Property valuation using hedonic price method
– procedure and its Application

Abstract
This paper discusses one of the property valuation methods, i.e. property valuation using the hedonic price method, which makes use of a classical linear regression model. The study characterises the calculation procedure of the selected method and indicates the fields of application of the hedonic approach in the construction sector. The operation of the hedonic price method is presented based on a valuation of flats in Lublin.

Keywords: property, property valuation, hedonic price method, classical linear regression model, hedonic regression

Streszczenie
Niniejszy artykuł traktuje o jednej z metod wyceny nieruchomości, tj. hedonicznej metodzie wyceny nieruchomości, która wykorzystuje klasyczny model regresji liniowej. W pracy scharakteryzowano procedurę obliczeniową wybranej metody oraz wskazano obszary stosowania podejścia hedonicznego w budownictwie. Działanie metody hedonicznej wyceny przedstawiono na przykładzie wyceny lokali mieszkalnych w Lublinie.

Słowa kluczowe: nieruchomość, wycena nieruchomości, metoda cen hedonicznych, klasyczny model regresji liniowej, regresja hedoniczna
1. Introduction

The property is defined as a controlled asset constituting a specific and defined space of the land together with its natural and anthropogenic (created by man) elements [1].

In the academic literature, property is considered in three main aspects as:

▶ a technical object analysed in terms of its components that comprise land and building structures together with equipment subject to wear and tear, which in turn determine the need to take actions related to maintenance of their technical efficiency and usability, their safety of use and investment with a reconstruction and development character [2],

▶ economic assets regarded as the subject of capital allocation, capital asset, utility asset and market asset [1–3],

▶ a legal object that constitutes a basis for classification of the property in legal terms, as well as for transfer of rights and property management [2].

Due to the above-mentioned economic aspect and conditioning inherent to scientific research work, the authors of the hereby publication focused their attention on the aspects related to the generation of economic benefits obtained by an investor in the form of the market value of the property.

The main purpose of this paper is to present one of the property valuation methods, i.e. property valuation using hedonic price method, which makes use of the classical linear regression model. In their study, the authors characterise the calculation procedure of the method and indicate its fields of application in the construction sector. The operation of this method is illustrated on the basis of an exemplary valuation of flats.

2. Hedonic method of property valuation

2.1. General characteristics of the hedonic method of property valuation

Methods for valuation of assets are divided into direct ones as well as indirect ones. The hedonic price method, also called the hedonic regression method, is a kind of indirect valuation method. It belongs to the family of methods based on surrogate or replacements markets [4]. The property is valuated by means of evaluation of its specific features and attributes with regards to the whole market. Analysing the regularities statistically, and more specifically, the relationship between the price and features of the property, a model is made based on that relationship, on the basis of which it is possible to estimate the price of the given property.

The purpose of using the hedonic price method is to specify the influence of the features of the property on its value. Hedonic regression is a method, thanks to which we may define the impact of the specific features of the property on its value [5]. Based on the regression method, a function determining the relation of its features to the price is created.

Table 1 presents the advantages and disadvantages of the hedonic approach to valuation.
Table 1. Advantages and disadvantages of the hedonic approach

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freedom of model creation based on available features</td>
<td>Necessity to provide substantial amount of data</td>
</tr>
<tr>
<td>(at the stage of model creation)</td>
<td>(at the stage of model creation)</td>
</tr>
<tr>
<td>Universality and versatility (in case when the model is</td>
<td>No possibility to make allowance for external factors, i.e. interest rate or</td>
</tr>
<tr>
<td>constructed we may valuate many properties using it)</td>
<td>political situation</td>
</tr>
<tr>
<td>Possibility of model updating and making corrections</td>
<td>No possibility to make allowance for the features identified as unit features</td>
</tr>
<tr>
<td>(the originally developed equation may be theoretically still valid)</td>
<td>or features of unique nature</td>
</tr>
<tr>
<td>Ease of use</td>
<td>Necessity to possess knowledge within the field of mathematical statistics</td>
</tr>
<tr>
<td></td>
<td>to be able to build and verify the model, as well as to interpret the results</td>
</tr>
</tbody>
</table>

Source: own study based on [5–7].

As results from the literature, the hedonic approach (although not without some disadvantages) may be treated by economists as more reliable, since it is based on valuations disclosed on the real market.

2.2. Definition and calculation procedure of the hedonic valuation model

The relationship between the price of the asset (property) $P_i$ and the set of its features (characteristics) [8, 9] is called a hedonic model or hedonic regression. In general, this relationship may be described with the use of the following function:

$$P_i = f(X_i, \alpha, \varepsilon_i)$$  \hspace{1cm} (1)

where:

- $X_i$ – vector of asset attributes,
- $\alpha_i$ – vector of model parameters,
- $\varepsilon_i$ – random variable.

In the hedonic price method, the valuation process consists in decomposition of the price of the asset into combinations of the specific characteristics, which reflect the importance when it comes to pricing. In fact, only the specific features of the asset are valuated, and not the asset itself.

The assumptions of the hedonic price method are similar to the assumptions known from the comparative approach, but in this case, they are more general. It results from the fact that, in the case of the hedonic approach, the importance of the specific features is shaped in relation to the total sample. It also means that the constructed model corresponds to all the observations and is universal.

The hedonic approach to property valuation is versatile, which means that, for example, we may encounter one model for property valuation within the precincts of one city or several models that have been constructed making a distinction for the specific districts. At the same time, the model composed of many models is much more detailed and gives a result, which is better adjusted to the data.
The calculation procedure of the hedonic price method may be divided into two basic stages, during which the following take place:

1) collection of data on transactions regarding the given assets (in this case property),
2) statistical estimation of the linear function that will describe the relationship between the value of the property (response variable) and the specific features having a potential impact on price (explanatory variables).

In order to make it possible to estimate the price of any property using the created function, we must note that the data on transitions regarding the property should include information both concerning the sales prices, location, as well as other significant measurable and immeasurable properties connected with the property (e.g. size, number of rooms, characteristics of the neighbouring areas, standard of finish, crime rate, environmental aspects, distance from shopping centres, etc.). Therefore, as early as at the stage of data collection, it is very important to properly process the collected data in terms of consistency, coherency and completeness of information [8].

2.3. The fields of application of the hedonic approach in the construction sector

Hedonic methods are used in the construction sector inter alia for:

▶ property valuation in case of which it is possible to estimate the price of the specific properties based on the constructed models,
▶ valuation of the specific elements (attributes), e.g. balcony in the given flat or the fact of using advanced technologies in the building,
▶ valuation of the qualitative aspects, e.g. through examination of changes between the values of the specific features that are not binominal variables [10],
▶ valuation of environmental losses or benefits on the degraded areas that may be subject to recultivation measures [9],
▶ research of market preferences to determine the importance and the level of their importance during the course of shaping the prices of properties [8],
▶ valuation of the potential health risks accompanying dangerous professions that consists in specifying the amount of additional pay for a worker for bearing additional health or life-threatening risks during work [6].

3. Characteristics of a single-equation classical linear regression model

3.1. Definition of a linear regression model

The basic model describing the relationship between the phenomena (variables) is a single-equation linear model, also called descriptive econometric model, the form of which is as follows [11]:

\[ y_i = \alpha_0 X_0 + \alpha_1 X_1 + \ldots + \alpha_k X_k + \epsilon_i \]  

(2)
where:
\[ y_i \quad – \quad \text{endogenous variable (response variable)}, \]
\[ X_{0}, X_{1}, \ldots, X_k \quad – \quad \text{exogenous variables (explanatory)}, \]
\[ \alpha_0, \alpha_1, \ldots, \alpha_k \quad – \quad \text{structural parameters}, \]
\[ \varepsilon_j \quad – \quad \text{random variable}. \]

The above-mentioned equation for a response variable is composed of a sum of two components, i.e. certain linear combination of explanatory variables \( X_{0}, X_{1}, \ldots, X_k \) and a random variable \( \varepsilon_j \). The explanatory variables have a significant influence on shaping \( y_i \), whereas the random variable \( \varepsilon_j \) (random disturbances) makes allowance for the total impact of other factors (excluded after analysis), not present in the model, which have an impact on shaping the response variable \( y_i \), however, of a random (secondary) nature. The \( \varepsilon_j \) coefficient also provides for any possible, non-systematic, random errors of measurement of variables, as well as any deviations from the adopted analytical form of the model from the actual relationship between them [12].

3.2. Model estimation

The \( \alpha_0, \alpha_1, \ldots, \alpha_k \) coefficients that are present in the equation (2) as the so-called structural parameters are not known in the linear regression model. Their task is to make a quantitative description of the impact of the explanatory variables (next to which they are located) on the response variable. Most frequently, one of the exogenous variables (usually the first one) is defined to be identified as \( X_0 \equiv 1 \). Thus, \( \alpha_0 \) is called an absolute term of the model or regression constant [12]. Hence estimation (determination) of structural parameters \( \alpha_j \) \((j = 0, 1, \ldots, k)\) based on observation of the variables \( y_i, X_{1}, \ldots, X_k \) becomes the main task in the construction of a linear regression model.

Due to the available \( n \)-element observation sequence for all the variables or in other words vector sequence \((y_t, x_{1t}, \ldots, x_{kt}) \ (t = 1, \ldots, n)\), every realization \( y_t \) of the response variable \( y \) may be presented (pursuant to the assumed model) as a sum of a linear combination \( \alpha_0 + \alpha_1 x_{1t} + \ldots + \alpha_k x_{kt} \) of the adequate realization of explanatory variables and unobservable realization \( \varepsilon_t \) of a random component \( \varepsilon_t \) [11, 12]. Hence we obtain the following arrangement constituting a starting point for estimation of model parameters:

\[
\begin{align*}
    y_t &= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \\
    X_t &= \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{nt} & \cdots & x_{nk} \end{bmatrix}, \\
    \alpha_j &= \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix}, \\
    \varepsilon_t &= \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_k \end{bmatrix},
\end{align*}
\]

The determination of model parameters takes place on the basis of a structural and statistical form of the model that is presented in the arrangement (3). Except that, the probability conditions must be met, so that estimation of parameters makes sense.

The assumptions for the linear model are as follows [11, 12]:

\[
\]
1) \( y_i = X_i \alpha + \varepsilon_i \), i.e. every observation \( y_i \) is a linear function of \( x_{ij} \) observation and the random component \( \varepsilon_i \).

2) \( X \) is a non-random matrix, thus the explanatory variables are non-random variables and their values are derived from observations for \( t = 1, 2, \ldots, n \).

3) \( r(X_i) = k + 1 \leq n \), which means that the observation matrix \( X_i \) has got a full column rank, namely the observation matrix vectors \( X_i \) (columns) are linearly independent and no collinearity of the explanatory variables occurs, and moreover the number of the explanatory variables together with the absolute term \( \alpha_0 \) is lower than the number of observations,

4) \( E(\varepsilon_j) = 0 \), i.e. the value expected from the random component equals 0, which means that the multidirectional disturbances are reduced,

5) \( D^2(\varepsilon_j) = \sigma^2 \), i.e. the variance of random components is constant for the entire sample, whereas random components of the observation are not correlated with each other and \( \sigma^2 < +\infty \),

6) \( \varepsilon_j \sim N_n \), which means that the random component \( \varepsilon_j \) is characterised by an \( n \)-dimensional normal distribution.

The above-mentioned assumptions form the so-called classical model of a normal linear regression, which is most often based on the so-called classical least squares method. In order to be able to use this method for estimation of the structural parameters \( \alpha_j \) (\( j = 0, 1, \ldots, k \)) on the basis of observation of variables \( y_i, X_i \), all the six assumptions described above must be absolutely fulfilled [11, 12].

### 3.3. Model verification

Model verification consists in assessment of adaptation of the model to empirical data and the quality of the estimations of structural parameters [11, 12].

The quality of calculation, namely the fact how well the regression line reflects the reality (to what extent it corresponds to the observations), may be verified inter alia by means of \( R^2 \) determination factor. The value of the factor specifies what part of variation of the response variable \( y \) has been explained by the variation of all the explanatory variables \( x_1 + x_2 + \ldots + x_n \) [12].

We should note that the \( R^2 \) determination factor assumes values within the range \(< 0; 1 >\). It means that if \( R^2 = 1 \), then the linear regression function explains in 100% the variation of the response variable \( y \). If \( R^2 = 0 \), we should assume that the linear regression equation doesn't describe or explain the variation of the endogenous variable.

Other measurement methods, thanks to which we may find out how well the regression line corresponds to empirical observations include:

- residual variance that is based on examination of residual component variance [13],
- standard deviation of residuals that reflects the mean difference between the observed values of the described variable and the theoretical values [14],
- significance tests are conducted to determine, whether the given parameter assigned to the factor has got a real impact on the examined response variable [11, 13].
4. An example of application of the classical linear regression model in the hedonic method for valuation of flats

4.1. Data and assumptions adopted for the model

Functioning of the hedonic regression model will be presented based on the example of theoretical prices for 1 m$^2$ of usable area of a flat in Lublin.

Based on data collected from an expert (property valuer) an initial, 12-element set of features has been prepared, which constitute a set of analysed variables. The set referred to above is presented in Table 2.

<table>
<thead>
<tr>
<th>No.</th>
<th>Name of feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>date of transaction</td>
<td>an exact date of entering into agreement on the transfer of ownership (from the 1st quarter of 2009 to the second quarter of 2013)</td>
</tr>
<tr>
<td>2</td>
<td>area, district</td>
<td>one of four districts of Lublin, for which the data have been collected (Czuby, LSM, Rury, Śródmieście)</td>
</tr>
<tr>
<td>3</td>
<td>type of right acquired</td>
<td>a binominal feature referring to the right to the property (0 – cooperative ownership right, 1 – ownership)</td>
</tr>
<tr>
<td>4</td>
<td>floor</td>
<td>a feature divided into 6 floors (0 – ground floor, 1–1st floor, …, 4–4th floor, 5 – over 4th floor)</td>
</tr>
<tr>
<td>5</td>
<td>usable area</td>
<td>a measurable feature informing on the usable area of the flat</td>
</tr>
<tr>
<td>6</td>
<td>total flat price</td>
<td>a measurable feature informing on the transaction price for the flat together with belonging premises (if any)</td>
</tr>
<tr>
<td>7</td>
<td>price for 1 m$^2$ of usable area</td>
<td>a feature that has been created based on combination of features related to the price and floor area (as a result of dividing the first one through the second one)</td>
</tr>
<tr>
<td>8</td>
<td>storey</td>
<td>directly connected with the floor feature</td>
</tr>
<tr>
<td>9</td>
<td>number of overground floors</td>
<td>a feature characterising type of building (one-storey, multi-storey), related to the floor feature</td>
</tr>
<tr>
<td>10</td>
<td>year of construction</td>
<td>a feature described by a number of time intervals (0–50-ties and 60-ties, 1–70-ties, …, 5 – years from 2010 to 2019)</td>
</tr>
<tr>
<td>11</td>
<td>seller</td>
<td>a feature characterising the market, described binominally (0 – developer, primary market, 1 – private person, secondary market)</td>
</tr>
<tr>
<td>12</td>
<td>belonging premises</td>
<td>a binominal feature informing, whether a basement belongs to the flat (0 – no, 1 – yes)</td>
</tr>
</tbody>
</table>

Source: own study.

As a result of the analysis of variables, one response variable $y_i$ has been selected, i.e. the price of 1 m$^2$ of usable area of a flat and six explanatory variables, i.e. area (district), type of a right
acquired, floor, year of construction, belonging premises and seller. The date of transaction has been arbitrarily deemed to be insignificant because the information about date of entering into agreement on the transfer of ownership was included in the short four years’ time interval for all observations. The following has also been excluded from the final set of features adopted for the model:

- usable area and total flat price, since both these features are included the response variable $y_i$ (price for 1 m$^2$ of usable area of a flat),
- storey and number of overground floors, since both these features are directly related to the explanatory variable for the floor.

The following assumptions have been adopted to construct a model:
3) the created model is based on 1211 observations that have been properly prepared and processed,
4) all the qualitative variables have been expressed in qualitative terms,
5) it has been verified for each observation, whether the random component has got a normal distribution with a mean equal to 0 and a standard deviation $\sigma$, and whether it is independent from the random components related to all other observations,
6) it has been verified for each observation, whether the random component (error) is independent from other random errors (fulfilment of this limitation is not significant from the point of view of model construction, but enables to use $t$ and $F$ tests at the stage of verification of model operation),
7) $y$ response variable has been expressed as logarithm data to normalise the distribution of residuals, which is a condition for use of a least squares method as an estimator,
8) to calculate the price for 1 m$^2$ of usable area of a flat exponential function has been used, i.e. $\exp(y)$.

4.2. Estimation and verification of a hedonic regression model

The least squares method was used to determine all the structural parameters $\alpha_j$ ($j = 0, 1, \ldots, 8$) based on observations of $y_i$, $X_{i,j}$, $X_s$ variables.

Due to strong collinearity, one of the variables related to the Śródmieście area has been omitted at the very beginning. Collinearity means that explanatory variables are strongly correlated with each other. It turned out in the constructed model that the Śródmieście variable was so strongly correlated with other variable that it allowed for its exclusion from the model because it provided no additional information on $y_i$ variable [13].

Table 3 presents the results of estimation of structural parameters of the model for the following variables: Czuby, LSM, Rury, type of right acquired, floor, year of construction, belonging premises and seller. GRETL statistical package has been used in calculations.

The test for the normality of distribution of the residuals has confirmed that the random component is characterised by normal distribution. The value of the test $\chi^2$-square $\chi^2 = 381.906$ has been obtained, with the value of test probability $p = 1.1755e-083$. The $R^2$ determination factor has amounted to 0.2395666, which at the assessment of model adaptation means that, the $y_i$ variable has been explained by the factors included in the model.
only in ca. 24%. The standard deviation of residuals, illustrating the mean difference between the observed values of the response variable and the theoretical values, has amounted to \( s = 0.172223 \). The quotient of standard deviation and the arithmetic mean of the dependent variable has resulted in the value of the factor of the random variation of the residuals \( V_ε = 0.02 \), which is tantamount with the fact that the random component decides about the price for 1 m\(^2\) of usable area of a flat calculated with the used of the model only in 2%.

The performed statistical analysis has indicated that all the variables are significant, apart from the \( X_5 \) variable (floor), for which \( α_5 = -0.00238505 \). Taking into consideration that the standard error amounts to 0.00320693 for the given variable, and that the absolute value of the testing statistic barely 0.7437 (no grounds for rejecting the null hypothesis that means no significant impact of this factor on the response variable), hence we can exclude this variable from the constructed model.

The final form of the hedonic regression model is as follows:

\[
\ln (y_t) = 8.13521 - 0.138665 \cdot X_1 - 0.139852 \cdot X_2 + 0.0474448 \cdot X_3 + \\
- 0.0638818 \cdot X_4 + 0.0645932 \cdot X_6 - 0.0685445 \cdot X_7 + 0.277866 \cdot X_8
\]  

(4)

Table 3. The results of estimation of parameters for the dependent variable \( ln(Price) \)

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Structural parameter</th>
<th>Standard error</th>
<th>t-student</th>
<th>Value p</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_0 ) ≡ 1</td>
<td>( α_0 )</td>
<td>( ε_0 )</td>
<td>0.0337872</td>
<td>240.7776</td>
</tr>
<tr>
<td>( X_1 ) Czuby</td>
<td>( α_1 )</td>
<td>( ε_1 )</td>
<td>0.0249623</td>
<td>-5.5550</td>
</tr>
<tr>
<td>( X_2 ) LSM</td>
<td>( α_2 )</td>
<td>( ε_2 )</td>
<td>0.0323819</td>
<td>-4.3188</td>
</tr>
<tr>
<td>( X_3 ) Rury</td>
<td>( α_3 )</td>
<td>( ε_3 )</td>
<td>0.0193003</td>
<td>2.4582</td>
</tr>
<tr>
<td>( X_4 ) type of right acquired</td>
<td>( α_4 )</td>
<td>( ε_4 )</td>
<td>0.0202477</td>
<td>-3.1550</td>
</tr>
<tr>
<td>( X_5 ) floor</td>
<td>( α_5 )</td>
<td>( ε_5 )</td>
<td>0.0032069</td>
<td>-0.7437</td>
</tr>
<tr>
<td>( X_6 ) year of construction</td>
<td>( α_6 )</td>
<td>( ε_6 )</td>
<td>0.0038175</td>
<td>16.9203</td>
</tr>
<tr>
<td>( X_7 ) belonging premises</td>
<td>( α_7 )</td>
<td>( ε_7 )</td>
<td>0.0138514</td>
<td>-4.9486</td>
</tr>
<tr>
<td>( X_8 ) seller</td>
<td>( α_8 )</td>
<td>( ε_8 )</td>
<td>0.0173859</td>
<td>15.9822</td>
</tr>
</tbody>
</table>

| arithmetic mean on the dependent variable | 8.430903 | standard deviation of the dependent variable | 0.196843 |
| sum of the squared residuals | 35.65216 | standard error of the residuals | 0.172223 |
| \( R^2 \) determination factor | 0.239566 | \( R^2 \) corrected | 0.234505 |
| result of \( F \) statistical test | 47.33458 | \( p \)-value probability for the \( F \) test | 1.68e-66 |
| logarithm of likelihood | 416.2902 | Akaike information criterion | -814.5804 |
| Schwarz Bayesian criterion | -768.6876 | Hanna-Quinn criterion | -797.3007 |

Source: own study.
4.3. Assessment of the quality of the model based on property valuation

15 properties located in Lublin have been valued by means of the model. The formula (4) and the collected input data on properties have been used for that purpose, connected with transaction prices for 1 m² of usable area of a flat, location and information on the type of right acquired, year of construction, belonging premises and seller. The obtained results have been presented in Table 4.

Table 4. Comparison of transaction and theoretical prices obtained based on the model

<table>
<thead>
<tr>
<th>No.</th>
<th>District</th>
<th>( P_{\text{TRANS}} ) (PLN/m²)</th>
<th>( \ln(P_{\text{THR}}) )</th>
<th>( P_{\text{THR}} ) (PLN/m²)</th>
<th>AB. DIFF.*** (PLN)</th>
<th>DIFF.**** (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Czuby</td>
<td>4.809.29</td>
<td>8.2556292</td>
<td>3.849.23</td>
<td>960.06</td>
<td>19.96</td>
</tr>
<tr>
<td>2</td>
<td>Rury</td>
<td>4.336.75</td>
<td>8.3731945</td>
<td>4.329.44</td>
<td>7.31</td>
<td>0.17</td>
</tr>
<tr>
<td>3</td>
<td>Rury</td>
<td>5.035.55</td>
<td>8.5897072</td>
<td>5.376.04</td>
<td>340.49</td>
<td>6.76</td>
</tr>
<tr>
<td>4</td>
<td>Rury</td>
<td>4.601.23</td>
<td>8.3926877</td>
<td>4.414.67</td>
<td>186.56</td>
<td>4.05</td>
</tr>
<tr>
<td>5</td>
<td>Śródmieście</td>
<td>3.808.07</td>
<td>8.2806497</td>
<td>3.946.76</td>
<td>138.69</td>
<td>3.64</td>
</tr>
<tr>
<td>6</td>
<td>Rury</td>
<td>4.523.03</td>
<td>8.4612322</td>
<td>4.727.88</td>
<td>204.85</td>
<td>4.53</td>
</tr>
<tr>
<td>7</td>
<td>Śródmieście</td>
<td>5.297.23</td>
<td>8.3491942</td>
<td>4.226.77</td>
<td>1070.46</td>
<td>20.21</td>
</tr>
<tr>
<td>8</td>
<td>Śródmieście</td>
<td>4.161.46</td>
<td>8.3491942</td>
<td>4.226.77</td>
<td>65.31</td>
<td>1.57</td>
</tr>
<tr>
<td>9</td>
<td>Rury</td>
<td>5.839.42</td>
<td>8.5864673</td>
<td>5.358.65</td>
<td>480.77</td>
<td>8.23</td>
</tr>
<tr>
<td>10</td>
<td>Czuby</td>
<td>3.877.89</td>
<td>8.4035974</td>
<td>4.463.09</td>
<td>585.20</td>
<td>15.09</td>
</tr>
<tr>
<td>11</td>
<td>Czuby</td>
<td>4.256.14</td>
<td>8.3397156</td>
<td>4.186.90</td>
<td>69.24</td>
<td>1.63</td>
</tr>
<tr>
<td>12</td>
<td>Czuby</td>
<td>5.064.83</td>
<td>8.4681906</td>
<td>4.760.89</td>
<td>303.94</td>
<td>6.00</td>
</tr>
<tr>
<td>13</td>
<td>LSM</td>
<td>4.409.94</td>
<td>8.4024104</td>
<td>4.457.80</td>
<td>47.86</td>
<td>1.09</td>
</tr>
<tr>
<td>14</td>
<td>LSM</td>
<td>5.232.20</td>
<td>8.5315968</td>
<td>5.072.54</td>
<td>159.66</td>
<td>3.05</td>
</tr>
<tr>
<td>15</td>
<td>LSM</td>
<td>4.441.04</td>
<td>8.4024104</td>
<td>4.457.80</td>
<td>16.76</td>
<td>0.38</td>
</tr>
<tr>
<td>MEAN</td>
<td></td>
<td>4,646.27</td>
<td>–</td>
<td>4,523.68</td>
<td>309.14</td>
<td>–</td>
</tr>
</tbody>
</table>

where: \( P_{\text{TRANS}} \) – transaction price, \( P_{\text{THR}} \) – theoretical price, AB. DIFF. – absolute difference, DIFF.**** – difference

Table 4 presents predictive and real values resulting from observations. On such a basis, it is possible to assess the quality of the hedonic model for property valuation. As can be seen, 6 among 15 estimations (40%) differ from transaction prices by more than 5%, which in case of process of flats is very significant. Moreover, in case of 3 estimations the error has exceeded 15% of the difference, which is unacceptable. The mean for the absolute difference amounts to PLN 309.14 m, which is a deviation amounting to ca. 6.65% from the mean transaction price. The fact that 5 properties have been valued with accuracy below 2% should be deemed to be a positive aspect.
While considering the set of properties presented in table 4 through comparison of mean values from both price groups (mean transaction price PLN 4,646.27 and theoretical price PLN 4,523.68), the difference between them is absolutely 2.64% (PLN 122.59), which constitutes a value nearly below half the statistical error (to 5%).

5. Conclusions

The constructed model, despite the acceptable value of standard deviation of residuals, coefficient of residual variation and statistical significance of nearly all the variables, is not satisfactory. The low value of the determination factor \( R^2 = 0.239566 \) initially proves poor representation of reality by the model (insufficient description of the relationships between the prices and property features), as a result, of which the values estimated with the use of the model may be substantially different from the transaction prices.

It should be noted that the constructed model does not take into account features, such as impact of the standard of interior finish of particular properties and the neighbourhood conditions. The direct reason for this is the high level of diversity that has been observed for the districts, which could have a very significant effect on the results of the calculations (especially for two districts Śródmieście and Czuby).

In practice, it was found that the statistical error (up to 5%) resulting from maladjustment of the model is moderate, and in 60% of the performed valuations, the model has performed well. However, it is not sufficient for the determination of property prices due to their substantial value. A few percent differences between the theoretical and transaction prices would translate into differences of several – dozen thousand PLN and this will be very important in the case of drawing up estimated operates in the economy.

References


