Automated Valuation Models

Mass Appraisal And Its Application In Romania

Wiliam Bognar

Bucharest University of Economic Studies

Abstract

This paper focuses on the application of automated valuation models, in essence mass appraisal, on a series of data taken from the real estate market in Bucharest. Its purpose is to investigate the costs and benefits of designing a multiple regression analysis in order to appraise large real estate portfolios. It takes into account the main characteristics of different types of real estate properties and focuses on evidencing their influence on the value of the property through using regression analysis. It concludes with the main indicators of three types of properties on the market and with a SWOT analysis of what it would mean to apply a mass appraisal model at company level. It concludes that mass appraisal may lead the way in the following years and makes a series of recommendations in what regards better automated valuation models, such as neural networks, which are more able to deal with this kind of information.

I. Introduction

This study presents an overview of the basics of automated valuation models, with an emphasis on mass appraisal through multiple regressions and illustrates its use in real estate valuation of mass property for tax assessment. The problem of mass appraisal in Romania has started with NBR's Regulation 18/2009, which requires banking institutions in Romania to appraise their portfolios every three years in order to measure the adequacy of their capital. Banks are required to test the fair values of their portfolios of real estate guarantees and to report any modification in its value.

This paper aims to provide a research on what automated valuation models are, how they work applied on the Romanian market, with an emphasis on the real estate market in Bucharest and what real estate companies should look at when deciding

to implement such models. The methodology used is based on the scientific article published by Irina Ana – Maria Mihailescu and Anamaria Ciobanu, as it applies a multiple regression analysis on the Bucharest market, on sample data made of apartments, houses and land plots. Its aim is to identify the accuracy of this kind of analysis and to create the background in order to make it easier for appraisers to implement such a model. In this sense, qualitative analysis has been made, such as the map containing the characteristics of each area in Bucharest. Moreover, a SWOT analysis has been undertaken, taking into consideration the valuation model used by one of the top appraisers on the market, Colliers International, and comparing it to the multiple regression analysis.

II. MASS APPRAISAL – A HISTORICAL PERSPECTIVE

Property taxes may be a centuries-old form of taxation, but the principles and procedures used to assess them are not nearly as fixed as one might expect. The understanding of basic terms and the real estate appraisal practices and principles used for valuation continue to evolve. For example, the modern definition of market value was still being refined as recently as two decades ago. In the nineteenth century, rent capitalization served as the predominant method of value determination. (Appraisal Institute, 2002)

During the twentieth century, two more approaches to value, cost (initially referred to as scientific appraisal) and market comparison, were defined and their use became widespread. Although it may not be generally recognized, all valuation methods are rooted in the concept of the present value of future benefits, whether those benefits are in the form of housing services or an income stream. The foundation of property appraisal theory is economics (Gloudemans, 1999).

On March 28, 1874, a groundbreaking paper discussing methods of real estate valuation for taxation purposes was presented to the Social Science Association of Philadelphia. It proposed that standard procedures be established to arrive at methods for the just and equal distribution of local taxation. The 24-page paper outlined the problems associated with the existing methods of rent multipliers and price paid (acquisition value). An early definition of market value appears in the reading, "… to wit: what they would sell for separately and singly at a bona fide or fair sale, after public notice." (Cochran 1874, 13) The paper also used phrases such as "full value" and "equalization" and seemed to promote use

of a market approach. From the text of this paper, it appears that the cost approach was unknown at the time. The phrase "cost to construct" was mentioned only once, and it was not used in the context of valuation. It is noteworthy that Cochran's paper was presented only 98 years after the first published work on economic theory; only 48 years after von Thünen (1826), a German landowner, developed the first serious treatment of spatial economics, relating it to the theory of rent; and 19 years before Marshall (1893) set forth the economic theory for real estate appraisal. (Cochran, 1874; Smith, 1776; Thunen, 1826; Marshall, 1893)

The concept of spatial economics is an important topic of current property valuation research. It takes the form of location value response surface analysis (LVRSA), which provides "a means of adjusting smoothly for location in MRA [multiple regression analysis], feedback, or other valuation models." (Gloudemans, 1999, 193)

The writings of Ricardo (1817) provide another classic influence on modern appraisal theory. His work was based on the premise that given free competition in trade, the exchange value of commodities would be determined by the amount of labor expended in production. *Ricardian rent* is a type of economic rent created by variation in resource quality, which bears directly on valuation theory.

III. LITERATURE REVIEW

Multiple regression analysis (MRA) has traditionally been used as the main method of mass assessment of residential real estate property values. Methodological problems associated with MRA have been known for some time and they include non-linearity, multicollinearity, function form misspecification, and heteroscedasticity. Several AI methods, such as neural networks, have been introduced into mass assessment research to address these problems in MRA. The most commonly studied such methods are neural networks (NN)-based. (Guan, Zurada, and Levitan, 2008; Nguyen and Cripps, 2002)

Some studies have reported that NN-based approaches produce better results when compared with those obtained with MRA (Nguyen and Cripps, 2002; Peterson and Flanagan, 2009), while others have reported comparable results using NN-based methods but have not found NN-based methods to be superior (Guan and Levitan, 1997; Limsombunchai, Gan, and Lee, 2004).

Authors of other studies, however, are more skeptical of the potential merits of the NN-based approaches (Limsombunchai, Gan, and Lee, 2004). The main criticisms include the black box nature of NN-based methods, lack of consistency, and difficulty with repeating results. Worzala et al. (1995) find that NN-based methods do not produce results that are notably better than those of MRA except when more homogeneous data are used. McGreal et al.'s (1998) study leads their authors to express concerns similar to those by Worzala et al. (1995). Rossini (1997) finds MRA yields consistent results, while NN results are unpredictable.

In addition to NN-based methods, other AI methods have also been explored in real estate valuation, including fuzzy logic, MBR, and adaptive neuro-fuzzy inference system (ANFIS) (Bagnoli and Smith, 1998).

IV. METHODOLOGY

My database comprises of three individual databases, relating to apartments, houses and land plots. In my analysis, I have used the databases in order to reach conclusions about the behaviour of the buyer or the seller on the real estate market by looking into what influences the value of these properties in the eyes of the consumers. To begin with, I set a minimum threshold for every property type. Similar research and classes were made on a sample with a minimum size of 15 inputs. However, for the sake of the study and in order to reach more conclusive results, I have resorted to using 30 data inputs. In order to do that, I have research the best known real estate websites in Romania to find sellers and their real estate ads. I have gathered information mainly from www.imobiliare.ro, but I have also used www.mercador.ro and www.tocmai.ro, the websites in this field with the highest traffic according to www.traffic.ro.

First, I decided before getting into talking with the sellers and the real estate agents what kind of characteristics I was looking for. For a better understanding of the market, I have spoken to our colleague who is in charge with the Residential Department, as part of her consultancy work. Together, we came up with the following characteristics which influence the value of a certain property. For an apartment, the following were find to be important:

- Total built area
- Number of bathrooms
- Number of rooms
- Detached or not
- Interphone
- Parking space or not

- Boiler
- Air conditioning
- Distance to downtown
- Asking price

The same rationale was applied in looking at the other two types of properties, the houses and the land plots.

House – important charactertistics

- Number of rooms
- Total built area
- Total land area
- Construction year
- Finishes – with a grade of 1 to 10, awarded by me, based on the pictures attached and my discussion with the owners or agents
- Distance to downtown
- Asking price.

Similarly, important characteristics for a land plot are:

- Terrain type
- Use – residential, industrial, commercial
- Distance to downtown – km
- Minutes to downtown – minutes
- Area
- Frontage
- Access – asphalt road, stone road, soil road
- Utilities – water, gas, electricity, sewage
- Asking price

Now, having all the relevant information in place, I went on to check the ads. For each individual database, I selected a series of 100 random ads. I then proceded to filter the information according to my needs. The most important principles were the accuracy and completeness of the characteristics in the ads. After selecting the most complete, I went on to talk to the owner or real estate agents in order to check the validity of the data. I encountered a big surprise when I realised there were a lot of wrong ads with incomplete or wrong information, which lead me to believe that the verification stage was a must.

In order to verify the data, I went on to call the owners or their representatives. I impersonated a buyer, as there were very few prepared to give me information for a study that I was conducting. I verified each and every ad I selected and then I comprised the databases that can be found in the Appendix.

V. CASE STUDY – MASS APPRAISAL APPLIED TO BUCHAREST PROPERTIES

First of all, my analysis is focused on Bucharest solely. It is done in such a way because, as an appraiser at Colliers International, our expertise is mostly focused on Bucharest. Being the capital, there are numerous transactions in the area, which is a highly complex and large area. I believe that, in order to understand the nature of the transactions with different types of residential properties, may they be land plots, houses or apartments, one must first look at an economic analysis of Bucharest and its macroeconomic conditions, as these are the main factors that influence the prices of properties by influencing the supply and demand of properties.

Bucharest is the capital city of Romania, the most populated city and the most important economic centre of the country. It is located in the south, on the banks of the river Dambovita, within Ilfov county, whose official residence it is, without actually being part of the territory administered by it. Bucharest has a special status in the country, being the only city that does not belong to any county. Ilfov county is bordered by the following counties: Prahova to the north, Ialomita to the north – east, Calarasi to the south – east, Giurgiu to the south and south – west and Dambovita to the north – west.

The net average monthly salary for Bucharest has increased from October 2012 to October 2013 by approximately 6.8%, up to a level of 500 EUR. In 2012, the net average salary in Bucharest was 488 EUR, 41.8% over the national average (i.e. 344 EUR). In the first half of 2013, the net average salary in Bucharest registered a value of 512 EUR, higher than the national average of 365 EUR.

The unemployment rate in October 2012 – October 2013 period has remained constant, at 2.1%. The average unemployment rate for 2012 was 2%, situated at a significantly lower level than the national average of 5.2%. The average unemployment rate for H1 2013 was of 2%, compared to the 5.9% registered at national level.

In implementing the mass appraisal model for the three sets of data, I have gone through the following methodology in testing and analysing the sample data. To begin with, I have analysed the data in terms of descriptive statistics in order to better understand it better. Then, I have tested the correlation between each pair of variables (each independent variable coupled with the dependent variable), to analyse how strong are they correlated and what is the direction of that correlation (positive or negative). Then, I used a

scatter plot and for each set in order to identify outliers and to see if there are any properties with abnormal pricing or characteristics, compared to the other in the sample data. Then, I introduced all the variables in a multiple regression and I have analysed the results, from a quantitative point of view, but also from a qualitative perspective, analysing what seemed to influence the price on the market and what seemed to drive buyers and sellers to invest in a certain property. All these tests and methods have been tested in IBM SPSS Statistics.

LAND PLOTS ANALYSIS

The sample data I used for applying mass appraisal to land plots valuation contains 30 entries. The average land plot in my analysis has roughly 740 sqm (as I mostly focused on residential land plots, so the areas are suitable for these kind of developments), a frontage of 19 meters (which I find sufficient for a residential project), are located about 4 km away from the downtown (in my analysis, I considered the downtown as being "KM 0", located in Sfantu Gheorghe Square), are roughly 11 minutes away from the city centre (driving distance) and located at under 15 minutes walk from the nearest subway station. The average asking price for these properties has been around 570,000 Euros (which is the currency mostly used on the local real estate market).

The multiple regression results reflect the relationship between the factors and the price. To begin with, it has an R-square of 60.6%, with an adjusted R-square of 52,4%. In terms of regression results, an R-square over 50% is a great result in this case, keeping in mind that only 30 entries were used in my sample data.

Reviewing the ANOVA tests, it is clear that the regression is significant, given the value of F (7.389). The coefficients also draw a good picture. The surface of the land plot (in this case, SqrtSurface) and the frontage are strong indicators of the price of the land plot. Moreover, although at a lower significance level, the time it takes to walk to the nearest subway station is also a good coefficient in my relationship.

In terms of the buyer's behaviour, it is clear that they are analysing the time to the nearest subway station when making a decision. On the other hand, it does not have such a high importance as the area of the land plot and its frontage. This can probably be explained that most people deciding to buy a land plot for a residential project most likely own a car and use it to drive around the city. The fact that they do not use the subway on a regular basis

makes it less important for them, but they still keep it in mind when making a decision. In terms of the time to downtown and the distance to the downtown, things are a bit different. These factors are not so relevant in my analysis for one of two reasons. First of all, there are expensive areas where housing project are developed (such as Primaverii), where the buyers are looking more to be in the area and have their privacy than being closer to the downtown. Secondly, the city's centre is not a desired destination for houses, as there are almost no land plots left in the area, the blocks of flats are tall so there is no privacy and a lot of properties on the market have architectural monuments built on their area, which makes them useless for someone who wants to build a new house.

After adjusting the model and calibrating it, the regression equation becomes:

LnPrice = 11.526 + 0.0386 x SqrtSurface + 0.0359 x Frontage

Testing the model was made with three random properties from www.imobiliare.ro, after calling the real estate agents and verifying the published information. As it can be seen, the land plots valuation mass appraisal model is quite powerful. In the three tests made, the average error of the model is around 16%, which I consider to be a very good percentage error. It must be kept in mind that this model is only based on a sample data of 30 properties, but still proves to provide quite relevant information about what the price may be. Moreover, it must be kept in mind that the practice on the local market is that each price is negotiable up to around 10%, which would make the error even smaller.

HOUSE ANALYSIS

The sample data analysed consists of 30 houses from Bucharest, with an average total built area of about 300 sqm, another 300 sqm of land free of constructions, with around 6 to 7 rooms, built around 2007, with superior finishing, located at about 7 kms away from the city's centre and priced around 290,000 Euros.

The multiple regression results reflect the causality between the independent and the dependent variables. In this case, the R-square is 92.8%, with an adjusted R-square of 82.4%. This suggests a very good regression, reflecting a clear relationship between the factors taken into consideration and the price of the properties.

Reviewing the ANOVA tests, it is clear that the regression is significant, given the value of F (over 23). The coefficients of

the regression equation are mostly significant. The strongest are the built area, the finishing and the distance to downtown. However, the land area is still a very good variable, with a significance level of 0.1, while the construction year and the number of rooms do not seem to matter in the pricing of a house.

The behaviour of buyers in this market seems to be how one would expect it to be. The buyer is not concerned with the number of rooms, probably because a house can be refurbished quite fast and the rooms can be reorganised, and definitely not concerned with the construction year of the house. This latter is probably due to the fact that a buyer is definitely focusing on the quality of the finishing and the total built area, so it is important how big and well decorated the house is, while the age is not of paramount importance. The land area is another strong indicator, because people moving away from apartments to houses are looking forward to having a garden for the kids or pets to play. However, an interesting fact is that in this equation, the distance to downtown becomes an important factor. This may be because as most house owners drive to work, it is important for them to know how far away the city centre is, especially that a lot of the spare time activities are concentrated around the Old Centre.

As a result of my previous analysis, the regression equation that resulted is:

LnPrice = 9.443 + 0.101 x SqrtBuiltArea + 0.04 x SqrtLandArea + 0.095 x Finishes – 0.03 x DistanceToTown

APARTMENTS ANALYSIS

The actual equation created as a result of the regression analysis sheds more light on the behavior of the consumer of this particular market. I am taking into consideration as being relevant and significant those variables with a p-value of less than 0.1. In this particular model, I found as significant the total area of the apartment and the distance to downtown. It is also true that the presence of a boiler, the presence of a videophone are quite significant, but below my target confidence level. My expectation was that the number of bathrooms and the floor number will also be significant, but apparently the buyer on the market is not too concerned with the floor where the apartment is or with how many bathrooms it has. The detachment is also not a factor, but to be honest, most of the apartments on the market in Bucharest are detached. The videophone and the parking spot appear not to be an issue probably especially that a lot of blocks now have a videophone, and only very few have a dedicated parking spot per apartment. However, it is

clear that people looking for apartments in Bucharest do look at how far away they are located in comparison to the city's centre. This is probably due to the fact that most jobs are around downtown, while most of the entertainment is based there. The Old Town is probably the most popular area for pubs, bars and clubs in Bucharest, so it follows naturally that people want to be as close as possible to the place where they go out.

As a result of my previous analysis, and taking into account the significant variables, the regression equation that resulted is:

LnPrice = 9.95 + 0.142 x SqrtTotalArea - 0.036 x DistToDowntown

VI. CONCLUSIONS

The application of statistics in the real estate appraisal sector could lead the way to corporations changing their appraisal methods in order to adapt to the new market conditions. Automated valuation models are able to deal with large quantities of data, which is a characteristic of the present big data era we live in. Being able to comply with large amounts of market series and variables will soon be an advantage of any appraising firm on the market.

However, as concluded from my analysis, big data also requires a big investment. The investment relates to computing power, to teams of people well prepared in fields like statistics and big data analysis and time to be able to develop a model and to implement it successfully. All these require quite large sums of money, which may be a problem for smaller companies, but may lead to a competitive advantage for large companies. It follows naturally that banks' portfolios are becoming larger and larger, so the ability to appraise them and to deliver within a shorter time will be an asset.

Having analysed the background of mass appraisal in Romania and having tested already a simplified model, I can only conclude that this field has only laid its cornerstone and its development is in the making. It will be down to companies like Colliers International, JLL, Knight Frank The Advisers, CBRE and other large corporations if they will follow this path or simply stick to applying the classic market comparison applied to large series of data, like Colliers International does today.

All in all, I believe that this massive investment may not be the way to go, but this depends on the country and the market there. In Romania, where appraisals are priced a lot less than in the Western

economies, managers may wish to maximise their ROI and go along with the older methods in order to gain profits and turnover, than to embrace the new technologies, such as multiple regression analysis, neural networks and hybrid models, and lead the way into innovation and change. This latter will probably come in the following years from the West, when new procedures and methodologies will be adopted at the regional offices of these firms.

VII. BIBLIOGRAPHY

Appraisal Institute. 2002. *The dictionary of real estate appraisal*, 4th ed. Chicago: Appraisal Institute.

Bagnoli, C. and H. C. Smith, The Theory of Fuzzy Logic and Its Application to Real Estate Valuation, *Journal of Real Estate Research*, 1998, 16:2, 169-200.

Bonissone, P. P. and W. Cheetham, Financial Applications of Fuzzy Case-Based Reasoning to Residential Property Valuation, Proceedings of Sixth International Conference On Fuzzy Systems (FUZZ-IEEE'97), Barcelona, Spain, 1997, 37-44.

Byrne, P., Fuzzy Analysis: A Vague Way of Dealing with Uncertainty in Real Estate Analysis?, *Journal of Property Valuation and Investment*, 1995, 13:3, 22-41.

Cochran, T. 1874. Methods of valuation of real estate for taxation. *Penn Monthly* May.

Do, A. Q. and G. Grudnitski, A Neural Network Approach to Residential Property Appraisal, *The Real Estate Appraiser*, 1992, 58:3, 38–45.

Eckert, J., ed. 1990. *Property appraisal and assessment administration.* With R. Gloudemans and R. Almy, tech. eds. Chicago: International Association of Assessing Officers.

Gloudemans, R.J. 1999. *Mass appraisal of real property.* Chicago: International Association of Assessing Officers.

Guan, J. and A. S. Levitan, Artificial Neural Network Based Assessment of Residential Real Estate Property Prices: A Case Study, *Accounting Forum*, 1997, 20:3/4, 311-26.

Guan, J., J. Zurada and A. S. Levitan, An Adaptive Neuro-Fuzzy Inference System Based Approach to Real Estate Property Assessment, *Journal of Real Estate Research*, 2008, 30:4, 395-420.

Gonzalez, M. A. S. and C. T. Formoso, Mass Appraisal with Genetic Fuzzy Rule-Based Systems, *Property Management*, 2006, 24:1, 20-30.

Kauko, T. and M. D'Amato (2008), *Mass Appraisal Methods*. Oxford :Wiley-Blackwell.

Limsombunchai, V., C. Gan and M. Lee, House Price Prediction: Hedonic Price Model Vs. Artificial Neural Network, *American Journal of Applied Sciences*, 2004, 1:3, 193-201.

Mark, J. and M. Goldberg, Multiple Regression Analysis and Mass Assessment: A Review of the Issues, *Appraisal Journal*, 1988, 56:1, 89-109.

McGreal, S., A. Adair, D. McBurney and D. Patterson, Neural Networks: The Prediction of Residential Values, *Journal of Property Valuation and Investment*, 1998, 16, 57-70.

Marshall, A. 1893. On rent. *Economic Journal* vol. 3. http://socserv2.socsci.mcmaster.ca/~econ/ugcm/3l l3/marshall/rent [accessed 15th May, 2014].

Marshall, A. 1920. *Principles of economics*,8th ed. Repr., London: Macmillan for Royal Economic Society, 1961.

Nguyen, N. and A. Cripps, Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks, *Journal of Real Estate Research*, 2002, 22:3, 313-36.

Peterson, S. and A. B. Flanagan, Neural Network Hedonic Pricing Models in Mass Real Estate Appraisal, *Journal of Real Estate Research*, 2009, 31:2, 147-64.

Ricardo, D. 1817. *The principles of political economy and taxation*. London: Murray.

Rossini, P., Artificial Neural Networks Versus Multiple Regression in the Valuation of Residential Property, *Australian Land Economics Review*, 1997, 3:1, 1-12.

Smith, A. 1776. *An inquiry into the nature and causes of the wealth of nations.* London: Strahan & Cadell. Repr., Danbury, CT: Grolier, 1984.

Von Thünen, J.H. 1826. *Der isolierte staat in beziehung auf landwirtschaft und nationalökonomie.* Germany: unknown binding.

Worzala, E., M. Lenk and A. Silva, An Exploration of Neural Networks and Its Application to Real Estate Valuation, *Journal of Real Estate Research*, 1995, 10, 185-202.