is independent of location. The neighborhood attributes for each area used as the independent variables are:

· Population density
· Single house density
· Rate of turnover

Three accessibility measures are moreover used for each flat. They are measured as the accessibility to:

· The CBD
· The university
· Nearest major shopping centre

To enhance the understanding of the data, the city of Umeå, and to give a deeper sense for the problem descriptive statistics are presented in Table 1 below.

Table 1: Descriptive Data for Co-operative Flats in Umeå 1998/99.

VARIABLE	MEASURE	MEAN	ST.DEV	MIN	MAX
$PRICE$ (p)	SEK	145 000	151 889.9	1	1 195 000
$AREA$ (m^2)	m^2	79.5	20.9	17.5	128
FEE	SEK/month	3 800	1 302.2	650	6 800
$POPDENS$	# Inhab/1 000 m^2	4.8	3.9	0.12	9.4
$SHDENS$	# Single.h/1 000 m^2	0.6	0.4	0	1.0
TO	# Sold/stock	19.7	8.9	8	35
CBD	m	4 350	1 147.8	600	5 700
$UNIV$	m	3 300	2 126.1	1 100	8 200
$SHOP$	m	3 550	2 500.7	1 200	9 300

The dependent variable p/m^2 consist of two parts. The first part, $PRICE$, measures the realized sell price and ranges from 1 to 1 195 000 SEK.[3] The fact that the study is based on *realized market prices* gives a benefit compared with other studies that often use stated preference or the advertised sell price, see for example Borukhov et.al. [10] and Werner (2000)[36].

The second part of the dependent variable is the variable $AREA$, which measures the floor size of each flat. The smallest flat in the sample has a floor size

[3]The co-operative flats that are sold for 1 SEK (6 observations) belong to a fairly new co-operative that have high monthly fees. The other co-operative flats sold in this co-operative are also sold at a low price.

of 17.5 m^2 while the largest are 128 m^2. In the estimations, $AREA$ will not be used as an independent structural variables alone but as part of the variable fee per square meter, FEE/m^2. The variable FEE is the monthly fee paid to the co-operative and ranges from 650 up to 6 800 SEK per month.

The neighborhood attribute population density, $POPDENS$, has a mean of 4.8 inhabitants per 1 000 m^2 for the area in which the co-operative is located. The maximum is nine inhabitants per square meter. A figure that indicates that the multifamily houses in Umeå is dominated by four to five floor buildings. The density of single-family houses in the area nearby the co-operative, $SHDENS$, is the second neighborhood attribute. The variable has it's minimum at 0 since one area in the sample are without any single-family houses. Both variables are computed from data provided by the municipality of Umeå and dates back to 1996.

The third neighborhood attribute is the turnover in the co-operatives. In our sample the average turnover, TO, is 19.7% each year. The youngest co-operative has the highest turnover, 35%. This may be explained by the age but also that those flats are relatively cheap to buy, although coupled with high fees. Thus they are most near the flats with right of tenancy in their attributes.

The accessibility in Umeå is quite good in all directions, and a fair approximation for it ought to be the travelled distance by car from each observation to the major nodes of attraction. The distance to CBD, the variable CBD, from the flats are at minimum 600 meters and at most 5 700 meters. The university area is located at the east of the city and the maximum travel distance given in the variable, $UNIV$, to that campus is 8 200 meters. The city of Umeå is graced with two external shopping centres and the average distance to them, given through the variable $SHOP$ is 3 550 meters.

Before we continue to the econometric part of the paper it is appropriate to discuss what sign that may be expected between the independent variables and the depended variable. In Figure 2 the p/m^2 is plotted against FEE/m^2. Each dot represents a single observation while the straight lines are fitted between observations within each co-operative, in order to illustrate the differences between co-operatives.

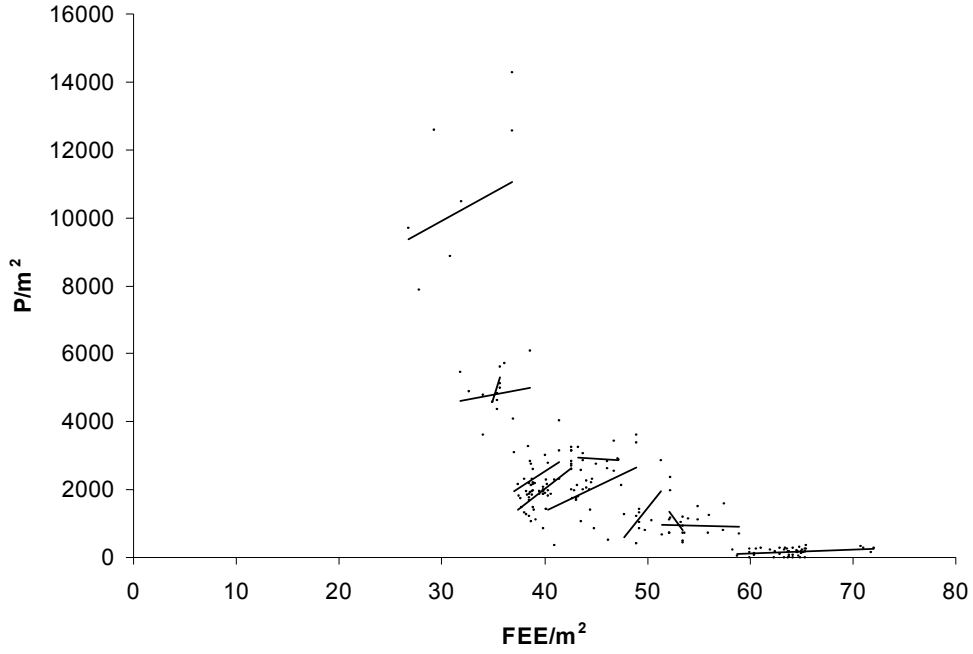

Figure 2: P/m^2 and FEE/m^2 for co-operative flats in Umeå 1998-1999

In the figure, one can see that p/m^2 generally is negatively correlated with FEE/m^2. This illustrates the fact that there is a trade off between price and fee in the co-operative flat market. We can also detect that this relation not is linear. To take this into account in the estimations it was decided to use the inverse of FEE/m^2, that is $(FEE/m^2)^{-1}$, as an independent variable.

For the individual co-operatives the figure indicates that it may be the other way around (the black lines). For these, a higher FEE/m^2 generates higher prices per square meter! A deeper analysis of the data reveals that the positive relation is connected with the size of the sold flats within each co-operative. Clearly there is a scale or market impact so that smaller flats within a co-operative both have higher price and a higher fee per square meter. To take care of this scale and market impact a structural dummy for one room flats $(1 - ROOM)$ was introduced. Smaller flats are more expensive to purchase per square meter than larger ones since some minimum standard, such as a kitchen and plumbing in all flats gives a minimum fee per square meter. Flats with a single room attracts other household groups compared to the other flats i.e. there are different markets for co-operative flats. The sign of the single room flat variable is thus expected to be positive.

The monthly fee that the household pays to the co-operative is moreover in itself determined by factors directly related to information currently not available but important for the price, e.g. the standard of the flat and the maintenance status. The fee thus also is a proxy for such structural characteristics. A dominating factor is that younger flats have higher standard and thus higher fees. Given that the single room dummy is introduced to take care of scale impacts, this leads to the hypothesis that $(FEE/m^2)^{-1}$ the sign for this variable ought to be negative.

The group of neighborhood characteristics consists of the two density variables and the variable for turnover. The idea behind the inclusion of $POPDENS$ is to check whether households finds it attractive to live in a densely populated neighborhood. The city centre is one of those, but there are also other parts of the city that have an equal or higher density, so the variable should not be considered as demand for "city life". This demand is taken care of through the CBD variable explained below. A to high density may also be negative for the attractivity. It is for this reason difficult to have an a priory hypothesis regarding the sign of this variable.

In a similar way $SHDENS$ measure the attractiveness of areas with a high degree of single family houses. This measure may be stronger related with the price and thus have a positive sign. A set of households perceive a mixture of buildings as attractive. Detached houses in the neighborhood may indicate the school quality and the neighborhood income.

Finally the turnover variable TO is defined as the percentage of the flats that has been sold, in each co-operative, during the year 1999. The interpretation of the variable is also twofold. It could be a proxy for the liquidity of the asset. A higher liquidity would accordingly yield a higher price per square meter. The other way to interpret the variable is from the household stability point of view. A stable co-operative with long time owners may put more interest in the maintenance and thereby increase the value. The expectation is here is that low turnover/high stability yields a higher price/m^2.

The last group of characteristics is the accessibility measures. Since due to confidentiality the exact position of the observations within each block is unknown all observations within a co-operative are given the same accessibility. We have chosen only to include three accessibility measures here. They are defined as the squared inverse distances to the CBD ($INCBD^2$), to the university area ($INUNIV^2$) and to the nearest of the two external shopping centres[4] ($INSHOP^2$). These measures were computed from estimated travel distance, using the city map Stad-

[4]We assume that each household generally uses the nearest of two shopping areas.

skartan, Lantmäteriet (1997)[27]. If these nodes are attractive then the price per square meter should increase for locations in the vicinity of the nodes. We expect a positive sign for all three variables.

4. The empirical examination

In this chapter, the sample is first examined by the Moran's I test statistic and the two G-statistics. However, before this a normality test for the dependent variable has to be made. An OLS regression follows next in order to evaluate if any autocorrelation has been taken care of by the regression parameters or if it is still present. If the latter is the case it gives an hint on what estimation method to use in order to make an appropriate estimation.

Hewnce, the first question to consider is whether spatial dependence (autocorrelation) exists or not by use of Moran's I test. As mentioned earlier, Moran's I test statistic depends on the chosen weight matrix. Five weight matrices are tested in this paper. They are based on the distances between observations within the ranges of 1 000, 1 500, 2 000, 2 500, and 3 000 meter and named d_1000, d_1500, d_2000, d_2500 and d_3000.

The elements of row r in the weight matrices are set to one for all observations within the specified distance from observation r and zero otherwise, including the diagonal element which is zero by convention. The reason for choosing the weights as binary and not as actual distance between all observations per se is the lack of exact knowledge of the observation location within the co-operative. So instead of guessing the locations within the co-operative and thereby impose some incorrect distances between them we chose to make every flat within the co-operative a neighbour (they are after all within all distance ranges). The matrices have then been row standardized, i.e. each element have been divided by the row sum. The reason for this is an urge that every co-operative flat should be influenced to the same degree regardless of if it has ten or forty neighbors.

To asses wether or not the dependent variable is normally distributed the asymptotic Wald test is used. It is distributed as χ^2 with 2 degrees of freedom.

$$\begin{aligned} W &= N\left[b^2/6 + (c-3)^2/24\right] \qquad (15) \\ b &= mean = \mu = \frac{\sum_r x_r}{N} \\ c &= variance = \sigma^2 = \frac{\sum_r (x_r - \mu)^2}{N} \end{aligned}$$

The dependent variable p/m^2 gives a W-value of 1189.79 and the probability 0.000 to reject the null hypothesis of a normal distribution. This means that Moran's I test must be performed using the permutation approach. That is, randomly reshuffling the observed values over all locations and by re-computing the I statistic for each new sample. A reference distribution is created to compute the first two moments. The results from the Moran's I tests for all weight matrices are presented in Table 2.

Table 2: Moran's I test for Spatial Autocorrelation
in the Co-operative Flat Data
(empirical pseudo-significance based on 999 random permutations)

VARIABLE	WEIGHT	I	MEAN	ST.DEV	PROB
p/m^2	d_1000	0.82	-0.005	0.023	0.001
p/m^2	d_1500	0.72	-0.005	0.020	0.001
p/m^2	d_2000	0.52	-0.005	0.017	0.001
p/m^2	d_2500	0.50	-0.005	0.015	0.001
p/m^2	d_3000	0.19	-0.005	0.011	0.001

The null hypothesis is that spatial autocorrelation does not exist for the co-operative flats. The dependent variable, p/m^2, shows a significant positive spatial autocorrelation on the 0.1 percent level for all five spatial weight matrices. This indicates that observations located nearby tend to have a more similar price per square meter than could be expected purely by chance. The Moran's I-value is 0.82 for the d_1000, which is a high degree of spatial autocorrelation. This is a good indicator that the OLS regression presented later in this paper will violate the assumptions of normal errors and is thus incorrect. For the other weight matrices the Moran's I value is also positive but decreasing with the increased influence bandwidth from r, which is quite reasonable. For d $\rightarrow \infty$, this means that every flat is a neighbor to all other flats, something that we think is unsuitable. There ought to be some critical distance where the influence expires. Which weight matrix should then be utilized in the next step, the OLS regression? The high level of spatial autocorrelation for all weight matrices (except for d_3000) makes it difficult at this stage to select a specific weight matrix in front of the others.

Another test for spatial dependence is the G-statistic presented in chapter two.

Table 3: G-statistic for Spatial Association

VARIABLE	WEIGHT	G	MEAN	ST.DEV	z-VALUE	PROB
p/m^2	d_1000	0.14	0.14	0.01	0.35	0.73
p/m^2	d_1500	0.21	0.29	0.03	-2.54	0.01
p/m^2	d_2000	0.34	0.38	0.04	-0.94	0.35
p/m^2	d_2500	0.35	0.39	0.04	-1.02	0.31
p/m^2	d_3000	0.34	0.55	0.04	-0.13	0.90

This test does not, at first glance, give such clear cut results as the Moran's I test did. All except the second weight matrix is insignificant. The reason for the first weight matrix not to be significant ought to be due to the location of the co-operatives. Since the distances between co-operatives often is larger than 1 000 meters many of the observations only have surrounding neighbors in their own co-operative. This implies that the difference between observations within this first band width is not significant. Only the weight matrix for the distance bound $0 - 1\ 500$ meters show sign of positive spatial dependence. This should then be interpreted that now has the radius of influence increased and we can detect some important patterns.

We have also tested the G-I* statistic. This test indicates to what extent each observation is surrounded by high or low values for a given distance band. The observations with the highest z-values are located in the two co-operatives north of the CBD. The most negative z-values can be found in the co-operative at the east of the city, near the lake. This is also the co-operative with the lowest prices. All five weight matrices presents the same pattern with some minor differences.

The next step is as mentioned the OLS regression which consists of the following elements.

$$p/m^2(\mathbf{z}) = \boldsymbol{\beta} \cdot \mathbf{s} + \boldsymbol{\eta} \cdot \mathbf{n} + \boldsymbol{\psi} \cdot \mathbf{a} + \boldsymbol{\varepsilon} \tag{16}$$

The reason for performing the OLS despite all indications of its unsuitability is first to see if this is indeed true, but if so is it is still able to give us guidance towards a proper model specification. The OLS model explains 91 percent of the variance in the p/m^2 measured as R^2-adjusted and the variance and standard deviation is 427 552 and 653.9 respectively. The parameter values are presented in table 5 below since it may be compared with the regression results presented later in the paper.

The Moran's I tests earlier indicated that there is spatial autocorrelation present in the data. A drawback of the Moran's I test is that it does not specify what kind of dependence that exists. To analyze if there still exist problems with

spatial autocorrelation when we have introduced the explanatory variables, and in that case what kind of spatial autocorrelation, four Lagrange multiplier tests based on the OLS regression results are used. These are presented in Table 4 for the five weight matrices.

Table 4: Diagnostics for spatial dependence. Probability in brackets.

Weight matrix	LM, error	Robust LM, error	LM, lag	Robust LM, lag
d_1000	34.44 (0.000)	27.30 (0.000)	8.12 (0.004)	0.97 (0.325)
d_1500	42.08 (0.000)	35.25 (0.000)	6.83 (0.011)	0.00 (0.965)
d_2000	49.31 (0.000)	46.59 (0.000)	2.95 (0.086)	0.23 (0.629)
d_2500	1.36 (0.243)	2.23 (0.135)	0.86 (0.354)	1.73 (0.188)
d_3000	1.96 (0.161)	0.00 (0.958)	26.84 (0.000)	24.88 (0.000)

The table should be interpreted in the way that if we include an error correction and the result is a significant value (the two columns to the left) this means that the autocorrelation is taken care of. The robust case is a test for the case that if we instead of an error term put a lag on the dependent variable in the model . If it would yield an insignificant value this would mean that the lag correction should be used instead, and the model would not benefit from an error correction. This is so to say a test of the first LM test. The opposite applies if we look at the tests for lag dependence (the two columns at the right). In case all four LM tests are significant, the thumb rule is then that the two with the highest values determines how to continue.

As shown, both tests for error dependence are highly significant, while the robust LM for the lag is insignificant for d_1000. This is interpreted as that if a spatial error correction is introduced in the equation it solves the problem of spatial autocorrelation. The two following weight matrices give the same indication, but the value for d_2500 drops dramatically from the value of d_2000, and does not give any signals of spatial dependence. This is probably an indication that this is as far as we should go in our search for a proper weight matrix. In the case of d_3000 the opposite applies. Here, there is indication that a lag on the dependent variable should be introduced to receive a proper model. Since the weight matrices with a more restrictive distance band width indicated error dependence, both d_2500 and d_3000 are ruled out as inappropriate to use in the subsequent analysis.

A further problem that we encountered from the use of the OLS regression was a problem with the non-normal distribution of the error terms. To test this we used the Jarque-Bera test statistic[5]. A critical value of 133.15 (prob=0.000)

[5] χ^2 distributed with 2 degrees of freedom.

clearly rejects the null hypothesis that the residual terms are normally distributed.

The next step in the process is then to incorporate the spatial dependence in the model. The fact that we have non normal errors rules out ML-estimation. To solve both the problem with spatial autocorrelation as well as the non-normal distribution of the error terms we suggest a spatial autoregressive model estimated by General Method of Moments using the weight matrices presented earlier.

The SAR-GMM model was put forward by Kelejian and Prucha in their 1999 article[25]. One of the benefits from using the GMM estimation technique is the acceptance of non-normality, which is needed for the following estimation. An alternative to GMM-estimation would have been to use IV-estimation. This then means that we would have to determine proper instruments in a rather ad-hoc fashion so we decided to go along with the SAR-GMM.

The model to estimate has the following form:

$$\begin{aligned} p/m^2(\mathbf{z}) &= \boldsymbol{\beta} \cdot \mathbf{s} + \boldsymbol{\eta} \cdot \mathbf{n} + \boldsymbol{\psi} \cdot \mathbf{a} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda \mathbf{W} \boldsymbol{\varepsilon} + \boldsymbol{\xi} \end{aligned} \tag{17}$$

The weight matrices are the three first specified above. The first two SAR-GMM models converges after 9 iterations and the third after 8 iterations. For the regression with the weight matrix d_1000 Sig-sq is 342253 (585.02), or 85 000 lower than in the OLS case. In other words the estimates are improved. d_1500 and d_2000 improve this further. The last weight matrix stands out with some insignificant parameters.

Table 5: The SAR-GMM regression results.
*=Insignificant at 5% level

VARIABLE	d_1000	d_1500	d_2000	OLS
$CONSTANT$	3971.27	4163.71	3931.17	620.94*
$1 - ROOM$	665.69	654.43	657.55	1068.52
$(FEE/m^2)^{-1}$	-167753	-174351	-168094	8251.97*
$INCENT^2$	4.19E+09	4.25E+09	4.29E+09	3.43E+09
$INUNIV^2$	4.46E+09	4.12E+09	4.33E+09	2.93E+09
$INSHOP^2$	-1.64E+09	-1.60E+09	-1.21E+09*	-7.09E+09*
$POPDENS$	109.83	130.38	106.44	6.35*
$SHDENS$	2489.03	2446.31	2568.36*	1897.92
TO	-94.26	-96.86	-97.52	-62.90
λ	0.62	0.67	0.74	
R2	0.89	0.90	0.92	0.92
R2(Buse)	0.76	0.79	0.84	
Sq. Corr	0.91	0.91	0.91	
Sig-Sq	342 253	337 604	337 720	427552
Sig	585.02	581.07	581.14	653.77
Iterations	9	9	8	

As hypothesized the parameter $(FEE/m^2)^{-1}$ has a negative sign saying that the price per square meter increases if $(FEE/m^2)^{-1}$ decreases and more so for small co-operative flats. The dummy variable is positive and significant indicating that small flats are more expensive to purchase per square meter than larger.

The two accessibility measures $INCBD^2$ and $INUNIV^2$ are both significant and positive. The interpretation of this should be that Umeå has a structure with two attractive nodes with different characteristics, the CBD and the university area. External shopping centres has a negative impact on price per m^2. It should not necessarily be seen as negative per se to be close to a shopping centre, but their existence could induce large traffic in the area and other negative external effects that is shown in the regression results.

Two of the neighborhood variables are positive and significant. To some extent a high population density, but especially a large percentage of single-family houses drives the price upwards. The variable TO is negative meaning that high turnover has negative impact on price per square meter, i.e. stability in contrast to liquidity is valued positively.

Finally the autoregressive parameter λ is positive, thus the error terms in

locations nearby tend to coincide more than purely by chance. Unfortunately there is no test to confirm if this parameter is significant or not. However, given the indications by Moran's I and the Lagrange Multiplier tests earlier, the parameter probably is significant.

To illustrate the predicted values for p/m^2 and a smoothing across the city of Umeå a map of this is presented in Figure 3. There the reader can clearly detect a concentration of high values around the CBD and the university area. Apparent low prices are found for the co-operative flats at the east and especially near the lake.

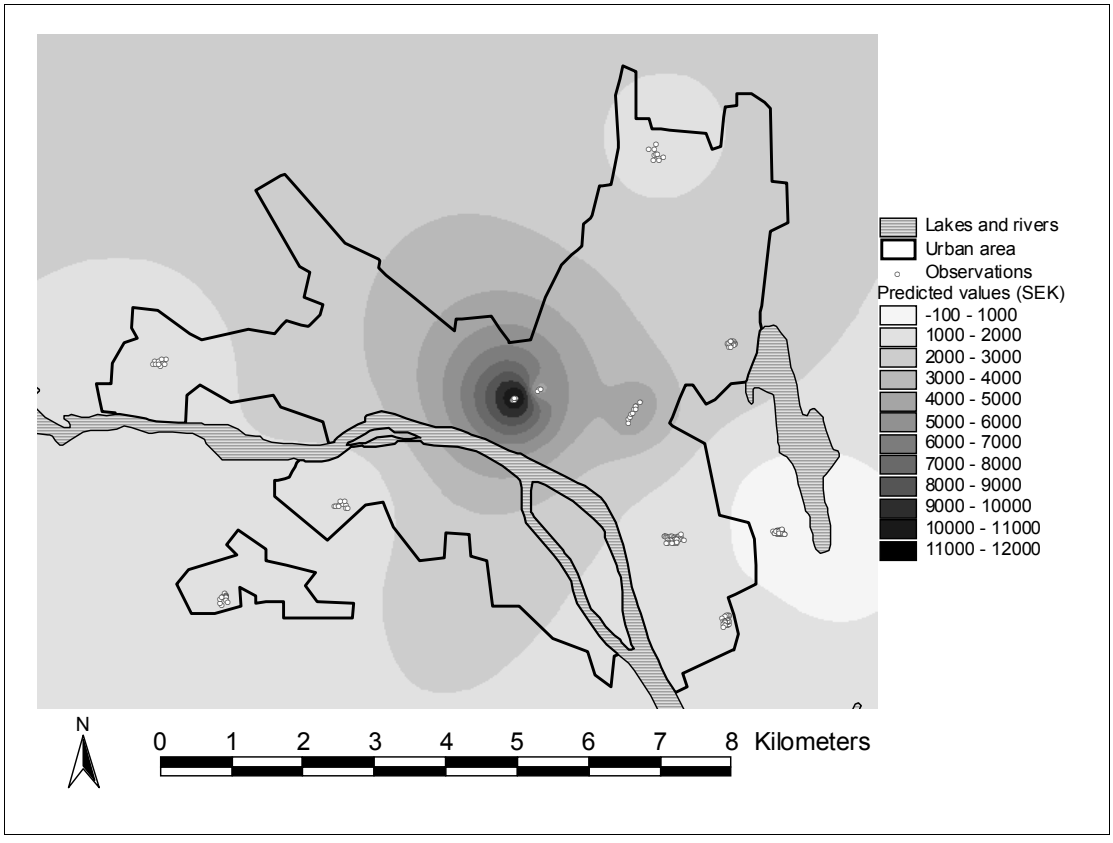

Figure 3: The prediction on p/m^2 for co-operative flats in Umeå

To further illustrate the results and the possible use in the future we present a table for prediction of some new sites that the municipality of Umeå consider for exploitation in the near future[35]. The potential sites are indicated in Figure 4 as darker areas.

Future 4: Future construction areas

The parameter values from the regression with the weight matrix d_1500 are used in the calculations. To be able to compare different locations we consider a hypothetical co-operative flat. We assume a 70 m^2 flat with a monthly fee of 3 500 SEK. The area for each site is approximately as given in the figure. The turnover is assumed to be 20 percent and the population is set to three times the number of flats prepared to be built at each site[35]. The number of single family houses are assumed to be 0.6/1 000 m^2. This gives the prices per square meter indicated in Table 6.

Table 6: Calculated values of p/m^2 in tentative new sites for Co-operative flats in Umeå.

SITE	p/m^2
Hamrinsberget	15 000
Öbacka 2	5 400
Dragonfältet	4 600
Bryggeriet	2 500
Olofsdal	2 100
Tomtebo 1b	1 400
Lundåkern	1200
Tomtebo 1a	1 100
Tomtebo 2	900
Nydala	900
Västra Umedalen	800

As indicated the most valued location for future development is the area named Hamrinsberget, close to the university area and on the way to the city centre. Compared to the values in Figure 3, this is quite high. This is mainly due to the closeness of the university. The second highest valued area is the Öbacka 2 area, close to the river. Since no objects with pure waterfront location exists in the cooperative flat market in Umeå this impact could not be investigated. If this was possible, Öbacka 2 would possibly be valued a bit higher. It is neither a surprise to find the cheapest co-operative flats at the fringe of Umeå.

5. Conclusions

The purpose for this paper was to use hedonic price theory to assess some important attributes for the price determination of co-operative flats in the city of Umeå. Especially the paper has been considered with the treatment of spatial dependence in this type of markets. The econometric analysis showed that OLS was not applicable due to spatial dependence between the error terms. This was accounted for through introduction of spatial weight matrices based on observations with neighbors within 1 000 up to 2 000 meters and estimation with spatial autoregressive GMM.

The attribute that describes the fee for the flat $(FEE/m^2)^{-1}$ was negative and significant so an increased fee increases the square meter price. It was also clear that the market price for single room flats are significantly more expensive

per square meter compared with larger flats. This could also be a sign indicating that there are different markets for small and larger co-operative flats in the city.

Among the tested attributes, the "classical" influence of accessibility to centres on real estate prices has been confirmed. But in the case of Umeå, not only the CBD but also the university area influence the market price positively. Instead accessibility to external shopping centres has a negative impact on the price. The city of Umeå may then be described as having a multi nodal structure with the centres having different characteristics. The density attributes also matter in a positive direction.

The conclusion to draw from this paper for future urban development in order to enhance the attractiveness is to build co-operative flats in areas somewhere between the CBD and the university area and more closely to the former. In other places a mix with a high percentage of single-family houses improves the value of the flat. One should moreover try to establish robust and attractive co-operatives to secure a low turnover rate.

The fact that we in this paper have not used more attributes for the individual flats makes it difficult to say anything about internal attractiveness and we thereby cannot argue for a specific type of housing. Further studies are needed in this field to gain knowledge regarding the preferences held by the households in Umeå concerning their housing preferences. Another needed knowledge is the household's willingness to pay for waterfront location and public goods such as park areas. A value for such "unbuilt" areas may give a value on alternatives to exploitation for housing. It would also be interesting to do a similar study for sigle-family houses to see if the preferences are similar for the two different types of housing.

References

[1] Adelman, I. and Griliches, Z. (1961) On an Index of Quality Change. Journal of the American Statistical Association 56, 535-548

[2] Andersson, D. E. (1997) Hedonic prices and Center Accessibility: Conceptual Foundations and an Empirical Hedonic Study of the Market for Condominium Housing in Singapore. Ph. D. Dissertation, Royal Institute of Technology.

[3] R. Andersson (1998) Attraktiva Städer - en samhällsekonomisk analys. Stockholm, Byggforskningsrådet Svensk Byggtjänst.

[4] Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht, Kluwer Academic.

[5] Anselin, L. (1995) SpaceStat Version 1.80 User's Guide

[6] Archer, W. R. et al.(1996) Measuring the Importance of Location in Housing Price Appreciation. Journal of Urban Economics 40, 334-353

[7] Blomquist, G. C. (1998) Valuing Urban Lakeview Amenities Using Implicit and Contingent Markets. Urban Studies 25, 333-340.

[8] Blomquist, G. C. et al (1998) New Estimates of Quality of Life in Urban Areas. The American Economic Review 78, (1), 89-107.

[9] Bodson, P. and Peeters, D. (1975) Estimation of the coefficients of a linear regression in the presence of spatial autocorrelation. An application to a Belgian labour-demand function. Environment and Planning A 7, 455-472.

[10] Borukhov, E. (1978) Housing Prices and Housing Preferences in Israel. Urban Studies 15, 178-200.

[11] Buck, A. J. et al. (1991) A Von Thünen Model of Crime, Casinos and Property Values in New Jersey. Urban Studies 28, (5), 673-686.

[12] Can, Y. (1992) Specification and estimation of hedonic housing price models. Regional Science 22, 453-474.

[13] Cheshire, P. and Sheppard, S. (1995) On the Prise of Land and the value of Amenities. Economica 62, 247-267.

[14] Cliff, A. and Ord, K. (1972) Testing for Spatial Autocorrelation Among Regression Residuals. Geographical Analysis 4, 267-284.

[15] Cliff, A, and Ord, K. (1973) Spatial Autocorrelation. London, Pion Limited.

[16] Colwell, P.F. (1990) Power Lines and Land Value. The Journal of Real Estate Research 5, (1), 117-127.

[17] Colwell, P. F. and Dilmore, G. (1999) Who Was First? An Examination of an Early Hedonic Study. Land Economics 75(4) 620-626.

[18] Court, A. T (1939) Hedonic price indexes with automobile examples, in: The dynamics of automobile demand. General Motors Company, New York, 99-117.

[19] Englund, P. et al. (1998) Improved Price Indexes for Real Estate: Measuring the Cource of Swedish Housing Prices. Journal of Urban Economics 44, 171-196.

[20] Eriksson, B. J. (1997) Bostäder och Hedoniska Priser. Beställning av flerfamiljshus 1967-1987. Ph. D. Dissertation, Royal Institute of Technology.

[21] Gavlefors, U. and Roos, H (1992) Prisbildning på bostadsrätter : en undersökning om prisbildningen på bostadsrätter i flerbostadshus samt vilka faktorer som bör beaktas vid värdering av enskilda bostadsrätter. LMV-rapport 1992:6, Lantmäteriverket, Gävle.

[22] Getis, A and Ord, J. K. (1992). The Analysis of Spatial Assosiation by Use of Distance Statistics. Geographical Analysis 24, 189-206.

[23] Haas, G. C. (1922) A Statistical Analysis of Farm Sales in Blue Earth County, Minnesota, as a Basis for Farm Land Appraisal. Master Thesis, the Univesity of Minnesota.

[24] Heikkila, E. (1989) What happened to the CBD - distance gradient?: land values in a policentric city. Environment and Planning A 21, 221-232.

[25] Kelejian, H. H. and Prucha, I. R. (1999) A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. International Economic Review 40, No 2, 509-533

[26] Lancaster, K. (1966) A new approach to consumer theory. Journal of Political Economy 74, 132-157

[27] Lantmäteriet. (1997). Stadskartan. Kartförlaget, Gävle

[28] Lindgren, P. and Rosberg, G. (1992) Vem bor till vilket pris? Boverket, Karlskona

[29] Matérn, B. (1947) Metoder att uppskatta noggranheten vid linje- och provytetaxering. Meddelanden från Statens Skogsforskningsinstitut 36, 1-117.

[30] Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy. 34-55

[31] SABO (1997) Vad värderar vi i vårt boende. En sammanfattning av två undersökningar. Stockholm

[32] Sheppard, S. (1999) Hedonic Analysis of Housing Markets. Handbook of Regional and Urban Economics III. Amsterdam Elsevier.

[33] Sivitanidou, R. (1996). Do Office-Commercial Firms Value Access to Service Emlpoyment Centers? A Hedonic Value Analysis within Polycentric Los Angeles. Journal of Urban Economics 40, 125-149

[34] SOU 2000:33 Bruksvärde, förhandling och hyra - en utvärdering. Finansdepartementet Stockholm

[35] Översiktsplan (1998) öpl 98 - för en hållbar utveckling. Umeå Kommun

[36] Werner, I. B. (2000) Spelar kvalitet någon roll för priset? - En studie av bostadsrättsköp i StorStockholm. Ph.D. Dissertation, The Royal Institute of Technology.

[37] Wigren, R. (1987). House Prices in Sweden: The Significance of Attributes. Scandinavian Housing and Planning Research 4: 243-261.