

Mass Appraisal of Property Tax Base based on Multivariate Linear Regression—Empirical Study on Four Districts of Jinan

Shan TIAN

School of Economic Management, Shandong Women's University, Jinan, Shandong, China

Tianshanhaoxingqing@163.com

Keywords: Property tax base; mass appraisal; multivariate linear regression

Abstract. In this paper, a multivariate linear regression analysis is made to the effective transaction data of 226 housing stocks in four administrative regions (Tianqiao District, Shizhong District, Lixia District, and Huaiyin District) of Jinan, a mass appraisal model of real estate values is established, and also the transaction prices of 226 housing stocks are compared with the appraised prices. In this study, it is thought that an efficient, low-cost, objective, and accurate mass appraisal can be made to property tax base through the establishment of a unified real estate information database. In the mass appraisal system, geography information technology should be integrated so as to improve the objectivity and accuracy of the appraisal results.

Introduction

Since the uniformly standardized collection of the property tax was proposed in the Third Plenum of the 16th CPC Central Committee on October 14, 2003 and relevant charges were canceled, a lot of fundamental work has been made by China's government to the collection of property tax. However, there are many problems necessary to solve, and property tax base and tax rate should be primarily considered because they directly play a decisive role in the amount of payable tax and also has a tie with the fairness of the taxation. Tax base is the basic amount of the payable tax for calculation, so it is the basis of tax assessment. The main basis for determining the base of the property tax collected in the transferring process of real estate lies in the appraisal value of property. Therefore, accurately estimating the appraisal value of real estate tax is the foundation of determining property tax base and deciding payable property tax amount, and also the key to the success of collecting property tax. The appraisal scope of real estate value for the purpose of taxation generally includes all real estates in the country, so the amount in appraisal is huge. Appraisal time-point generally refers to the unified appraisal base date provided in the tax law, and the time points are required to be consistent. To ensure the fairness of the taxation, appraised objects should be periodically assessed. Therefore, time and cost should be considered seriously in the appraisal, and mass appraisal techniques are necessarily introduced to the appraisal of property tax base.

Mass appraisal is a new computer technology based appraisal method system, in which a corresponding asset appraisal model is established with three principles of asset appraisal and also the appraisal model is repeatedly corrected and evaluated with mathematical statistics technology and other related technologies. Mass appraisal rose in foreign countries in the 1970s, but has developed into a new appraisal mode of great influencing power currently. At present, property tax base is commonly assessed with mass appraisal in the countries such as the United States, Britain, Germany, and Finland using market value as tax assessment basis.

As early as 1919, scholars in the western world had applied the multivariate regression analysis in statistics to price appraisal. Due to the absence of computer technology at that time, early technical analyses on mass appraisal were manually completed by economists using multivariate regression analysis, so a lot of time and energy were required to cost. Later, with the popularization of information technology and computer technology, computer-assisted mass appraisal (CAMA) has been widely applied to property tax base appraisal, public administration value appraisal, and the mass real estate value appraisal in business activities. Now, this technique has been used in the

countries having collecting property tax to appraise property tax base.

A most critical technology in the application of mass appraisal is automated valuation model (AVM). AVM is a series of computerized econometric models that are used for appraising the value of properties. More specifically, in the appraisal of property tax base, AVM, by relying on the feature information such as transaction price, location, area, and orientation collected in the real estate market, implements and also generates the appraisal value of mass real estate. In the implementation of mass appraisal, computer-assisted mass appraisal emerged when AVM was closely combined with the computer system. CAMA is a method applying statistics theory and unified programs to value mass real estate. In CAMA system, the tasks such as market analysis and valuation model establishment can be completed using regression technology, and the appraisal value can be automatically output after real estate related information is input.

In this paper, using mass appraisal based on multivariate linear regression analysis, the mass appraisal of the real estate value in four administrative regions (Tianqiao, Shizhong, Lixia, and Huaiyin) of Jinan is empirically studied.

Data collection and classification

The sample data of this study is about the real transaction cases of 248 residential housing stocks in Jinan of Shandong, and mainly sourced from the successful transactions provided by Century 21 China. The transactions of these cases were made throughout 2012. According to the sample contents, it mainly included building address, house number, the nature of property rights, total sale price, single sale price, completion date, building area, floor area, the number of floors, the total number of floors, housing types, decoration, facilities, rooms, etc. And the absent items in several transactions were supplied by the author through field research. In these samples, 71 transaction cases were in Tianqiao District, 70 were in Huaiyin District, 65 were in Shizhong District, and 42 were in Lixia District; 22 samples were randomly chosen for testing the model, and the remaining 226 were used for the establishment of the model. All transaction cases in the samples sourced from Jinan real estate market, so they were in the same market area.

12 variables in table 1 were involved in the establishment of the model. Different from previous studies, a new characteristic value (facilities) is introduced in this study because the city where samples were collected is located in north China. This characteristic variable is used for verifying if there are central heating equipments and pipeline natural gas available in the property, namely if there are double-gas source facilities available. In general, this variable is positively correlated with the housing sales price.

The establishment and analysis of mass appraisal model

At present, the techniques mainly used in the establishment and analysis of mass appraisal model include time trend analysis, artificial neural network analysis, multiple regression analysis, feedback, etc. Learning from foreign experience, multivariate regression analysis model has been praised highly because of its characteristics such as easy-to-use, easy-to-understand, and high accuracy results, etc. Therefore, multivariate regression analysis was chosen by the author for establishing the model, and this is conducive to the specific implementation of property tax base appraisal. The specific steps are as follows:

First, the following multivariate regression model was established by setting the sales price P of the property as dependent variable and the other 11 variables in table 1 as independent variables:

$$P = \beta_0 + \beta_1 \times R + \beta_2 \times GL + \beta_3 \times B + \beta_4 \times LI + \beta_5 \times LA + \beta_6 \times W + \beta_7 \times RS + \beta_8 \times AGE + \beta_9 \times DS + \beta_{10} \times I + \beta_{11} \times L \quad (1)$$

The first regression result as follows was obtained using SPSS18.0: the P values of rooms number, halls number, toilets number, housing years, and floors number did not pass the check, while the P values of other explanatory variables were very significant. Therefore, a multivariate regression model was necessarily established again for improving the explanation ability of real estate appraisal.

Five variables (the number of rooms, the number of halls, the number of toilets, and the years of

floors) were removed from the model, and then the second regression result was obtained as shown in table 2.

After five variables were removed in turn, a result was obtained as follows: model's determining coefficient was $R^2=0.731$ and adjustment's determining coefficient was 0.723, suggesting this model is with high higher goodness of fit on the whole; model $F=98.952$ and concomitant probability value <98.952 suggest that the model is obviously of statistical significance; the tolerance of each variable in the model was greater than 0.1 before and after transformation, suggesting the problem of co-linearity does not exist in this model. All variables in the model passed the check, and an appraisal model was obtained as follows according to the result of regression analysis:

$$P=-59506.228+58080.342 \times GL-35112.641 \times B+6972.190 \times LA+29134.010 \times RS+25098.127 \times DS+105639.321 \times I \quad (2)$$

Checking the Model

After model was established, the quality of the appraisal should be checked. Checking the quality of the appraisal should be implemented from appraisal level and consistency. Appraisal level should be checked through median ratio and mean ratio; consistency should be checked through discrete coefficient and price-related difference.

First, the appraisal level was measured. It was in 0.9-1.1 according to *Standard on Ratio Studies* made by IAAO (1999). In this study, the median ratio was 0.99 and average ratio was 0.99, which were in the reasonable range, so the appraisal level accorded with the standard.

Second, the appraisal consistency was checked.

Discrete coefficient COD was calculated first. The consistency of the appraised values was checked by COD through measuring if the difference of payable tax amounts is caused by improper appraisal work.

$$COD = \frac{\sum |AR_i - M|}{nM} \times 100 \quad (3)$$

Where,

$AR_i = A_i / S_i$;

A_i was the appraisal value of the i th property;

S_i was the market value;

AR_i was the ratio between the appraisal value and market price of the i th property;

M was the median of AR_i

Then, the related difference PRD of the appraisal value was calculated again. PRD is the index determining the fairness of the taxation. It is used for checking if the payable tax amount of the high-value property is lower than that of the low-value property.

$$PRD = \frac{\sum AR_i}{\sum \frac{A_i}{S_i}} \quad (4)$$

According to *Standard on Ratio Studies* made by IAAO (1999), the value of COD should be in 0.00~10.0 and the value of PRD should be in 0.98~1.03 in the appraisal areas of the uniform housing types. This appraisal result was $COD=7.48$ and $PRD=1.01$, which were reasonable. After the check, the consistency of the mass appraisal in this study also conformed to the standard, as shown in table 3.

Conclusion and prospect

Through the multivariate linear regression analysis on the transaction data of the housing stocks in Tianqiao District, Shizhong District, Lixia District, and Huaiyin District of Shandong Jinan in 2012, it is known that the geographic position, the number of verandas, building area, housing types, decoration, and facilities (double-gas source) play a very significant effect on the total value of real estate. The number of verandas has a negative effect on real estate prices, and the other five factors have a positive effect on real estate prices and the effects of building area, geographic position, and

facilities are the most significant. The result of model checking shows that the mass appraisal based on multivariate linear regression model can be chosen by Jinan in the future property tax base appraisal.

Accurately quantifying the influencing factors of real estate prices is a guarantee for the accuracy of mass appraisal result. If the factors are quantified using the traditional method, not only a lot of time is cost, and also many factors are difficult to estimate and the appraisal is subjectively made by appraiser, so that the appraisal result is not accurate. In the future practice, CMAM and GIS (geographic information system) can be combined for improving the accuracy of the results, so as to ensure the fairness of the taxation.

GIS, emerged in the 1960s, is a new subject combining geographic information science with computer. The functions of GIS include the collection, storage, analysis and visualized expression of spatial information. GIS can provide visualized graphics and images in addition to text-based data, and also can implement input/output, editing, modifying, query and dynamic management on the objects of the spatial attributes. Under the support of GIS, the accuracy of the appraisal result can be significantly improved if the local fundamental geographic database of real estate is applied and also the surrounding environmental facilities and properties in these transaction cases are analyzed, evaluated and scored using all query and analysis functions of GIS. However, a large amount of real estate transaction data is required by mass appraisal, so the improvement of the real estate registration system and the establishment of a unified real estate information database are urgently demanded in China, aiming at providing a technical support for the collection of property tax.

References

- [1] Lin'na Zhou. Research on the Mass Appraisal System based on the Purpose of Collecting the Property Tax [D]. Xiamen University, 2007.
- [2] Zhimin Qiao, Defeng Li, Wenjun Zou, Zhanqiang Zhou. Appraisal of Immovable Property Values based on the Collection of the Property Tax—Empirical Analysis on the four districts of Beijing [J]. Urban Development Research, 2009, (03): 81-86.
- [3] Jiang Cui. Research on the Mass Appraisal Method for the Base of the Property Tax [D]. Nanjing University of Finance, 2010.
- [4] Daojing Han. Research on the Problem of Mass Appraisal for the Tax Base Valuation of Property Tax [D]. Northeastern University, 2009.
- [5] Qin Guo, Jinghong Wei. About partition to construct China's property tax base assessment model of design thinking [J]. Research of Finance and Accounting, 2010,(06):27-28.

Table 1: The variables of the model

The Name of Variable	The Symbol of Variable	The Type of Variable	Definition
Housing Sales Price	P	Numerical Variable	The total sales price of each property (Unit: 1RMB)
The Number of Rooms	R	Numerical Variable	The number of the rooms in the property of each transaction case (Unit: room)
Geographic Position	GL	Classified Variable	Tianqiao District=1; Huaiyin District=2; Shizhong District=3; Lixia District=4
The Number of Verandas	B	Numerical Variable	The number of verandas in the property of each transaction case (Unit: veranda)
The Number of Halls	LI	Numerical Variable	The number of drawing rooms and dining rooms in the property of each transaction case (Unit: room)

Building Area	LA	Numerical Variable	The building area of each property (Unit: m ²)
The Number of Toilets	W	Numerical Variable	The number of the toilets in the property of each transaction case (Unit: room)
Housing Type	RS	Classified Variable	multi-storey housing (≤ 7)=0; Small high-rise housing (8~12, including 12)=1; High-rise housing (>13)=2
Years of Housing	AGE	Numerical Variable	Base period is the year of 2012, and the year of the housing in 2012 is 0 (Unit: year)
Decoration	DS	Classified Variable	Blank or clear-water=1; simple decoration=2; medium decoration=3; refined decoration=4; luxury decoration=5
Facilities	I	Dummy Variable	Non double-gas source facilities=0; double-gas source facilities=1
The Number of Floors	L	Numerical Variable	The number of floors in the property of each transaction case (Unit: floor)

Table 2: Coefficients

Model	Non-standardized Coefficients		Standardized Coefficients	t	Sig.	Co-linearity Statistics	
	B	Std. Error				Tolerance	VIF
(Constant)	-59506.228	50917.800		-1.169	0.244		
Geographic Position	58080.342	8630.594	0.239	6.730	0.000	0.979	1.021
The Number of Verandas	-35112.641	14006.309	-0.092	-2.507	0.013	0.914	1.094
Building area	6972.190	363.181	0.713	19.198	0.000	0.893	1.120
Housing type	29134.010	11631.539	0.099	2.505	0.013	0.782	1.279
Decoration	25098.127	8085.780	0.110	3.104	0.002	0.973	1.028
Facilities	105639.321	26027.895	0.158	4.059	0.000	0.816	1.226

Table 3: The ratio between market value and appraisal value

No.	Market value Si (RMB)	Appraisal value Ai (RMB)	ARi=Ai/Si	ARi-M	No.	Market value Si (RMB)	Appraisal value Ai (RMB)	ARi=Ai/Si	ARi-M
1	800,021	739,821.96	0.92	0.5011	12	850,000	883,447.32	1.04	0.0043
2	850,000	794,543.01	0.93	0.1002	13	850,000	889,152.54	1.05	0.0111
3	876,000	841,314.04	0.96	0.0746	14	830,000	874,848.08	1.05	0.0190
4	860,000	831,851.35	0.97	0.0677	15	850,000	901,573.25	1.06	0.0257
5	850,000	826,027.18	0.97	0.0632	16	850,000	907,278.47	1.07	0.0324
6	849,966	830,460.43	0.98	0.0579	17	960,000	1,028,254.20	1.07	0.0361
7	850,000	833,375.02	0.98	0.0546	18	870,000	934,555.47	1.07	0.0392
8	860,000	852,767.92	0.99	0.0434	19	820,000	904,048.47	1.10	0.0675
9	850,000	863,206.52	1.02	0.0195	20	850,000	939,015.44	1.10	0.0697
10	830,000	858,033.77	1.03	0.0012	21	900,000	999,559.56	1.11	0.0756
11	850,000	882,180.35	1.04	0.0029	22	850,000	983,060.66	1.16	0.1215
Median						850000	878514.21		
Median ratio							1.03		
Mean						854,818	881,744.32		
Mean ratio							1.03		
Discrete coefficient									6.54
Price-related difference									1.00