The prediction of the chance of selling of houses as the factor of financial stability

Ceyhun Abbasov
Department of World Economy
Qafqaz University
Baku, Azerbaijan
ceyhunabbasoglu@gmail.com

***Abstract* - This paper has devoted to the estimating of the selling chance of houses in market. It is a clear that housing market is a very important for the financial stability. The database of our approach consist of the price of the houses and their quality indicators as the repair level, distance from the city centre, projects of the building, the place of the flat in a building and so on. In the base of this database, we calculated G score and found that if the computed G value is equal to 0.39 and greater than that one it is classified as the high chance of selling the house, but in the case of being less than 0.39 it is classified as none selling chance. We hope that this calculation can be useful for the banks. Because, they always face the question about what is a probability of selling of houses encumbered under loan agreement in banks. So, banks are able to estimate the probability of the realization of mortgaged houses.**

***Index Terms* - Selling chance of houses, I and II type errors, Hedonic price method, Edmister's approach, Kappa test.**

I. Introduction

The bank system is very sensitive to housing market. There are kind of loans provided under encumbrance of houses. It raises an important question in the process of lending: What is the probability of selling of houses encumbered under loan agreement. If the granted loan is not repaid by the borrower, then the bank is going to sell the same house charged against the case of non-payment. In this context it must be very interesting to the bank to know the probability of selling the house before lending money. The bank needs to have the statistical data base for assessing the probabilities. By the standpoint of the more efficiency the qualitative properties of houses should be included in this data base. From the end of the previous century till now the house market has been developing in our country, especially in Baku city. The development of house market is explained by several factors: accelerated urbanization process, increasing investment in houses, real estate direction of funds earned abroad by residents etc. As heterogeneous goods a house has three major peculiarities unlike other goods. First, it includes many other goods and meets different requirements of the family. Second, the house is an immovable. That is why the location of house is playing very important role in forming its price. Third, the demand in the house market is more stable than in the market of other goods, which could be explained by the fact that house cannot be substituted by other goods. First section of working paper has been devoted to form the quality statistical data base. The second section describes the methodology of our approach and the related literatures. Third section is about the empirical result of the statistical analysis. Also this paper has conclusion and reference sections.

II. Statistical data base

The qualitative statistics of the housing market is understood as the information about the properties of the houses. For example, the repair level, the distance from the city centre, the level ratio of the floor in the building, gas in a building, metro station etc. can be introduced as the qualitative statistic of the housing. Table 1 describes this data base. This database consists of about 200 variables. These variables are qualitative properties of the houses. Some of them have been described in table 1. In total there are 1700 observations in this database. We collected these data from the housing market in Baku, Azerbaijan. There are only 6 variables have been used in our statistical analysis. These are: to be furnished with gas, the distance from the city centre, the price of per sq m., to be near the metro station, availability of the documents of building and to be without repair. We have taken a log of the distance between the city centre and a log of the price of per sq. m. The first column of the table 1 describes the score for the quality statistical data base. This score is coded as the 0 when the house hasn't been sold during the period, otherwise is coded as 1.

III. Methodology and related literatures

Our approach consists of two steps. In the first step we have defined the independent variables of our model. In the literatures the characteristics of houses have introduced as the creator of their prices. Usually this method called hedonic price method. Hedonic price method assists in defining the influence of these properties on the price of houses. It is first cornerstone of our approach. (J)Ceyhun Abbasov (2014) had investigated the influence of these properties of houses on their price in Baku city. In this work we changed the structure of the independent variables of our model. In literature there are many research works which have been devoted to investigate the relationship between the price of houses and their

qualitative properties. For example, Ercan Baldemir (2007) noted in his work that first time Haas (1922) investigated the relationship between the price of house and its properties like "far from the city centre" and "scale of the city". Court (1941) analyzed the price of automobiles as a function of their characteristics in his research work and had defined the marginal cost of each characteristic. In this work He explained that the marginal cost of the characteristics generates the prices of automobiles. Robert J. Hill (2011), Erwin Diewert, Saeed Heravi, and Mick Silver (2007) etc. have also investigated the price of houses as a function on their characteristic properties. We also will use this major methodology. But the price of houses isn't a dependent variable in our analysis. It will be independent variable as the price of per sq. m. together other qualitative and quantitative variables. So we are able to take the qualitative properties of the houses as independent variables for statistical analyses. For to do this we refer to some empirical research works. So, the independent variables of our approach will be to be furnished with gas, the distance from the city centre, the price of per sq m., to be near the metro station, availability of the documents of building and to be without repair.

TABLE 1: A SAMPLE TO THE QUALITY STATISTICAL DATA BASE OF HOUSES

Score	Price of per sq. m (AZN)	Without repair	Parquet	To be near to metro station	Distance from the centre[1]	Level ratio of the floor[2]	Availability of the documents of building[3]	Furnished with gas
0	700.0	1	0	0	11.9	0.76	0	0
1	700.00	0	0	1	11.8	0.19	0	0
0	1192.6	0	0	1	11.8	0.88	0	0
1	1099.6	1	0	0	9.0	0.65	1	1
0	1173.9	0	1	0	17.2	0.15	0	1
1	507.63	1	0	0	13.2	0.33	0	0
1	855.27	0	0	0	12.4	0.21	1	0
0	776.39	1	0	0	12.4	0.79	0	0
0	1629.1	0	0	1	6.4	0.63	0	0
0	1300.0	0	1	0	3.4	0.38	0	0

The second cornerstone of our approach is to take the score which has been given in the first column of the table 1 as the dependent variable of our approach. In this step we refer to Edmister's (1972) approach. Edmister (1972) had used 7 financial ratios in his insolvency model. While establishing z-score values as a dependent variable in the model he coded 0 for those companies which considered as default case, otherwise it was coded as 1. In accordance with this model the companies are classified to be not default case if z-score of these ones equal to 0.530 and more. Whereas the cases less than 0.530 it is expected to exposure the default. So, the dependent variable of our approach will be the G-score (see the first column of the table 1) which is coded as the 0 when the house hasn't been sold during the period, otherwise is coded as 1.

These two approaches let us to construct our model as the following:

$$G = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6$$

Where,
β = the parameters of the model
X_1 = to be furnished with gas
X_2 = log of the distance from the city centre
X_3 = log of the price of per sq. m.
X_4 = to be near the metro station
X_5 = availability of the documents of building
X_6 = to be without repair

G = the dependent variable of our model will be the score which is coded as the 0 when the house hasn't been sold during the period, otherwise is coded as 1

IV. EMPIRICAL RESULTS

We have 1700 observations about the price and characteristic properties of the houses. But we used 234 observations for estimating the equation (1) because to take only 3-rooms houses in new buildings which have been built after the Soviet Union time.

G-score is the dependent variable which is coded as the 0 when the house hasn't been sold during the period, otherwise is coded as 1. The results of the estimating of equation (1) have been given in table 2.

TABLE 2: THE ESTIMATING OF THE EQUATION (1)

Independent variables	**Coefficients**
to be furnished with gas[4]	0.18880
ln_the distance between the city centre[5]	-0.12051
ln_the price of per kv. m.[6]	-0.72534
to be near the metro station[7]	0.13618
to be documents in a building[8]	0.43440
to be without repair[9]	-0.13691
C[10]	5.61938

[1] This distance was measured by the "Google maps". We took "Icheri Sheher" metro station as the starting point in "Google maps".
[2] These ratios present the level of the flats. For example, suppose that the first flat (house) in table 1 have been observed in the 13th floor of the 17 floors building then we can calculate this ratio as 13/17=0.76. So, if the ratio is equal to 0.5 or is close to 0.5 (mining that 0.4, 0.6) it means that the flat is on the middle of the building. If the ratio is close to 1 it means that the flat is on the top of the building. If the ratio is close to 0 meaning that the flat is on the lower floors of a building.
[3] As the mentioned above all houses which have been observed are located in new buildings and some these ones don't have the important official documents which affirm the private ownership.

[4] The coefficient is significant at the 0.02 confidence level
[5] The coefficient is significant at the more than 0.01 confidence level
[6] See foot-note 5
[7] The coefficient is significant at the 0.03 confidence level
[8] See foot-note 5
[9] The coefficient is significant at the 0.1 confidence level
[10] See foot-note 5

So, if the computed G value found by using the equation (1) based on the results of the table 2 is equal to 0.39 and greater than that one it is classified as the high chance of selling the house, but in the case of being less than 0.39 it is classified as none selling chance. Note that 5 of 93 houses those had remained in none selling status, classified as the high chance of selling. It means that we have I type error. But 14 of 141 houses those had sold, classified as none selling chance. It means that we have II type of error (see the table 3).

TABLE 3: CLASSIFICATION ANALYSIS

	Selling	*Non-selling*	***Total***
Selling	127	14	**141**
Non-selling	5	88	**93**
Total	**132**	**102**	**234**
Selling %	90.1	9.9	100,0
Non-selling %	5.4	94.6	100,0

In statistical analysis I type error always is more dangerous than II type error. Because classification of houses as none selling chance house isn`t bad for the market participants. But of course in that case they may face to lose its good dealing. Here one question is it possible to make forecast the selling case in the term of existing I type error and II type error? In order to reply this question, it is necessary to define the possibility of making I type error and II type error.

I Type error = 5/234*100 = 2.14%

II Type error= 14/234*100 = 5.98%

Total occurred error = 2.14%+ 5.98% = 8.12%

We can say that the forecasting will be correct with 91.88 % probability. But with 8.12 % probability it will be resulted in error. Actually the predictability power is seen very highly (91.88%). But we are in need of to check it. For checking the predictability power Cohen's Kappa test statistics is useful. The suggested Cohen's Kappa test statistics is as follows:

The results of the table 3 will be used for this test statistics

$$k = \frac{w_1 - w_2}{N - w_2}$$

Where,

$w_1 = 127 + 88 = 215$

$w_2 = \left\{\frac{127+5}{234} * \frac{127+14}{234} * 234\right\} + \left\{\frac{88+5}{234} * \frac{88+14}{234} * 234\right\} = 79.58 + 40.5 = 120.1$

So,

$k = \frac{215 - 120.1}{234 - 120.1} = 0.833$

There are 6 significance intervals for the decision about the result of kappa test. These intervals are as follows:

$k \leq 0$ (Very small compliance)

$0 < k \leq 0.2$ (Small compliance)

$0.2 < k \leq 0.4$ (Acceptable compliance)

$0.4 < k \leq 0.6$ (Medium strong compliance)

$0.6 < k \leq 0.8$ (Significant compliance)

$0.81 < k$ (Almost perfect compliance)

In order to use of this classification in forecasting none selling chance and the high chance of selling of houses, the result of test should be more than 0.81. Otherwise there is no any significance to use of this classification in making the prognosis. Our estimating value (0.83) is greater than critical value of the kappa test statistics (0.81). It means that we have obtained perfect compliance. So, these classifications can be used in forecasting none selling chance or the high chance of selling of houses.

V. CONCLUSION

Finally, it would be usefully that the main points of the research are reviewed. The first point is about the qualitative statistical data base. This data base is a universal data base. The average price for the all properties of houses can be calculated by using of this data base. For example, you can choose any property of house say that like "to be documents in a building" and can define average price, max or min prices, number of houses which were offered to a market etc. over this property. Secondly this data base can be used for the different estimating that ours is one of them. We constructed the classification for the housing market. This classification lets us to forecast none selling chance or the chance of selling of houses. Thus we found that if G-score (which is calculated by using the equation (1) based on the results of the table 2) of any house is equal or greater than 0.39 it means that this house has the high chance of selling. But if G-score (which is calculated by using the equation (1) based on the results of the table 2) of any house is less than 0.39 than it is classified as none-selling chance house. This result can be used both the market participants and the banks. The market participants always want to know that how much selling chance their houses have. On the other hand, it is interesting the banks to know how much selling chance the houses which have been taken as the collateral by the banks if any borrower unable to give back the money. The latter one is very important for maintaining the financial stability.

REFERENCES

[1] A.T.COURT. (n.d.) "Hedonic Price Indexes with Automotive Examples" http://www.econ.wayne.edu/agoodman/research/PUBS/Court_Hedonic.pdf.

[2] Aasen M. R. (2011) Applying Altman's Z-Score to the Financial Crisis an Empirical Study of Financial Distress on Oslo Stock Exchange.

[3] Ercan Baldemir, C.Yenal Kesbich, Mustafa İnci (24-25 May 2007). Estimating hedonic demand parameters in real estate market: the case of Mugla. IIIV Turkey Econometrics and Statistics Conference. Malatya: İnonu University.

[4] Erwin Diewert, Saeed Heravi, and Mick Silver (2007). Hedonic Imputation versus Time Dummy Hedonic Indexes. IMF Working Paper- WP/07/234.

[5] Edward I. Altman, Anthony Saunders (1998). Credit risk measurement: Developments over the last 20 years. Journal of Banking & Finance

[6] Grice, J. & (2001). Tests of the generalizability of Altman's bankruptcy prediction model. Journal of Business Research, 54(1), 53-61.

[7] Hill R (2011), "Hedonic Price Indexes for Housing", OECD Statistics Working Papers, 2011/01, OECD Publishing. http://dx.doi.org/10.1787/5kghzxpt6g6f-en

[8] Robert O. Edmister (1972). An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. The Journal of Financial and Quantitative Analysis, Vol. 7, No. 2 Supplement: Outlook for the Securities Industry (Mar., 1972), pp. 1477-1493

[9] (J)Ceyhun Abbasov (2014). Hedonic price method for housing. International Journal of Business and Social Research (IJBSR), Vol 4, No 8, pp. 8-16.