



Sebastian Gnat

 <https://orcid.org/0000-0003-0310-4254>

Institute of Econometrics and Statistics
Faculty of Economics and Management
University of Szczecin, Szczecin, Poland
sebastian.gnat@usz.edu.pl

Measurement of entropy in the assessment of homogeneity of areas valued with the Szczecin Algorithm of Real Estate Mass Appraisal

Accepted by Editor Ewa Ziemia | Received: January 14, 2019 | Revised: May 8, 2019; June 5, 2019 | Accepted: June 19, 2019.

Abstract

Aim/purpose – General real estate taxation is a process regulated, inter alia, by the Real Estate Management Act. It is intended to establish a tax base for real estate in the event of a change in real estate tax base. General taxation is one of several applications of mass valuation of real estate, which enables valuation of many properties at the same time and with a uniform approach. One of the methods of mass valuation of real estate already applied in practice is the Szczecin Algorithm of Real Estate Mass Appraisal (SAREMA). One of the immanent features of general taxation and the algorithm itself is the division of a selected area into possibly homogeneous areas called taxing zones within the general taxation terminology and, more broadly, elementary areas, according to the nomenclature used in the SAREMA. The paper presents the results of the studies on the measurement of elementary areas homogeneity on the example of land plots located in Szczecin. It is important to assess whether the designated sub-areas of valuation cover properties similar to each other in terms of their specific characteristics. If so, it will help to obtain more accurate mass valuation results.

Design/methodology/approach – The paper proposes to use a modified entropy measure to establish whether the designated areas are homogeneous in terms of the specified properties of real estate. The database of real estate includes more than 1.5 thousand urbanised land plots located in Szczecin. The measurement of entropy will be preceded by the specification of elementary areas. The available methods include the application of an expert approach, under which land boundaries will be indicated by property valuers.

Findings – The main conclusion of the study is that a modified measure of entropy ensures a better indication of the degree of indefiniteness of valued sub-areas and thus it offers a better way of supporting the delimitation of these sub-areas in comparison to the classical measure of entropy.

Research implications/limitations – The delimitation of valuation sub-areas constitutes an important element of mass valuation. Proper execution of this process enables obtaining much more precise valuations. An objective measure of homogeneity gives a chance to compare different approaches to the creation of the above-mentioned sub-areas and to choose the best of them.

Originality/value/contribution – The main achievement of the study is a proposal to modify the classical entropy measure, thanks to which it better reflects the specificity of the assessment of homogeneity of the areas valued in terms of property market analysis.

Keywords: property mass valuation, entropy, property market analysis.

JEL Classification: C51, R52.

1. Introduction

In the practice of real property valuation two main trends can be distinguished: individual and mass appraisal. In the process of an individual appraisal the entity valuing real estate focuses on one real property or on a small number of properties. Whereas in the case of mass appraisal (e.g. Hozer, Kokot, & Kuźmiński, 2002; Kuryj, 2007; Telega, Bojar, & Adamczewski, 2002):

- 1) the subject of an appraisal involves a large number of real properties of one type,
- 2) an appraisal may be carried out with a uniform approach resulting in cohesive results,
- 3) all the real properties subject to an appraisal are valued taking into account the same dates on which a real estate condition is considered and at the same price level.

In practice and above all in the theory of real property mass appraisal, many models and algorithms can be differentiated (Jahanshiri, Buyong, & Shariff, 2011). The process of accepting an adopted model of real property mass appraisal ought to be based on results reliability, which is intended in particular to prevent complaints, justified from the standpoint of the involved parties, regarding the manner of appraisal as well as the obtained results themselves. One of the fundamental elements of many mass appraisal models is the division of an appraised area into sub-areas, which in the Polish law is sanctioned by the Ordinance of the Council of Ministers dated 29 June 2005 on general real property taxation (Journal of Laws No. 131, item 1092). The specification of the zones

constitutes one of the key problems from the point of view of the accuracy of the achieved appraisal results (Cellmer & Kuryj, 2011; Hozer, Kokot, Foryś, Zwolankowska, & Kuźmiński, 1999). On the one hand, a taxation zone is an area in which a certain number of real properties subject to an appraisal demonstrate the same impact of the location on their value. In other words, all the real properties located in a given elementary area (a taxation zone) formally do not differ in terms of their locality. On the other hand, the Ordinance introduces the concept of a representative real property (§ 4.1), which “is selected by taking into consideration the properties characteristics for a given type of real property in a given zone”. Furthermore, the Real Estate Management Act specifies, when referring to taxation zones, that ‘in order to define a cadastral value of land, unitary values are determined for a land surface located in the zones designated on the grounds of similar properties affecting their market value’ (Article 166.1). These are the key statements. They mean that real properties belonging to a given zone (sub-area) ought to be similar to one another. Hence the differences between them resulting from the characteristics that define them should be as small as possible. The aim is to achieve a situation in which representative real properties will reflect to the greatest degree possible the set of the real estate from which they originate, which is meant to enable the extrapolation of representative real properties values onto the entire zone as accurately as possible. Incorrect determination of elementary zones may have far-reaching consequences. An insufficient number of elementary zones, and consequently too vast zones, will result in qualifying into them real properties that are clearly different in terms of the market appeal of their locality and in terms of other characteristics, which will obviously translate into burdening the appraisal results with excessively great errors. In turn, designating too great a number of elementary zones, i.e. elementary zones of a relatively small area, might not endanger the result with inaccuracies, but it may unnecessarily lengthen the appraisal process and raise its cost. The paper proposes to use a modified entropy measure to establish whether the designated areas are homogeneous in terms of the specified properties of real estate. The main aim of the paper is to verify the possibility of using a measure of information entropy to assess the homogeneity of designated real estate subareas.

The paper consists of seven parts. Section 2 describes a scientific problem related to the segmentation of real estate markets. The next part is a literature review of the research problem. Next, the research methodology used in the field of mass property valuation and the assessment of the homogeneity of the speci-

fied sub-areas is described. Parts five and six contain the results of the research and a discussion of the findings. The paper ends with a conclusions containing the potential implications of the research results and future research objectives.

2. Problem specification

Designating appraisal sub-areas entails an evaluation of the similarities existing between the real properties located within their limits. In many studies on the structure and functioning of local markets (not only real estate markets) there is a conviction that in order to effectively understand the functioning of a highly complex and spatially heterogeneous system, submarkets need to be properly distinguished, as it is on these submarkets that statistical regularities are revealed (Keskin & Watkins, 2017). The main benefits of such an approach include expected greater predictive accuracy of models (Adair, Berry, & McGreal, 1996). In addition, benefits may include greater usefulness of the results for decision-makers (Galster, 1996) and assistance in improving the decision-making process of real estate operators (e.g. Peng & Thibodeau, 2013; Rae, 2015). Different methods may be used to distinguish submarkets, such as statistical analyses (Bourassa, Hamelink, Hoesli, & MacGregor, 1999), methods using cross-price elasticity (Pryce, 2013), hedonic price theory (Rosen, 1974) or spatial analyses (Wu & Sharma, 2012). Previous sub-market research has focused on grouping properties to create better models with the price as dependent variable (e.g. Manganelli, Pontrandolfi, Azzato, & Murgante, 2014). In contrast, studies on a real estate grouping for tax purposes do not refer to an *ex ante* assessment of grouping. The assessment of clustering mainly focuses on whether the clustering method has allowed better models to be constructed (e.g. Wu, Wei, & Li, 2019). In order to make such an assessment it is necessary to have information about prices or valuations of real estate. In this study, the focus is on whether specified areas contain homogeneous properties and not on assessing the quality of grouping for market modelling. The paper will present an approach in which a measure of entropy (Doszyń, 2008) will be employed in order to define the differences between real properties in specified sub-areas. A modification of a classic entropy measure will be proposed, which will enable a better presentation of the real property market specificity. It will be demonstrated that the new measure allows for a superior evaluation of homogeneity of designated zones. A geographic area of the study covers the northern part of Szczecin (the largest city in the north-west of Poland). Over 1,600 plots of land constitute the subject of the appraisal, within which elementary zones are created.

3. Literature review

In the literature, the issue of determining elementary zones as well as taxation zones is chiefly related to mass appraisal conducted for fiscal purposes, i.e. to general taxation (cf. Bird & Slack, 2002; Brown & Hepworth, 2001; Kopyściańska, 2016; Mika, 2016). Isolating taxation zones constitutes one of the fundamental activities in the procedure of general real property taxation. Taxation zones should be designated on the basis of the so-called real estate price-determining characteristics, and a function assigned in a zoning plan is one such fundamental characteristic (Sawilow, 2009). The existing approaches to determining taxation zones entail, apart from the techniques based on zoning plan provisions, the use of aerial, satellite photographs and expert intervention into the boundaries of the areas designated on the grounds of those criteria (Surowiec & Malczewska, 2001). At the same time, district and residential estate borders, plot borders, streets, roads, rivers, railway lines and other artificial and natural features may also be used (Dedkova & Polyakova, 2018). On the basis of the concept of spatial planning, defined elementary zones, understood as dense areas sharing the same zoning function, may be linked with the value level of the real properties covered by such zones. On the one hand, a real property function in a zoning plan as well as a real property location (belonging to a particular elementary area) are significant factors determining a real property value, on the other, a real property value may constitute the basis for determining the limits of such an area (Cymerman, Fiedorowicz, Grabowski, Gwiazdzińska, & Kurowska 2000). There are also postulates for taking into account the so-called physiognomic features, or even legal ones (ownership structure of a real property) apart from the terrain functions (Kotkowski, 1999). A discussion is chiefly conducted with respect to why taxation zones ought to be determined and what criteria should be taken into account for their formation. However, there are no precise proposals that would support the process of evaluating whether elementary zones (taxation zones) have been properly designated. Various methods can be employed for such a specification and a measure that would enable their objective comparison is undoubtedly required.

4. Research methodology

As was previously mentioned, there are a number of methods of mass appraisal. An example of such a method is the Szczecin Algorithm of Real Estate Mass Appraisal (SAREMA). One of the algorithm stages involves the designation of elementary areas and it is this algorithm that will be used to determine

a real property value, and a proposal of modifying the entropy measure will be used to evaluate the homogeneity of the real properties in the designated terrains. The Szczecin Algorithm of Real Estate Mass Appraisal assumes the following form (Hozer et al., 1999, 2002):

$$w_{ji} = wwr_j \cdot pow_i \cdot c_{baz} \prod_{k=1}^K \prod_{p=1}^{k_p} (1 + a_{kp}) \quad (1)$$

where:

w_{ji} – the market (or cadastral) value i -th real property in j -th elementary terrain,

wwr_j – the market value coefficient in j -th elementary terrain ($j = 1, 2, \dots, J$),

pow_i – the surface of i -th real property,

c_{baz} – the price of 1m^2 of the cheapest land (without the utility infrastructure) in the appraised area,

a_{kp} – the influence of p -th category of k -th attribute ($k = 1, 2, \dots, K; p = 1, 2, \dots, k_p$),

K – the number of attributes,

k_p – the number of categories of k -th attribute.

Algorithm (1) has a multiplicative form. The point of reference for an appraisal with the use of the algorithm is provided by a base price. It is the price of 1m^2 of the cheapest land without any utility infrastructure in the appraised municipality. It may be assumed that it is a unitary price of a real property featuring the worst categories of attributes, which includes the impact of attributes of the appraised real properties. The impact (a_{kp}) may be defined with an expert approach by real estate appraisers.

For each elementary zone market value coefficients (wwr_j) are determined, which reflect the impact of a widely understood location.

A market value coefficient for j -th elementary zone is a geometric mean of the quotients of real properties actual values and their hypothetical values:

$$WWR_j = \sqrt[n_j]{\prod_{i=1}^{n_j} \frac{w_{ji}^{rz}}{w_{ji}^h}} \quad (2)$$

where:

w_{ji}^{rz} – the value of i -th real property in j -th elementary terrain defined by a real property appraiser,

w_{ji}^h – the hypothetical value of i -th real property in j -th elementary terrain,

n_j – the number of representative real properties valued by real property appraisers in j -th elementary terrain.

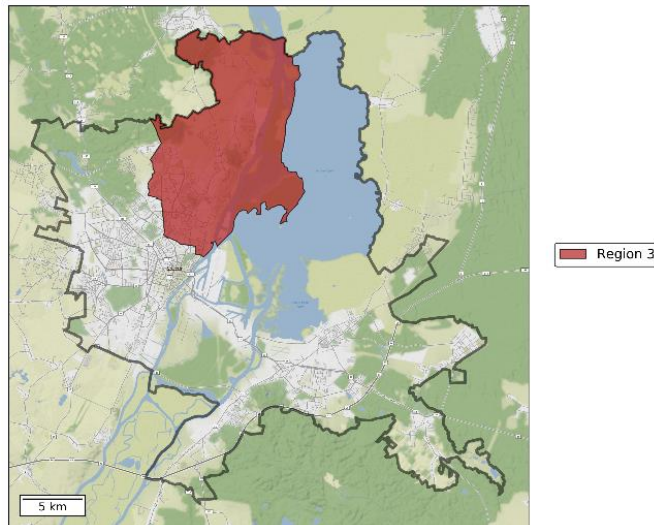
Hypothetical values (w_{ji}^h) are calculated on the basis of formula (1), omitting the market value coefficients:

$$w_{ji}^h = pow_i \cdot c_{baz} \prod_{k=1}^K \prod_{p=1}^{k_p} (1 + a_{kp}) \quad (3)$$

If the values of the drawn representative real properties (w_{ji}^{rZ}), the categories of attributes and their impact are known, along with the base price (C_{baz}) and surfaces, then market value coefficients may be evaluated for each elementary area as a geometric mean calculated from the quotients of actual and hypothetical real properties values.

The described SAREMA procedure will be used for the appraisal of approximately 1,600 land plots located in the northern part of Szczecin (known as region 3 in line with the existing geodetic division), which is the capital of the West Pomeranian province, one of 16 Polish provinces. The real properties constitute a set for which an update of annual perpetual usufruct charges was conducted. The real properties did not constitute one cohesive area. They were located in clusters of various numbers of real properties in the entire area of region 3. The area within which the appraised real properties lie is shown in Figure 1. Region 3 comprises the terrains of very different functions. There are both residential and industrial terrains, chiefly located alongside the Oder River.

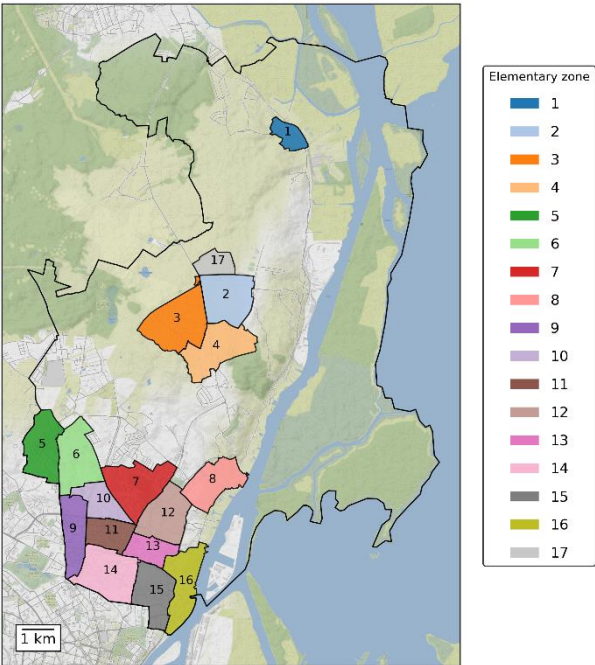
Figure 1. Location of Region 3 within Szczecin boundaries



The research team analysing the location aspect in the context of similar level of transactional prices specified 17 elementary zones. Figure 2 presents the locations of elementary zones in Szczecin’s region 3 (zones are marked with consecutive natural numbers).

When attempting to designate elementary zones, it is beneficial to have a certain preliminary division of the area in which the appraised real properties are located. Then the designation of elementary zone borders may be carried out through an adequacy analysis of that division. The division into surveying districts may constitute such a preliminary demarcation, and the research team used precisely such a division. As a result of creating the zones, some of those zones constituted one surveying district, while in other cases a specified area comprised one surveying district and a fragment of another one. Furthermore, there were cases in which an elementary zone was comprised of several surveying districts and in which one elementary zone constituted only a part of one surveying district.

Figure 2. Locations of specified elementary zones



In this study the focus was placed on assessing whether the entropy measure is suitable for determining the homogeneity of real properties within the scope of individual zones and on presenting a proposal for the modification of a classic measure in such a way that the evaluation of homogeneity could better correspond to the specificity of a real property market analysis. The entropy of distribution of a given variable enables calculating the degree of determinateness (definitiveness) of that distribution (e.g. Raschka & Mirjalili, 2017: 90), and thus the possibility of predicting different types of events or behaviours. Entropy ratio (4) is normalised and it reaches the values from the range of (0,1). A high value of the ratio indicates a significant degree of definiteness of the analysed system (the system does not demonstrate an inclination). Entropy as a measure of homogeneity has been used in a regional analysis. Early papers originate from the late 1960's and early 1970's and they include studies regarding settlement patterns (Medvedkov, 1967) and regional modelling (Wilson, 1970). More recent research (Czyż & Hauke 2015) provides evidence that Shannon entropy is useful for measuring inequality of regions. When it comes to research conducted at a city level, Cabral, Augusto, Tewolde, & Araya (2013) investigate an urban sprawl phenomenon using entropy. Boeing (2019) uses the entropy measure to reveal city's streets' order and disorder. Studies go also beyond classic entropy and draw the distinction among three interrelated definitions of thermodynamic, figurative, and statistical information entropy (cf. Purvis, Mao, & Robinson, 2019). Property market analysis utilising entropy (cf. Ge & Du, 2007) demonstrates that it is a useful tool for measuring the main variable influence on property market value. In this study, the entropy measure, which was initially used for measuring the diversification of real properties in elementary zones assumed a classic Shannon form (Doszyń, 2008):

$$H = - \sum_{i=1}^k p_i \log_k p_i, \quad (4)$$

where:

p_i – the percentage of real properties belonging to i -th class,

k – the number of classes.

In order to understand the concept of class used in this case an explanation is required. For the purpose of evaluating entropy, converted information on the variants of three market properties was used, namely:

- utility infrastructure,
- surroundings,
- transportation availability.

The variants of these characteristics were encoded in the form of natural numbers (the worst condition – 0, average condition – 1, the best condition – 2, and in the case of transportation availability, which was an attribute occurring in four states – 3) and they were combined into a three digit code. Each code value constitutes a combination of the variants of market characteristics, which was understood as a class.

As will be presented, formula (4) has a fundamental limitation in terms of evaluating real property homogeneity in elementary zones. Which is why, it was proposed that the entropy of elementary areas should be assessed in the following manner:

$$H_z = \frac{(-\sum_{i=1}^k p_i \cdot \log_k p_i) \cdot k}{L}, \quad (5)$$

where:

k – the number of classes (combination of variants of market characteristics occurring in a given elementary terrain),

L – the number of all possible combinations of the variants of market characteristics.

The modification involves a change in the evaluation of entropy depending on the number of classes featured in individual zones. In a classic approach, the zones of two and ten classes and a uniform percentage of classes will demonstrate complete entropy. However, in the case of the analysis of the real property market, these two examples need to be assessed differently. Two classes of a combination of market characteristics, in a situation when the specified terrain comprises several dozen or more real properties, mean a high degree of homogeneity, even if the percentage of both classes is close to 50%. H_z measure will distinguish between levels of entropy depending on the percentage of classes as well as their number. Measure (5), similarly to measure (4), falls within the range from 0 to 1, where 0 stands for complete real property homogeneity in a given terrain. In turn, value 1 is interpreted differently. In the case of measure (4) we deal with complete entropy when for each of n classes existing in a given zone its percentage is equal to $1/n$. Measure (5) assumes the value of one in the same case as measure (4) with an additional condition that n will mean all the possible classes of real properties that can occur in a given set.

5. Research results

Real properties forming a set that is subject to appraisal include real properties of various uses. The three most numerous uses are:

- residential use,
- industrial use, bases, storage houses,
- commercial uses.

Real properties of the above-specified uses constituted the subject of this study. The last terrain use found in the data base included rural and green areas. However, since there were only four plots to which such functions were assigned, they were not taken into consideration in the study of entropy. Furthermore, it needs to be mentioned that entropy was studied for each of the uses separately. In Figure 3 and 4 the distributions of entropy measures are presented in line with formulae (4) and (5). In the case of measure (4), the distributions are either bi- or multi-modal. A comparison of distributions of both measures of entropy for real properties assigned with particular functions shows that they differ significantly from one another. There are zones with the entropies close to the maximum value of 1, which means similar percentages of classes of real properties in a given elementary zone. The distributions of measure (5) are more symmetrical (for residential and industrial real properties). In the case of commercial real properties, the evaluation of entropy is not extremely different, as in the case of entropy measured with a classic formula.

Figure 3. Distributions of H measure taking into account the use of real property

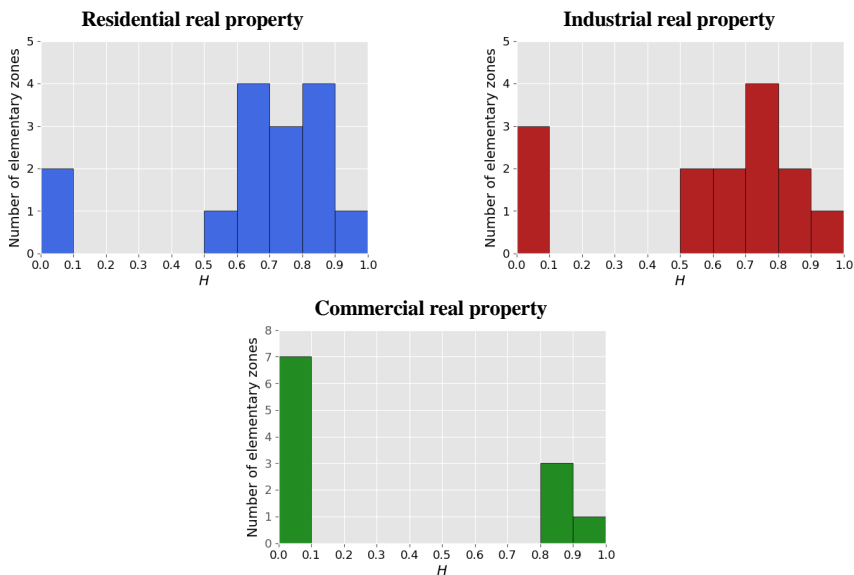
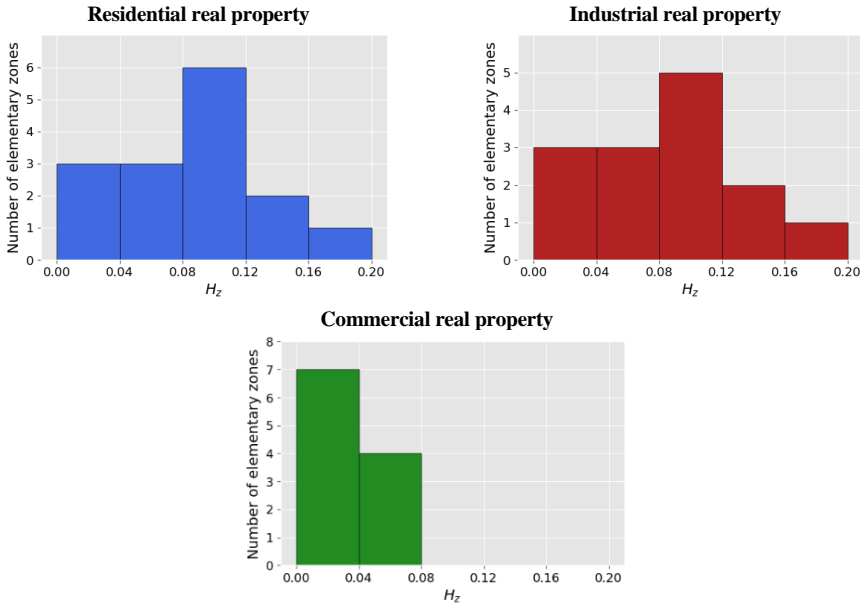


Figure 4. Distributions of H_z measure taking into account the use of real property

6. Discussion

As was demonstrated in Figures 3 and 4, the distributions of both entropy measures differ. In the last stage of the study, it will be presented how both measures describe selected elementary zones and what informational value is derived from the application of the proposed measure (5). In Figure 5, a relationship was presented between entropy measures (4) and (5) for residential real properties in individual elementary zones. Particular attention needs to be paid to the elementary terrains designated as no. 10 and no. 15, which constitute the most representative example of the usefulness of entropy measure modification. Both zones feature a high value of a classic entropy measure. However, the modified H_z measure leads to a different conclusion, which results from the fact that both zones demonstrate nearly uniform distributions of real property classes, but they still differ in terms of the number of those classes. In the case of zone 10, there are fewer of those classes (as shown in detail in Figures 6 and 7). There are only two combinations of real property market characteristics in zone 10. Even though the percentages of those two classes are close, measure (5) assumes low values. It is due to the fact that there are altogether 36 possible classes (three states of the following attributes: utility infrastructure and transportation availa-

bility as well as four states of the surroundings attribute $3 \times 3 \times 4 = 36$). The zone which contains only two classes from the standpoint of the real property market analysis is considered to be highly homogeneous. In the case of zone 15 we deal with 7 classes. Such a zone features higher entropy than zone 10, within the borders of which two combinations of variants of specified market characteristics were identified. These two examples demonstrate how the evaluation of entropy changes when it is performed with the use of a classic and modified measure. Taking into consideration the number of classes, the appraisal provides a more accurate evaluation of homogeneity because as was previously mentioned two classes, even of similar percentages, do not translate into high entropy in that context.

Figure 5. Relationship of H_z measure and H for residential real properties

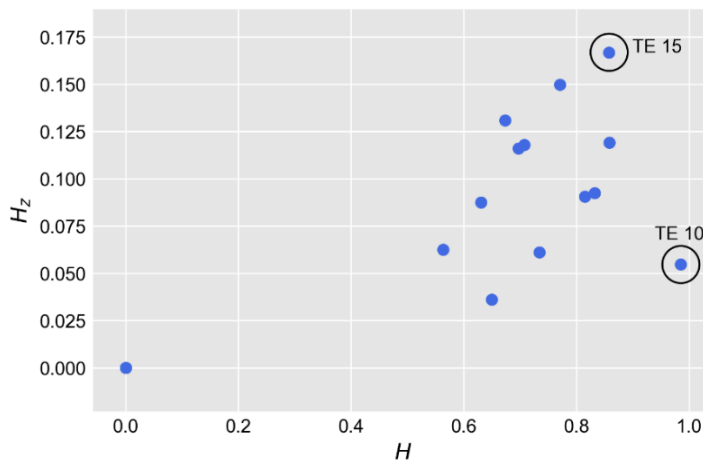


Figure 6. Elementary zone no. 10 featuring high H value and low H_z value



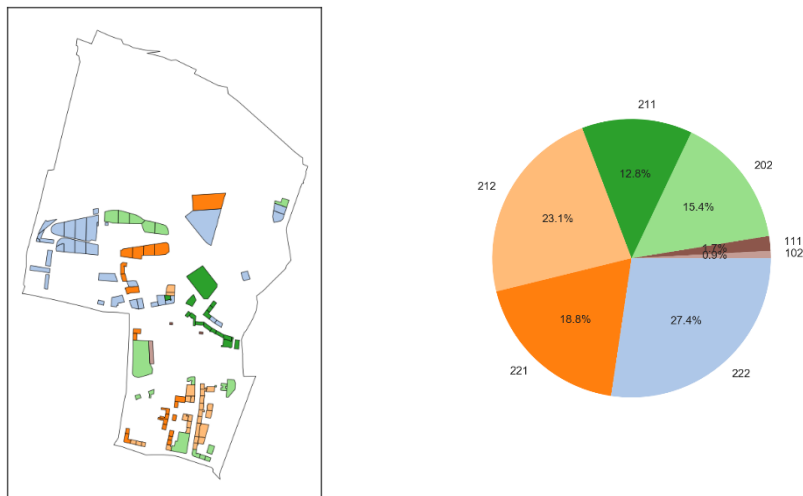
Figure 7. Elementary zone no. 15 featuring high both H and H_z value

Table 1 presents the results of entropy measurement for both analysed measures, as well as the number of classes and the number of properties located in particular zones (zones 1 and 17 did not include residential properties). The position of the zone due to entropy is given in brackets.

Table 1. H and H_z measure for residential properties, number of classes and properties count in particular zones

Zone number	H	H_z	Number of classes	Number of properties
2	0 (14)	0 (14)	1	3
3	0.771 (6)	0.150 (2)	7	216
4	0.832 (4)	0.092 (7)	4	50
5	0.697 (9)	0.116 (6)	6	178
6	0.631 (12)	0.088 (9)	5	62
7	0.673 (10)	0.131 (3)	7	316
8	0.563 (13)	0.063 (10)	4	35
9	0.815 (5)	0.091 (8)	4	66
10	0.985 (1)	0.055 (12)	2	35
11	0.708 (8)	0.118 (5)	6	49
12	0 (14)	0 (14)	1	1
13	0.859 (2)	0.119 (4)	5	25
14	0.734 (7)	0.061 (11)	3	35
15	0.858 (3)	0.167 (1)	7	117
16	0.650 (11)	0.036 (13)	2	6

Analysing other specified zones, it can be seen that for zones 2 and 12 both the classical and the modified measure indicate the lack of entropy. Both zones contain real estate with only one class of combination of market characteristics. Similar conclusions can be drawn from zone 8. It has 4 classes, but one of them covers 25 out of 35 plots. In the case of both measures (H and H_z) this zone is characterised by quite low entropy (positions 13 and 10, respectively). A slightly less spectacular change than in the case of zone 10 occurred in zone 7. Due to the largest number of classes, the indication of its heterogeneity increased significantly.

7. Conclusions

In the paper the issues of homogeneity of mass appraised sub-areas (elementary zones, taxation zones) are presented. Homogeneity is a very important aspect in terms of the effectiveness of mass appraisal, in this case conducted with the application of the Szczecin Algorithm of Real Estate Mass Appraisal (SAREMA). It was proposed that a modified entropy measure be used for the evaluation of homogeneity, which would more effectively take into consideration the specificity of real property appraisal. This kind of modification (to the best knowledge of the author) has not been proposed before. H_z measure takes into account the number of combinations of characteristics and thanks to that it enables differentiating between elementary zones of relatively homogeneous distributions of the real properties featuring given combinations of the variants of market characteristics, but with a different number of those combinations. What is more, the review of previous studies shows that the assessment of the division of the valued area is usually made *ex post*. The assessment of the homogeneity of submarkets carried out before the implementation of the target study was not subject to intensive consideration. Entropy is not a phenomenon derived from the real estate market. The formula proposed to measure can be used for the measurement of entropy in different areas as well. Especially when one deals with data subsets (subregions or submarkets) with a various number of features in them. A similar way to achieve the assessment of data subsets is to use a concentration coefficient. An entropy coefficient, however, is a better option because it does not need the data to be sorted, which is especially significant for data on a nominal scale. When there is a need to assess homogeneity of data subsets, the proposed formula can be utilised. The proposed measure will be used in the process of evaluating the specification of elementary zones. Various

studies show that both in the context of property valuation and other types of analyses, the specification of homogeneous sub-areas increases the quality of results, which in turn allows more rational economic decisions. Further studies will be carried out on the verification of property valuations in the context of whether valuations from areas with a higher degree of homogeneity are characterised by more accurate valuations.

Acknowledgements

The research was conducted within the framework of a project financed by the National Science Centre, Project No. 2017/25/B/HS4/01813.

References

- Adair, A. S., Berry, J. N., & McGreal, W. S. (1996). Hedonic modelling, housing submarkets and residential valuation. *Journal of Property Research*, 13(1), 67-83. <https://doi.org/10.1080/095999196368899>
- Bird, M., & Slack, E. (2002). Land and property taxation around the world: A review. *Journal of Property Tax Assessment and Administration*, 7(3), 31-80.
- Boeing, G. (2019). Urban spatial order: street network orientation, configuration, and entropy (Preprint). Retrieved from <https://arxiv.org/pdf/1808.00600>
- Bourassa, S. C., Hamelink, F., Hoesli, M., & MacGregor, B. D. (1999). Defining housing submarkets. *Journal of Housing Economics*, 8(2), 160-183. <https://doi.org/10.1006/jhec.1999.0246>
- Brown, P. K., & Hepworth, M. A. (2001). *Study of European land tax systems* (Working Papers). Cambridge, MA: Lincoln Institute of Land Policy.
- Cabral, P., Augusto, G., Tewolde, M., & Araya, Y. (2013). Entropy in urban systems. *Entropy*, 15(12), 5223-5236. <https://doi.org/10.3390/e15125223>
- Cellmer, R., & Kuryj, J. (2011). Określanie stref o podobnej cenności gruntów z wykorzystaniem metod geostatystycznych [Designation of the zones of similar land value with the use of geostatistical methods]. *Studia i Materiały Towarzystwa Naukowego Nieruchomości*, 19(3), 7-18.
- Cymerman, R., Fiedorowicz, E., Grabowski, R., Gwiazdzińska, M., & Kurowska, K. (2000). Wartość planistyczna jako wskaźnik w planowaniu przestrzennym [Planning value as an indicator in spatial planning]. In *Materiały konferencyjne VIII Konferencji Naukowej Towarzystwa Naukowego Nieruchomości*, „Koncepcje wartości w teorii i praktyce wyceny nieruchomości” (pp. 101-108). Olsztyn: Towarzystwo Naukowe Nieruchomości.

- Czyż, T., & Hauke, J. (2015). Entropy in regional analysis. *Quaestiones Geographicae*, 34(4), 69-78. <https://doi.org/10.1515/quageo-2015-0037>
- Dedkova, O., & Polyakova, I. (2018). Development of mass valuation in Republic of Belarus. *Geomatics and Environmental Engineering*, 12(3), 29-39. <https://doi.org/10.7494/geom.2018.12.3.29>
- Doszyń, M. (2008). *Statystyczno-ekonometryczna analiza skłonności ludzkich* [Statistical and econometric analysis of human propensities]. Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Galster, C. (1996). William Grigsby and the analysis of housing submarkets and filtering. *Urban Studies*, 33(10), 1797-1805. <https://doi.org/10.1080/0042098966376>
- Ge, X. J., & Du, Y. (2007). Main variables influencing residential property values using the entropy method – the case of Auckland (Paper presented at the proceedings of the 5th International Structural Engineering and Construction Conference. Shunan, Japan). Retrieved from <https://www.asres2007.um.edu.mo/papers/041%20-%20PAPER.pdf>
- Hozer, J., Kokot, S., & Kuźmiński, W. (2002). *Metody analizy statystycznej rynku w wycenie nieruchomości* [Methods of statistical analysis of the market in real estate appraisal]. Warszawa: PFSRM.
- Hozer, J., Kokot, S., Foryś, I., Zwolankowska, M., & Kuźmiski, W. (1999). *Ekonometryczny algorytm masowej wyceny nieruchomości gruntowych* [Econometric algorithm of mass land real estate appraisal]. Szczecin: Uniwersytet Szczeciński, Stowarzyszenie „Pomoc i Rozwój”.
- Jahanshiri, E., Buyong, T., & Shariff, A. R. M. (2011). A review of property mass valuation models. *Pertanika Journal of Science & Technology*, 19, 23-30.
- Keskin, B., & Watkins, C. (2016). Defining spatial housing submarkets: Exploring the case for expert delineated boundaries. *Urban Studies*, 54(6), 1446-1462. <https://doi.org/10.1177/0042098015620351>
- Kotkowski, P. (1999). Propozycja nowej klasyfikacji terenów miejskich [Proposition of a new classification of urban areas]. *Acta Universitatis Lodzensis, Folia Geographica Socio-Oeconomica*, 2, 115-124.
- Kopyściańska, K. (2016). *Koncepcja podatku katastralnego w Polsce na tle doświadczeń wybranych krajów* [Concept of a cadastral tax in Poland against the experience of selected countries]. Wrocław: Prawnicza i Ekonomiczna Biblioteka Cyfrowa.
- Kuryj, J. (2007). *Metodyka wyceny masowej nieruchomości na bazie aktualnych przepisów prawnych* [Methodology of mass valuation of real estate on the basis of current legal regulations]. *Wycena – Wartość – Obrót – Zarządzanie Nieruchomościami*, 4(81), 50-58.
- Mika, M. (2016). Implementacja europejskich trendów podatku katastralnego w warunkach polskich [Implementation of European trends in cadastral taxation in Polish conditions]. *Acta Scientiarum Polonorum Administratio Locorum*, 15(4), 99-110.

- Manganelli, B., Pontrandolfi, P., Azzato, A., & Murgante, B. (2014). Using geographically weighted regression for housing market segmentation. *International Journal of Business Intelligence and Data Mining*, 9(2), 161-177. <https://doi.org/10.1504/IJBIDM.2014.065100>
- Medvedkov, Y. V. (1967). The concept of entropy in settlement pattern analysis. *Regional Science Association Papers*, 18(1), 165-168.
- Ordinance of the Council of Ministers dated 29 June 2005 on general real property taxation. *Journal of Laws*, 131, item 1092.
- Peng, L., & Thibodeau, T. (2013). Risk segmentation of American homes: The case of Denver. *Real Estate Economics*, 41(3), 569-599. <https://doi.org/10.1111/reec.12005>
- Pryce, G. (2013). Housing submarkets and the lattice of substitution. *Urban Studies*, 50(13), 2682-2699. <https://doi.org/10.1177/0042098013482502>
- Purvis, B., Mao, Y., & Robinson, D. (2019). Entropy and its application to urban systems. *Entropy*, 21(1), 56. <https://doi.org/10.3390/e21010056>
- Rae, A. (2015). Online housing search and the geography of submarkets. *Housing Studies*, 30(3), 453-472. <https://doi.org/10.1080/02673037.2014.974142>
- Raschka, S., & Mirjalili, V. (2017). *Python machine learning*. Birmingham-Mumbai: Packt Publishing.
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55. <https://doi.org/10.1086/260169>
- Telega, T., Bojar, Z., & Adamczewski, Z. (2002). Wytyczne przeprowadzenia powszechnej taksacji nieruchomości [Guidelines for conducting universal real estate taxation]. *Przegląd Geodezyjny*, 6, 6-11.
- Sawilów, E. (2009). Analiza wybranych metod modelowania wartości katastralnych nieruchomości [Analysis of selected methods of real estate cadastral value modeling]. *Acta Scientiarum Polonorum Geodesia et Descriptio Terrarum*, 8(2), 27-38.
- Surowiec, G., & Malczewska, A. (2001). Automatyzacja procesu wyodrębniania stref izowartościowych nieruchomości w oparciu o zdjęcia lotnicze [Automation of the process of isolating isovalue real estate zones based on aerial photographs]. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 11, 1-29-1-35.
- Wilson, A. G. (1970). *Entropy in urban and regional modelling*. London: Pion Press.
- Wu, Ch., & Sharma, R. (2012). Housing submarket classification: The role of spatial contiguity. *Applied Geography*, 32(2), 746-756. <https://doi.org/10.1016/j.apgeog.2011.08.011>
- Wu, Y., Wei, Y. D., & Li, H. (2019, May). Analyzing spatial heterogeneity of housing prices using large datasets. *Applied Spatial Analysis & Policy*, 1-34. <https://doi.org/10.1007/s12061-019-09301-x>