# The Impact of Urban Amenities on Apartment Prices in Münster – How Important is Location within a Bike-City?

Norbert Hiller*

*Unpublished draft, April 2013. Please do not cite or quote!*

**Abstract**: Beside the specific object attributes, location is the fundamental determinant for real estate. But how important is location within a city? Contrary to popular belief our results suggest that this parameter explains only a small proportion of price variation, whereas object attributes are crucial. “Location, location, location” does not capture the essence of an apartment value. In a bike-city it is only an explanation part which has relative little impact on price variations. Therefore, we suggest “size, age, location” as comprehensive determinants for local apartment prices.

**Keywords**: Amenities; Hedonic Method; Real Estate and Urban Economics; Factor Analysis, Relative Importance

## 1. Introduction

The quality of a location can be valued by its amenities, such as shops, restaurants or a lake. Therefore primarily amenities determine where people want to live. At the macro level, researchers underline their positive consumption effect, which increases the demand to live in cities with generous amenities. Particularly the role of restaurants, museums, movie theaters, bowling alleys and hotels are compared between metropolitan areas (Glaeser et al. 2001). At an urban micro level, their role has only partly been investigated, although many hedonic pricing models (which measure its effects) have been developed. For example, Noonan (2007), and Ahlfeldt & Maenning (2010) highlight the price effects of famous landmarks on house prices. Furthermore, road noise or the distance to the nearest green space and body of water have a significant impact on residential property prices (Brandt and Maenning 2011). Sirmans et al. (2005) provide a broad overview of potential determinants of hedonic pricing models in different studies. The coverage of classic macro level amenities is very low, so that further research is clearly necessary.

In this empirical paper, we measure the impact of location characteristics on apartment prices with a long vector of local amenities. We use data for the medium-sized German city of Münster (Westphalia) for three reasons. Firstly, the role of amenities in major German major cities is unexplored at a micro level. Secondly, in the Second World War, 91% of the historic center of Münster was destroyed and 63% the whole city. Compared to other German cities, the historical center has largely been rebuilt and a substantial amount of built heritage exists in the central business district (CBD). Thus, amenities that can be regarded as historic buildings are measureable. Thirdly, Münster is a “bike-city”. There are round about 500,000 bikes for 291,754 inhabitants (University of Münster 2005). This is a rare characteristic, compared with other major cities in developed nations.

* Institute of Regional and Urban Economics, University of Muenster, Am Stadtgraben 9, Germany, 48143 Muenster, e-mail: hiller@insiwo.de

In the following Section 2, we give a short overview of the data. To measure the micro level impact of amenities on housing prices, we construct a basic hedonic regression model, which is described in Section 3. Therefore, we expand our model with a neighborhood matrix to take spatial dependence into account. In Section 4, we measure the relative importance of the variables to determine the strength of dimensions leading to apartment price variations in Münster.

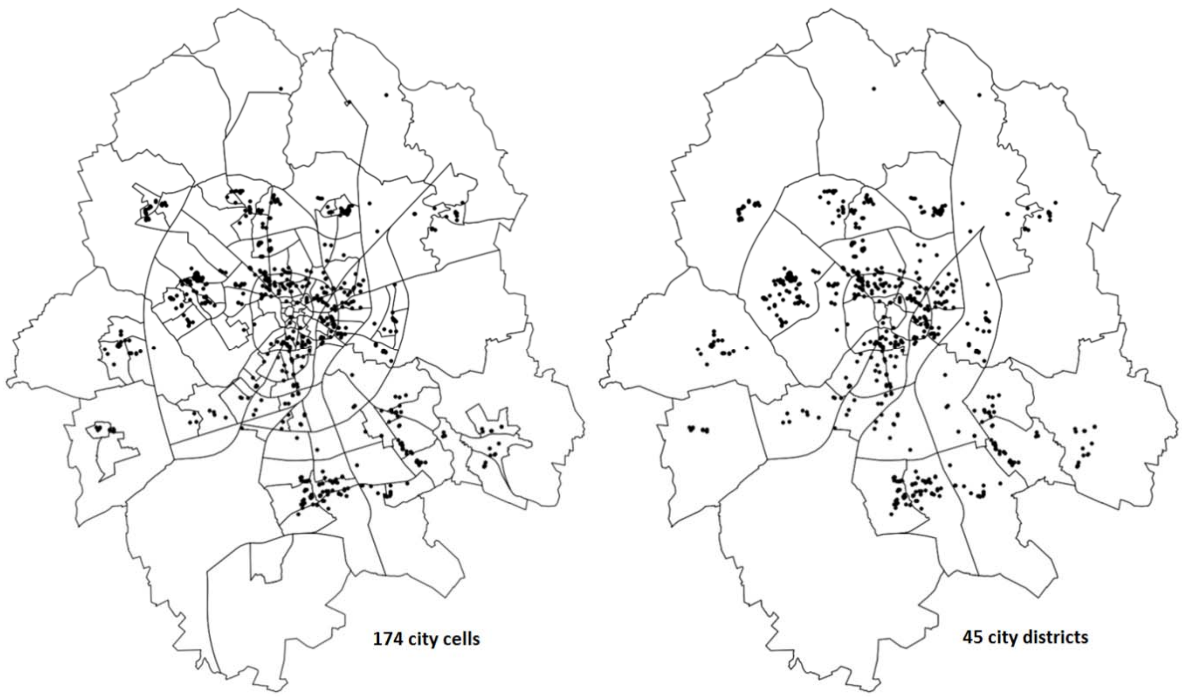
## 2. Data

The study covers the entire city of Münster, which has an area of 303 km$^2$ and a population of 291,754 (December 31st, 2011). As in other central European cities, the apartment market is an important submarket. With its 45,843 students (2012), the role of apartments in Münster as an investment object is very attractive.

We considered 633 apartment sales by “F+B GmbH” between 2008 and 2012. The data includes the usual object parameters (e. g. *rooms*, *age*, *kitchen*, *bathroom*, *balcony*, *garden*, *elevator*, *apartment type*). Furthermore, we observe different amenities (e. g. *restaurants*, *bakeries*, *pubs*, *sport clubs*, *building heritage*), landscape parameters (e. g. *building area*, *forest area*, *farmland*, *high-traffic area*, *commercial area*) and social neighborhood parameters (e. g. *unemployment rate*, *income per taxpayer*, *share of foreigners*, *basic service area in sqm*) from the year 2011. Additionally, we use *street noise* data which can be negatively related to apartment prices (Ball 1973, Andersson et al. 2009). The amenity data is collected by “Stadt Münster” (City of Münster).

The apartment- and amenity-data is based on geographic coordinates, so we can measure the Euclidean distance between each apartment and the located amenity. The social neighborhood parameters are based on 174 statistical districts (called *city cells*), the landscape parameters on 45 statistical districts (called *city districts*). Figure 1 shows the location of the apartments and the boundaries of the different district types.

**Figure 1:** Apartments in the statistical districts of Münster



The complete data set consists of the transaction price, 23 object attributes, 11 neighborhood parameters, 5 landscape parameters, 25 (dis-)amenity values and 5 time dummies. A statistical description of the data set is shown in Table 1.

**Table 1:** Variable names, definitions and summary statistics

| Variable | Definition | Mean |
|---|---|---|
| *Dependent variable (P)* | | |
| Price | Transaction price of property | 131502.20 |
| *Object attributes(O)* | | |
| Rooms | Number of rooms | 2.59 |
| Size | Living area in sqm | 74.79 |
| Age | Age of property in years | 30.07 |
| First time in use | 1 if apartment is used first time, 0 otherwise | 0.12 |
| Kitchen | 1 if property has a kitchen, 0 otherwise | 0.38 |
| Loft | 1 if property is a loft, 0 otherwise | 0.15 |
| Balkony | 1 if property has a balcony, 0 otherwise | 0.56 |
| Elevator | 1 if property has an elevator, 0 otherwise | 0.08 |
| Parking place | 1 if property has a parking place, 0 otherwise | 0.74 |
| Fireplace | 1 if property has a fireplace, 0 otherwise | 0.02 |
| One floor heating | 1 if property has a one floor heating, 0 otherwise | 0.10 |
| Basement | 1 if property has a basement, 0 otherwise | 0.65 |
| Quiet | 1 if property is in a quiet area, 0 otherwise | 0.26 |
| Central heating | 1 if property has a central heating, 0 otherwise | 0.63 |
| Garden | 1 if property has a garden, 0 otherwise | 0.20 |
| Renovated | 1 if property is renovated, 0 otherwise | 0.09 |
| Luxurious | 1 if property is luxurious, 0 otherwise | 0.11 |
| Brightly | 1 if property has lots natural light, 0 otherwise | 0.25 |
| Floor heating | 1 if property has a floor heating, 0 otherwise | 0.06 |
| Loggia | 1 if property is a loggia, 0 otherwise | 0.09 |
| New building | 1 if property is a new building, 0 otherwise | 0.06 |
| Maisonette | 1 if property is a maisonette, 0 otherwise | 0.06 |
| Parquet floor | 1 if property has a parquet floor, 0 otherwise | 0.09 |
| *Neighborhood (N)* | | |
| Population density | Population density in city cells (174) | 44.49 |
| Households with childr. | Households with children in city cells (174) | 111.73 |
| Ratio of family househ. | Ratio of family households in city cells (174) | 0.12 |
| Ratio free housing | Ratio of sponsored housing in city districts (45) | 0.05 |
| Playground in sqm | Playground in sqm in city districts (45) | 10141.36 |
| Basic service in sqm | Basic service in sqm in city districts (45) | 2317.6 |
| Ratio foreigners | Ratio of foreigners in city districts (45) | 0.21 |
| Ratio elderly | Ratio of elderly (> 65 years) in city districts (45) | 0.22 |
| Ratio unemployment | Ratio of unemployment in city districts (45) | 0.06 |
| Income per taxpayer | Income per taxpayer in city districts (45) | 29276.39 |
| Ratio house reserve | Ratio of house reserve in city districts (45) | 0.04 |
| *Landscape (L)* | | |
| Ratio building area | Ratio of building area in city districts (45) | 0.43 |
| Ratio industrial area | Ratio of industrial area in city districts (45) | 0.003 |
| Ratio traffic area | Ratio of traffic area in city districts (45) | 0.15 |
| Ratio farm area | Ratio of farm area in city districts (45) | 0.23 |
| Ratio forest area | Ratio of forest area in city districts (45) | 0.08 |

**Table 1:** continued

| Variable | Definition | Mean |
|---|---|---|
| *(Dis-)Amenities (A)* | | |
| Dist CBD | Distance to the central business district | 3977.66 |
| Dist built heritage | Distance to the next built heritage | 376.19 |
| Dist restaurants | Distance to the next restaurants | 440.94 |
| Dist sport club | Distance to the next sport club | 1433.50 |
| Dist canteen | Distance to the next canteen | 1434.62 |
| Dist amusement place | Distance to the next amusement place | 1752.61 |
| Dist ice cream parlor | Distance to the next ice cream parlor | 727.21 |
| Dist coffee bar | Distance to the next coffee bar | 1250.15 |
| Dist disco | Distance to the next disco | 2597.34 |
| Dist kiosk | Distance to the next kiosk | 1031.59 |
| Dist bakery | Distance to the next bakery | 594.52 |
| Dist bakery self-service | Distance to the next bakery self-service | 1146.45 |
| Dist bakery and coffee | Distance to the next bakery and coffee bar | 459.57 |
| Dist bus station | Distance to the next bus station | 135.84 |
| Dist hotel/inn | Distance to the next hotel/inn | 922.68 |
| Dist snack bar | Distance to the next snack bar | 326.11 |
| Dist pub | Distance to the next pub | 1029.65 |
| Dist other pub | Distance to the next “alternative” pub | 1590.01 |
| Dist central station | Distance to the central station | 4021.83 |
| Dist cinema | Distance to the next cinema | 2627.15 |
| Dist day-care center | Distance to the next day-care center | 635.91 |
| Dist grocery store | Distance to the next grocery store | 381.00 |
| Dist museum | Distance to the next museum | 1782.63 |
| Dist primary school | Distance to the next primary school | 507.76 |
| Noisy | 1 if noise level is above 55 DB, 0 otherwise | 0.13 |
| *Time dummies (T)* | | |
| Dummy date 2008 | Time dummy for 2008 | 0.22 |
| Dummy date 2009 | Time dummy for 2009 | 0.20 |
| Dummy date 2010 | Time dummy for 2010 | 0.19 |
| Dummy date 2011 | Time dummy for 2011 | 0.25 |
| Dummy date 2012 | Time dummy for 2012 | 0.15 |

The relationship between the value and *age* of the apartment is expected to follow a U-shaped curve. Very old houses can earn a premium due to their historic character, and thus, the *age* and *square of age* are included in the model. The variable *quiet* captures the object attributes done by “F+B GmbH”, whereas the *noise dummy* is generated by the street-noise level collected by the “Stadt Münster”. The distances to the next amenity are linear. Furthermore, the *ratio of free housing* represents social housing apartments which are subsidized by the government. Its residents do not have to pay for rent or energy. The *basic service* describes the sales area for basic food in the 45 city districts. Yearly *time dummies* are also added to the model.

Since the amenity data is observed for only one year, we must assume that they do not change over the observed period (2008 to 2012). With only 633 observations, the number of 69 independent variables is very high. In a regression, the coefficient of determination could be overestimated, but the number of variables will decrease, for reasons we mention in Section 3.1.

## 3. Model

### 3.1 Basic Approach

We assume that at equilibrium, the attractiveness of real estate is fully capitalized into property prices. The attractiveness can depend on object attributes (O), amenities (A), landscape parameters (L) and neighborhood parameters (N). Furthermore, we use time dummies (T) to control for price changes over the observation period. The implicit prices in these parameters are estimated using a standard hedonic approach (Rosen 1974, Muellbauer 1974, Li and Brown 1980, Blomquist and Worley 1981). For the semi-logarithmic form, the basic model can be written as:

$$log(P)_i = \alpha + \beta_{ij}O_{ij} + \gamma_{ik}L_{ik} + \delta_{il}N_{il} + \theta_{im}A_{im} + \eta_{in}T_{in} + \varepsilon \quad (1)$$

where *P* is the transaction price; *j*, *k*, *l, m* and n represent the number of attributes of object *i*; *α*, *β*, *γ*, *δ, θ* and *η* are coefficients and ε is an error term. Log-linear specifications allow non-linearity and are intuitively interpretable. The coefficients give the percentage impact of changes in attribute values on property prices. Furthermore, we assume that all attributes are homogeneous. There are no quality differences between restaurants, highways or historical monuments.[2] Since Alonso (1964), a large body of literature has appeared on real estate values and their decline with increasing distance from the CBD. Therefore, we include this distance to avoid bias, which can result from congestion.[3] In a first regression (model 1) we identify high multicollinearity, especially between the distance measures. The distance is calculated as linear, although the multicollinearity appears by using decaying or logarithm distances. Table 2 shows the variance inflation factors of all variables, where significant variables (10% and better) appear in bolt.[4] In general, if the variance inflation factor is higher than 10, the level of multicollinearity appears to be a problem.

[2] The impressive castle of Münster is treated as a geographical data point like any other national heritage. There is no weighting.

[3] We define the CBD near the cathedral of Münster at Domplatz 1.

[4] The variables “Ratio of building area” and “Dummy date 2008” are not included, because they are a linear combination of other dependent variables (perfect multicollinearity).

**Table 2**: Variance inflation factors of model (1)

| Variance inflation factor | | | |
|---|---|---|---|
| **Rooms** | **2.615053** | **Ratio house reserve** | **5.392912** |
| **Size** | **2.603307** | Ratio industrial area | 21.554964 |
| **Age** | **7.527693** | Ratio traffic area | 9.754702 |
| **Age (squared)** | **5.792492** | Ratio farm area | 24.450225 |
| **First time in use** | **1.773360** | Ratio forest area | 20.069555 |
| Kitchen | 1.336391 | **Dist CBD** | **191.341625** |
| Loft | 1.245640 | Dist built heritage | 4.986515 |
| Balkony | 1.227519 | Dist restaurants | 3.875663 |
| Elevator | 1.187053 | Dist sport club | 6.133198 |
| **Parking place** | **1.510472** | Dist Canteen | 15.413397 |
| **Fireplace** | **1.171064** | Dist amusement place | 16.042240 |
| **One floor heating** | **1.643958** | Dist ice cream parlor | 10.258949 |
| Basement | 1.404974 | Dist coffee bar | 18.189568 |
| Quiet | 1.169566 | **Dist disco** | **4.656709** |
| Central heating | 1.548470 | Dist kiosk | 20.195550 |
| **Garden** | **1.259665** | Dist bakery | 5.727338 |
| **Renovated** | **1.313616** | Dist bakery self-service | 10.530710 |
| **Luxurious** | **1.397434** | Dist bakery and coffee | 5.157992 |
| Brightly | 1.219271 | **Dist bus station** | **1.383534** |
| Floor heating | 1.290559 | Dist hotel/inn | 5.211585 |
| Loggia | 1.126296 | Dist snack bar | 2.518550 |
| New building | 1.445130 | Dist pub | 11.749564 |
| **Maisonette** | **1.176760** | **Dist other pub** | **22.122599** |
| Parquet floor | 1.209619 | Dist central station | 80.221640 |
| **Population density** | **4.032483** | **Dist cinema** | **58.050392** |
| Households with children | 4.660571 | Dist day-care center | 5.261692 |
| Ratio familiy households | 4.721573 | Dist grocery store | 3.060511 |
| Ratio free housing | 16.352002 | **Dist museum** | **14.328708** |
| Playground in sqm | 6.919873 | **Dist primary school** | **2.083016** |
| Basic service in sqm | 7.486790 | **Noisy** | **1.699894** |
| Ratio foreigners | 17.676349 | Dummy date 2009 | 1.774063 |
| Ratio elderly | 5.700676 | **Dummy date 2010** | **1.883338** |
| Ratio unemployment | 24.510488 | **Dummy date 2011** | **1.994120** |
| Income per taxpayer | 3.314479 | **Dummy date 2012** | **1.795211** |

Particularly the variance inflation factors suggest that many neighborhood and distance measures are correlated with each other. Thus, the estimators have a wide confidence interval and are may not be significant. However, there are different strategies for handling the problem, namely elimination of variables, using more data – which is not available[5] – or merging correlated variables by a principal component analysis. We prefer merging than eliminating variables, so as to avoid omitted variable bias and to benefit from all available location information.

[5] E. g. we cannot control for „crime“, because the data privacy level in Germany is too restrictive.

### 3.2 Principal Components

We use principal component analysis to extract orthogonal factors from the landscape, neighborhood and amenity data. After the analysis, we will have different amenity-indices which represent different location benefits. The landscape and neighborhood variables are based on 45 city districts. We first define the principal neighborhood components (PCN) for this spatial unit, and then we define principal components for the distance-based amenities (PCA).

For an effective principal component analysis, we need a useful correlation matrix. The goodness of the spatial unit and the amenity correlation matrix can be controlled by Bartlett's test and the Kaiser-Meyer-Olkin measure of sampling adequacy (Dziuban and Shirkey 1974). The Bartlett test hypothesizes that the correlation matrix is equal to the identity matrix. This hypothesis can be declined for both correlation matrices with a p-value of 0. The Kaiser-Meyer-Olkin measure of sampling adequacy represents the ratio of the squared correlation between variables to the squared partial correlation between variables. A value close to 1 indicates that patterns of correlations are relatively compact, so that principal component analysis should yield distinct and reliable factors. Both matrices pass the test. We further use the Kaiser-Guttman-Criteria, where the number of eigenvalues > 1 determines the number of components. In addition, the cumulative variance should be higher than 75%. We do not follow these criteria strictly. Overall, the components should describe the data in an interpretable manner. Tables 3 and 4 show, we argue, the most suitable results of principal component analysis for both correlation matrices.

**Table 3**: Principal component analysis of 45 city district variables

| Component | Percent variance explained | Cumulative | Eigenvalues |
|---|---|---|---|
| A | 31 | 31 | 4.88 |
| B | 28 | 59 | 3.04 |
| C | 16 | 75 | 1.84 |

| Component | Name | Variable loadings (>\|0.50\|) | |
|---|---|---|---|
| A | PCN low income | Ratio free housing | 0.77 |
| | | Ratio industrial area | 0.83 |
| | | Playground in sqm | 0.58 |
| | | Ratio foreigners | 0.88 |
| | | Ratio unemployment | 0.94 |
| B | PCN high income | Ratio house reserve | 0.66 |
| | | Ratio traffic area | -0.88 |
| | | Ratio farm area | 0.90 |
| | | Ratio forest area | 0.72 |
| | | Ratio building area | -0.96 |
| | | Income per taxpayer | 0.44 |
| C | PCN shopping | Basic service in sqm | 0.82 |
| | | Ratio elderly | -0.70 |

**Table 4**: Principal component analysis of amenity distances

| Component | Percent variance explained | Cumulative | Eigenvalues |
|---|---|---|---|
| D | 28 | 28 | 8.83 |
| E | 10 | 37 | 2.89 |
| F | 9 | 47 | 2.20 |
| G | 9 | 56 | 1.41 |
| H | 9 | 65 | 1.34 |
| I | 8 | 72 | 1.23 |
| J | 6 | 78 | 0.99 |
| K | 4 | 82 | 0.87 |

| Component | Name | Variable loadings (>\|0.50\|) | |
|---|---|---|---|
| D | PCA centrality | Dist coffee bar | 0.90 |
| | | Dist other pub | 0.84 |
| | | Dist disco | 0.84 |
| | | Dist cinema | 0.84 |
| | | Dist kiosk | 0.83 |
| | | Dist CBD | 0.82 |
| | | Dist canteen | 0.78 |
| | | Dist central station | 0.77 |
| | | Dist amusement place | 0.63 |
| E | PCA food | Dist bakery and coffee | 0.80 |
| | | Dist grocery store | 0.78 |
| | | Dist snackbar | 0.55 |
| | | Dist ice cream parlor | 0.55 |
| F | PCA culture | Dist museum | 0.91 |
| G | PCA recreation | Dist sport club | 0.84 |
| | | Dist built heritage | 0.75 |
| H | PCA convivial | Dist bakery | 0.88 |
| | | Dist hotel/inn | 0.63 |
| | | Dist pub | 0.62 |
| I | PCA childcare | Dist day-care center | 0.62 |
| | | Dist bakery self-service | 0.57 |
| | | Dist primary school | 0.54 |
| J | PCA restaurant | Dist restaurants | 0.81 |
| K | PCA accessibility | Dist bus station | 0.93 |

The landscape and neighborhood components account for at least 75%, and the amenity components for 82% of their (correlation matrix) variation. These components are rotated using a VARIMAX procedure to produce uncorrelated factors. Overall, we identify groups of variables, which determine the value of different amenities.

Firstly, the neighborhood variables (Table 3) can be separated into three different components: *Low income*, *high income* and *shopping* neighborhoods. The low income component is characterized by a high *ratio of subsidized households*, *foreigners* and *unemployment*. Furthermore, large *playgrounds* can be observed, because there is nowhere for children to play, in contrast to areas with free-standing houses. A high ratio of *industrial areas* can be interpreted as a disamenity (pollution, noise, high traffic) which has a negative impact on apartment prices. Overall, the low income component explains 31% of the variation. High income neighborhoods in a city are characterized by a many natural amenities (a high *ratio* *of forest* and *farm area*, a low *ratio of high-traffic* and *building areas*) and high *income per taxpayer*. Thus, the *housereserve* (developing area) is large, so only taxpayers with a high income will build and live in this part of town. The component explains 28% of the variation. The shopping component is characterized by a high grade of *basic services* and a low *ratio of elderly* people. Therefore, it forms part of the business district, which explains 15% of the variation.

Secondly, the amenity variables (Table 4) are converted into eight components, where the centrality variables (e. g. *CBD*, *central station*, *cinema*, *discos*) explain, at 28%, most of the variation. Although the University of Münster has no central campus, its *canteens* are distributed around the central city and therefore form part of the centrality component. The other components remain close to 10%, where the accessibility (distance to the next *bus station*) has the lowest variance explanation at 4%. By using principal components, our OLS-model takes the following form:

$$log(P)_i = \alpha + \beta_{ij}O_{ij} + \delta_{ik}PCN_{ik} + \theta_{il}PCA_{il} + \eta_{im}T_{im} + \varepsilon \quad (2)$$

where *P* is the transaction price; *j*, *k*, *l and m* represent the number of attributes of object *i*; *α*, *β*, *γ*, *θ* and *η* are coefficients and ε is an error term. The attractiveness of a apartment depends on object attributes (O), neighborhood components (PCN), amenity components (PCA) and the time dummies (T). Before we use the component indices as OLS repressors, we have to consider spatial dependence, which can affect our results in various ways.

### 3.3 Spatial Dependence

If spatial autocorrelation arises from error terms and is therefore not independently distributed across space, the autocorrelation can lead to biased and inefficient standard errors. As Brasington and Hite (2005) show, spatial methods increase efficiency and consistency and reduce the bias of parameter estimates in hedonic analysis. To find the most likely spatial model for explaining the data, the LM-tests of Anselin (1988) and the robust LM-tests of Anselin et al. (1996) are used. A positive test suggests rejecting the regular OLS-model and using a Spatial Error Model (SEM) or a Spatial Lag Model (SAR) to account for spatial autocorrelation. Although LeSage and Pace (2009) and Elhorst (2010) recommend a Spatial Durbin Model (SDM) for interpretation purposes, we cannot include a large number of lag variables in our model, because we have only a few observations.

To control for spatial dependency, a neighborhood matrix is necessary, and its definition affects the test result. The distance between apartments is essential for the neighborhood matrix or rather the spatial model. Not every neighboring house has the same weight in the model. It would make more sense to have the weights decrease with distance. This is more robust to variability, due to houses receiving equal weight because they enter a neighborhood. Thus, we use the decay distance parameter 1/d. For example, if the apartment is 100 (200) meters away, it takes the value 0.01 (0.005). By doubling the distance, their weight will be halved. Based on Moran’s I (p-value = 0.000, statistic = 0.0368), we cannot reject spatial dependence. The Lagrangian multipliers suggest a spatial lag model (SAR).[6] Therefore, we

| [6] **Lagrange Multiplier** | **p-value (t-statistic)** |
|---|---|
| LM Spatial Error: | 0.000 (20.0609) |
| LM Spatial Lag: | 0.000 (75.778) |
| Robust LM Spatial Error: | 0.697 (0.1516) |
| Robust LM Spatial Lag: | 0.000 (55.8682) |

include the weighted neighborhood matrix W in our OLS-model to obtain spatial autocorrelation. The SAR-Model (3) shows that the price of apartments is related to the price of neighboring apartments.

$$log(P)_i = \alpha + \beta_{ij}O_{ij} + \delta_{ik}PCN_{ik} + \theta_{il}PCA_{il} + \eta_{im}T_{im} + \rho W log(P) + \varepsilon \quad (3)$$

The model induces a global form of spillovers, where the price setting depends partly on the neighborhood price settings (Anselin 2002). Thus, for example, the object attributes of neighboring apartments will have an effect on the observed apartment prices. To measure the spatial behavior of the data, we also construct a SEM-Model (4).

$$log(P)_i = \alpha + \beta_{ij}O_{ij} + \delta_{ik}PCN_{ik} + \theta_{il}PCA_{il} + \eta_{im}T_{im} + u \quad (4)$$

with $u = \lambda W u + e$

In comparison to our SAR-Model, the SEM-Model includes the neighborhood matrix in the error term which, follows a spatial autoregressive process. Thus positive externalities like tree-lined roads or a low crime rate, which are omitted in our data, are accounted for the model.

## 4. Results

### 4.1 Regressions

Table 5 provides the coefficient results of the three models. The estimated signs generally meet theoretical expectations. The *number of rooms*, the *size* and the *age squared* are positively related to the sale price. *Age* itself is negatively related to the sale price, confirming the hypothesized U-shaped relationship between price and age. Other quality and condition aspects of the apartments are also important. A *fireplace* or a garden has a significant positive impact on prices. If the apartment is *renovated* (or is *luxurious*, *has lots natural light* or is a *maisonette*), the price will increase. *Noise* is negatively related to the price. Only in our spatial models are the *balcony* or *one floor heating* (each floor has only one heating unit) significant variables. Thus, the observed apartment will benefit from the neighborhood, because all nearby apartments share the same pleasant view (*balconies*) or benefit from new housing technology (*one floor heating*).

Two of three neighborhood components are significant. If the neighborhood is characterized by *low income*, the value of the apartment is lower. *High income* does not have a significant impact on prices. A high basic supply of sales area in sqm (*shopping*) is positively related to the apartment prices, because its residents save transport costs in the context of food purchases. Six of eight amenity components are significant and negatively related to the price. The highest coefficient is related to the *centrality* and *culture* component. This confirms the relationship between price and the distance to the central business district, which is generally measured in hedonic house price analysis. Large distances (in a view of a higher component value) to food-outlets, recreation, childcare and restaurants are also significantly negatively related to prices.

The *time dummies* suggest that, compared to 2008, the prices in 2011 and 2012 increase significantly. This is associated with the housing bubble debate that commenced, only two years ago in German research and the media. Real estate is referred to as “Betongold” (concrete gold) as an alternative to real gold, for instance, due the Euro-crisis. In 2008, the low interest rate set by the European Central Bank and the Euro crisis of confidence lead to higher apartment prices in Münster.

Table 6 compares the three models. For the spatial models, we document log-likelihoods, as well as the likelihood ratio and Wald statistics for testing joint significance of the spatial parameters. They support the need to take spatial dependence into account. The inclusion of a spatial autocorrelation, as lag or error terms, leads to an increase in the $R^2$ (from 0.7964 (OLS) to 0.8132 (SEM) to 0.8224 (SAR)). The AIC and BIC decline as well. Given a Moran's I value of +0.036, the null hypothesis of spatially uncorrelated error terms has to be rejected, a result that is confirmed by the LM-test for error dependence.

**Table 5**: Regression coefficient results

| | OLS-Model (2) | | SEM-Model (4) | | SAR-Model (3) | |
|---|---|---|---|---|---|---|
| Variables | coefficient | t-statistic | coefficient | t-statistic | coefficient | t-statistic |
| Constant | **10.7700*** ** | 129.811 | **10.7740*** ** | 90,2139 | **3.2312*** ** | 3.5049 |
| Rooms | **0.1800*** ** | 12.514 | **0.1826*** ** | 13,5578 | **0.1804*** ** | 13.6688 |
| Size | **0.0045*** ** | 9.984 | **0.0041*** ** | 9,7370 | **0.0042*** ** | 10.0070 |
| Age | **-0.0119*** ** | -8.02 | **-0.0118*** ** | -8,2254 | **-0.0111*** ** | -8.1444 |
| Age (squared) | **0.0001*** ** | 6.952 | **0.0001*** ** | 7,2380 | **0.0001*** ** | 7.1862 |
| First time in use | **0.2329*** ** | 4.829 | **0.2345*** ** | 5,2027 | **0.2414*** ** | 5.4549 |
| Kitchen | -0.0189 | -0.666 | -0.0086 | -0,3252 | -0.0038 | -0.1448 |
| Loft | -0.0503 | -1.358 | -0.0351 | -1,0131 | -0.0348 | -1.0250 |
| Balkony | 0.0411 | 1.521 | 0.0333 | 1,3188 | **0.0416** ** | 1.6784 |
| Elevator | 0.0767 | 1.541 | 0.0435 | 0,9360 | 0.0437 | 0.9552 |
| Parking place | 0.0426 | 1.265 | 0.0443 | 1,3976 | 0.0426 | 1.3801 |
| Fireplace | **0.3060** ** | 3.199 | **0.2860** ** | 3,1097 | **0.2826*** ** | 3.2184 |
| One floor heating | 0.0709 | 1.409 | **0.0865* ** | 1,8325 | **0.0797* ** | 1.7262 |
| Basement | 0.0407 | 1.355 | 0.0365 | 1,3203 | 0.0328 | 1.1906 |
| Quiet | 0.0182 | 0.612 | 0.0066 | 0,2343 | 0.0097 | 0.3548 |
| Central heating | 0.0041 | 0.131 | 0.0131 | 0,4538 | 0.0104 | 0.3654 |
| Garden | **0.0893*** ** | 2.617 | **0.0922*** ** | 2,9258 | **0.0875*** ** | 2.7953 |
| Renovated | **0.1003** ** | 2.076 | **0.0914** ** | 2,0030 | **0.0941** ** | 2.1225 |
| Luxurious | **0.1470*** ** | 3.251 | **0.1178*** ** | 2,7738 | **0.1211*** ** | 2.9136 |
| Brightly | **0.0504* ** | 1.659 | **0.0473* ** | 1,6758 | **0.0538* ** | 1.9273 |
| Floor heating | 0.0274 | 0.46 | -0.0035 | -0,0624 | -0.0214 | -0.3888 |
| Loggia | -0.0024 | -0.053 | -0.0070 | -0,1622 | -0.0073 | -0.1724 |
| New building | 0.0487 | 0.771 | 0.0423 | 0,7193 | 0.0475 | 0.8198 |
| Maisonette | **0.1664*** ** | 2.953 | **0.1734*** ** | 3,2501 | **0.1825*** ** | 3.5274 |
| Parquet floor | 0.0243 | 0.529 | -0.0002 | -0,0056 | -0.0012 | -0.0280 |
| Population density | -0.0006 | -1.092 | -0.0006 | -1,0976 | -0.0004 | -0.7997 |
| Households with children | 0.0000 | -0.039 | 0.0001 | 0,4188 | 0.0001 | 0.6664 |
| Ratio family households | 0.1583 | 0.533 | 0.1320 | 0,3969 | -0.2702 | -0.9763 |
| Noisy | **-0.1376*** ** | -3.316 | **-0.1300*** ** | -2,9477 | **-0.0915** ** | -2.3712 |
| PCN low income | **-0.1043*** ** | -5.232 | **-0.0950*** ** | -4,0711 | **-0.0659*** ** | -3.4819 |
| PCN high income | -0.0293 | -0.87 | -0.0281 | -0,7550 | -0.0153 | -0.4954 |
| PCN shopping | **0.0491*** ** | 3.095 | 0.0260 | 1,2052 | **0.0312** ** | 2.1324 |
| PCA centrality | **-0.1353*** ** | -5.486 | **-0.1351*** ** | -4,3856 | **-0.0861*** ** | -3.6533 |
| PCA culture | **-0.1146*** ** | -5.705 | **-0.0893*** ** | -3,5394 | **-0.0614*** ** | -3.1671 |
| PCA food | **-0.0497*** ** | -3.418 | **-0.0566*** ** | -3,4621 | **-0.0419*** ** | -3.1303 |
| PCA recreation | **-0.0775*** ** | -4.773 | **-0.0656*** ** | -2,9791 | -0.0223 | -1.3595 |
| PCA convivial | -0.0179 | -1.124 | -0.0251 | -1,2791 | -0.0126 | -0.8593 |
| PCA childcare | **-0.0422** ** | -2.384 | -0.0279 | -1,2864 | -0.0218 | -1.3253 |
| PCA restaurant | **-0.0618*** ** | -3.98 | **-0.0578*** ** | -3,1161 | **-0.0396*** ** | -2.7348 |
| PCA bus station | 0.0137 | 1.033 | 0.0223 | 1,6213 | 0.0141 | 1.1594 |
| Dummy date 2009 | -0.0229 | -0.562 | -0.0097 | -0,2559 | 0.0011 | 0.0301 |
| Dummy date 2010 | 0.0606 | 1.427 | 0.0545 | 1,3665 | 0.0612 | 1.5730 |
| Dummy date 2011 | **0.0930** ** | 2.349 | **0.0935** ** | 2,5342 | **0.0916** ** | 2.5231 |
| Dummy date 2012 | **0.1550*** ** | 3.523 | **0.1685*** ** | 4,0572 | **0.1639*** ** | 4.0598 |

Note: The symbols *, **, and *** denote significance at the 10, 5, and 1% levels, respectively.

**Table 6**: Regression comparison

| Indicators | OLS-Model (2) | SEM-Model (4) | SAR-Model (3) |
|---|---|---|---|
| $R^2$ | 0.7964 | 0.8132 | 0.8224 |
| AIC | 363.165 | 339.944 | 308.23 |
| BIC | 563.436 | 544.666 | 512.952 |
| Moran's I of residuals | 0.0368*** | 0.0066 | -0.0063 |
| Log-likelihood | | -123.9723 | -108.1152 |
| Likelihood ratio | | 25.221*** | 57.149*** |
| Wald statistics | | 120.2*** | 69.649*** |

Furthermore, the likelihood ratio and Wald test suggest taking spatial dependence into account. The SEM and SAR models capture spatial dependence, as indicated by LM-tests for error dependence and Moran's I for the residuals. Furthermore, the coefficients have a low level of multicollinearity (all variance inflation factors are below 8, as Table A. 1 in the appendix suggests). Summarizing the SAR model is most suitable for explain the data.

### 4.2 Relative Importance

Until now, we have gathered a lot of information about three explanatory dimensions: object attributes, location and time. In summary, we can explain around 80% percent of the price variation. The spatial models improve our results marginally (in view of our $R^2$, AIC and BIC). The signs of the coefficients do not change, only the p-values are partly lower. Altogether, we know a lot about interrelated signs of the coefficients and the behavior of different variables, as well as their impact on real estate prices, but not their relative importance. Johnson and Lebreton (2004) define relative importance as

> *"the proportionate contribution each predictor makes to $R^2$, considering both its direct effect (i.e., its correlation with the criterion) and its effect when combined with the other variables in the regression equation".*

Therefore, we consider the relative importance of location. To investigate this, we focus on the $R^2$ of our OLS-Model. For spatial models, there is no valid method for computing the relative importance. However we believe that the results will only change marginally for the spatial models. Following Grömping (2009), there are four criteria for decomposing $R^2$. Firstly, the model has to be decomposed into shares, which are in sum, the model variance (proper decomposition). Secondly, all shares have to be non-negative (non-negativity). Thirdly, regressors with a $\beta = 0$ should have a zero share (exclusion) and finally, a regressor with $\beta \neq 0$ should receive a non-zero share (inclusion). Grömping (2009) recommends two computer-intensive methods: LMG (developed by Lindeman, Merenda and Gold 1980) and PMVD (Proportional Marginal Variance Decomposition by Feldman 2005). LMG allocates for instance to *rooms*, the average over all allocations to *rooms* from all possible orderings of the regressors. This gives each order of regressors the same weight, which is data-independent. PMVD calculates β-weighted averages of the regressors. Each order of regressors receives a data-dependent weight (Feldman 2005). We prefer the PMVD method, because it complies with the criteria of exclusion and inclusion (Grömping 2009). As a result, some regressors are more dominant. Moreover, we use the LMG-results to control our PMVD-results.

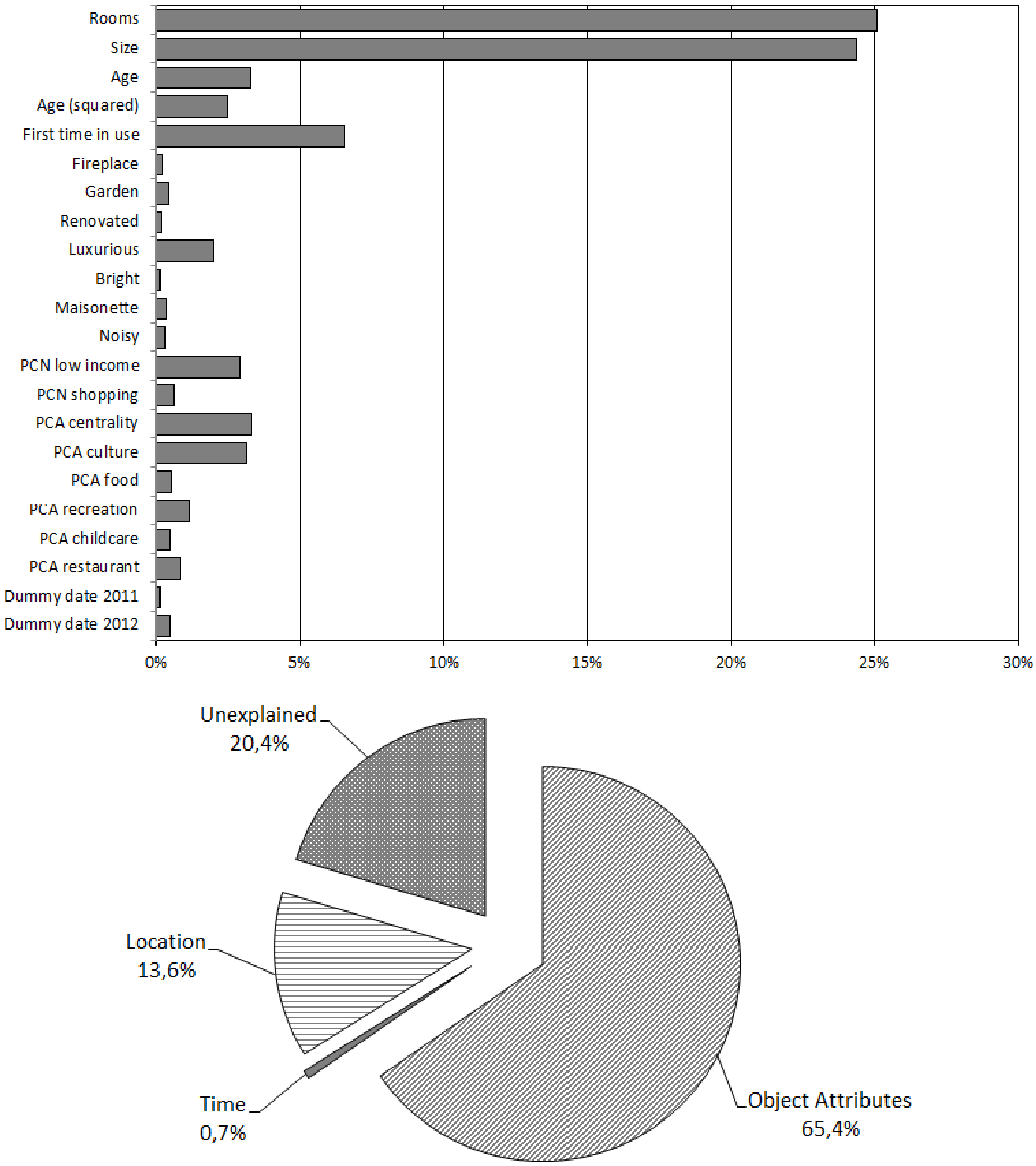
To measure the relative importance, it is essential that the correlation between each variable be low, and that only significant variables are used (Johnson and Lebreton 2004). The principal components lead to low variance inflation factors, so that our 22 significant variables are suitable for this method. Figure 2 shows the relative importance of the regressors for apartment prices based on PMVD.[7]

Some variables explain a high proportion of $R^2$. Especially *rooms*, *size*, *age*, *first time in use*, *luxurious*, *poor neighborhood*, *centrality* and *culture* explain most of the price variations. But what is the difference between a coefficient and the relative importance? For example, *a fireplace* has a positive and significant coefficient of 0.3060, whereas *first time in use* has only a coefficient of 0.2329. Yet, the relative importance of *fireplace* is only 0.24% and *first time in use* explains 6.60% of $R^2$. The residents do not consider the incremental amount of a *fireplace*, while holding other attributes constant. They consider simultaneously all aspects that are important to them and implicitly weight each aspect relative to the others in determining their overall willingness to pay.

Altogether, the location-based parameters do not have a high explanatory value. In sum, they explain only 13.6% of the price variation, while the object attributes explain most with 63.2%. Hence, “location, location, location” does not in fact reflect the important determinants of apartment prices. The ranking is more like “size, age, location”, but even this is only half the story. We only focus on the micro level of Münster. Rural apartments are not observed in the data, hence location parameters are only assigned for the real estate market of the city of Münster. Therefore, location parameters (like the distance to the CBD) can be explained more comprehensively, if we compare apartments between rural and urban regions. If a large-scale regional data set can be explained by one model, the role of location may increase. An investor, who wants to buy an apartment in Münster, should evidently not overestimate the role of location.

[7] For LMG, we observe marginal differences. Location explains 15,5%, time 0,6% and the object attributes 63,6% of the variation. For more details, see Figure A.1 and Table A.2 in the appendix.

**Figure 2**: Shares of relative importance for apartment prices (method PMVD)




Note: The $R^2$ of the OLS with (only) significant variables is 79.64

## 5. Conclusion

In our hedonic models, we outline the role of different location parameters for real estate prices. Our results underline the negative relationship between a large distance to amenities and apartment prices. Furthermore, location is not the main price determinant. Object attributes like size and age explain most of price variations. Thus, we suggest “size, age, locaction” instead of “location, location, location” as the fundamental price determinants within a bike-city. However, the relative importance of location parameters may rise, if rural data is included. Therefore more research on these issues is clearly necessary.

## Acknowledgements

We would like to thank F+B Forschung und Beratung für Wohnen, Immobilien und Umwelt GmbH, particularly Bernd Leutner, for providing us with a dataset of apartment prices of Münster. We would also like to thank Martin Zarth and Hans Uhlit at the City Administration of Münster for supplying us with an extensive georeferenced dataset. Also we like to thank Ulrich van Suntum, Edzer Pebesma and Ulrike Grömping for their helpful comments. Thanks too are due to Brian Bloch for his English editing.

## References

Ahlfeldt, G. M., Maenning, W. 2010. Substitutability and Complementarity of Urban Amenities: External Effects of Built Heritage in Berlin, *Real Estate Economics 38(2)*, pp. 285-323.

Alonso W. 1964. Location and Land Use Harvard University Press, Cambridge.

Andersson, H., Jonsson, L., Ögren, M. 2009. Property prices and exposure to multiple noise sources: hedonic regression with road and railway noise, *Environmental Resource Economics 45(1)*, pp. 73-89.

Anselin, L. 1988. Spatial Econometrics: Methods and Models, Dordrecht, Kluwer.

Anselin, L. 2002. Under the hood – Issues in the specification and interpretation of spatial regression models, *Agricultural Economics 27*, pp. 247-267.

Anselin, L., Bera, A. K., Florax, R. and Yoon, M. J. 1996. Simple diagnostics tests for spatial dependence, *Regional Science and Urban Economics 26(1)*, pp. 77-104.

Ball, Michael J. 1973. Recent Empirical Work on the Determinants of Relative House Prices, *Urban Studies 10*, pp. 213-233.

Blomquist, G., Worley, L. 1981. Hedonic prices, demands for urban housing amenities, and benefit estimates, *Journal of Urban Economics 9(2)*, pp. 212-221.

Brandt, S., Maenning, W. 2011. Road noise exposure and residential property prices: Evidence from Hamburg, *Transportation Research Part D 16*, pp. 23-30.

Brasington, David M., Hite, Diane 2005. Demand for environmental quality: a spatial hedonic analysis. *Regional Science and Urban Economics 35*, pp.57-82.

Dziuban, Charles D., Shirkey, Edwin C. 1974. When is a correlation matrix appropriate for factor analysis? Some decision rules, *Psychologial Bulletin Vol 81(6)*, pp.358-361.

Elhorst, J. Paul 2010. Applied Spatial Econometrics: Raising the Bar, *Spatial Economic Analysis 5(1)*, pp. 9-28.

Feldman B. 2005. Relative Importance and Value, manuscript version 1.1, ( http://www.prismanalytics.com/docs/RelativeImportance.pdf , 04-01-2013).

Glaeser, Edward L., Kolko, Jed; Saiz, Albert 2001. Consumer City, *Journal of Economic Geography 1(1)*, pp. 27-50.

Grömping, Ulrike 2009. Variable Importance Assessment in Regression: Linear Regression versus Random Forest, *The American Statistician 63(4)*, pp. 308–319.

Johnson, J.W., Lebreton, J. M. 2004. History and Use of Relative Importance Indices in Organizational Research, *Organizational Research Methods 7*, pp. 238–257.

LeSage, J., Pace, R. K. 2009. Introduction to Spatial Econometrics, Taylor & Francis Group, LLC.

Li, Mingche M., Brown, H. James 1980. Micro-Neighborhood Externalities and Hedonic Prices, *Land Economics 56(2)*, pp.125-141.

Lindeman, R. H., Merenda, P. F., Gold, R. Z. 1980. Introduction to Bivariate and Multivariate Analysis, Glenview, IL: Scott, Foresman.

Muellbauer, J. 1974. Household Production Theory, Quality, and the "Hedonic Technique", *American Economic Review* 64(5), pp. 977-994.

Noonan, Douglas S. 2007. Finding an Impact of Preservation Policies: Price Effects of Historic Landmarks on Attached Homes in Chicago, 1990-1999. *Economic Development Quarterly 21(1)*, pp. 17-33.

Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy 82(1)*, pp. 34-55.

Sirmans, G. S., Macpherson, D. A., Zietz, E. N., 2005. The composition of hedonic pricing models, *Journal of Real Estate Literature 13(1)*, pp. 3-43.

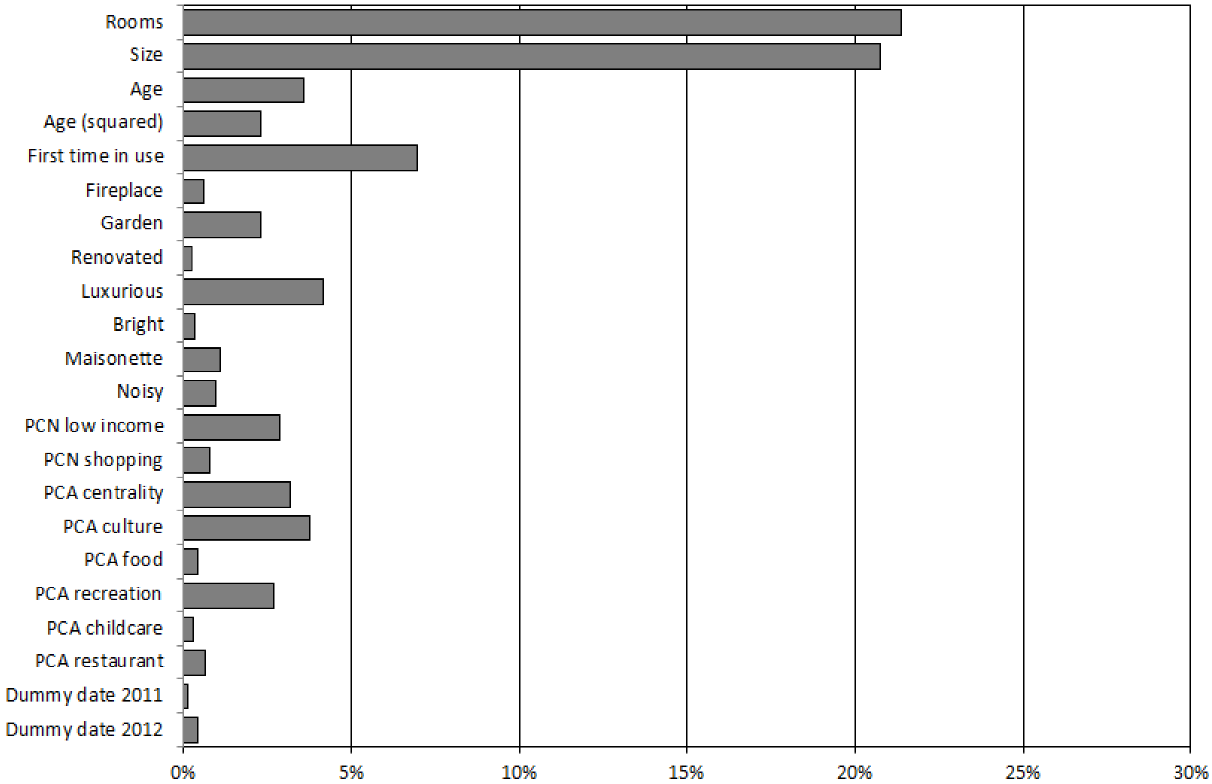
University of Münster 2005. Politbarometer Münster Häufigkeitsauszählung 3, ( http://bema.uni-muenster.de/barometer/pdf/haeufigkeiten/2005-3_Haeufigkeiten.pdf , 01-13-2013).

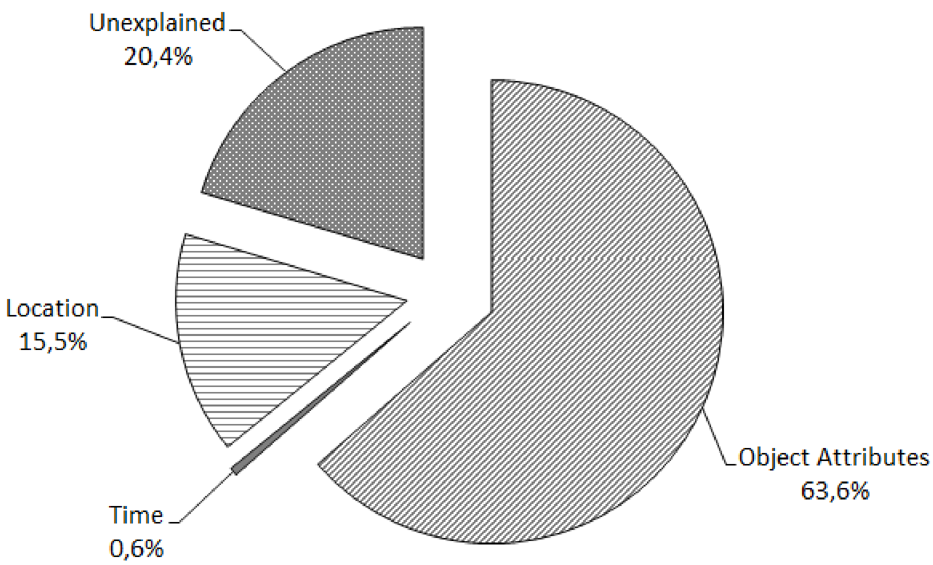
## Appendix

**Table A.1**: Variance inflation factors of model (2)

| Variance inflation factor | | | |
|---|---|---|---|
| ***Rooms*** | ***2.517467*** | ***Maisonette*** | ***1.113504*** |
| ***Size*** | ***2.528619*** | Parquet floor | 1.161690 |
| ***Age*** | ***6.909526*** | Population density | 2.942633 |
| ***Age (squared)*** | ***5.210640*** | Households with Children | 3.300579 |
| ***First time in use*** | ***1.641588*** | Ratio family households | 3.209623 |
| Kitchen | 1.246547 | ***Noisy*** | ***1.268530*** |
| Loft | 1.161656 | ***PCN low income*** | ***2.593662*** |
| Balkony | 1.172207 | PCN high income | 7.382646 |
| Elevator | 1.133247 | ***PCN shopping*** | ***1.640002*** |
| Parking place | 1.430739 | ***PCA centrality*** | ***3.968922*** |
| ***Fireplace*** | ***1.112031*** | ***PCA culture*** | ***2.632975*** |
| One floor heating | 1.523426 | ***PCA food*** | ***1.377366*** |
| Basement | 1.342086 | ***PCA recreation*** | ***1.721568*** |
| Quiet | 1.115499 | PCA convivial | 1.659065 |
| Central heating | 1.463963 | ***PCA childcare*** | ***2.040492*** |
| ***Garden*** | ***1.197209*** | ***PCA restaurant*** | ***1.573055*** |
| ***Renovated*** | ***1.250988*** | PCA bus station | 1.144212 |
| ***Luxurious*** | ***1.346797*** | Dummy date 2009 | 1.700901 |
| ***Brightly*** | ***1.140309*** | Dummy date 2010 | 1.761961 |
| Floor heating | 1.212307 | ***Dummy date 2011*** | ***1.901951*** |
| Loggia | 1.077047 | ***Dummy date 2012*** | ***1.655086*** |
| New building | 1.400862 | | |

**Figure A.1**: Shares of relative importance for apartment prices (method LMG)




Note: The $R^2$ of the OLS with (only) significant variables is 79.64

**Table A.2:** Variance inflation factors of 22 significant variables for the PMVD/LMG method

| Variance inflation factor | | | |
|---|---|---|---|
| **Rooms** | **2.385525** | **Noisy** | **1.133975** |
| **Size** | **2.381177** | **PCN low income** | **1.835748** |
| **Age** | **5.410911** | **PCN shopping** | **1.330356** |
| **Age (squared)** | **4.646479** | **PCA centrality** | **1.132052** |
| **First time in use** | **1.431833** | **PCA culture** | **1.100446** |
| **Fireplace** | **1.070660** | **PCA food** | **1.047428** |
| **Garden** | **1.120334** | **PCA recreation** | **1.307264** |
| **Renovated** | **1.185718** | **PCA childcare** | **1.569857** |
| **Luxurious** | **1.187413** | **PCA restaurant** | **1.205348** |
| **Brightly** | **1.083022** | **Dummy date 2011** | **1.129814** |
| **Maisonette** | **1.065476** | **Dummy date 2012** | **1.142102** |

If we omit *rooms* and *age (squared)*, the variance inflation factors are all below 2. This reduced OLS-Model explains 70.69% ($R^2$) of the price variation (method LMG). The object attributes explain 53.86%, the location parameters 16.37%, and the time dummies 0.46%.