# FUNCTION MODELLING OF THE MARKET AND ASSESSING THE DEGREE OF SIMILARITY BETWEEN REAL PROPERTIES - DEPENDENT OR INDEPENDENT PROCEDURES IN THE PROCESS OF OFFICE PROPERTY VALUATION<sup>1</sup>

# Anna Barańska, D.Sc., Ph.D.

Faculty of Mining Surveying and Environmental Engineering AGH University of Science and Technology *e-mail: abaran@agh.edu.pl* 

# Dominika Nowak, M.Sc., Eng.

Faculty of Mining Surveying and Environmental Engineering AGH University of Science and Technology e-mail: nowak.dominika.1990@gmail.com

# Abstract

Referring to the developed and presented in previous publications (e.g. BARAŃSKA 2011) two-stage algorithm for real estate valuation, this article addresses the problem of the relationship between the two stages of the algorithm. An essential part of the first stage is the multi-dimensional function modelling of the real estate market. As a result of selecting the model best fitted to the market data, in which the dependent variable is always the price of a real property, a set of market attributes is obtained, which in this model are considered to be price-determining. In the second stage, from the collection of real estate which served as a database in the process of estimating model parameters, the selected objects are those which are most similar to the one subject to valuation and form the basis for predicting the final value of the property being valued. Assessing the degree of similarity between real properties can be carried out based on the full spectrum of real estate attributes that potentially affect their value and which it is possible to gather information about, or only on the basis of those attributes which were considered to be price-determining in function modelling. It can also be performed by various methods. This article has examined the effect of various approaches on the final value of the property obtained using the two-stage prediction. In order fulfill the study aim precisely as possible, the results of each calculation step of the algorithm have been investigated in detail. Each of them points to the independence of the two procedures.

**Keywords**: *real estate valuation, multidimensional functional modelling, assessing similarity of real properties, independence of qualitative and quantitative methods.* 

**JEL Classification**: C1, C3, C4, C5, C8, R3.

**Citation**: Żelazowski K., 2015, Function Modelling of the Market and Assessing the Degree of Similarity Between Real Properties - Dependent or Independent Procedures in the Process of Office Property Valuation, Real Estate Management and Valuation, Vol. 23, No. 3, pp. 36-46.

DOI: 10.1515/remav-2015-0023

<sup>&</sup>lt;sup>1</sup> The task was carried out within the scope of statutory research in the Department of Geomatics, Faculty of Mining Surveying and Environmental Engineering, AGH Krakow, Poland



## 1. Introduction

Referring to the developed and presented in previous publications two-stage algorithm for real estate valuation, this article addresses the problem of the relationship between the two stages of the algorithm. An essential part of the first stage is the multi-dimensional function modelling of the real estate market. As a result of selecting the model best fitted to the market data, in which the dependent variable is always the price of a real property, a set of market attributes is obtained, which in this model are considered to be price-determining. In the second stage, from the collection of real estate which served as the database in the process of estimating model parameters, the selected objects are those which are most similar to the one subject to valuation, and form the basis for predicting the final value of the property being valued. Assessing the degree of similarity between the real properties can be carried out based on the full spectrum of real estate attributes that potentially affect their value, and which it is possible to gather information about, or only on the basis of those attributes which were considered to be price-determining in function modelling. It can also be performed by various methods. This article will examine the effect of various approaches on the final value of the property obtained using the two-stage prediction.

# 2. Database of office properties

A set of data which were the basis for the conducted analyses, included information on transactions involving office properties in the Krowodrza district of Cracow. Office space is marketed much less frequently than, for example, residential dwellings. In the analyzed period of five years (2009–2014), it was possible to capture information about 65 transactions, which means an average of only one transaction per month. Each of the offices subjected to market turnover was described by means of 23 potentially price-determining attributes, usually graduated in 4- and 5-grade scales, unless they were of numerical nature. At the same time, some attributes were assigned even 8- or 9-grade scales. Some efforts were made to avoid dichotomous (bivalent) scales when they were not the only possible ones, for example, for the *form of transaction*, in which information about transactions concluded on the primary or secondary market were considered.

Despite the relatively long period of studying transactions, due to the lack of an apparent regularity of price volatility over time (as illustrated in Figure 1), it was not possible to separate the influence of the transaction date on the price of the properties from it being influenced by other attributes. Therefore, all the analyses were performed with adopting the transaction price as the dependent variable in the valuation models and the graduated transaction date was considered as an independent variable, at par with other attributes.



Fig. 1. Dependence of the prices of office properties on the transaction date in the years 2009-2014. *Source*: own study.

Such an approach is also supported by the general condition of the office-space market, associated with the crisis in the considered period of time, which entails the stagnation in the market. It is even noticeable based on the data regarding the vacancy rates for office space in major Polish cities (Table 1), despite the fact that demand for such properties there is, for obvious reasons, the greatest.

All of the 22 examined attributes (besides price) could make up a set of independent variables, as they successfully passed the verification of independence, considered in pairs of *each one with each one*, at the initial stage of the analysis. The following attributes of office properties were taken into account: the location, access, public transport, parking, surroundings, floor, condition of the building, type of building, exposure of the office, social facilities, number of rooms, source of data, availability to customers, auxiliary premises, mortgage, office area, form of transaction, land surface area, development plan, right to land, right to the property and date of the transaction.

# Table 1

Supply and vacancy rates for office space in the largest Polish cities (the fourth quarter of 2013)

City	Supply [m <sup>2</sup> ]	Vacancy rates [%]
Warsaw	4 112 800	11.7
Cracow	519 500	4.4
Wrocław	458 900	12.9
Trójmiasto	350 000	13.4
Poznań	265 900	15.0
Łódź	236 200	16.5
Katowice	218 300	8.6
Szczecin	92 800	28.0
Lublin	89 500	22.1

Source: Colliers International (2014).

# 3. Modelling the office-space market

Modelling of the analyzed office-space market was performed in two variants: assuming its approximate homogeneity, estimating the parameters of the multiple linear regression model, and without such an assumption, estimating the parameters of the non-linear multidimensional model, whose general form is as follows:

С

$$= a_0 + \sum_{i=1}^{\infty} f_i(X_i) \qquad \mathbf{c} = \mathbf{a}_0 + \sum_{i=1}^{\infty} f_i(X_i)$$
(1)

where:

*c* – unit price of office property,

 $a_o$  – intercept term of the model,

 $X_i$  – attributes of the property,

 $f_i$  – function of the dependence of price on attribute i,

*m* - number of analyzed, potentially price-determining attributes of properties.

The forms of the  $f_i$  function were selected on the basis of a scatter plot of prices relative to the individual attributes of the real properties. Estimation of model parameters was carried out in several steps in order to obtain the best fitting to the market data. This fitting was measured, inter alia, by the coefficient of determination  $R^2$  (tested for statistical reliability), the *p*-value for each of the estimated parameters of the model (which illustrates their statistical significance), and the measure of the potential outliers for the individual cases from the hyper-surface formed by a multi-dimensional function model. The initial value of the squared multiple correlation coefficient  $R^2$  in the case of a linear model was at a high level of 0.73 (up to 73% of the variation in the prices for office space in the selected market was explained by this model). This is a much better result than the one achieved with the residential market in the same area ( $R^2$  of about 50%). This demonstrates that the office-space market is much more homogeneous than the residential market. At the same time, for such  $R^2$ , more than half of the estimated parameters turned out to be statistically significant at the significance level of 0.05 and 0.10 (12 out of 22), and no transaction was considered to be an outlier according to the adopted criteria for the residual value of the dependent variable (the price).

In the case of a non-linear model whose initial form contained as many as 43 parameters, after the gradual rejection of the least important ones - 21 parameters remained, including 13 independent variables (attributes of the properties). The model finally took the following form:

 $\begin{aligned} FRICE &= a_0 + a_1 \cdot (lacation)^3 + a_2 \cdot (lacation)^2 + b_2 \cdot (access) + f_1 \cdot (flaar)^2 + \\ + g_1 \cdot (cond. of the building)^3 + g_2 \cdot (cond. of the building)^2 + g_3 \cdot (cond. of the building) + \end{aligned}$ 



 $\begin{array}{l} +h_2 \cdot (type \ af \ the \ building)^2 + h_2 \cdot (type \ af \ the \ building) + j_2 \cdot (sacial \ facilities) + \\ +m_1 \cdot (nature \ af \ the \ premises) + a_1 \cdot (martgage) + p_2 \cdot (usable \ area)^2 + p_4 \cdot (usable \ area) + \\ +s_1 \cdot (land \ surface \ area)^2 + s_2 \cdot (land \ surface \ area) + u_1 \cdot (right \ to \ the \ land)^2 + \\ +y_1 \cdot (auxiliary \ premises)^2 + y_2 \cdot (auxiliary \ premises) + z_1 \cdot (time)^2 \end{array}$ 

For such a model form, the percentage of the variance of the part explained by the regression model in a general spread of the dependent variable was at a high level of  $R^2 = 82\%$ . No cases of an outlier were observed, and the vast majority of model parameters proved to be statistically significant: as many as 17 out of 21, at a significance level of 0.05, and as many as 19 out of 21 at a significance level of 0.10. When studying the significance of the high values of the coefficients of determination, and also the significance of the final sets of parameters of the two models - the values of *F*-statistic were calculated in order to verify the following hypotheses:

$$H_0: R^2 = 1 - R^2$$
  

$$H_0: R^2 > 1 - R^2$$
(3)

The values of this test statistic for the linear and non-linear models were respectively:

$$F = \frac{R^2}{1-R^2} \cdot \frac{u-u}{u-1} = \frac{0.59}{1-0.59} \cdot \frac{69-29}{28-1} = 5.16$$
(4a)

$$F = \frac{\pi^2}{1 - \pi^2} \cdot \frac{u - u}{u - 1} = \frac{0.022}{1 - 0.02} \cdot \frac{0.02 - 21}{21 - 1} = 10.02$$
(4b)

where:

 $R^2$  – coefficient of determination,

*n* – number of real properties in the database,

*u* – number of estimated parameters of the model.

The corresponding critical values in the Fisher-Snedecor distribution for the significance level of 0.01 are as follows:  $F_{0,01;22,42} = 2.30$  for the linear model, and  $F_{0,01;20,44} = 2.32$  for the non-linear model. In both cases, the critical values are exceeded by the values of the test functions, which points to the necessity of rejecting the null hypothesis  $H_0$  in favour of the alternative hypothesis  $H_1$ . This corresponds to the high reliability of the two coefficients of determination and the models themselves, modelling the local office-space market, especially due to the lower than standard (0.05) level of significance. This stringent criterion for statistical significance allows for a great reliability relative to the high fitting levels to the real data of the estimated valuation models.

Based on these models, point estimation of model value was performed for a selected office property. This office was also the subject of a transaction and, therefore, its market price was known. The results have been summarized in Table 2.

Model value of office property

Table 2

	model value of office property				
Model	Price of the office	Model value w <sub>M</sub>	StandardCoefficientdeviationof variationR $\sigma(w_M)$ $\lambda = \frac{\sigma(w_M)}{w_M}$ t		Residual of the model
	$[zl/m^2]$	$[zl/m^2]$	$[zl/m^2]$	[%]	$[zl/m^2]$
linear regression	5 522 15 -	4 965.54	593.51	12.0	567.61
non-linear regression	- 5 555.15 -	5 841.05	290.92	5.0	-307.90

Source: own study.

The results contained in Table 2 demonstrate the advantage of the non-linear model, as it results in a more than twice more accurate estimate of the model value and to the smaller deviation (model residual), which confirms its better fitting to the actual fluctuation in the market, which was shown earlier.

#### 4. Evaluating degree of similarity between office properties

The modelling stage is followed by the assessment of the degree of similarity between the properties.

DE

In order to draw general conclusions about the dependence or independence of this process from the function modelling itself, several different methods were used to assess the degree of similarity, so as to select the properties out of the full set of market data which are the most similar to the valued object.

In order to assess the degree of similarity, three qualitative methods were used: the number of compatible attributes, relative comparison analysis and scheduling analysis. A full description of the last two can be found in the publication (CZAJA, PARZYCH 2007). The first one, on the other hand, recognizes the object to be the most similar if at least half of the considered attributes are identical to the reference point. Each of these qualitative methods can be used taking into account all the relevant attributes of a real property in a given model (11 and 13 in the two models, respectively), or taking into account only those which proved to be important in the modelling process and, at the same time, at the stage of the statistical analysis of the market (7 and 6, respectively). Such variants were examined in the first of the three methods of assessing the degree of similarity. This was the major aspect of the research, which verified the independence of quantitative methods (market modelling stage) and qualitative methods (similarity assessment stage) in the prediction of the real property value based on the most similar objects.

Table 3 contains the results of the selection of objects most similar to the valued office property. The subsequent variants in the Table refer to:

Ia - the number of conforming attributes, taking into account all the attributes significant in a model,

Ib - the number of conforming attributes, taking into account all the attributes significant in a model and significant in the market,

II - relative comparison analysis for all the attributes significant in a model,

III - scheduling analysis taking into account all the attributes significant in a model.

For methods Ia, II and III, the starting point was a set of seven attributes which proved to be significant in both function models. At the same time, this set represented approximately half of all independent variables, which in each of the models were considered statistically significant. On the other hand, in method Ib, each of the function models was considered separately, mainly due to the fact that by selecting only the attributes which were significant in both models, and also generally significant in the local market, the research would be narrowed down to only four variables, and this number was considered too divergent from the other methods. In contrast, when analyzing the models independently, similarities were evaluated in terms of six or eight attributes, in linear and non-linear models, respectively. In addition to this, an additional variant extension of the range of the conducted research was also obtained in this way.

Table 3

Variant	Number of selected real estate	Real estate numbers	Percentage of the entire database	Percentage of the selected in a minimum of two methods
Ia	14	<b>2</b> , <b>17</b> , <b>24</b> , <b>25</b> , <b>37</b> , <b>38</b> , <b>40</b> , <b>44</b> , 47, <b>55</b> , 58, 61, 62, 63	22%	64%
Ib – lin	7	2, 17, 24, 25, 37, 38, 40	11%	100%
Ib – nlin	9	2, 17, 24, 25, 37, 38, 40, 44, 49	14%	89%
II	11	1, 31, 33, <b>40</b> , <b>44</b> , 48, <b>49</b> , 51, <b>55</b> , 56, 60	17%	36%
III	8	11, 15, 18, <b>25</b> , 29, 30, 32, 41	12%	12%

Summary of the office properties most similar in four variants

#### Source: own study.

In Table 3, the numbers of properties in bold were considered similar in at least two of the considered methods for the assessment of the degree of similarity. It should be noted at this point that the method for assessing the degree of similarity which is based on confirming that at least half of the attributes of the objects are identical, led to almost the same results, regardless of the form of the function model (variants Ib) and, to a large extent, to the convergent results, taking into account all significant attributes of the office properties (Ia). This is the first sign of the independence of the

method for assessing the degree of similarity from the quantitative procedures involving the statistical analysis of the market or its modelling.

Another observation refers to the comparison of algorithms I, II and III. The results which are the most distinct from the rest of the qualitative methods are yielded by the scheduling analysis of real properties: only 12% of the selected objects were considered most similar also in other methods. This raises doubts as to the reliability of this method for assessing the degree of similarity and the criteria formulated in it. However, it may prove to be valuable when making conclusions about the studied independence of the procedures.

#### 5. Predicting the market value of office property

Referring to the algorithm for the two-stage real estate valuation, which was mentioned in the introduction and described several times, for example in articles published by the Real Estate Scientific Society, such as (BARAŃSKA 2010,BARAŃSKA 2011), the final prediction of the market value of a selected office property was performed by adding to its model value from Table 2 a random correction calculated as the average (weighted or arithmetic) from the residuals of a given model for the offices considered to be the most similar in one of the methods (I, II or III) according to the following formula:

$$w_{R} = w_{M} + w_{L} \tag{5}$$

**DE GRUYTER** 

OPEN

DE

where:

 $w_{\mathbb{R}}$  – market value of office property,  $w_{\mathbb{M}}$  – model value of office property,  $w_{\mathbb{L}}$  – random value (random correction).

The equation (5) originates from the known relationship  $\delta = C - W \Leftrightarrow C = W + \delta$ , in which the residual  $\delta$  is the difference between the empirical value (e.g. the market price of the real property *C*) and the model value (e.g. the model value of the real property *W*).

All the calculations were carried out with the full analysis of accuracy:

$$\sigma(w_R) = \sqrt{\sigma^2(w_M) + \sigma^2(w_L)} \tag{6}$$

where:

 $a^{2}(w_{M}), a^{2}(w_{L})$  – variances of model value and random correction of office property.

Standard deviation  $\sigma(w_M)$  in the formula (6) is calculated by means of an analysis of variance of the estimated function model.

#### 5.1. Random correction as the weighted average of residuals

Random correction as the weighted average results from taking into account a weighing matrix as the inverse of the covariance matrix of residuals of the office properties which were considered to be the most similar (7). Therefore, it is not about weighing residuals due to the degree of similarity of the offices to the valued objects which these residuals correspond to - in this way we avoid an intentional dependency of the modelling procedures and the procedure for the assessment of the degree of similarity.

$$w_{\mathbf{L}} = [\underline{1P1}^{T}]^{-1} [\underline{1P}] [\mathcal{S}_{w}]$$

$$\tag{7}$$

where:

[1] – vector of one of the dimensions  $(1 \times k)$ ,

k – number of real properties considered to be the most similar to the subject property,

[4] – vector of residuals of the model corresponding to the office properties which are the most similar,

 $[P] = Cov^{-1}[\delta_{uv}]$  – weight matrix  $(k \times k)$ , the inverse of the covariance matrix of residuals  $[\delta_{uv}]$ .

Variance of the random correction is calculated from the following formula:

$$\sigma^{2}(w_{L}) = \sigma^{2}_{\theta_{W}} \cdot \left[\underline{1}P\underline{1}^{T}\right]^{-1} = \frac{\left[\delta^{T}_{W}P\delta_{W}\right] - w_{L}[\underline{1}P\delta_{W}]}{k-1} \cdot \left[\underline{1}P\underline{1}^{T}\right]^{-1}$$
(8)

where:

a. – residual variance estimator for *k* office properties most similar to the valued one.

Table 4 demonstrates the results of the point estimation of random corrections to the model value in the form of a weighted average and the final prediction of the market value of the selected office property for both types of estimated function models, and for the four variants of assessing the degree of similarity between the office properties.

Table 4

Prediction of the market value - random correction as the weighted average of residuals

	method for assessing the degree of similarity			
	Ia	II	III	Ib
transaction price [PLN/m <sup>2</sup> ] C	5 533.15			
estimated value		LINEAR	MODEL	
model value [PLN/m <sup>2</sup> ] $w_M \pm \sigma(w_M)$	4 965.54 ± 593.51			
random correction [PLN/m <sup>2</sup> ] $w_{L} \pm \sigma(w_{L})$	-255.06±110.97	394.01±155.02	-103.07±213.94	-358.48±207.79
market value [PLN/m²] w <sub>R</sub> ± $\sigma(w_R)$	4 710.48±603.79	5 359.55±613.42	4 862.47±630.89	4 607.07±628.83
coefficient of variation $\lambda = \frac{\sigma(w_{be})}{w_{be}} [\%]$	12.8	11.4	13.0	13.6
deviation <b>C</b> – <b>w</b> <sub>R</sub> of the market value from the transaction price [PLN/m <sup>2</sup> ]	822.67	173.60	670.68	926.08
		NON-LINE	AR MODEL	
model value [PLN/m <sup>2</sup> ] $w_M \pm \sigma(w_M)$	$5841.05 \pm 290.92$			
random correction [PLN/m <sup>2</sup> ] $w_{\Sigma} \pm \sigma(w_{\Sigma})$	-49.63±90.85	163.74±142.65	324.41±172.35	11.25±175.71
market value [PLN/m²] Wa ± a(Wa)	5 791.42±304.78	6 004.79±324.01	6 165.46±338.14	5 852.30±339.35
coefficient of variation $\lambda$ [%]	5.3	5.4	5.5	5.8
deviation <b>C</b> – <b>W</b> <sub>R</sub> of the market value from the transaction price [PLN/m <sup>2</sup> ]	-258.27	-471.64	-632.31	-319.15

Source: own study.

#### 5.2. Random correction as the arithmetic mean of the residuals

Point estimation of the random correction to the model value of the office property in the form of the arithmetic mean of residuals was performed according to the following formula:

$$w_{\underline{\lambda}} = \hat{w}_{\underline{\lambda}} \perp \sigma(\hat{w}_{\underline{\lambda}}) = \frac{4}{k} \sum_{l=1}^{k} \sigma_{w_{\underline{l}}} \perp \sqrt{\frac{16 c_{W_{\underline{l}}} (\hat{\sigma}_{w_{\underline{l}}}) \pm^{T}}{k^{2}}}$$
(9)

with denotations same as in formulas (6) and (7).

Table 5 is equivalent to Table 4 for random correction to the model value of the office property in the form of the arithmetic mean of residuals, estimated according to formula (9).

Based on the results presented in Tables 4 and 5, we can see a clear advantage of the non-linear model over the linear model, both in terms of a more accurate estimation of the final prediction (over twice smaller coefficient of variation  $\lambda$ ) as well as regarding a greater coherence of the results in relation to the market price of the valued office property (mean deviation  $|\mathbf{C} - \mathbf{w}_{\mathbf{R}}|$  for the linear model it is 648.26 or 624.38 [PLN/m<sup>2</sup>] depending on the type of the correction, and for the non-linear model it is 420.34 or 384.31 [PLN/m<sup>2</sup>]). Due to the fact that this advantage is independent of both the method of calculating the random correction as well as the method of assessing the degree of similarity between the office properties, it would appear therefrom that the most important factor ensuring



greater accuracy (reliability) of market value predictions is the quality of fitting the function model to the market data. This is yet another indication that allows to claim that the function modelling of the real estate market is a procedure independent from the assessment of the degree of similarity between real properties.

#### Table 5

Prediction of the market value - random correction as the arithmetic mean of residuals

	method for assessing the degree of similarity			
	Ia	II	III	Ib
transaction price [PLN/m <sup>2</sup> ] C	5 533.15			
estimated value		LINEAR	MODEL	
model value [PLN/m <sup>2</sup> ] $w_M \pm \sigma(w_M)$	4 965.54 ± 593.51			
random correction [PLN/m <sup>2</sup> ] $w_{L} \pm \sigma(w_{L})$	-249.65±138.83	462.09±168.21	-77.23±229.61	-362.30±151.67
market value [PLN/m <sup>2</sup> ] $w_{R} \pm \sigma(w_{R})$	4 715.89±609.53	5 427.63±616.89	4 888.31±636.38	4 603.25±612.58
coefficient of variation $\lambda = \frac{\sigma(m_M)}{m_{Per}} [\%]$	12.9	11.4	13.0	13.3
deviation of the market value from the transaction price [PLN/m <sup>2</sup> ]	817.26	105.52	644.84	929.90
		NON-LINE	AR MODEL	
model value [PLN/m <sup>2</sup> ] $w_M \pm \sigma(w_M)$	5 841.05 ± 290.92			
random correction [PLN/m <sup>2</sup> ] $w_{2} \pm \sigma(w_{2})$	-110.89±114.31	326.60±143.79	151.60±174.14	-61.68±159.18
market value [PLN/m²] w <sub>R</sub> ± σ(w <sub>R</sub> )	5 730.16±312.57	6 167.65±324.51	5 992.65±339.06	5 779.37±331.62
coefficient of variation $\lambda$ [%]	5.5	5.3	5.7	5.7
deviation of the market value from the transaction price [PLN/m <sup>2</sup> ]	-197.01	-634.50	-459.50	-246.22

Source: own study.

#### 6. Study of independence by means of parametric tests of significance of differences

A tool allowing to draw the most reliable and objective conclusions regarding the independence of the two stages of the implemented valuation algorithm are parametric tests of the significance of differences between the estimations of the value of an office property in different variants and their accuracy parameters. This is discerned by a number of authors involved in research on the comparative approach to the valuation of property (e.g. CUPAL 2014). To compare market values, one of the following statistics was used: *T*-Student (10a), *C*-Cochran-Cox (10b) and *Z*-normal distribution (10c), depending on the size of the random sample and the test result comparing variances. On the other hand, to compare the accuracy of the market values, the *F*- Fisher-Snedecor statistic (11) was used, comparing the variances.

$$T = \frac{w_1 - w_2}{\sqrt{w_1 + w_2 R_1 - q^2 (w_1) + k_2 - q^2 (w_2)}}$$
(10a)

$$\mathcal{L} = \frac{w_1 - w_2}{\sqrt{\sigma^2 (w_1) + \sigma^2 (w_2)}}$$
(10b)

$$Z = \frac{w_1 - w_2}{\sqrt{\sigma^2 (w_1) + \sigma^2 (w_2)}}$$
(10c)

43

(11)





$$F = \frac{\sigma^2 \langle w_1 \rangle}{\sigma^2 \langle w_2 \rangle} \cdot \frac{k_2}{k_1}$$

where:

 $w_1, w_2$  – compared market values,  $\sigma^*(w_1), \sigma^*(w_2)$  – variances of the compared market values,  $k_1, k_2$  – numbers of the degrees of freedom corresponding to the compared values,  $n_1, n_2$  – random sample sizes, which determined  $w_1, w_2$  respectively.

The comparisons were carried out with regard to the increased significance level of 0.10 to tighten the criterion of the absence of a statistically significant difference between the compared values, and thus lend credibility to the conclusions which were drawn.

# 6.1. Studying the significance of differences between the accuracies of market values

The use of one of the statistics (10a-c) for testing the significance of differences between the market values may depend on the significance of the difference between their variances. Therefore, in the first place, a comparison of the accuracy of predictions was performed.

Table 6 illustrates the calculated statistics (11). Critical values, derived from statistical tables which the values of the statistic were compared with, for all the existing combinations of the numbers of the degrees of freedom satisfy the condition  $F_{0,10,R_1,R_2} > 1,30$ . It can be noticed that all the values of the test functions are lower than the critical value of the test, i.e. there are no grounds to reject the hypothesis of equal accuracy of the compared estimations regarding market values of an office property. Hence the conclusion on the independence of market modelling procedures and assessment of the degree of similarity between office properties in their effect on the accuracy of the final estimations of the value of a selected office property.

Table 6

	compared variants of calculations	values of the <i>F</i> -statistic		values of the Z- statistic	
		linear	non-linear	linear	non-linear
	Ia-II	1.03	1.13	-0.754	-0.480
correction	Ia-III	1.09	1.23	-0.174	-0.822
w₂ as the → weighted ⊥ average _	Ia-Ib	1.04	1.24	0.119	-0.133
	II-III	1.06	1.09	0.565	-0.343
	II-Ib	1.05	1.10	0.857	0.325
	III-Ib	1.01	1.01	0.287	0.654
correction <sup>W</sup> 2 as the arithmetic mean	Ia-II	1.02	1.08	-0.821	-0.971
	Ia-III	1.09	1.18	-0.196	-0.569
	Ia-Ib	1.01	1.12	0.130	-0.108
	II-III	1.06	1.09	0.609	0.373
	II-Ib	1.01	1.04	0.948	0.837
	III-Ib	1.08	1.04	0.323	0.450

Studying the significance of differences between point estimations of market values

Source: own study.

#### 6.2. Studying the significance of differences between the predictions of market values

The last two columns of Table 6 contain values of the Z-statistic, calculated as the results of comparisons of market value predictions of the selected office property, determined by the use of two forms of function models, two ways of determining the random correction to the model value and four variants of selecting office properties which are the most similar to the estimated one. Also in this case none of the test function values belongs to the critical region, which in this test is the following sum of ranges:  $(-\infty; -1.645) \cup (1.645; +\infty)$ . We can, therefore, conclude that the function modelling of the market is independent from assessing the degree of similarity between office properties, also in relation to market value predictions.

## 7. Conclusions

The aim of the research, formulated at the beginning of this article, was an attempt to determine whether in a two-stage algorithm of property valuation, the quantitative stage of function modelling of the local market is dependent on or independent from the qualitative stage of the assessment of the degree of similarity between objects in this market. To resolve the stated problem as precisely as possible, the results of each calculation stage of the algorithm were investigated in detail. Each of them points to the independence of the two procedures.

This can be noticed for the first time when analyzing the effects of applying various qualitative methods to select office properties which are the most similar to the one being valued (Table 3). The method for assessing the degree of similarity, which is based on confirming that at least half of the attributes of the objects are identical, led to almost the same results, regardless of the form of the function model (variants Ib) and, to a large extent, to the convergent results, taking into account all significant attributes of the office properties (Ia).

Another signal of independence can be observed when evaluating "at first glance" the results of the point estimation of the final predictions of office market values (Tables 4 and 5), which exhibit a great coherence within each of the types of function models separately, regardless of the method of calculating the random correction and the method for assessing the degree of similarity used.

The study of the significance of differences of both the predictions of market values as well as their accuracy, carried out at the end of the study, confirmed the previous assumptions of independence, especially considering that the parametric tests of significance used here adopted a more stringent significance level of 0.10. In the absence of grounds for rejecting the null hypothesis of an insignificant difference between the compared values, such a high significance level greatly increases the power of the test.

The conducted studies allow to conclude that, indeed, the real property selection process in the two-stage algorithm of their valuation remains independent from statistical procedures which the local market where these properties come from is subjected to. It is worth noting that the study was carried out for two models that differed significantly not only in the function form, but also by about 10% in the degree of fitting to the market data. However, this did not affect the conclusions regarding independence. In practice, this means that a real estate appraiser who would like to specify the model value of a real property using the two-stage algorithm in order to approximate the market price, , can use the methodology which he or she finds convenient to assess the degree of similarity without reference to the constructed function model, or use the opinions of others about similar objects, for example those contained in the division into valuation units established for mass appraisal purposes.

#### 8. References

- ADAMCZEWSKI Z., 2011, Elementy modelowania matematycznego w wycenie nieruchomości, Podejście porównawcze, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa (Elements of mathematical modeling in real estate valuation, comparative approach, Publishing House of the Warsaw University of Technology, Warsaw).
- BARAŃSKA A., 2010, Statystyczne metody analizy i weryfikacji proponowanych algorytmów wyceny nieruchomości, WYDAWNICTWA AGH, KRAKÓW (STATISTICAL METHODS FOR THE ANALYSIS AND VERIFICATION OF THE PROPOSED ALGORITHMS IN REAL ESTATE APPRAISAL, PUBLICATIONS AGH, KRAKOW).
- BARAŃSKA A., 2011, Urealnienie wartości modelowej nieruchomości (Reliability of the real estate model value), STUDIA I MATERIAŁY TOWARZYSTWA NAUKOWEGO NIERUCHOMOŚCI = JOURNAL OF THE POLISH REAL ESTATE SCIENTIFIC SOCIETY, Vol. 19 No. 3, OLSZTYN.
- CHIRILA V., CHIRILA C., 2014, THE IMPACT OF ECONOMIC AND FINANCIAL CRISIS ON THE REGIONAL DISPARITIES IN ROMANIA AND EUROPEAN UNION, PROCEDIA SOCIAL AND BEHAVIORAL SCIENCES 109 (2014) 502-506.
- COLLIERS INTERNATIONAL, LUTY 2014, *Przegląd rynku nieruchomości. Podsumowanie 2013 r.,* WARSZAWA (OVERVIEW OF THE REAL ESTATE MARKET. SUMMARY OF 2013. WARSAW). HTTP://WWW.COLLIERS.COM/-/MEDIA/1BBF047A5AE143D5B3B96AE08D8B2A79.ASHX, 10.02.2015.
- CUPAL M., 2014, *The Comparative Approach theory for real estate valuation*, Procedia Social and Behavioral Sciences 109 (2014) 19-23.
- CZAJA J., PARZYCH P., 2007, Szacowanie rynkowej wartości nieruchomości w aspekcie międzynarodowych standardów wyceny, Stowarzyszenie Naukowe im. Stanisława Staszica, Kraków (The estimation of

the market value of the property in aspect of international valuation standards, Science Society Stanislaus Staszic, Krakow).

- GREŃ J., 1978, *Statystyka matematyczna, modele i zadania,* Wydawnictwo Naukowe PWN, Warszawa (Mathematical Statistics, models and tasks, Scientific Publishers PWN, Warsaw).
- Kokot S., Bas M., (2013), "Evaluation of the applicability of statistical methods in studies on price dynamics on the real estate market", Real Estate Management and Valuation, Vol. 21, No. 1: 49-58.
- Sawiłow E., 2012, Zastosowanie skalowania wielowymiarowego do wyceny nieruchomości (The use of multidimensional graduating in the pricing of real estate), Studia i Materiały Towarzystwa Naukowego Nieruchomości = Journal of The Polish Real Estate Scientific Society, Vol. 20 No. 1.
- SEYA H., TSUTSUMI M., YOSHIDA Y., KAWAGUCHI Y., 2011, Empirical Comparison of the Various Spatial Prediction Models: in Spatial Econometrics, Spatial Statistics, and Semiparametric Statistics, Procedia – Social and Behavioral Sciences 21 (2011) 120-129.
- UYSAL F., TOSUN O., 2014, Multi criteria analysis of the residential properties in Antalya using TODIM method, Procedia Social and Behavioral Sciences 109 (2014) 322-326.
- Wiley J. A., Liu Y., Kim D., Springer T., 2014, *The Commercial Office Market and the Markup for Full Service Leases*, Journal of Real Estate Researche, Vol. 36 No. 3, USA.
- ZUJO V., CAR-PUSIC D., ZILESKA-PANCOVSKA V., 2014, Cost and Experience Based Real Estate Estimation Model, Procedia – Social and Behavioral Sciences 119 (2014) 672-681.