



Multiple Regressions in Analysing House Price Variations

Aminah Md Yusof and Syuhaida Ismail

Universiti Teknologi Malaysia, Johor, Malaysia

Abstract

An application of rigorous statistical analysis in aiding investment decision making gains momentum in the United States of America as well as the United Kingdom. Nonetheless in Malaysia the responses from the local academician are rather slow and the rate is even slower as far as the practitioners are concern. This paper illustrates how Multiple Regression Analysis (MRA) and its extension, Hedonic Regression Analysis been used in explaining price variation for selected houses in Malaysia. Each attribute that theoretically identified as price determinant is priced and the perceived contribution of each is explicitly shown. The paper demonstrates how the statistical analysis is capable of analyzing property investment by considering multiple determinants. The consideration of various characteristics which is more rigorous enables better investment decision making.

Keywords: Property, analysis, hedonic and decision making.

Introduction

Property investment as other investment, involves the decision of putting aside for the future benefits. The main distinguishing feature between property investment and other paper investment is it involves huge amount of initial capital which certainly the investor will not take risk by omitting proper decision making process. The decision to develop or to invest in property, for example lies in the hand of the developers. The developers are responsible to decide what developments will be profitable as well as permissible on that particular price of land. An additional feature of property investment is its market is characterized by low demand and high supply or vice versa. These are realities in the property business. Improper analysis could lead to overbuilding or glut in the market. Overbuilding becomes a topic of discussion and it is impossible to turn the clock backward. The adjustment process will take place either at favorable or undesirable pace to stabilize the market.

Thus the development undertaken has to be continued without a guarantee of being taken up or it might end up like a hot cake.

Economic recession in Malaysia had caused collapse in property sector despite the common belief that property investment is hedge against inflation. What went wrong has been left without explanation. The impact of the recession may be regarded as a normal phenomenon in business cycle, which as a result helps consolidate the economy. The decision to invest or to develop should be accompanied by thorough analysis. The analysis with an aid of contemporary approach should be employed to enhance the reliability of the decision making.

The recent mode to property investment is via real estate investment trust. Compared to 'paper investment', direct property investment involves the rights of ownership to a price of land, typically with a building. Property as an asset class exhibits distinctive characteristics, such as fixed location, heterogeneity, high unit

value, illiquidity and often, the use of valuations rather than prices for market information. The diversity and the complexity of the property investment demand extensive research into the property investment sector.

Property investment management has changed substantially in the last 30 years with the increasing integration with that of the other investment classes. This resulted in an expanded from the selection of individual buildings to include a portfolio perspective in which portfolio structures are set relative to a benchmark and property is traded more frequently. As a result, analysis has developed in the context of capital market and the wider economy. The integration has also made some of the property distinctive features become more problematic. Dissatisfaction has developed about illiquidity, large lot size and the use of valuations to measure returns. This has been one factor in the development of indirect vehicles to gain exposure to property market. The management of property investment has also been affected by the trends of globalization in economic activity and investment.

The 1980s witnessed a dramatic expansion of property research to meet the needs of major investors who sought a portfolio perspective for the management of their assets to link property market to the economy and to capital markets. The expansion of research brought new skills from economic and finance to property market, which enriched the forecasting ability in property research. In addition to simple trend analysis, the availability of sufficient data of the market enabling advanced statistical analysis to be implemented in property analysis. The statistical tool is capable of modeling property investment performance by taking into consideration the unique features of the investment. The potential of the statistical application to enhance the traditional approach to evaluating property investment hence guide the property investment decision making has been demonstrated in many works.

This paper aims at demonstrating an application of statistical tool to property investment analysis. The focus of the review will be on micro or property-specific factors and to link these factors to price and rental. The application of multiple regression analysis in a data set of property investment data, attempt to explain or model variation in house price which demonstrates good examples of the strategic application of mathematical tool to aid analysis hence decision making in property investment.

Property Price Determinants

An old maxim related to the price and rental determination is 'location, location and location'. Nonetheless, time and technology changes have shift the old price paradigm. Price is better explained by the utility function possessed by each commodity, i.e. the price paid for its utility function. The higher the function, the higher price paid for it. It is, therefore important to identify these functions. As these function identified, the application of statistical tools enabling these function be evaluated. Generally there are two forms of price: implicit and explicit. The combination of implicit price summed up explicit price. Studies undertaken are commonly agreed that there are three major common determinants that explain or model price variation. Chau et.al (2001) and So.et.al (1995) suggest that residential properties are multidimensional commodities characterized by durability, structural inflexibility and spatial fixity hence these characteristics determine price. Although it is quite impossible to get perfect information the decision to buy a house which involves a substantial capital outlay will be driven by an effort of gathering as much information as possible. The most commonly available information that are normally rely on purchasers can be broadly categorized into structural (physical characteristics), locational and neighborhood attributes. The similar proposition applies to the office building. The decision to rent will be based on similar characteristics but differ in relation to perceive importance of each

characteristics from house buyers to office tenants.

Ball (1973) suggests that a house desirable attributes reflected in a higher market price. The structural or physical characteristics such as floor area, numbers of room were strongly related to price. The heterogeneity of the houses has resulted in different attributes been used such as numbers of rooms (Fletcher, et.al.2000, Li and Brown, 1980), the number of bathrooms (Garrod and Willis, 1992 and Linneman, 1980). Other attributes include age which shown by Kain and Quigley (1970) as negatively correlated to house price as a result of increase in maintenance cost as well as a decrease in usefulness due to changes in technology and post-fordism. Among other attributes suggested are lot size, the existence of basement, garage, patio and heating system (Forrest, Glen and Ward, 1996). This is particularly true for office especially the demand will be higher for green office building and building which is free from sick building syndrome.

The importance of location has long been recognized as main property price determination. The effect of location on property price was examined by Gelfand et.al. (2004). The work highlights the importance of spatial component in explaining prices. When the location is taken into consideration, homogeneity affects patterns of variation across space of time. The location of a property has been regarded in terms of fixed and relative locational attributes. This nonetheless is applicable when the location is anticipated to be influential. On one hand, some aspect of location can be accurately quantified, for example Follani and Jeminoz (1985) and Orford (1988) used fixed locational attributes with respect to the whole area and pertained to some form of accessibility measures. On the other hand some surrogate measures such as socio-economic class, social composition, and aesthetics attributes, pollution levels and proximity to local amenities are used by Dubin and Song (1990). Accessibility, which is normally measured in terms of access to the central business district have

some influences on housing prices. The positive effect of transport accessibility on housing price has been demonstrated in Plamquist (1992), Ridker and Henning (1968), Adair et.al (2000) and So et.al (1996). So et.al (1996), for example show that buyers are willing to pay more for properties with easy accessibility to public transportation.

In addition to the two above factors, neighborhood attributes may affect price or rental of a property. The difficult task will be to determine and ascertain the quality of the house which could be implicitly valued through hedonic pricing. Neighborhood specific-factors has been identified as explaining 15-50% of the standardized variation in valuation and inducing differential valuations as large as 100% between the structurally identical sites. Amongst neighborhood attributes that are considered are socio-economic factors and the occupation of the habitant. Works that considered neighborhood qualities includes Garrod and Willis, 1992 and Roe.et.al (2004) which include neighborhood amenities on housing values and residential growth.

In summary, three determinants that are normally associated with price and rental are factors related to structural, locational and neighborhood. These factors will be considered in property investment decision making.

The Anatomy of Decision Making

As we are aware that all decision makers based their decision on what they believe about all the foregoing element of a decision. The process of decision-making involves the synthesizing stage. Synthesis of candidate, development or investment alternatives has to be undertaken by developers or owners. The set of alternative candidates (a, b, c...) is a fundamental element of a decision situation. The identification and description of the candidate alternatives are designated as synthesis.

Next stage involves an analysis of all activities that may be used for the

estimation of outcomes. Estimation of profit requires an understanding of how much time required to perform various tasks are influenced by various levels of resource allocations. Since the actual consequences of the decision have not occurred when the analysis is performed, uncertainty is inherent in the estimates. This can be done although the degree may not be explicitly stated.

Third stage involves an evaluation of candidate development or investment. The transformation of estimates of outcomes into relative ratings of alternatives is known as evaluation. The quantitative representation of evaluation relationships and the explicit accomplishment of the evaluation activity are major contribution of the decision analysis.

Finally, the selection of a course of action been made. The selection will therefore critically depending on the analysis and evaluation perform earlier. Hence the reliability of method of evaluation is critical. As it is necessary for investors to make choice between the different investment media, decision-makers require investment analyses for property, which are comparable with those available for other investments. The decision makers are responsible for investment decisions in the major investment institutions.

One of the important stages in the above chronology is the analysis technique upon which the decision is based on. The analysis can be used to evaluate potential development, project or investment. The techniques especially employed in property investment are basically adopted from financial and economic. The techniques can easily be categorised as property valuation and property analysis. The valuation methods address the need to state the worth of property whilst property analysis aims at explaining the pricing mechanism. The advanced analysis techniques are, however, based on performance measurement and incorporation of statistical tools. The approach allows more factors and situations to be taken into consideration. The approach involves comparing property

performance, either cross-sectionally or longitudinally. The most widely used methods to explain and forecast property performance is Multiple Regression Analysis (Chaplin, 1998). The technique is then extended to Hedonic Regression and utilised in Simultaneous Equation System (Thompson and Tsolacos, 2000).

Multiple Regression Analysis (MRA)

Multiple regression analysis is based on the correlation analysis. Generally correlation and regression can be used to conduct multivariate analysis on fairly small samples. Steven (1992) points out that the strength of the MRA lies primarily in its use as a means of establishing the relative importance of independent variable on the dependent variable (Dependent variable is the phenomenon to be explained and independent variables are factors used to explain dependent variable). Assuming a linear relationship, the model for the regression analysis may take in the following form:

$$Y = a + bx$$

Simple regression analysis)(1)

$$Y = a + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n$$

(Multiple Regression Analysis)(2)

The benefits of MRA can only be appreciated with an understanding of basic concepts of MRA. MRA in its classic form is used for prediction. However, in recent practice, MRA is normally used for an explanation of the subject under investigation. Cohen and Cohen (1983) use MRA as a causal model to explain changes in independent variables to explain changes in dependent variables as well as to assess the relative importance of each independent variable. Each regression coefficient estimates the amount of change that occurs in the dependent variable for a unit change in the independent variables in the equation. The following steps are necessary in the formation of the MRA model:

- a) Conceptualization of the model: The consideration of the relevant theory, which will help in determining the

relevant factors or independent variables to explain the dependent variable in the model

- b) Model Estimation and development: model can be estimated by statistical tests to ensure the model is rigorous
- c) Model Testing: this will involve an application of relevant statistics tests. The quality of the model is evaluated using coefficient of determination, which is denoted as R^2 . The higher the R^2 , the better the model.
- d) Forecasting/explaining: The model can then, be tested and cross validate on the other sample.

Methodology

The methodology of the paper is based on two main approaches: an exploration of theoretical rationale on the issue under investigation and the empirical works which are guided by the findings of the literature search. The empirical enables the problems which were identified in the earlier works to be addressed and new problems identified and could be taken into consideration in related future works. The aim of this paper is to explain what actually Malaysians are paying for their offices and house(s). Therefore, modeling house price as dependent variables using independent variables seek to segregate the impact or contribution of each independent variables in price variation. The methods involved are:

- An identification of house price determinants: this is very important to the house price model development: Literature review carried out enables the identification of variables that are capable of explaining house price and offices generally.
- Data collection: To limit the scope of the study, data collection and analysis are confined to a particular type of house in the City of Kuala Lumpur, Malaysia. Double storey terrace has been selected due to the fact that the house is most commonly/frequently transacted in

Malaysia. The data collected included floor area, land area, age of the house, and numbers of rooms, distance, and the quality of amenities, neighborhood quality and type of holding. The information obtained from the dataset available in the National Property Information Centre (NAPIC) and some are observed through the site visits. Five locations represented each region: south, north, upper north, central and east have been selected. This is expected to provide exclusive explanation for the issue under investigation. All factors are related to locational, structural and neighborhood characteristics. A total of 1500 houses have been selected from the database.

- Model development: Models for houses in Malaysia have been developed. The stepwise method is selected as it provides better justification for the variables included in the model. Each variable will be assessed and reassessed for its significance in the model or equation. The other important exercise in model development includes model testing and validating. The models are observed for an existence of multicollinearity and other problems that might affect the stability of the models developed.
- Assimilation: findings are discussed and conclusion is drawn from the analysis.

Practical Application of MRA

In order to appreciate the application of MRA in property analysis, an example is discussed. The MRA is intended to explain price variation for houses in Malaysia. The analysis reveals what actually Malaysians are paying for their house. The likely factors that influenced price are gathered. The analysis began with close examination of correlation between independent variables (factors related to site, structural and neighborhood and variation in House Price (HP) in Malaysia. A testable form of equation related to price variation (HP) with a standard cross-sectional hedonic model (Rosen, 1974). The equations are as follows:

$$HP_{it} = a + b_1x_1 + b_2x_2 + \dots + b_nx_n + e \dots \dots \dots (3)$$

Where the above office rental and house price is a function of building and site related factors for the *i*th houses. The intercept 'a' represents that portion of house price variation for each house that may be attributed to the overall level of house price. The coefficients on *b*₁ to *b*_{*n*} are allowed to change over time. Any unexplained variation is captured by the random error 'e'. Two models have been developed for house prices for two times 2000 and 2007. Each is represented as follows:

Model 1

$$HPL_{00} = 213147.9 - 0.5(DIST) + 0.365(LD_AREA) + 0.154(BLD_AREA) - 0.260(AGE) + 0.284(LOCALITY) - 0.164(PUB_TRANS) + 0.149(NEIGH) \dots (4)$$

$$HP_{07} = 63231.6 + 0.284(LD_AREA) + 0.191(BLD_AREA) + 0.430(LOCALITY) + 0.109(HOLD) - 0.235(AGE) - 0.173(AMN) + 0.131(NEIGH) - 0.387(DIST) (5)$$

The models have been tested for any existence of multicollinearity and the explanatory power of each model is

considered as adequate. The summary of each model is presented in Table 1.

Finding and Discussion

To better illustrate the above scenarios, each model will be discussed separately. Two models for year 2000 and 2007 using information mentioned earlier have been developed. Each explains price variation for houses in Malaysia. An explanatory power is quite high (as depicted by the R²) thus can be reliable to explain the underlying of price variation for houses in Malaysia. As illustrated in Table 1 Locality has been the most influential factor in price. This mean that price paid for location, thus proved that location is still relevant. In this case, it can be derived that location, denoted as 'locality' in this case explain 50.3% and 63% of price variation for 2000 and 2007 accordingly. The proportion reflects the perceive importance of location to the buyers in the 'price' that they paying for the house.

The model also explains that second factor that is significant in price variation is Building Area. Model 2000 indicates that 16.3% of price variation is explained by Building Area whilst in Model 2007, building area contributes 14.1% in price variation.

Table 1: A Summary of Models Developed

Variables	HP ₀₀	HP ₀₇
R ²	77.2%	83.6%
ADJ R ²	83.6%	83.3%
LOCALITY	42 (50.3%)	52.5 (63%)
PUB_TRANS	0.8	-
LD_AREA	2.3	5.5
BLD_AREA	14.1 (16.9%)	8.9 (14.1%)
ACCES	-	-
DIST	8.6	6.6
PRO_TYPE	-	-
AGE	8.0	5.8
HOLD	-	-
NUM_ROOM	-	-
NEIGH	0.9	5.3
AMMEN	-	0.9

(Note: Only significantly contributed variables or factors will be discussed)

Other factors which are significant in the model are land area, distance from the city centre, age of the building and property and the quality of neighborhood.

Generally, this paper seeks to explain the composition of price paid for a double storey terrace houses in Malaysia. A hedonic analysis utilizing multiple regression analysis has been performed on data specific to the housing attributes. The hedonic analysis explains reasonably well the phenomenon and the impact of each determinant are segregated via modeling process. Based on the assumption that houses are priced based on utilities-bearing by the houses itself, the analysis shows that the price is dominated by functional area. The finding indicates that physical attributes in terms of building area become very important feature or factors paid for a house. Higher price is paid for land area and building area. This finding is quite consistent for all locations with an exception for Kuala Lumpur. The Kuala Lumpur's housing is dominated by preference for location and location-related factors. This finding indicates that although there are tendencies to pay for other features such as amenities and neighborhood qualities, it is less significance as compared to functional areas. The location-related factors such as distance from the city centre is also presented but at low level hence signifies the lesser emphasis in price paid for. The models show that different attributes are priced differently when combined with other attribute and different factors pose different level of influence on price.

Conclusion

This paper demonstrates how the statistical analysis could be utilized to better analyze investment. With an application of hedonic multiple, the contribution of each price determinant for the overall price of a house can be determined. The analysis thus, aids the decision making process. In this case, the potential investors or developers will be able to identify importance factors to be taken into consideration when developing and buying a house(s). As the contribution of each variable could be quantified, it is

possible to determine the significance of each variable. The application of multiple regression analysis in a house data set explains or models variation in house price which demonstrated good examples of the strategic application of mathematical tool to aid analysis hence decision making in property investment.

Acknowledgement

The author wishes to thank NAPREC, and Universiti Teknologi Malaysia for their support in the undertaking this research and enabling the output of the research presented in numerous international conferences and seminars

References

- Abd Rahman, M. T. (1992). ' Investment Advantages of an Intelligent Buildings,' *Real Estate* 1992, 26-31.
- Abel, J. (1994). "What Tenants Want and What They Will Not Compromise on When Looking for New Premises – Considerations Influencing Relocation," *Property Management*, Vol 12(1), 1994, 28-30.
- Abraham, J. M., Goetzmann, W. N. & Wachter, S. M. (1994). "Homogenous Groupings of Metropolitan Housing Markets," *Journal of Housing Economics*, Vol 3, 186-206.
- Adair, A., McGerald, S., Smyth, A., Cooper, J. & Riley, T. (2002). "House Prices and Accessibility: The Testing of Relationships within the Belfast Urban Area," *Housing Studies*, vol. 15 (5), 699-716.
- Adair, A. S., Berry, J. N. & McGreal, W. S. (1996). "Hedonic Modeling, Housing Submarkets and Residential Valuation," *Journal of Property Research*, vol. 13, 67-83.
- Alonso, W. (1964). *Location and Land Use*, Harvard University Press, Cambridge, MA.
- Ashworth, J. & Parker, S. C. (1997). "Modelling Regional House Prices

in the UK," *Scottish Journal of Political Economy*, Vol 8. 44,3.

Ball, M. (1973). "Recent Empirical Work on the Determinants of Relative House Prices," *Urban Studies*, vol. 10, pp. 213-233.

Barot, B. & Yang, Z. (2002)."House Prices and Housing Investment in Sweden and the United Kingdom Econometric Analysis for the period 1970-1998," *Review of Urban and Regional Development Studies (RURDS)*, Vol 14(2).

Basu, S. & Thibodeau, T. G. (1998)."Analysis of Spatial Autocorrelation in House Prices,"*Journal of Real Estate Finance and Economics*, 17, pp. 61-85.

Birch, J. W. & Sunderman, M. A. (2003)."Estimating Price Path for Residential Real Estate,"*Journal of Real Estate Research*, Vol 25(3).

Brown, R. L., Durbin, J. & Evans, J. M. (1975)."Techniques for Testing the Constancy of Regression Relationships Over Time,"*Journal of the Royal Statistical Society*, series B, 37, 149-172.

Can, A. (1992)."Specification and Estimation of Hedonic Housing Price Models," *Regional Science and Urban economics*, 22, 453-474.

Can, A. & Megbolugbe, I. (1997)."Spatial Dependence and House Price Index Construction," *Journal of Real Estate Finance and Economics*, 14, 203-222.

Chin, T.- L. & Chau, K. W. (2003)."A Critical Review of Literature on the Hedonic Price Model,"*Journal of Housing Science*, Vol 27(2), 145-165.

DiPasquale, D. & Wheaton, W. C. (1994)."Housing Market Dynamics and the Future of Housing Prices,"*Journal of Urban Economics* 35, 1-27.

Dubin, R., Pace, R. K. & Thibodeau, T. G. (1999)."Spatial Autoregression Techniques for Real Estate Data,"*Journal of Real Estate Literature*, 7(1), 79-5.

Elder, H. W. & Zumpano, L. V. (1991)."Tenure Choice, Housing Demand and Residential Location,"*Journal of Real Estate Research*, Vol 6 (3): 314-356.

Fik, T. J., Ling, D. C. & Mulligan, G. F. (2003)."Modeling Spatial Variation in Housing Prices : A Variable Interaction Approach," *Real Estate Economics*, Vol 31(4), 633-646.

Freeman, A. M. (1979)."The Hedonic Price Approach to Measuring Demand for Neighborhood Characteristics,' in: D. Segal (Ed.) *Studies in Urban Economics*, 52-77. New York : Academic Press.

Gabriel, S. A., Mattey, J. P & .Wascher, W. L. (1999)".(House Price Differentials and Dynamics: Evidence from the Los Angeles and San Francisco Metropolitan Areas ", *FRBSF Economic Review 1999 (1.(The Property Market Report (Annually), Ministry of Finance, Malaysia.*

Gatzlaff, D. H. & Ling, D. C. (1994)."Measuring Changes in Local House Prices: An Empirical Investigation of Alternative Methodologies,"*Journal of Urban Economics* 35: 221 -224.

Giussani, B. & Hadjimatheou, G. (1992)."House Prices, an Econometric Model for the U.K," *Netherlands Journal of Housing and the Built Environment*, vol. 7, 31-58.

Goodman, A. C. & Thibodeau, T. G. (1998)."Housing Market Segmentation,"*Journal of Housing Economics*, 7, 121-143.

Green, R. K & Hendershott, P. H. (1993)."Demographic Factors and Real Home Prices," *National Bureau of Economic Research*.

Guisanni, B. & Hadjimatheou, G. (1991)."Modeling Regional House Prices in the United Kingdom," *The Journal of the RSAI* 70, 2 : 201- 219.

Kain, J. F. & Quigley, J. M. (1970). "Measuring the Value of Housing Quality," *Journal of the American Statistical Association*, 45, 532-548.

Roy A. Gruver & Randall E. Wambold (1992). "Turning the Vision into Reality," *Real Estate*, 1992: 22-25.

Kelley Pace, R., Barry, R., Clapp, J. M. & Rodriguez, M. (1998b). "Spatiotemporal Autoregressive Models of Neighborhood Effects," *Journal of Real Estate Finance and Economics*, 17, 15-33.

Kelley Pace, R., Barry, R. & Sirmans, C. F. (1998a). "Spatial Statistics and Real Estate," *Journal of Real Estate Finance and Economics*, 17, 5-13.

Kutty, K. N. (1996). 'Indicators of Housing Quality,' *Housing Science*, vol 20(3), 151-166.

Lerman, S. R. (1979). 'Neighbourhood Choice and Transportation Services,' D. Segal (Ed.) *Studies in Urban Economics*, 129-153. New York: *Academic Press*.

Md Yusof, A. & Omar, I. (2002). 'Government Intervention in House Provision in Malaysia,' *Proceeding of ENHR*, Vienna, Austria.

Ozanne, L. & Thibodeau, T. (1983). "Explaining Metropolitan Housing Price Differences," *Journal of Urban Economics*, 13, 51-66.

Ridker, R. G. & Henning, J. A. (1967). "The Determinants of Residential Property Values with Special Reference to Air Pollution," *Review of Economics and Statistics*, 49, 246-257.

Roe, B., Irwin, E. G. & Hazel, A. & Morrow-Jones, H. A. (2004). "The Effects of Farmland, Farmland Preservation, and Other Neighborhood Amenities on Housing Values and Residential Growth," *Land Economics*: Feb 2004. 80(1): 55-75.

Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34-55.