ISBN - 978-989-9121-08-9
43rd PATTAYA International Conference on "Advances in Science, Engineering & Information Technology" (PASEIT-22)
Pattaya (Thailand) Aug. 25-26, 2022

Comparative Study on Regression Analysis of Housing Defect Repair Cost

Junmo Park[1] and Deokseok Seo[1*]

[1]School of Architecture, Halla University, 28 Halla university-gil, Wonju-si, Gangwon, 26404, South Korea

***Abstract*:** *Defect repair cost is a major concern in disputes over housing quality. Therefore, predicting and controlling the defect repair cost is vital. A regression model was used in the previous studies to examine how helpful the regression model was for predicting the cost of repairing house defects. One hundred cases of housing defect lawsuits in Korea were analyzed using the regression model of previous studies. Results indicated that even in the regression model with a high coefficient of determination, there was a multicollinearity problem, and some independent variables' coefficients were not significant. Analysis results suggested that it is necessary to establish and compare standard models for in-depth comparison in the future.*

Keywords: *housing defect repair cost, regression model, total floor area, elapsed period, lawsuit period*

1. Introduction

Management, including cost and construction period, seems essential in the housing industry's growth. On the other hand, there are standard specifications and performance evaluation systems for the primary construction materials of a house, but the quality of the finished house is undefined. As a result, disputes over quality between homebuyers and developers are on the rise. This dispute is called a housing defect dispute. The biggest concern in a housing defect dispute is whether there are many or few defects. Moreover, it is directly settled on the scale of the defect repair cost. Therefore, predicting the repair cost is essential to researchers in related fields, decision-makers, and those in charge in the field.

According to prior studies, the regression model was mainly used to predict the cost of repairing housing defects. The regression model is an analysis method that assumes linearity between the dependent and independent variables that affect it. This study compared how helpful the regression model used by previous major studies was in calculating the cost of repairing housing defects. One hundred housing defect lawsuits filed in Korea were collected for this study. Whether the regression model of previous studies was excellently explained was assessed using these data.

2. Literature Study

Prior studies which predicted the defect repair cost of housing and building with the regression model were briefly reviewed.

Kang et al. [1] proposed a regression model that assumed the defect repair cost using the housing defect lawsuit cases. The independent variables used for this purpose were the period of use from the completion of the house to the litigation (Elapsed Period; EP), the time taken until the litigation was completed after the defect litigation was filed (Lawsuit Period; LP), the value of the subject-matter in the litigation (VL) at the time of filing a lawsuit, and the home warranty deposit (HWD) to guarantee the defect repair.

Seo and Lee [2] used the total floor area (TFA), the number of households (HouseHold; HH), and the main building's quantity (MB) as independent variables to measure the housing defect repair cost.

Meanwhile, Choi [3] used deposit, total floor area, and number of households as independent variables to calculate the defect repair cost.

Forcada et al. [4] used the value obtained by taking the natural logarithm of the defect repair cost as a dependent variable. As the independent variable, they adopted values obtained by taking the natural logarithm of the construction cost (CC) and a nominal variable that classified the regional difference between the location of the headquarters of the construction company and the location of the building to be constructed into 4 classes.

In the study by Kim [5], the total floor area, households, and number of the highest floor of main buildings (HF) were adopted as independent variables.

3. Case Study

3.1. Outline

This study compared how helpful the regression model for predicting housing defect repair costs of previous studies was by using the cases of housing litigation in Korea. The incidents involving 100 housing complexes that reached the defect lawsuit were collected as study cases. Korean courts have established standard procedures and standards for housing defect litigation to investigate defects and calculate repair costs. The repair cost in the case was based on the amount specified in the judgment of the lawsuit.

On the other hand, there were cases in which data could not be collected among the independent variables presented in the regression model in the previous studies. In most cases, the litigation value, security deposit, and construction cost were not mentioned in the judgment, so it was impossible to get the data. Therefore, some of the regression models of previous studies could not use some of the suggested independent variables.

In addition, among the independent variables, locations were classified according to whether the administrative district where the headquarters of the developer and the constructor were located and the primary administrative district of the project site were the same or different and whether the regional administrative districts were different.

Since the regression model has various independent variables, it was constructed as a multiple regression model. In addition, the analysis was performed using IBM SPSS ver. 21.

3.2. Result

The results of the regression model of each previous study are shown in Table I.

Table I compared the coefficients of determination (R2) of each model. In Kang et al.'s model, the coefficient of determination R^2 was analyzed to be 0.149. In Seo and Lee's model, the coefficient of determination R^2 was 0.398. Choi's model showed the coefficient of determination R^2 as 0.398. Forcada et al.'s model analyzed the coefficient of determination R2 to be 0.013. In Kim's model, the coefficient of determination R^2 was 0.401.

In summary, Kim's model was evaluated to be the best when judging based on the R^2 value of the coefficient of determination. However, linearity is a pre-requite for the regression model, and there should be no residual and multicollinearity problems. Also, even if independent variables were included in the model, the model was judged inappropriate if the coefficients were not significant.

TABLE I: COMPARISON OF R^2 VALUES IN DIFFERENT MODELS

Model	Kang et al.	Seo & Lee	Choi	Forcada et al.	Kim
R^2	0.149	0.398	0.398	0.013	0.401

According to the analysis results, linearity of all models was assumed. In addition, since the Dubin-Watson value was in the range of 1 to 3, there was no problem with the independence of the residuals.

On the other hand, as shown in Table II, the independent variable of some models has a VIF of 4 or more, which raises concerns about multicollinearity. In Seo and Lee's model, the VIF of the three independent variables was 4 or higher. Choi's model had a VIF of 4 or more for the number of households and total floor area, and Kim's model had a VIF of 5 or more for total floor area and number of households. Therefore, since it is anticipated that a multicollinearity problem exists in these models, it seems that some independent variables should be excluded from the model.

TABLE II: COMPARISON OF MODELS` VIF

Model	Kang et al.	Seo & Lee	Choi	Forcada et al.	Kim
VIF≥4	-	TFA, HH, MB	HH, TFA	-	TFA, HH

In addition, the model seems inappropriate since some models' coefficients of the independent variables are not significant (Table III). In Seo and Lee's model, the coefficients of the two independent variables were insignificant. In Choi's model, the number of household coefficients was not significant. In Kim's model, the coefficients of the number of households, the location, and the number of the highest floors were not significant.

TABLE III: COMPARISON OF COEFFICIENT FOR INDEPENDENT VARIABLES

Model	Kang et al.	Seo & Lee	Choi	Forcada et al.	Kim
Not significant independent variables		HH, MB	HH	Lo	HH, Lo, HF

4. Implication

According to the results of this study, which analyzed the regression model of previous studies that estimated housing defect repair costs, the following points were drawn.

First, it was difficult to ascertain whether the regression model was excellent because the coefficient of determination (R^2) was high. As seen from the above analysis results, Kim's model had the highest R^2, but among independent variables, the number of households and total floor area seemed to have a multicollinearity problem. Also, the coefficients within the model for the number of households, location, and number of the highest floor were not significant. Therefore, it was difficult to determine the superiority or inferiority of the regression model or whether to use it only with the R^2 value of the coefficient of determination.

Second, the independent variables suggested in some prior studies, such as litigation value, deposit, and construction cost, could not be used for analysis due to limitations in data collection. Also, in the case of locations, the variables suggested in the study could not be used as such.

Although preceding studies suggested various independent variables to estimate the repair cost, it was difficult to determine the most reasonable one. Therefore, in the future, either a model that can serve as a comparison standard should be proposed, or a comparison procedure after establishing a standard model that summarizes the preceding studies should be added.

5. Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (2019R1A2C1009913).

References

[1] Y. Kang, B. Kim, J. Park, J. Seo, and O. Kim, “Regression analysis on the dispute cost property in apartment housing claims,” in Proc. 2010 Spring Conference of Korean Institute of Building Construction., 2010, 10(1), pp. 225-228.

[2] D. Seo and U. Lee. (April 2015). Characteristic analysis of utilization of security deposit for repairing defects using statistical analysis. Journal of Korea Institute of Building Construction. [Online]. 15(2). pp. 209-215. Available: https://doi.org/10.5345/JKIBC.2015.15.2.209.

[3] J. Choi, “Evaluation of defect repairing bond ratio through defect lawsuit case study in apartment building,” Ph.D. dissertations, Dept. Architectural. Eng., Chungbuk National Univ., Cheongju, South Korea, 2017.

[4] N. Forcada, M. Gangolells, M. Casals and M. Macarulla. (March 2017). Factors affecting rework costs in construction. Journal of Construction Engineering and Management. [Online]. 143(8). id: 04017032. Available: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001324

[5] Y. Kim, “Regression Models for Defect Repair Cost Predictions Year·WetConstruction of Apartment House,” M.S. thesis, Dept. Architectural. Eng., Univ. Seoul, Seoul, South Korea, 2019.