Division of Building and Real Estate Economics
Department of Real Estate and Construction Management
Royal Institute of Technology, Stockholm, Sweden

Master of Science Thesis Number: 40

Modeling Hedonic Real Estate Price for Small Family Houses in Addis Ababa

Author
Sema, Negash Zerga

Supervisor
Mats Wilhemsson

Master of Science Thesis

Title: Modeling Real Estate Price for Small Family Houses in Addis Ababa
Author: Negash Zerga Sema
Department: Department of Real Estate and Construction Management
Division of Building and Real Estate Economics
Master Thesis number: 40
Supervisor: Mats Wilhemsson

Keywords: Hedonic modeling, Price index,

Abstract

The primary purpose of this research is modeling real estate price and identifying degree of significance of price determinants for the new housing developments in Addis Ababa. It also tests effect of five public holidays on housing price and finally develops a Quarterly Real Estate Price Index for five years period using hedonic method. The findings in model one shows that location has a significant effect in housing price. From the empirical result, a 10% increase in plot area increases the housing price by 5.5%. Furthermore, a ten percent increase in plot area around *CMC* will cause a 5.5% while the same increase in *Alemgena* moves upward the housing price by 4.5%. The result also shows no seasonal effect but house price increases by ten percent during Easter holiday period. Price of small family houses has increased by 46% in real prices between first quarters of 1997 and second quarter of 2001.

Acknowledgement

I would like to express my gratitude to my supervisor Professor Mats Wilhelmsson for his insightful comments and great assistance in enabling me to find the way to tackle the research question, despite his tight schedule. I am also deeply indebted to all the teachers and staffs in the department.

The analysis of this research is based on the data provided by *Gift*, *Sunshine*, and *Enyi* Real Estate companies and I am thankful for their help, special thanks goes to Mr. Tesfaye Misganaw who facilitated getting the data.

Finally, I would like to thank my parents for their selfless love and moral encouragements. I also offer my regards and blessings to all who supported me in any respect during my stay in Sweden.

Negash Zerga Sema
May 23, 2010

Abbreviations

PRCE: Transaction Price

PLTA: Plot Area

TFA: Total Floor Area

LVGA: Living Area

DNGA: Dining Area

KA: Kitchen Area

BDR: Number of Bed Rooms

BTHR: Number of Bath Rooms

MBA: Master Bed Room Area

Loc1: Location One

Loc2: Location Two

Contents

1. Introduction

A considerable difference exists on demand and supply of residential houses in Addis Ababa. High demand of residential land for a long period of time and shortage of housing supply from Government side makes housing price unaffordable for many citizens residing in Addis Ababa (Tesfaye, 2007). The Government is taking a series of measures to reduce this shortage for the last several years. One of the measures is encouraging private sector to participate in real estate development by supplying huge land in different exits of the city with additional incentives. The first developer commenced in 1989* Ethiopian calendar with the aim of developing 25,000 homes and handed over 4000 in its first phase.

For the last ten years private real estate development has been booming in the outskirt of Addis Ababa. Despite the fact that there is a tremendous increase in the production of small family houses year after year the housing price also increases. Currently there are more than 60 listed large to small scale developers competing to provide better homes. However; in a developing country like Ethiopia, the increasing cost of living and the weak earning capacity of population makes home ownership extremely difficult that makes the basic one most luxuries item for low income households. Private developers are unable to address low income households for many reasons, in order to minimize this gap the second measure the government has been taking within the last few years is, inner city renovation having a goal of changing the slums to high rise residential buildings and constructing condominiums in open areas. Even if many actions have been taken from the supply side, housing price is still increasing.

This paper tries to answer the question what are the main factors affecting housing price. It will investigate the price determinants by using hedonic method. Heterogeneous nature of house makes finding out house price determinant a bit tough. It also develops a real estate price index for five years. Due to lack of centralized institution which can provide real estate transaction data, we prefer to get the data from large scale Private real estate developers that are willing to do so. For this research paper we collect data of small family houses sold by private developers starting from 1997-2001*. The data will comprise of transaction price, plot size, number of bed rooms, and number of bath rooms, total floor area, date sold, location and

* *Ethiopian Calendar: All dates refereeing data and empirical results are in Ethiopian calendar.*

developer. Date sold is the contract agreement date when the property is sold. There are two locations, *Alemgena*(Loc1) and *CMC*(Loc2). The national holidays, Ethiopian New Year (*Enkwtatash*), Easter, Christmas, Eid Al-Fitr and Eid Al-Adha are taken to check whether holiday periods have an effect in housing price. This hypothesis test is required because of the fact that most of the customers of private developers are Ethiopians residing abroad and those customers go home commonly for the above holidays.

1.1 Aim

Many studies have been made on hedonic real estate price model and showed that the price of a house is determined by its structural, neighborhood and location attributes. The primary purpose of this research is modeling a hedonic real estate price by identifying the real estate price determinants and their degree of significance for the new housing developments in Addis Ababa. Location effect on implicit price of attributes is checked. It also tests the effects of holidays on housing price and finally develops a Quarterly Real Estate Price Index for the period of five years.

1.2 Methodology

There are three different types of transaction based methods used to analyze the price fluctuations in the housing market. The first one is average price, the second one is repeated-sales method and the third one used in this paper is hedonic method. Since housing market has a heterogeneous nature, Hedonic method is good in attempting to control for different types of homes sold over time. The statistical software STATA is used for empirical analysis.

1.3 Limitations

Due to lack of institutions established to collect housing transaction data in the country, the research data is gathered from willing developers. The reason we prefer to use only private developers' transaction is that they are relatively transparent and can provide the data required.

1.4 The layout

In the first chapter of the paper an introduction to the real estate market in Ethiopia will be discussed. The aim of the paper, limitations and method used will also be seen briefly in the introduction part. The second chapter is literature review which discusses the different approaches for Real Estate price modeling. It also gives detail clarification of the hedonic model and justification for using it in this research. Getting exact variables and functional form in our model pave a way to get the right specification of hedonic equation so that the last section of this chapter will talk about functional form. In the third chapter, we will discuss about the model specifications used in our study. Homoskedasticity and Multicollinearity problems encountered in hedonic models and methods used to solve them will be seen in detail. In the last section of this chapter we will see Serial Correlation which has an effect in efficiency of a model in time series analysis.

Descriptive statistics and correlation of the data is presented in the fourth chapter. The Empirical result chapter presents the different hedonic models developed and construction of Hedonic Real Estate Price Index. It also discusses the results of hypothesis tests, Homoskedasticity, multicollinearity, Serial Correlation, Seasonal effects and Autoregressive tests. Finally, conclusion will be presented in chapter six.

2. Literature Review

2.1 Real estate price models

An accurate real estate price measurement method is prominent in understanding the current boom in housing market and the construction industry in Addis Ababa. Financial economists noted that lack of precise measurement methods in real estate price as the main reason for lack of focus on the real estate market relative to other markets like stocks, bonds and foreign exchange (Richard 1997).

In real estate market, it is not easy to get many houses with similar quality. The facilities provided and the standard of construction may vary greatly between houses. Even if buildings have the same facilities, the rate of usage depends on the building's age. Besides, the real estate price has complex individuality (Shimizu 2007).The primary problem (Follain and Calhoun 1997) related with the evaluation of real estate price is its "heterogeneity". He also added that bigger heterogeneity is more complicated in valuing, due to variation in quality. As a result, we need to adapt a method in order to manage change in quality of real estate through time. One way to avoid these difficulties is to estimate quality differences using regression analysis (Bailey 1963).

2.2 Modeling approaches

Estimation of housing prices can be done in three ways. The first approach and the one applied for this study is hedonic which uses regression and explain price as a function of the house attributes. Freeman (1993) categorizes these variables in hedonic models in to three: structural, neighborhood, and environmental. The other alternative approach is repeats sales method which applies a trace in change of house selling price variation of the same home in various positions through time (Follain and Calhoun 1997). The third approach which we will not discuss in this paper is Hybrid approach. It is combination of elements in the two approaches and first recommended by Quigley (1995). The composition which he imposes result in more proficient factor estimates, if the hypothesize error is acceptable. However, only the first two approaches are commonly used.

2.3 Hedonic Price modeling

A regression model showing the housing price by using explanatory variable(Y) based on consumers' preference criteria, such as travel time from the nearest train/bus station; distance from CBD, number of bed rooms and age of the house is projected (Shimizu 2007). It is a well-liked method used for modeling house price, where property is assessed based on its different attributes. Implicit Property price is possibly estimated from such models according to the inherent attributes. Many studies have been undertaken in hedonic pricing to the property market for several years and a lot of improvements have been made on the theory.

Hedonic price analysis is a technique which studies the demand side of housing, assuming that a property is sold as a package of inherent attributes (Freeman 1979; Rosen 1974). The word 'hedonic' comes from a Greek word 'hedonism' which relates to pleasurable or painful feelings or results (Els and Fintel 2008). It is used to evaluate the price of a real estate by different characteristics such as number of bed rooms, plot area, number of bath rooms and age of the house, access to public transport, view and so on. It was widely used in housing studies for investigating problems such as racial discrimination, neighborhood change and accessibility to work (So 1997).

The basic hedonic equation is mostly represented by regressions of selling price of house as a function of different characteristics of the residence. The regression bases on the assumption that we know the attributes of selling price.

P = f (S, N, L, T), where *(Malpezzi 2002)*

P = selling price of house;

S = structural characteristics;

N = neighborhood characteristics;

L = location characteristics

T = the time.

Regressions are estimated for different markets or time periods that explain price of housing unit then an index is computed by evaluating the evaluation regression for a particular collection of characteristics in the market. The hedonic approach is commonly employed in the market for owner occupied housing (Follain 1997). Weakness of hedonic regression approach includes omitting functional forms and set of housing characteristics in the model (Meese and Wallace 1997).

"Researchers have dealt with the functional form issue by using either flexible parameterization of the model or nonparametric techniques. The second drawback is more difficult to deal with, as it plagues all research in empirical economics. Consistent estimates of implicit hedonic prices will rely on the other assumption that all omitted variables are orthogonal to those included in the analysis". (Meese and Wallace 1997)

However, several empirical studies show that hedonic approach is more preferable relative to other methods. Mark and Goldberg (1984) evaluate hedonic and repeat sales methods for neighborhoods in Vancouver and conclude that hedonic approach is better relative to repeat sales that become visible to undervalue increase in price. In recent times, (Meese and Wallace 1997) argues for the hedonic approach when they apply it for US. The reason is that it is less affected by sample-selection bias and not steadily inherent prices of housing determinants compared to repeat sales approach. Clapham et al (2004) also discover that indices developed by Swedish data set using hedonic approach are less prone to adjustments than repeat sales. Moreover, hedonic approach is less sensitive to tiny segmented markets (Francke, Vos and

Janssen 2000). A study by Wilhelmsson (2008) to construct a price index for the segmented housing market also shows that hedonic regression method is more suitable than the repeated sales method.

According to (Els and Fintel, 2008) the main advantage of hedonic price method is that it is based on real market selection by professionals and theoretical background. Besides, real estate market is vast, so sound contests is anticipated, large observations and on average correct assessments. Finally, real estate data tend to be easily available, as real estate is a huge investment in many countries which requires appropriate record keeping.
Limsombunchai (2004) also added that managing property attributes is one of the advantages of hedonic method. He further noted that this allows a person to differentiate the impact of the real property appreciation and changing observation composition. Even if this model has been widely used Limsombunchai argues that there are problems such as model specification procedures, multicollinearity, independent variable interactions, Heteroskedasticity, non-linearity and outliers which can affect the performance of hedonic price model in real estate appraisals.
Difficulty in identifying the neighborhood characteristics due to similarity in a specific neighborhood is the first drawbacks of hedonic price method (Els and Fintel, 2008). Moreover these characteristics are measured based on the willingness to pay of the buyer for that specific attribute provided that the buyer is aware of its existence. They further noted that hedonic method considers the combined characteristics of the property transacted. This method is more complicated than other methods and needs skill in statistics on interpreting variables and selecting functional forms which are not a direct process.

The second approach to estimate the housing price is repeat sales approach. As the name refers it mainly concerns with properties which has a repeated sales record. Commercial properties and rental apartments are mainly estimated by this model. According to (Shimizu 2007) the repeat sales approach focuses on sales of transactions which have occurred repeatedly in the market. Follain and Calhoun (1997) argue that due to lack of ample database about commercial and rental properties and their characteristics hedonic approach do not fit such properties. They further noted that heterogeneity of these type of real estate exceed that of owner- occupied market which encompasses relatively standard single- family housing units.

(Meese and Wallace 1997) said that drawbacks to the hedonic regression approach are omitting of functional forms and combination of house attributes in the analysis, which can result in conflicting estimates in the model. However the functional form issue is resolved by taking different techniques.

The main drawback to use repeat sale is availability of data, even in countries like Britain (Malpezzi 2002) getting data is the hardest part of the job. Malpezzi also added a second shortcoming of repeat sale, *"even at its best, the method only yields estimates of price changes. No information on price levels, or place-to-place price index, is derivable from the repeat sales method".*

Malpezzi also mentions other argumentative points about repeat sales approach. First is sample selection bias, houses that have a high frequency of transaction may not be a representative of the actual stock market. He further noted in this issue that the effect of the sample selection bias will be a matter of the output's use. Second the approach assumes no new supply is introduced in the market for the period under consideration both in quality and quantity which is far from the reality to some extent. Finally it also assumes that the coefficients in hedonic model would not change. Again it doesn't seem reasonable to think number of bath rooms' coefficient is the same for 2000 and 2007 without keeping family number constant if we have *citrus paribus* effect on other things. For the reason we could not discover a comprehensive database containing repeated sales transactions of similar houses. So, it is impractical for us to apply this approach to establish a housing price index (Quigley 1995).

2.4 Functional Forms

The next thing to look at is functional form of hedonic equation. Linneman (1980) explicitly noted that the exact variables and functional forms pave a way to get the right specification of hedonic equation. Henry (1995) also added that hedonic function should consist of house determinants having high utility and expensive. Linear functional forms widely used before Goodman (1977) rejected using a method proposed by Box Cox, a transformation method named after the authors Box and Cox (1964). Henry (1995) also applied this technique to search for the finest functional form of hedonic equation.

$$P(\lambda) = \begin{cases} \left(\frac{P^{\lambda}-1}{\lambda}\right) \text{ when } \lambda \neq 0 \\ \text{Log}(P) \text{ when } \lambda = 0 \end{cases}, \quad \text{Where,} \tag{1}$$

λ = 2, Square transformation

λ = 1, no transformation/linear

λ = 0.5, Square root transformation

λ = 0, Log transformation

λ= -1, Reciprocal transformation

P is transaction price and **λ** is a parameter used to make a Box-Cox transformation on price. In the study of housing markets, weight should be given to the condition of the explained variable rather than the explanatory variables (Linneman, 1980).Henry further noted that **λ** differs for each hedonic equation according to the housing market. Megboluge (1989) said that **λ** indicates how constrained the housing market is. Moreover, **λ** value equal to one (zero) indicates a "stable equilibrium" while a value greater than one and less than one shows a loose and "tight market "respectively. Henry (1995) said that a wide variety for **λ** shows a different price structure in many cities. He strengthened his view by refereeing Studies in US cities. **λ** is between 0.2 - 0.4 for United States cities (Linneman, 1980).

(Halstead et al, 1997) also noted that since its earliest application in real estate studies in the late 1960s, hedonic method has shown to be a major tool for estimating and forcasting the effects of changes in different attributes of the housing market. However, those authors argue that many of the previous studies dealt with the *problem specification, data collection, variable definition and selection, and welfare changes* but not the functional form to use. They pointed out that in recent years few researchers turned their eye on the issue of appropriate functional form selection for hedonic model and they work on using flexible functional forms using a case study of a certain locational attributes amenities /disamenities on houses in the area. The findings of their study show that there might be a problem in estimating benefit which comes from chosing inappropriate functional form.

The functional form in hedonic model is a big issue. In early times reserachers used to choose from the many functional forms (linear, logarithmic, or semi log) based on the basis of goodness -of-fit criteria (Freeman 1993 Cropper *etal* ,1988). However, Halstead *et al* (1997) noted that linear functional form restricts independence of the explanatory variables and the logarithmic functional form makes the price dependent on the level of other variables. Those authors argue that in any situation the functional form problem might be solved by the data itself. Freeman (1993) refers Goodman (1978) as the first in applying the Box-Cox transformation of the dependent variable in hedonic modeling.

Box-Cox transformation gives a method for choosing from the different functional forms. Goodman's method has some limitations for not considering the alternative forms for the independent variables. Halstead et al, (1997) says that recent researches use transformation of both the dependent variable and different transformation for each independent variable but Freeman(1993) argue that this is not feasible in estimating actual number of attributes. Greene (1993) also added that transforming both the dependent and independent variable differently is more than necessary. The other method recommended by Palmquist (1991) to avoid this problem is to prioritize the dependent variable and independent variable(s) of main interest. Milon *et al* (1984) used this method differently in *water-related* attributes but those authors used no transformation for attributes which are not related to water to effectively restrict the effect of their coefficients.

We use the *Box-Cox transformation in both sides with the same parameters* and all the λ coefficients are nearly close to zero. This tells us that we better use the Log transformation.

3. Model Specifications

3.1 Homoskedasticity

The first model consists of the structural variables plot area, number of bed rooms, kitchen area and number of bath rooms. In addition to this we tried to add one location variable. The data is from two places so we create dummy variables for them to check if there is location effect on the housing price that have a prominent effect in accessing local facilities such as public transport and community services. Combination of variables is used to make the above variation, if it exists. We tried to set up a linear equation but it suffers from Heteroskedasticity problem and the functional form we used to explain the implicit price estimate for small-family house is Log-Log model.

The assumption of Homoskedasticity states that for a given explanatory variables, the variance of the error (*u)* is constant. If this is not true, if the variance of u is different for different values of x then we say that we have Heteroskedasticity. Problem encountered by having Heteroskedasticity is that the standard errors of the estimates are biased and if the standard errors are biased, we cannot use the usual *t* statistics or *F* statistics for drawing inferences even if Ordinary List Square is still unbiased and consistent.

Homoskedasticity assumption

Let x stand for $(x_1, x_2 \ldots x_k)$

Assuming that $\mathrm{Var}(u|x) = s_2$

We re-specify the model again by transforming the variables and changing functional forms in order to make the model simple and easy for interpretation. The second way to deal with Heteroskedasticity is the *robust standard error* option. Since Ordinary List square assumes that errors are independent and similarly distributed; *robust standard errors* satisfies one or both of the assumptions. Hence when we come up with Heteroskedasticity *robust standard errors* is more trusted. This method doesn't change coefficient estimates, it only changes the standard errors and F statistics gives a precise P values. However; in this paper we use the first option by changing the functional form in to Log-Log model.

3.2 Multicollinearity

Multicollinearity is a situation where independent variables are highly correlated with each other. Even if we have high value of R-square, multicollinearity can cause high standard error and low *t value* when it exists in a model. It also shows unexpected changes in coefficient magnitudes or signs, or non-significant coefficients. It doesn't violate Ordinary List Square assumptions since it is not perfect so OLS estimates are still unbiased and *BLUE (Best Linear Unbiased Estimators)*. However; high multicollinearity causes high standard error. High multicollinearity also causes confidence interval of coefficients to be wide and *t value* to be small which makes the variables statistically insignificant. STATA drops perfectly correlated independent variables but if they are not perfectly correlated we need to make a multicollinearity test.

3.3 Time series analysis

In order to find out the empirical relationship formally, econometric models in this study use time series data which are ordered in time that not cross-sectional data are. It is important that we do not change the ordering because we need a model that allows the past to have an impact on the future but not vice versa. Since we have data ordered in time we need to make some assumptions about the error term and how it can behave over time. The time series data will make us able to construct the price index and will also help us to see if time has any effect in housing price. Serial correlation(see section 3.5) and non stationary are the problems we need to take care of in time series data. Durbin-Watson, AR (1)-Test and Dickey-Fuller Tests are tests made to check those problems in models.
Models in time series data could be static or dynamic.

Wooldridge (2006) defines static model as the independent variables(x) directly influence the dependent variable(Y)

Static Model: $y_t = b_0 + b_1 z_t + u_t$

Whereas in a dynamic model [finite distributed lag (FDL) model] allows the independent variables to influence the dependent variable(Y) with a lag.

Dynamic Model $y_t = d_0 + d_0 z_t + d_1 z_{t-1} + d_2 z_{t-2} + u_t$

Where d_0 *"the impact propensity"* it gives us the immediate effect on y given a change in z. We call $d_0 + d_1 + \dots + d_q$ "the long-run propensity (LRP)" which give us the long run change in y given a permanent change in z. We are not going to estimate the above models, but we will test for seasonal effect by including quarter dummies together with a trended variable. We are also going to test if the price index is an autoregressive model of order 1 AR (1), see section four.

3.4 Trended time series

Many economic time series data have a tendency of fluctuating through time, either increasing or decreasing according to Wooldridge (2006). Therefore we have to consider a series with time trend so that causal inference can be made. He also added that "ignoring the fact that two sequences are trending in the same or opposite directions can lead us to falsely conclude that changes in one variable are actually caused by changes in another variable. In many cases, two time series processes appear to be correlated only because they are both trending over time for reasons related to other unobserved factors". But we have to make the trend model relatively simple to capture wider movements in the dependent variable which are not explained by the independent variables in the model.

The problem with a trended series is that we cannot fulfill zero conditional mean assumption of OLS, [E (e/x) =0)] the expected value of the error given the independent variable is equal to zero. Besides, the estimated parameters will be biased and we cannot make both economical and statistical interpretation from the model. Since our data is sorted by date we create dummy variables for five years and 18 quarters taking the first one as a benchmark for both to add a linear trend and control seasonal variation in our model. It is common that time series data can contain seasonal variation in quarter. Such variations can be taken care of by creating dummy variables for quarter.

Example linear trend

$$yt = a_0 + a_{1t} + e_t, \qquad t = 1, 2, \ldots$$

3.5 Serial Correlation

Autocorrelation between units of observation is the main trend in hedonic research (Els et al, 2008). The spatial dependency model assumption tells us that spatial autocorrelation over time between independent variables is only due to the existence of similar area effect in the data set. However; this might be a drawback in analyzing unforeseen changes in the current period which will have an effect on the coming periods Baltagi et al (2003). *Serial correlation* stands for correlation of error terms on different time periods. It occurs in time series data, as error of given period carry over in to future periods. Measuring serial correlation precisely could help both home owners and developers in predicting future price movement *(Kuo 1996).*

Compared to other *liquid financial assets,* the real estate market transaction is less frequent due to transaction costs and taxes. *(Kuo 1996).* In real estate market where properties are sold rarely Real Estate Price Indexes are constructed to show price movements for longer period. Kuo (1996) noted that those indexes are not correct figures and are correlated. If we use the indexes as true values in autoregressive model then we will end up with inconsistent autocorrelation estimator.

Serial correlation has no effect in the unbiasedness or consistency of OLS estimators; however it can have an effect in its efficiency. If we have positive correlation, our OLS estimates of the errors will have a lower value that the actual values and this makes us to think that our estimates are accurate which are not in actual case. There will be a tendency of rejecting the null hypothesis when it should not be rejected.

The assumption of no serial correlation of the error terms is one of the key assumptions in OLS models. In a model where the regresses are strictly exogenous the assumption of no serial correlation is.

$$E(u_t/X)=0,$$

Where, X represents the repressors in all time. The assumption implies that expected value of the error term, u_t, is not correlated with the repressors in all time period.

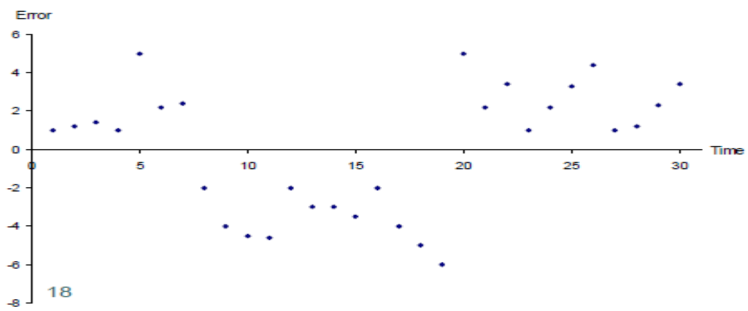
If the above assumption is not fulfilled: that is if $u_t>0$, the error term in the next period will also be positive. The first test and the one we use in this paper is t-test using first degree autoregressive model, AR (1).

AR (1) = Autoregressive model where the dependent variable is a function of the dependent variable lagged 1 year.

$$y_t = ry_{t-1} + e_t \text{ , } t = 1, 2,\dots$$

It is done by testing the null hypothesis H=0 in Test of AR (1) autocorrelation

$$u_t = ru_{t-1} + e_t, \; t = 2,\dots, n$$



Graph 1: Distribution of the error term over time

Test of autocorrelation of higher order [AR (q)] can be done in the same manner as for AR (1) by Including q lags of the error term in the regression and test the null hypothesis that all estimates are jointly significant different from zero. Instead of t- test we use the F-test.

An alternative test is the measure Durbin-Watson (DW) test, which many statistical programs estimate automatically. If DW statistics is 2, the hypothesis about autocorrelation can be rejected. If it is higher or lower than 2 then the hypothesis cannot be rejected. However, it is difficult to estimate the critical values which make the AR (1)-test simpler to use.

$$DW = \frac{\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n}\hat{u}_t^2}$$

- DW < 1, 5 autocorrelation
- DW = 2 no autocorrelation
- DW > 2, 5 autocorrelation

3.6 Hypothesis Test

One aim of this paper is to investigate the effects of holidays on housing transaction price. In order to make the test, first we develop a model containing holiday dummy variables for the selected ones'. There are five holidays, two Christian and Two Muslim and The New Year. Dummy variables are created in two ways. The first one is by taking holiday period one month prior to the holiday the other option is fifteen days before holiday and fourteen days after holiday which makes it one month including the holiday. Regression is made in both cases to see if there is any difference in choosing the holiday period. Only New Year and Christmas have a fixed day every year. The other three holidays have variable days each year; holiday period includes this variation in both cases.

4. Data Analysis

4.1 The Data

In modeling housing submarket for UK Midlands region using a dataset of nearly 18,000 transactions in 1994 Fletcher et al (2000) noted their concern with sample size for hedonic estimation. Those authors argue that in large sample size with wider location and property types there could be a difficulty in getting a stable coefficient representing the whole sample. Their estimates might not show the difference and result in biased coefficients due to wrong specification. However they noted that small number of sample size may not also be enough to correctly apply regression techniques. The data may consist of some attributes that can apply to only a few properties in the sample which gives inadequate information to precisely estimate the value of these attributes in the population.

The biggest problem in this study was getting the data. In a developing country like Ethiopia, where there is no well established real estate taxing system, it is difficult to get data from a single institution. The only option we had were collecting the required data from developers which are willing to provide us. The data used in this study contains 213 transactions of newly developed small-family-houses sold by real estate developers for the last five years (1997-2001) in Addis Ababa located around *CMC* and *Alemgena* area. From many developers found in the city, only three were willing to provide data required for this study. Unfortunately the third data provided by one of the companies is biased and will not be used in the analysis. 193 of the transaction data is from Sunshine Real Estate, Country's biggest developer and the rest 20 are from Enyi Real Estate, a growing real estate company. Sunshine homes are located in *CMC* while the others are located in the opposite exit of the city named *Alemgena.*

4.2 Descriptive Statistics

The table below shows descriptive statistics of the variables for 213 observations. It presents the average transaction price of newly built residences, standard deviation, maximum and minimum values of main structural attributes of the houses which can affect housing price mostly for the period of 1997-2001. It is not surprising to see that there is no big difference in between the min and max values in all variables used. Since the houses are developed by two developers, they mostly have many similarities in many characteristics compared to houses built privately. Each developer can have 5-10 different designs usually differ in external appearance but having closely related internal structure.

The variables used in the descriptive statistics are Transaction date, price, and living area, dining area, master bed room area including close-set area, kitchen area (only modern kitchen), number of bed rooms and number of bath rooms. The price here may not actually be the market price. Since any developer can set its price based on the construction cost and its own profit margin, we cannot be sure to say that it is market price. Besides, the transaction is not made in auction or publicly. It is a negotiation between the developer and each buyer so there might be a price difference between the same houses depending on how tough the negotiator/buyer is for the developer.

The date which the developer and customer signed an agreement is taken as the transaction date. Five holidays are also included here in the descriptive statistics as dummy variables for the hypothesis test of which holiday has an effect in house selling price. The holidays are Ethiopian new year (*Enkutatash*), Christmas (*Gena*), Id al Feter, Id al Adha and Easter (*Fasika*). Holiday periods are considered in two ways. The first one is taking the holiday period one month prior to each holiday and the second is fifteen days before and fourteen days after each holiday. It can be seen from the table below that the deviation from the mean for most of the variables is appropriate. No outlier effect is observed from this descriptive statistics.

Table 1: Descriptive Statistics

Variable |	Obs	Mean	Std. Dev.	Min	Max
-------------+	-----------	-----------	--------------	-----------	---------
PRCE|	213	1783662	598141.1	943485	3305218
PLTA|	213	422.5725	145.0286	250	838
T FA|	213	328.543	107.7767	171.45	489.44
LVGA |	213	27.74775	3.549843	20.86	32.5
DNGA |	213	16.78756	1.836325	13.65	20.1
-------------+	-----------	-----------	--------------	-----------	---------
MBA |	213	36.32014	7.558343	18.9	51.89
KA |	213	17.23193	1.803041	14.6	19
# BDR|	213	3.723005	0.4485684	3	4
# BTHR |	213	3.488263	1.007567	2	6

4.3 Correlations

Correlations between variables used in the analysis presented in table 2. It shows a rough estimate of interaction between the determinants and housing price. It is seen that housing price have a relatively strong positive correlation with the variables plot area, total floor area, living area, dining area and kitchen area. The correlation ranges from 0.44-0.54 which shows the importance of structural attributes of the house in determining its price. The independent variables are also highly correlated to each other. Plot area is positively correlated with dining area and master bed room, and total floor area is also positively correlated with area and master bed room area and dining area.

|	PRCE	PLTA	TFA	LVGA	DNGA	MBA	KA	#BDR	#BTHR

PRCE |	1								
PLTA|	0.448	1							
TFA |	0.536	0.91	1						
LVGA |	0.291	0.327	0.480	1					
DNGA|	0.487	0.737	0.848	0.785	1				
MBA |	0.262	0.811	0.839	0.515	0.7157	1			
KA|	0.312	0.574	0.544	-0.45	0.1104	0.308	1		
#BDR|	-0.096	0.445	0.510	0.724	0.5786	0.632	-0.17	1	
#BTHR|	-0.171	0.251	0.204	0.572	0.4413	0.542	-0.47	0.488	1

Table 2: Correlation between variables

5. Empirical results

5.1 Empirical Model

The equation shown below is our first model by using cross- sectional data. It is in logarithmic form and consists of combination variables with location.

$$\ln(PRCE) = a_0 + a_1 \ln(PLTA) + a_2 Loc_2 * \ln(PLTA) + a_3 \ln(MBR) + a_4 \ln(\# BDR) + a_5 \ln(\# BTHR) + a_6 Yr_2 + a_7 Yr_3 + a_8 Yr_4 + a_9 Yr_5 + e$$

PRCE= price of small family houses

PLTA= total plot area

Loc2=location of the development,

MBR=Area of master bed room

Yr2-5= Transaction Year from year two to year 5

#BDR= number of bed rooms,

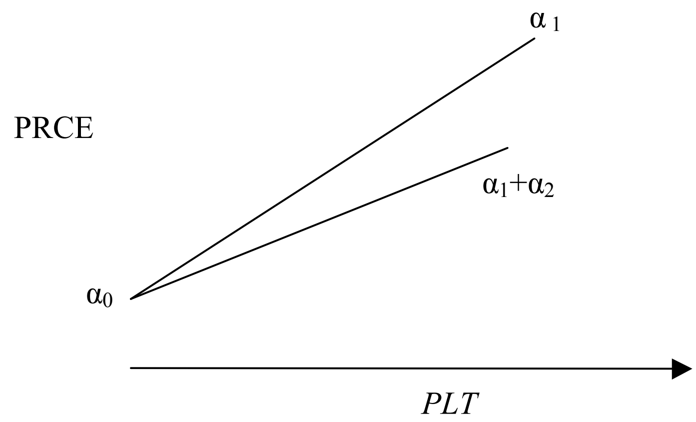
#BTHR= number of bath rooms

The model is tested for Homoskedasticity by Breusch-Pagan Test and gives f-value to be < 0.0758 a value greater than 0.05, which tells us that there is no problem of Heteroskedasticity *(See Appendix Two)*. Moreover, it gives a good power of explanation of the variation in house prices with an adjusted R-squared value of 0.9538(*See Appendix One*). The combined variable Location two and the natural logarithm of plot area have a decreasing effect on the price. All the variables except, the combined one, are positively related with the price and the variables in this model are significant at 95% confidence interval. The result shows that location has a significant effect in housing price. The implicit price for plot area which is for both locations is positive while the implicit price for plot area in location two is negative and this shows that location two has a lower implicit price for plot area.

Table 3: Econometric result

Ln(PRCE) |	Coefficients	t value
$Loc_2*Ln(PLTA)$ |	-0.10	-15.17
Ln(PLTA) |	0.55	22.07
Ln(#BTHR) |	0.11	4.69
Ln(#BDR) |	0.33	5.50
Ln(MBR) |	0.25	5.55
Yr2 |	0.14	10.16
Yr3 |	0.39	28.19
Yr4 |	0.33	9.23
Yr5 |	0.29	5.81
_cons |	10.02	106.11
R-squared	0.9557	
Adj R-squared	0.9538	

The following graph shows the change in the implicit price of plot area versus price. α_1 is implicit price for plot area while α_2 is implicit price for plot area in location two. Because α_2 is negative the summation becomes less in value than α_1 but still having positive slope.



Graph 2: *Location effect on implicit price of PLTA*

Test is done for multicollinearity by using a regress post estimation command*, estat vif*. Variance Inflation Factor (VIF) tells us how much the variance of the coefficient estimate is being inflated by multicollinearity. The square root of the VIF tells us how much larger the standard error is, compared with what it would be if that variable were uncorrelated with the other X variables in the model. All VIF values are less than ten so we don't have severe multicollinearity in our model. The table below shows the test result for multicollinearity and we see that all the variables have no impact on other variances.

Table 4: Multicollinearity Test result

Variable	VIF	1/VIF
-------------	+-----------	----------
$Loc_2*Ln(PLTA)$	7.77	0.129
Yr4	5.38	0.186

Ln(MBR)	4.17	0.240
Ln(PLTA)	3.16	0.317
Yr5	2.58	0.387
Ln(#BDR)	2.58	0.388
Ln(#BTHR)	1.84	0.542
Yr2	1.17	0.857
Yr3	1.14	0.880
Mean VIF	3.31	

The next thing to look at is effect of holidays. The table below shows regression result of a Log-Log model using cross sectional data for the first holiday period. The model gives good result with explanation power of 74.84%. The natural logarithm of plot area, living area, total floor area dining area, number of bath rooms and number of bed rooms are used and found that all are significantly different from zero at 95% confidence interval. The last three variables have a decreasing effect on the housing price whereas the others have an increasing effect. Considering one month prior to holiday as a holiday period only *Easter* (Fasika) is statistically significant having a t-value of 2.25. If a specific house is sold during this period the result shows that there will be nearly a 10% increase in price. Besides, the model shows that there is no variation in price when holiday approaches for the other holidays.

Table 5: Econometric result for holiday test

Ln(PRCE)	Coefficient	t value	Coefficient	t value
Ln(PLTA)	0.544	5.64	0.525	5.37
Ln(LVGA)	2.628	10.78	2.613	10.59
Ln(TFLA)	0.420	3.47	0.421	3.41
L(DNGA)	-1.644	-4.30	-1.611	-4.17
Ln(#BDR)	-2.086	-12.71	-2.025	-12.26
Ln(#BTHR)	-0.276	-4.45	-0.274	-4.36
NY30	0.053	0.94		
GNA30	0.037	0.69		
IDF30	-0.020	-0.53		
IDH30	0.022	0.48		
FA30	0.098	2.25		
_cons	7.641	18.18		
NY15			-0.031	-0.84
GNA15			0.010	0.22
IDF15			-0.012	-0.33
IDH15			-0.028	-0.56
FAS15			0.080	1.52
_cons			7.641	18.11
R-squared	0.7615		0.7569	
Adj R-squared	0.7484		0.7436	

The data used in model one is cross- sectional data, year is used in the model as dummy variable and gives us the coefficients which are significant in affecting the price. The result shows that year has an effect in the housing price which means that there is a linear trend between year and price. In a country like Ethiopia where political stability is a big concern, housing price might have been affected during the period under consideration. The dummy variables created for year can be used to see the variation of housing price each year in time series data analysis. Non homogenous character of the housing attributes and its fixed asset nature makes the housing value more dependent on location also. We make a regression using these location dummies combining with other structural attributes to see change in implicit price. The effect of seasonal variation is tested in the second model using quarter dummies and found that the first three quarters of year 1997 don't have any significant effect on the housing price whereas starting from the fourth quarter of 1997, time has an effect in housing price. The general election was held at the end of the third quarter in May 1997 and the result shows that it had an effect on housing market.

Our model is tested for autocorrelation using the autoregressive method and found that there is no autocorrelation problem. Residual term of the model is regressed by the residual lagged one year AR(1) and gives a t-value of 1.42, which is insignificant in explaining the residual*(See Appendix Four)*. Therefore, we can make the statistical and economical interpretations from the model.

Table 6: Econometric model with time trend(See Also Appendix Three)

Ln(PRCE) |	Coefficents	t value
-------------+	-------------	-----------
L(PLTA)	0.393	11.08
Ln(LVGA)	1.508	14.44
Ln(TFA)	0.545	10.97
L(DNGA)	-1.167	-7.74
Ln(#BDR)	-1.402	-14.44
Ln(#BTHR)	0.041	1.61
YQ2|	0.019	1.54
YQ3 |	0.030	1.44
YQ4 |	0.184	11.15
YQ5 |	0.183	8.17
YQ6 |	0.174	10.31
YQ7 |	0.177	9.35
YQ8 |	0.168	5.86
YQ9 |	0.398	23.09

YQ10 |	0.434	23.52
YQ11 |	0.417	18.76
YQ12 |	0.540	18.96
YQ13 |	0.425	12.05
YQ14 |	0.464	15.32
YQ15 |	0.358	10.12
YQ16 |	0.437	11.87
YQ17 |	0.452	10.34
YQ18 |	0.472	9.41
_cons |	8.757	49.27
R-squared	0.9746	
Adj R-squared	0.9716	

5.2 Real Estate Price Index Construction

Els *et al* (2008) used transaction data sourced from the Residential Property Price Ranger and consisting of 1930 house sales for the time period; from September 2004 to August 2007 to create a hedonic price model for Stellenbosch in South Africa. They noted that the two famous property market analysts in South Africa, ABSA and Rode Property Valuations use nearly the same approach to construct house price index. Rode House Price Index use median sales prices and ABSA House Price Index follows the average sales price of properties (Rode, 2005). Residential property transaction price in South Africa is becoming high since 2000 nearly 10% in nominal value but less for the late 2007 and 2008 (Rode, 2005& ABSA, 2008).

They further noted that the methods in practice now are not satisfactory to precisely analyze property market behavior, besides in contrast to the reality they consider it not as a heterogeneous asset. Els *et al* (2008) also argue that ABSA's and Rode's methods are not the best favorable empirical approach to come up with a concluding statement about the South African property price growth. By using microeconomic data they showed that hedonic models are able to separate the effect of measurable attributes by controlling for their individual effects over time, allowing '*pure property inflation'* to be measured.

The second part of our model is constructing a Real Estate Price Index . According to Meese and Wallace (1997) the construction of indices using hedonic approach models is a two step practice. First is estimating regression of house sales price on a bundle of house attributes. The second step of hedonic-regression approach is to use the estimates of the inherent attributes in constructing the index. Price index number typically aggregates a vast number of information into a single quantity which is used to compute inflation rates and for turning a time series measured in nominal values into real values. Some uses of real estate price index are important as macroeconomic indicator and as an input for CPI calculation. Besides, it can be used to calculate wealth as well as analysis of lenders exposure to insolvency and foreclosure. Furthermore, for the assessment of risk of housing price bubble according to Wilhelmsson *et al* (2009). There are three different methods to construct price index.

The first one is repeated- sales price index (Bailey *et al.* 1963), the second is average price index and the third one is hedonic price index (Wilhelmsson, 2009). Due to drawbacks in controlling different types of houses, the average method is not used for small family houses (Case and Schiller, 1987). This method doesn't take in to consideration that different types of properties are sold at different times. Since houses are not homogenous goods overtime there is change in sample over time for example large properties are sold in the boom and small properties are sold in the bust for such reasons the price decrease will be larger than the true price change. Meese and Wallace (1997) also noted that repeated sales method has problem in controlling influential observations and sample selection bias. Those authors recommend hedonic method for price index construction which is better needs controlling the omitted variable bias that is a drawback for hedonic method.

First specify a hedonic price equation as shown in model two then estimate the hedonic prices by including time dummies, year or quarter dummies. Time dummy variables account for shifts in the equation intercept. The estimated parameters (coefficients) concerning the time dummies are the price index. In our case we take coefficients of the YQ, which are quarters of each year. The table below shows the coefficients of quarters in the model and the real estate price index for each quarter taking quarter one of year 1997 as a benchmark. The last column shows the change in REPI from preceding quarters.

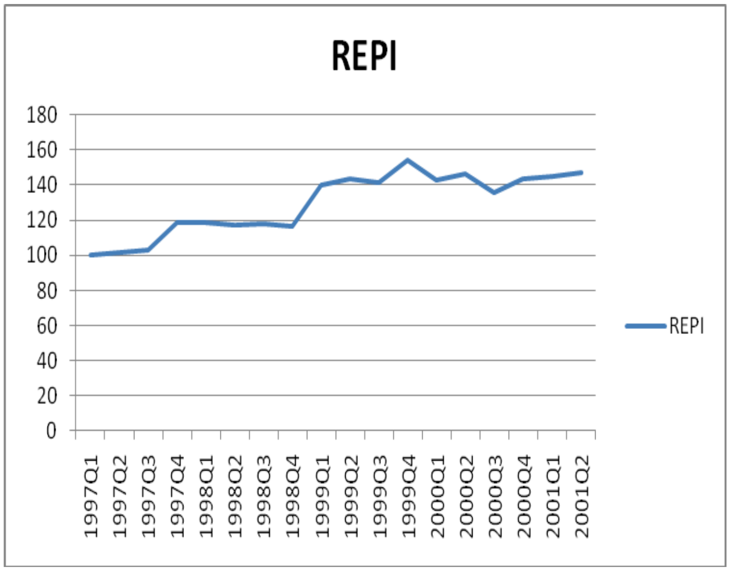
Table 7: Price Index

Year	Qr.	Coff.	REPI	% change
1997	Q1		100	

1997	Q2	0.02	102	2%
1997	Q3	0.03	103	1%
1997	Q4	0.18	118	15%
1998	Q1	0.18	118	0%
1998	Q2	0.17	117	-1%
1998	Q3	0.18	118	1%
1998	Q4	0.17	117	-1%
1999	Q1	0.4	140	20%
1999	Q2	0.43	143	2%
1999	Q3	0.42	142	-1%
1999	Q4	0.54	154	8%
2000	Q1	0.42	142	-8%
2000	Q2	0.46	146	3%
2000	Q3	0.36	136	-7%
2000	Q4	0.44	144	6%
2001	Q1	0.45	145	1%
2001	Q2	0.47	147	1%

Average price for second quarter of 2001 is 146% of price in first quarter of 1997. This tells us that price of small family houses have increased by 46% in real prices between first quarters of 1997 and second quarter of 2001. Maximum price change of 20% occurs in first quarter of 1999 and the second in fourth quarter of year 1997, which is just after the election. Yearly average

REPI starting from 1997 are 105.75, 117.5, 142 and 146(for two quarters of 2001). The following chart shows the Real Estate Price Index trend for the period graphically.



Graph Three: Real Estate Price Index

We made a test of seasonal effect for the Real Estate Price Index developed above.. The table below shows test result of seasonal effect which is done by including quarter dummies together with a trended variable and found that there is no seasonal effect on price (see also Appendix five).

Table 8: Test for seasonal effect

REPI	Coef.	Std.Err.	t
Q1	-1.316667	5.971414	-0.22
Q2	-2.25	5.957639	-0.38
Q3	-5.566667	6.292972	-0.88
T	2.933333	0.405366	7.24
_cons	103.9167	6.012549	17.28

We also test if the Real Estate Price Index is an autoregressive model of order one, AR (1). The current Real Estate Price Index is regressed by a preceding Real Estate Price Index and the result shows that it is an auto regressive model of order one. The previous period Real Estate Price Index has an effect on the current REPI.

Table 9: Price Index

Source	**SS**	**df**	**MS**	**Nr of obs**	**17**
				F(1, 15)	53.07
Model	3422.23584	1	3422.23584	Prob > F	0
Residual	967.293574	15	64.4862382	R-squared	0.7796
				Adj R^2	0.7649
Total	4389.52941	16	274.345588	Root MSE	8.0303

REPI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
REPI L1.	0.8279564	.1136543	7.28	0	0.58570	1.070205
_cons	24.87737	14.73719	1.69	0.112	-6.5342	56.28895

6. Conclusions

The primary purpose of this research is modeling a hedonic real estate price by identifying the real estate price determinants and their degree of significance for the new housing developments in the outskirt of Addis Ababa. It also tests the effects of holidays on housing price and finally develops a Quarterly Real Estate Price index for the period of five years. We try to figure out what the main determinants are in the housing price. The first model is without considering time trend but the second model takes time trend in to account. Model one is checked for Heteroskedasticity and Multicollinearity so that we can make statistical and economical interpretations. Serial correlation of variables is a problem mostly in time series data therefore we tried to fix the problem in model two. Appropriate functional form is used by making the Box-Cox transformation method for both sides of the model using the same variable. The second part of the thesis deals with constructing Real Estate Price Index for the first quarter of 1997 to second quarter of 2001.

Transaction data used in this study contains 213 transactions of newly developed small family-houses sold by real estate developers for five years (1997-2001) in Addis Ababa, *CMC* and *Alemgena* area. The price here may not actually be the market price. Since any developer can set its price based on the construction cost and its own profit margin, we cannot be sure to say that it is market price. Besides, the transaction is not made in auction or publicly. It is a negotiation between the developer and each buyer so there might be a price gap between the same houses depending on how tough the negotiator is to the developer. The date which the developer and customer signed an agreement is taken as the transaction date.

The findings in model one shows that location has a significant effect in housing price. From the empirical result, we can see that a 1% increase in plot area will increase the housing price by .55% which has the highest effect. A one percent upward change in number of bed rooms moves the housing price by 0.33% upward. A Plot found in *CMC* is more expensive than *Alemgena, nearly* by 0.1%. A one percent increase in plot area around *CMC* will cause a 0.55% while the same increase in *Alemgena* moves upward the housing price by 0.45%.

The effect of seasonal variation is tested in the second model using quarter dummies and found that the first three quarters of year 1997 don't have any significant effect but afterwards all the quarters are significant in determining housing price. The general election was held at the end of the third quarter in 1997. The increase in effect of time after the third quarter might be related to the election in such a way that people who were in need of buying house might have preferred to wait aftermath of the election before making decision.

One aim of this paper is to investigate the effects of holidays on housing price. Considering one month prior to holiday as a holiday period only Easter (*Fasika*) is statistically significant having a t-value of 2.25. If a specific house is sold during this period the result shows that there will be nearly a 10% increase in price. Besides, the model shows that there is no variation in price when holiday approaches for the other holidays. The average price in second quarter of 2001 is 146% of the price in first quarter of 1997. This tells us that price of small family houses have increased by 46% in real prices between first quarters of 1997 and second quarter of 2001.maximum price change of 20% occurs in first quarter of 1999 and the second in fourth quarter of year 1997, which is just after the election. Yearly average REPI starting from 1997 are 105.75, 117.5, 142 and 146(for two quarters of 2001). The REPI is an autoregressive model of order one, AR (1) without having no seasonal effect.

7. Reference

Andersson E.D.(2000),*Hypothesis testing in hedonic price estimation- On the selection of independent variables*, Ann Reg Sci, Vol.34:pp.293-304, The Annals of Regional Science: Springer-Verlag,The Institute of Economics, Academia Sinica, Taipei.

Bailey, M.J, Muth R.F and Nourse H.O. (1963), a Regression Method for Real Estate Price Construction, *Journal of the American Statistical Association Vol.*58, 933-942.

Baltagi B.H., Song S.H., Koh W. (2003), Testing Panel Data Regression Models with Spatial Error Correlation; *Journal of Econometrics*, 117, 123-150.

Branson, M. (2006), *The Australian Experience in Developing an Established House Price Index,* paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, November 6-7. Available on: http://www.oecd.org/dataoecd/2/26/37583386.pdf

Case, Pollakowski, and Wachter. (1997) frequency of transaction and house price modeling, *Journal of Real Estate Finance and Economics*, Vol. 14: pp. 173–187, Kluwer Academic Publishers.

Chau Et Al.(2005). Adjusting for Non-Linear Age Effects in the Repeat Sales Index, *the Journal of Real Estate Finance and Economics,* Vol. 3 No.2, pp.137–153, Springer Science + Business Media, The Netherlands.

Cheung U.S.L. Yau K.K.W and Hui Y.V. (2004). The effects of attributes on the repeat sales pattern of residential property in Hong Kong, *Journal of real estate finance and economics* Vol. 29, No. 3 pp.321-339, Kluwer Academic Publishers, The Netherlands

Chanel O, et al. (1996). The relevance of hedonic price indices: the case of paintings, *Journal of cultural economics* Vol.20, pp.1- 24, Kluwer Academic Publishers, The Netherlands

Clapham E. Eglund P. Quigley J. and Redfeam C. (2006). Revisiting the past and setting Score: Index revision for house price derivatives, *Journal of Real Estate Economics*, Vol. 34, No.2, pp. 275-302.

Clapp J.M. and Giaccotto C. (1998), price indices based on the hedonic repeat-sales method: application to the housing market, *Journal of Real Estate Finance and Economics,* vol.16:No.1, pp.5-26, Kluwer Academic Publishers, Boston

Clapp J.M. and Giaccotto C. (1992).Estimating Price Trends for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods, *Journal of Real Estate Finance and Economics,* Vol.5:pp.357-374, Kluwer Academic Publishers.

Cropper, M. L, Leland B. D, Kenneth E. (1988), On The Choice of Functional Form for Hedonic Price Functions, *Review of Economics and Statistics* 70,pp. 668–675.

Dombrow, Knight, and Sirmans.(1997),Aggregation bias in repeat-sales indices, *Journal of real estate finance and economics,* vol.14: pp.75–88, kluwer academic publishers.

Els M, Fintel D.V. (2008), *Residential Property Prices in a Sub-Market of South Africa:Separating Real Growth from Attribute Growth*; A 14/08 Working Paper Of The Department Of Economics and Bureau For Economic Research , University Of Stellenbosch File URL: http://www.ekon.sun.ac.za/wpapers/2008/wp142008/wp-14-2008.pdf

Follain, J.R, Calhoun, C.A.(1997), Constructing indices of the price of multifamily properties using the 1991 Residential Finance Survey, *Journal of Real Estate Finance and Economics* 14, 235–255.

Francke M.K., Vos G.A, Janssen JE.(2000), *Standardized price indices for the regional housing market;* a comparison between the fixed-sample index and the hedonic index, paper presented at the European Real Estate Society 48. Conference, Bordeaux.

Freeman, A.M., III. (1993). *The measurement of environmental and resource values. Theory and methods.* Resources for the future, Washington, DC.

Freeman A.M. (1979), *the hedonic price approach to measuring for neighborhood characteristics*. In: D. Segal, Editor, the Economics of Neighborhood, Academic Press, New York, NY (1979), pp. 191–218.

Gardini A. et al. (2003), Fundamentals and asset price dynamics, *Journal of statistical methods & applications* Vol.12: pp.211–226, Springer-Verlag. Department of Statistical Sciences, University of Bologna.

Goodman, A.C, (1978), Hedonic Prices, Price Indices and Housing Markets, *Journal of Urban Economics,* Vol.5, No.4, pp. 471-484.

Goodman, A.C, (1977). A comparison of block group and census tract data in a hedonic housing price model, *Land Economics*, Vol.53, pp.483-487.

Goodman A. C, Thibodeau, T. G, (2003), *Housing market segmentation and hedonic prediction accuracy,* Journal of Housing Economics, Vol.12, pp.181–201.

Halstead J., Bouvier R. A.(1997). On the Issue of Functional Form Choice in Hedonic Price functions: Further Evidence, *Journal of Environmental Management* Vol. 21, No. 5, pp. 759–765, Springer-Verlag, New York Inc.

Hansen, J. (2006). *Australian house prices: a comparison of hedonic and repeat-sales measures,* Economic Research Dept., Reserve Bank of Australia, Sydney. Downloaded from http://www.rba.gov.au/publicationsandresearch/rdp/rdp2006-03.html

Halverson R. Pollakowski, H. O. (1981), Choice of functional form for hedonic price equations, *Journal of. Urban Economics* Vol.10, pp. 37-49.

Henry (1995), *A* hedonic price model for private properties in Hong Kong, *Journal of Real Estate Finance and Economics,* Vol. 10 pp.37-48.

Hoffmann, J. Lorenz A. (2006), *Real Estate Price Indices for Germany: Past, Present and Future*, paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, November 6-7. http://www.oecd.org/dataoecd/31/20/37625451.pdf

Ioannidis C. Silver M.(1999). Estimating exact hedonic indexes: An application to UK television sets, *Journal of Economics,* Springer, vol. 69, No.1, pp. 71-94.

Jansen S.J.T. (2007), Developing a House Price Index for The Netherlands: A Practical Application of Weighted Repeat Sale*, Journal of Real Estate Finance Economics,* Vol.37:pp.163–186*,* Springer Science + Business Media.

John F. K., Quigley J. M. (1970), measuring the value of housing quality, *Journal of the American Statistical association*. Vol. 65, No. 330, pp. 532-548. Available at http://www.jstor.org/stable/2284565.

Kuo C.L.(1996), Serial Correlation and Seasonality in the Real Estate Market, *Journal of Real Estate Finance and Economics*. Vol. 12, No. 2 pp. 139-162, Kluwer Academic Publishers,

Leishman C.(2001*),* House building and product differentiation: An hedonic price approach, *Journal of Housing and the Built Environment* Vol.16: pp.131–152, Department of Building Engineering and Surveying, Heriot-Watt University, Riccarton, Kluwer Academic Publishers. The Netherlands..

Lennon H. T. Choy. Stephen W. K. Mak. Winky K. O. Ho (2007). Modeling Hong Kong real estate prices. *Journal of Housing and Built Environment* Vol.22 pp.359-368.

Limsombunchai, V. Gan, C. Lee, M.(2004). House Price Prediction: Hedonic Price Model vs. Artificial Neutral Network, *American Journal of Applied Sciences*, Vol. 1, No. 3, pp. 193-201, Asian Network for Scientific Information.

Linneman, P.(1980), Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market, *Journal of Urban Economics* Vol. 8, pp.47-68.

Malpezzi, S. (*2002), Hedonic Pricing Models: A Selective and Applied. Review,* University of Wisconsin; the Center for Urban Land Economics Research, Madison. Downloaded on Monday 19,0ct0ber 2009 from http://www.bus.wisc.edu/realestate/pdf/pdf/Hedonic%20Pricing%20Models%20Survey%20for%20Maclennan.pdf

Mark J.H. Goldberg M.A. (1986). A study of the impact of zoning on housing value overtime, *Journal of Urban Economics* Vol. 20, pp.257-273.

Meese, R.A, Wallace N. E(1997). The Construction of Residential Housing Price Indices: A Comparison of Repeat Sales, Hedonic Regression, and Hybrid Approaches. *Journal of Real Estate Finance and Economics,* Vol.14, No.1-2, pp. 51-73.

Milon, J.W, Gressel, J. and Mulkey D.(1984). Hedonic amenity valuation and functional form specification, *Land Economics*, Vol.60, No.4, pp.378-387.

Pakes. Ariel (2002). *A Reconsideration of hedonic price indexes with an application to PC's*. Harvard Institute of Economic Research Working Paper No. 1991. Available at SSRN: http://ssrn.com/abstract=373101 or DOI: 10.2139/ssrn.373101

Palmquist, R.B, (1982). *The effects of air pollutants on Residential Property Values*. Report to the office of air quality planning and standards. U.S. Environmental Protection Agency,

Parkhomenko. Alexander. Redkina. Anastasia. Maslivets, Olga(2007): *Estimating Hedonic Price Indexes for Personal Computers in Russia.*

Quigley, J.M.(1995). A simple hybrid model for estimating real estate price indexes. *Journal of housing economics.* Vol. 4, No. 1 pp 1-12.

Rosen S.(1974). Hedonic prices and Implicit markets:product differentiation in pure computation. The journal of political economy, Vol.82,No.1 pp.34-55 http://www.jstor.org/stable/1830899

Shimizu C, Nishimura K.G (2007). Pricing structure in Tokyo metropolitan land markets and its structural changes: pre-bubble, bubble, and post-bubble periods*, Journal of Real Estate Finance and Economics,* Vol.35, pp.476-495.

So, H.M., Tse, R.Y.C., Ganesan, S. (1997), Estimating the influence of transport on house price: evidence from Hong Kong, *Journal of Property Valuation & Investment*, Vol. 15 No.1, 1997. pp 40-47

Straszheim, M. (1974), Hedonic estimation of housing market prices: a further comment*, Review of Economic and Statistics.*

Tesfaye A. (2007). Problems and prospects of housing development in Ethiopia, *Journal of property management* Vol. 25, No. 1 pp.27-53, Emerald Group Publishing Limited, www.emeraldinsight.com/0263-7472.htm.

Wen Hai-zhen, JIA Sheng-hua. GUO Xiao-yu, (2004). Hedonic price analysis of urban housing: An empirical research on Hangzhou, China, *Journal of Zhejiang*.

Wilhelmsson, M. (2008). Construction and updating of Property Price Index series:The Case of the Segmented Market of Stockholm – *Journal of Property Management* vol. 27, No. 2 pp 119-137.

Wolverton, M.L. (1997), Empirical study of the relationship between residential lot price, size and view, *Journal of Property Valuation & Investment*, Vol. 15 No.1, pp.48-57.

Wooldridge, J.M. (2006), *Introductory Econometrics – A Modern Approach,* Thomson, South-Western, Mason, OH.

Quigly, J.(1995), *A simple hybrid model for estimating real estate price indexes*, *Journal of Housing Economics*, Vol.4, pp.1-12.

8. Appendix

8.1 Appendix One: Model One

A regression by stata software is made, taking transaction price in logarithmic form as dependent variable. The independent variables are the natural logarithms of master bed room area, plot area number of bath rooms, and number and form year two to year five. A combination of the natural logarithm of plot area and the dummy variable for location two (CMC) is used to test for location effect on transaction price.

reg LP LMBR LBD LBTH L2LLPA LPLA Yr2 Yr3 Yr4 Yr5						
Source	SS	df	MS			Number of obs = 213
=-------------	+-----------	-----------	--------			F(9, 203) = 487.14
Model	21.2593369	9 2.3	6214854			Prob > F = 0.0000
Residual	0.984350894	203 .00	4849019			R-squared = 0.9557
-------------+	------------	-----------	-------			Adj R-squared = 0.9538
Total	22.2436878	212 .10	4923056			Root MSE = .06963
--------------	------------	-----------	--------	----------	-----------------------	
LP	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
-------------+	------------	-----------	--------	----------	-----------------------	
LM BR	0.2528709	0.045531	5.55	0	.1630967	.3426451
LBD	0.3272102	0.0594805	5.5	0	.2099314	.4444889
LBTH	0.1137216	0.0242727	4.69	0	.0658627	.1615806
L2LLPA	-0.1047425	0.0069033	-15.17	0	0.1183539	0.0911311
LPLA	0.5473892	0.0247984	22.07	0	.4984937	.5962847
Yr2	0.1412901	0.0139119	10.16	0	.1138598	.1687205
Yr3	0.3868264	0.0137232	28.19	0	.3597681	.4138847
Yr4	0.3292022	0.0356523	9.23	0	.258906	.3994985
Yr5	0.2943004	0.0506515	5.81	0	.1944299	.3941709
_cons	10.02335	0.0944611	106.11	0	9.837104	10.20961
--------------	------------	-----------	--------	----------	-----------------------	

8.2 Appendix Two: Homoskedasticity test for model one

To Test for the presence of Heteroskedasticity in time series regression, we derive a residual and regress the square of the residuals on the dependent variables. We can see from the table that the value of *prob>F=0.0758,* which is greater than 0.05.

<table>
<tr><td colspan="7">reg e2 LMBR L#BDR L#BTHR L2LPLTA LPLTA Yr2 Yr3 Yr4 Yr5</td></tr>
<tr><td>Source</td><td>| SS</td><td>df</td><td>MS</td><td></td><td>Number of obs =</td><td>213</td></tr>
<tr><td>=------------</td><td>-+------------</td><td>-----------</td><td>-------</td><td></td><td>F(9, 203) =</td><td>1.77</td></tr>
<tr><td>Model</td><td>| .005288873</td><td>9 .000</td><td>587653</td><td></td><td>Prob > F =</td><td>0.0758</td></tr>
<tr><td>Residual</td><td>| .067402706</td><td>203 .000</td><td>332033</td><td></td><td>R-squared =</td><td>0.0728</td></tr>
<tr><td>-------------</td><td>+-------------</td><td>-----------</td><td>------</td><td></td><td>Adj R-squared =</td><td>0.0316</td></tr>
<tr><td>Total</td><td>| .072691579</td><td>212 .000</td><td>342885</td><td></td><td>Root MSE =</td><td>.01822</td></tr>
<tr><td colspan="7"></td></tr>
<tr><td>-------------</td><td>--------------</td><td>-----------</td><td>--------</td><td>---------</td><td colspan="2">-----------------------</td></tr>
<tr><td>e2</td><td>| Coef.</td><td>Std. Err.</td><td>t</td><td>P>|t|</td><td colspan="2">[95% Conf. Interval]</td></tr>
<tr><td>-------------</td><td>+-------------</td><td>-----------</td><td>--------</td><td>---------</td><td colspan="2">-----------------------</td></tr>
<tr><td>LMBR</td><td>| .0009826</td><td>0.0119144</td><td>0.08</td><td>0.934</td><td>-.0225091</td><td>.0244744</td></tr>
<tr><td>L#BDR</td><td>| -.0236891</td><td>0.0155646</td><td>-1.52</td><td>0.13</td><td>-.0543781</td><td>.007</td></tr>
<tr><td>L#BTHR</td><td>| -.0031485</td><td>0.0063516</td><td>-0.5</td><td>0.621</td><td>-.0156721</td><td>.009375</td></tr>
<tr><td>L2LPLTA</td><td>| .0058682</td><td>0.0018064</td><td>3.25</td><td>0.001</td><td>.0023065</td><td>.00943</td></tr>
<tr><td>LPLTA</td><td>| -.0060517</td><td>0.0064891</td><td>-0.93</td><td>0.352</td><td>-.0188465</td><td>.0067431</td></tr>
<tr><td>Yr2</td><td>| -.0032315</td><td>0.0036404</td><td>-0.89</td><td>0.376</td><td>-.0104094</td><td>.0039464</td></tr>
<tr><td>Yr3</td><td>| .0057319</td><td>0.003591</td><td>1.6</td><td>0.112</td><td>-.0013486</td><td>.0128124</td></tr>
<tr><td>Yr4</td><td>| .0245828</td><td>0.0093293</td><td>2.64</td><td>0.009</td><td>.006188</td><td>.0429777</td></tr>
<tr><td>Yr5</td><td>| .0241897</td><td>0.0132543</td><td>1.83</td><td>0.069</td><td>-.001944</td><td>.0503234</td></tr>
<tr><td>_cons</td><td>| .0372622</td><td>0.0247182</td><td>1.51</td><td>0.133</td><td>-.0114751</td><td>.0859995</td></tr>
<tr><td>-------------</td><td>--------------</td><td>-----------</td><td>--------</td><td>---------</td><td colspan="2">-----------------------</td></tr>
</table>

8.3 Appendix Three: Model two using time series data.

reg LPRCE LPLTA LTFLA LLVGA LDNGA L#BDR L#BTHR YQ2- YQ18

Source	| SS	df	MS		Number of obs	213
=------------	-+-----------	-----------	--------		F(23, 189	) = 315.82
Model	| 21.6796055	23 .94	2591542		Prob > F	0
Residual	| .564082327	189 .00	2984563		R-squared	0.9746
------------	+-----------	-----------	-------		Adj R-squared	0.9716
Total	| 22.2436878	212 .10	4923056		Root MSE	0.05463
------------	------------	-----------	---------	-------	---------------	----------
LPRCE	| Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
------------	+-----------	-----------	---------	-------	---------------	----------
LPLTA	| .3931341	0.0354679	11.08	0	0.3231702	0.463098
LTFLA	| .5452952	0.0496894	10.97	0	0.4472781	0.6433124
LLVGA	| 1.508192	0.1044281	14.44	0	1.302197	1.714186
LDNGA	| -1.167362	0.150776	-7.74	0	-1.464782	-0.8699417
L#BDR	| -1.402239	0.0971399	-14.44	0	-1.593856	-1.210621
L#BTHR	| .0414278	0.0256756	1.61	0.108	-0.0092198	0.0920754
YQ2	| .0186249	0.0121246	1.54	0.126	-0.005292	0.0425418
YQ3	| .0299711	0.0207786	1.44	0.151	-0.0110166	0.0709589
YQ4	| .1842623	0.0165289	11.15	0	0.1516575	0.2168671
YQ5	| .1830463	0.0224056	8.17	0	0.1388491	0.2272436
YQ6	| .1743896	0.0169083	10.31	0	0.1410363	0.2077429
YQ7	| .1770784	0.0189317	9.35	0	0.139734	0.2144229
YQ8	| .1675801	0.0286212	5.86	0	0.1111221	0.2240381
YQ9	| .3984206	0.0172556	23.09	0	0.3643822	0.432459
YQ10	| .4344026	0.0184703	23.52	0	0.3979681	0.4708371
YQ11	| .416951	0.0222289	18.76	0	0.3731025	0.4607996
YQ12	| .5398375	0.0284711	18.96	0	0.4836755	0.5959995
YQ13	| .4245221	0.0352374	12.05	0	0.355013	0.4940311
YQ14	| .4639865	0.0302793	15.32	0	0.4042577	0.5237153
YQ15	| .3583617	0.0354014	10.12	0	0.2885291	0.4281944
YQ16	| .4365334	0.0367749	11.87	0	0.3639915	0.5090753
YQ17	| .4515752	0.0436747	10.34	0	0.3654226	0.5377277
YQ18	| .4723604	0.0502068	9.41	0	0.3733228	0.571398
_cons	| 8.756692	0.1777244	49.27	0	8.406113	9.10727
------------	------------	-----------	---------	------	---------------	----------

	-	

8.4 Appendix Four: Test for Serial Correlation

reg e l1.e

Source	| SS	df	MS		Number of ob	s = 92
=------------	-+-----------	------------	-------		F(1, 9	0) = 2.0
Model	| .002618142	1 .002	618142		Prob > F	0.1592
Residual	| .116938341	90 .001	299315		R-squared	0.0219
-------------	+------------	------------	------		Adj R-square	d = 0.0110
Total	| .119556484	91 .001	313808		Root MSE	0.03605

-------------	-------------	------------	--------	---------	------------	-----------
e	| Coef.	Std. Err.	t	P>|t|	[95% Conf	. Interval]
-------------	+------------	------------	--------	---------	------------	-----------
e	|					
L1.	| .1429745	0.1007208	1.42	0.159	-0.057125	0.343074
_cons	| .0020279	0.0037671	0.54	0.592	-0.0054561	0.0095119
-------------	-------------	------------	--------	---------	------------	-----------

8.5 Appendix Five: Seasonal effect

reg REPI Q1 Q2 Q3 T

Source	SS	df	MS	Number of obs	18
				F(4, 13)	13.59
Model	4289.07778	4	1072.26944	Prob > F	0.0001
Residual	1025.36667	13	78.874359	R-squared	0.8071
				Adj R^2	0.7477
Total	5314.44444	17	312.614379	Root MSE	8.8811

REPI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Q1	-1.316667	5.971414	-0.22	0.829	-14.21712	11.58379
Q2	-2.25	5.957639	-0.38	0.712	-15.1207	10.6207
Q3	-5.566667	6.292972	-0.88	0.392	-19.16181	8.028474
T	2.933333	.405366	7.24	0	2.057593	3.809073
_cons	103.9167	6.012549	17.28	0	90.92734	116.906