CHAPTER 2: LITERATURE REVIEW

2.1 Hedonic house pricing models

The hedonic approach in multiple regression analysis applicable to housing may be attributed to Lancaster's consumer theory and the extension by Rosen (1974), into the residential housing market as a product differentiation for competitive pricing. Both of these approaches explained that prices of goods are imputed from their attributes based on the relationship between the observed prices of product characteristics and the number of attributes built into these products.

Pashardes and Savva (2009), investigated the impact of macro and micro variables on house prices in Cyprus from 1988-2008. They collected data from newspaper advertisements consisting of 4872 observations on various housing types like detached and semi-detached, number of bedrooms, size of building, geographical location and distance from the nearest city centre. They selected per capita GDP, interest and inflation rates, number of foreign workers, tourism arrivals, population, index of the Cyprus Stock Exchange, cost of materials and labour in construction, the Euro exchange rate to English pound and unemployment rate as their macro variables. They found that house prices are sensitive to population, cost of building materials and labour, GDP growth and the sterling-euro exchange rate. However, the number of foreign workers tends to restrain house price increase and the stock market was negative. Similarly, Apergis and Rezitis (2003) used macroeconomic variables of money supply,

employment, inflation, and housing loan rates. They found that housing loan rate caused the highest impact on housing price, followed by inflation and employment, while money supply does not seem to show any substantial impact. Bonnie (1998), had already concluded that employment growth and mortgage rate caused housing market to fluctuate. They found that economic variables have a different impact on the dynamic behavior of housing prices and the number of houses sold in regions at different time periods.

In a related macro economic study, determinants of house prices in the Asia-Pacific using regression was conducted by Glindro et al. (2008). They investigated the change in land supply index, change in construction index, real GDP, Mortgage to GDP trend, real mortgage rate and equity price trend. They reported the results for each country in the study.

As for micro variables in hedonic regression models, many researches used property characteristics, building structure, tenure, neighbourhood characteristics, location and environment. The commonly used variables are number of bedrooms and bathrooms (Li and Brown, 1980; Fletcher et al. 2000); detached or high rise buildings (Jim and Chen, 2007; Xu, 2008); age of building (Clark and Herrin, 2000; Goodman and Thibodeau, 1995); convenience to public transport system (Tse, 2002; Forrest, Glen and Ward, 1996); sea view and open space (Garrod and Willis, 1992; Chau and Chin, 2002).

Tse (2002) developed a stochastic approach which is able to correct autocorrelation bias in the hedonic function to estimate the neighbour effects in

house prices. He had used age of buildings, net floor area ratio, floor level with different views, transport accessibility and amenities in Hog Kong. He found that presence of sea view and accessibility to the Mass Transit Railway have significant effects on house prices in Hong Kong. He obtained an inconsequential 7 per cent reduction in sum-of squared errors using his stochastic approach. Similar use of property characteristics like property type, building area, land area, structure type and number of floors were used by Calhoun (2003), in Thailand. In addition, he added roof type, wall type and fence type. His paper was an initial effort to estimate hedonic property valuation models and house price index (HPI) for the provinces of Thailand. His results were summarised by application of graphical information system (GIS) software to illustrate differences in regional appreciation rates.

Kestens, Theriult and Rosiers (2006), used two sets of hedonic models with 761 single-property transactions carried out in Quebec City between 1993 and 2001. They used Casetti-type interactive terms and the Geographically Weighted Regressions. The variables used included living area, income of households, age profiles of buyers, house quality, in-ground pool and local tax rate. Each model can explain at least 84% of the house price variation from the combinations of the variables without collinearity effects. Their main finding is that buyer's household income, the previous tenure status and age have a direct impact on property prices. Likewise, Hui and Gu. (2009) considered household income and housing price in their study of Guangzhou house price bubble. (Casetti-type refers to a regression expansion model developed by Casetti

(1970) to analyse population growth across spatial drift in geographical research.)

Price volatility in housing was studied by Zhou and Haurin (2010) where they used data from two panels of the American Housing Survey from 1974 to 2003. They found that house values variance arises from quality, atypical house causes greater differences, there is premium for land leverage and minority households have greater variance than those of whites.

Using the hedonic modelling method, Xu (2008), studied the interaction behaviour between property specifics with location coordinates and buyers' characteristics. He used the natural log of sale price as the dependent variable and the semi-log equation for the hedonic regression procedure. The variables were grouped into transaction attribute, property specifics, location attributes and buyer's socio-economic characteristics. Dummy variables were added to control location of properties between discrete areas. His regression results for different models were with R-squares ranging from 0.7028 to 0.8544. Unlike western cities, high-rise buildings are valuable in Shenzhen as they command a better view, improved environment and a high social status.

In further consideration of the accuracy of using hedonic property market predictions, Bourassa, Hoesli and Peng (2003), looked into the segmentation of submarkets using 8421 transactions database from Auckland, New Zealand. The data related to location, detached and attached types, age, floor area, wall material and condition, quality of the structure and externalities of distance to

CBD, water view, landscaping, number of garages and quality of neighbourhood. Their research methodology included the use of principal component analysis with VARIMAX rotation, the new principal components and the factor scores calculated to ensure non-collinearity. Factor scores are then used in cluster analysis to construct homogeneous submarkets. The dependent variable, house price, was transformed into natural log. Their hedonic pricing models results were adjusted R-square of 0.69 to 0.72 with most variables at the significant level of 1%. Fletcher, Gallimore and Mangan (2000), examined the heteroscedasticity of age and external area of the property which gave a smaller forecast error than ordinary least squares estimates.

Hedonic property pricing has even been extended to the political economy of housing prices in Singapore as reported by Sue and Wong (2010). They conducted an empirical estimate of the value of publicly provided local goods and services in the constituencies of the ruling party relative to those of the opposition parties. To improve control for omitted variables that change smoothly over space, they used a regression discontinuity design to restrict the sample of houses that are near the electoral boundaries. Their data are the resale prices of HDB flats which are quite homogeneous in interior, floor level, floor area, age, distance to good performance schools, bus interchange, nearest MRT and industrial estate covering the periods from 2001-2006. They found that flats located in the neighbouring constituencies appear to enjoy a 3.2% price premium over those in the opposition constituencies. Likewise, Forrest., Glen and Ward (1996), used the semilog hedonic method to study the impact of light rail on the

structure of house prices. They used the property and neighbourhood characteristics together with location as variables. The result was negative as housing prices in Manchester was stable.

2.2 Housing price indexes development

Gourieroux and Laferrere (2006, 2009) contributed significantly to the understanding of housing price indexes through their extensive research in France. Their works have concluded that the only way to deal with the three problems that make housing price index compilation difficult are through the managing of hedonic housing price indexes. The three compilation difficulties identified are: (1) the quality adjustment problem, since dwellings have many attributes and some of these can change over time, (2) the fact that no transaction prices are observed for most dwellings over long periods of time, and (3) the fact that the 'basket' of dwellings to follow has to be precisely defined, since both the mix of transactions, and the stock of dwellings are changing over time. They also described the French institutional setting where real estate transactions are recorded by a network of notaries who have a monopoly and collaborated with the national statistical agency to provide a centralised quarterly housing hedonic price indexes since 1994. The French hedonic housing price index utilises a reference basket as defined as all dwellings sold during a 4-year period (1990-2001) in each of the 296 elementary zones/strata. The size of the basket in each zone is on average 4,000 dwellings, representing about 1,220,000 dwellings for the whole stock. This is indeed a huge database for reference and

they are fortunate that France has such an excellent system which Malaysia must definitely learn from it.

Another impressive study is the work of Robert and Melser (2008). Their findings were based on using data for 14 regions in Sydney over a six year period. Their data were obtained from Australian Property Monitors and consisted of prices and characteristics of houses sold in 198 postcodes in Sydney for the years 2001-2006. Out of a total 750,000 observations (i.e. house sales), information on characteristics were available for 173,329 or 21.4 percent of their sample. The characteristics for each property were time of sale (quarter/year), type, number of bedrooms, number of bathrooms and land area. They found clear evidence of bias in the region-time-dummy results as well as in simple average measures such as the median that fail to adjust for quality change. They concluded that they were in favour of the hedonic imputations approach. Their project was funded from the Australian Research Council Discovery Grant Program and Linkage Grant Program in collaboration with the Australian Bureau of Statistics. Prior to this project, both the gentlemen had collaborated in 2007 to study the Hedonic imputation and the price index problem: an application to housing as published in the Economic Inquiry.

It is interesting to read the publication of Dorsey, Haixin, Mayer and Wang (2010, 75-93), "Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle." In USA, the repeat-sales price index is common and the most well known being the Case-Shiller™ Index on housing which is the

mantra quoted by housing economists, investments analysts, risk management gurus, financiers, and anyone who has an interest in property prices. Dorsey et al. (2010, 75-93) uses a new extensive data set to construct constant-quality housing indexes. They stated that,

During the boom-bust housing cycle of the last several years, releases of price indexes by the National Association of Realtors (NAR) the S&P/Case-Shiller™ and Federal Housing Finance Agency (FHFA) (formerly Office of Federal Housing Enterprise Oversight) have been and continue to be followed closely. Interpretation of these indexes is often difficult, however, because they are computed infrequently and have unique attributes.

Their sample consisted of almost 1.1 million transactions with detailed property characteristics that occurred in Los Angeles and San Diego metropolitan areas from 1999 to 2008. They had drawn their sample from the National Collateral Database™ (NCD) operated by a mortgage technology firm, FNC, Inc. The extensive data allowed them to account for regional and temporal heterogeneity by constructing monthly indexes at the zip code level from coefficients that vary over time. Their variables were age of house, gross living area, number of bedrooms, number of bathrooms, lot size, property tax and buyer's loan to value ratio.

Another piece of work of computing price index is done by Englund, Quigley and Redfearn (1999), They had used the temporal disaggregation in the estimation of housing prices and volatilities. Their samples covered three cities in

Sweden from 1981-1993 with a total of 270965 transactions. The variables chosen were different from other studies apart from size they used number of garages, tile bath, sewer connection, sauna, stone/brick, fireplace, laundry room, waterfront location, quality of house and insulation. They recommended caution in the use of government produced price indices or those produced by private firms based on repeat sales model.

Filho and Bin (2005), took the normal regression which assumes parametric specifications a step further in their research to create a model of hedonic price function for housing using nonparametric model. The housing attributes used are, again, area, lot size and building age. However, they included environmental attributes of distance to central business district (CBD), nearest wetland and nearest commercial district. They found that sales price and nearest to wetland have a negative relationship, proximity to park has little impact on sales price, and industrial zones are bad for housing price, while positive relationship exists for nearest to commercial zone and CBD. Likewise,   Clapp (2004), proposed a semi-parametric method, Local Regression Model (LRM), for estimating local house price indices using 49511 sales from Q1 1972 to Q2 1991 in Fairfax County, Virginia. They concluded that normal price indices omitted noise in the estimates. The out-of-sample prediction errors demonstrated that LRM added significant information to the hedonic model.

2.3 Hedonic regression limitations

As with any asset pricing model there are bound to be limitations and it is good for us to learn of some of these. As can be observed from the three articles reviewed above, it showed that main limitation of hedonic regression must be the data requirement in order to smoothen out the heterogeneity of housing into more homogeneity. Limsombunchai., Gan and Lee (2004), cautioned that, "While the hedonic technique is an acceptable method…, it is generally unrealistic to deal with the housing market in any geographical area as a single unit." They also found that hedonic price coefficients of some attributes are not stable between locations, property types and age. In his evaluation of the housing market models, Gibb (2003), noted that the complexities associated with models are reflected in the different approaches. There are generally two main approaches using the mono-centric spatial model and the filtering process, which focus on the process and the dynamics of the housing market. Both models demonstrate some strength and also weakness. Therefore, extension of models and redefining markets remain a continuous process. The imperfect property market characteristics, demographic existence in the location influencing buyers' behaviour, existence of sub-sector and other heterogeneous attributes create problems in valuation of property market. Watkins (1999) suggested that regression models would be subjected to bias if they fail to accommodate the existence of the housing markets.

Despite the caution of Limsombunchai et al. (2004) further researches have been conducted on housing prices with hedonic modelling by incorporating macro economics factors (Pashardes and Savva, 2009; Glindro et at., 2008) and

22

studies continue to examine the impact of housing attributes from micro analysis of property specifics, location, spatial coordinates, temporal changes and even buyer's characteristics (Xu, 2008; Bitter and Gordon, 2007; Hamid, 2006).

2.4 Malaysian housing price index studies

Azhari (1991) used a sample of 322 double storey houses in Johor Bahru to construct a model to derive a housing price index based on the measurement of different floor types. He considered that the variety of floor finishes in Malaysian houses has a value to it. Floor ratings were assigned against the floor finishes.

Azahri and Ghazali (1994) conducted a case study on the construction of land value maps using GIS and MRA of residential properties in Johor Bahru, Johore. Dzurlkarnain, Thiruselvam and Ibrahim (1996), developed a mass appraisal valuation model using multiple regression technique to assess the micro and macro economic impact and the dynamics of the property market in rating valuation in Johore Bahru. These were important pieces of work that led to the development of changes in the manner in which the Malaysia House Price Index (MHPI) were constructed with more detailed information for users.

Chau and Chin (2002) reviewed the hedonic pricing model and applied it into the pricing of condominium market in Penang. They managed to obtain from Raine and Horne International Zaki & Partners Sdn. Bhd, a real estate company, a sample of 120 condominium transactions. They used the locational attributes

comprising accessibility to Centre Business Districts, views from the condominium and land tenure. For structural attributes they chose number of rooms, floor area, quality, age and facilities. They also included three neighbour attributes of socio-economic status of residents, proximity to shopping centres and environmental qualities. The results showed that all the coefficients are highly significant and of the expected sign.

One study done in Sabah using macro economic variables was by Mulok, and Kogid (2008). They had used the stepwise regression analysis to study the impact of government allocation, population, GDP and time dummies for three separate crisis periods from 1980-1998. They found that government allocation for low-cost housing projects is very important and population has a positive influence in the policy.

According to Norhaya et al. (2008), they highlighted that interest to adopt the multiple regression analysis in property market valuations were done by Fadilah and Fauzi (1991); Azhari (1987) and that integration of hedonic pricing models with GIS were done by Hamid (2006); Suriatini (2006) and Eboy et al. (2006).