Report on Hedonic Pricing Model and Price Index for Tsuen Wan Centre

Prepared by HKU Bank

NOV 2015

Chan Lok Yiu
Kong Ling Sang
Ko Wing Yan
Tang Chi Fung
Tom Tsz Hin
So Kai Yin
Sun Lai Ling

EXECUTIVE SUMMARY

A hedonic pricing model for Tsuen Wan Centre is designed in this project. The pricing trend and historical pricing level can also be known from the price index set up. As commissioned by the HKU Bank, the pricing model and index will assist the bank and the public in making decisions upon the purchase of flats.

In this study, different factors are chosen to be the attributes for the hedonic price model. Common ones include building age, Gross Floor Area (GFA) and floor level. Additional features like the provision of roof and balcony are also comprised in the models. The effect of orientation and the provision of green view are also considered as they are usually posing uncertain effects on the property prices.

Altogether three models are proposed in the study and the one with the highest R^2 was finally chosen for building up the price index. Detailed explanations on selecting the best-fit model and the actual effects of each attribute will also be discussed.

As transaction data can be obtained from the EPRC website, utilizing the information over there and constructing a pricing model can help the public in knowing and evaluating the prices of flats with desired elements and attributes.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1

OVERVIEW OF THE STUDY

1.1 Introduction

Tsuen Wan Centre, which is located at Tsuen King Circuit in Tsuen Wan, is the largest private housing estates in the district. It includes 19 buildings which are all developed by Sun Hung Kai Properties between year 1979 and 1982.

Figure 1 Overviews of Tsuen Wan Centre

There are two parts namely Tsuen Wan Centre I and II, altogether providing 4400 flats in the estate. The estate is surrounded by mountains and located far away from the district centre. From 2003 to 2007, Tsuen Wan Centre has undergone renovations, including the building façade, lobbies and pipework.

There are different facilities and amenities provided, including shopping mall, private podium garden, car park, estate office and minibus terminal, allowing residents to enjoy different services within the region.

1.1 Introduction (Cont’d)

Although there are not many schools and institutions near Tsuen Wan Centre, one can easily travel to the nearby core areas by buses and minibuses.

Figure 2 Accessibility of Tsuen Wan Centre. Icons in red: schools and institution

1.2 Objectives

1. To build up a hedonic price model for residential units in Tsuen Wan Centre

2. To build up a price index for residential units in Tsuen Wan Centre

3. To study the attributes affecting the price of Tsuen Wan Centre

CHAPTER 2

METHODOLOGY

2.1 Hedonic Price Model

A hedonic price model recognizes both complexity and heterogeneity of a housing product explicitly. A relationship between the property price and housing attributes is postulated (Mok, Chan & Cho, 1995).

The price of a property is a function if its own physical characteristics. The basic Hedonic Price Model of property market price is P=f (L, S, N), where L, S and N are locational, structural and neighborhood attributes respectively. Locational attributes refer to the physical distance from surrounding infrastructure and vital spots, such as social and civic venters, MTR stations, etc.; structural attributes refer to variables like the age of building and square footage; neighborhood traits refer to the overall neighborhood quality (Mok, Chan and Cho,1995).

Implicit prices of each characteristic is given by the regression estimates. The models are estimated as single-stage equations, which estimates the effect of the characteristics only, without examining the structural parameters of them. They are also estimated a variety of ways with regards to the property price.   Variation in characteristic price across a number of price ranges is allowed (G., Macpherso & Zietz, 2005).

2.2 Ordinary least square method

As a method for estimation of unknown parameters in linear regression models, ordinary least square (OLS) is used to minimize the differences between observed responses in a dataset and that predicted by the linear approximation, i.e. the sum of vertical distances between each point of data in the dataset and the corresponding point of the regression line.

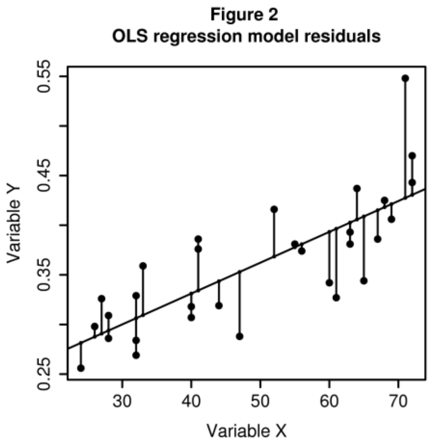

Figure 3
OLS regression model showing the difference between data points and regression line

The smaller the differences are, the better the regression model is. An estimator will be resulted and expressed by a simple formula.

OLS regression, which relies on the assumption of the absence of error in the measurement of explanatory variables, assumes that errors are all confined to the property price, i.e. the dependent variable (Leng, Zhang, Kleinman & Zhu, 2007).

2.3 Design of the Research

Source of data -- EPRC database

EPRC is an online real estate transaction database set up in 1991. (EPRC, 2015) Since then it has been collecting data not limited to residential unit transactions but also all sorts of real estate trades, through the Land Records by the Land Registry and other means. It takes up around 95% of overall market share and is widely used by different parties such as major property developers, banks, surveying firms and security company. (EPRC, 2015)

For our project, it provided us with data like instrument date, completion date, block number, floor, price, gross floor area (GFA), saleable floor area (SFA) and price per GFA and per SFA. Yet, the file we got from Moodle included incomplete set of data as well as information not needed, which meant that we had to clean it before in use. The cleaning process was strictly following the example illustrated in the Moodle. Also, we included other variables on our own, namely number of bedrooms, views, RVD price index, orientation, external features like roof and balcony. The descriptions and justifications regarding the variables of choice in use will be discussed later.

Sample

Transaction records of residential unit flats in Tsuen Wan Centre from April 1991 to August 2015 are taken as data of our sample.

Descriptions and justifications of choices of variables

The basis of the pricing method is that the price of a marketed property closely relates to its characteristics. The variations in housing prices are affected by a variety of attributes and by that we can value the individual characteristics of the property by simply seeing how the market price of the property will change when its characteristics change. A number of property characteristics are identified as variables and listed in the following section.

2.4 Choice of variables

1. Dependent Variable - Nominal Price (NP)

All transaction prices retrieved from the EPRC database are done at different time. That means, the prices cannot be directly comparable because they are not real prices but nominal prices. Nominal price cannot reflect the current situation since it is affected by the atmosphere of the market in accordance with time. Therefore, time dummies is added, with a view to including the overall changes from time to time.

2. Independent Variable - Building Age (Age)

Older age of buildings results in more dilapidation, which requires more maintenance. It is reasonable to predict that more cost have to be paid for maintenance of building if the building is older in the perspective of purchasers. In that sense, normally people would be eager to pay more for a newer building thanks to less maintenance fee assumed. That is also why this variable should be taken into account, and the age of building is reflected by the completion date of it.

3. Independent Variable -Transaction date

Transaction date is a variable because of mainly two reasons. First, the difference between the completion date and transaction date is the age of building at the time the transaction is done. As discussed beforehand, the age of building affects the price the purchasers are willing to pay. Second, as during different times there are different market conditions, including transaction date can allow us to refer to that respective market conditions at the specific time when the transaction was done.

4. Independent Variable -Floor level (FL)

Floor level should also be a variable because it is linked with a number of factors contributing to the experience of occupants in the flat, for example, the view, noise from the street, air pollution problem, etc. Flats in higher level may suggest a better outside environment since there is nothing near outside but air. The view is more “open” as well. On the other hand, for low level case, they are nearer to ground floor so that occupants may have a sense of oppression. In additions, as shown in the map, there are roads surrounding the properties. Flats in low levels may be facing roads and/or rubbish chamber, which could possibly give rise to noise problem and air pollution, causing disturbance to occupants. Normally, buyers tend to purchase flats in higher level.

2.4 Choice of variables (Cont'd)

5. Independent Variable -Gross floor area (GFA)

Size of a flat is a major physical attribute that determines the flat size. It should be a variable since it determines the extent of to which the flat can be utilized. Larger space is always preferred in normal cases. In the view of occupants, larger size leads to more usable space, which increases the flexibility of use. Comfortability of living increases with flat size. Besides, larger flats can accommodate more family members so large families prefer larger flats.

6. Independent Variable -Views (Green view denoted as Green and Open View denoted as Oview)

Views of the flats are also a vital determining variable. The occupants' comfortableness are influenced by what view they can see from their apartments. Sea views, greenery views, open views, street views, etc. should be identified and ranked in our model. For example, greenery views may give a feeling of refreshing while views blocked by buildings may give a compact feeling. In the case of TW CTR, most units face towards greenery mountain views and open city views.

7. Independent Variable -Roof (Roof)

Roof level affects the price of units on the top levels. The inclusion of roof level may affect the price positively or inversely. On the bright side of having a roof, there would be broader view outside and more usable area to the occupants. Nevertheless, there are cases that people found water leakage or overheating problem at the unit below the roof floor. Such problems would cause annoyance to occupants and require extra sum of money from occupants in order to have them solved. Since some of these problems still persist nowadays, there may be chances that the unit down below the roof floor for older buildings may be of lower selling price whilst the one of newer buildings may be at a higher price because those above-mentioned problems are less likely to appear in newer buildings.

2.4 Choice of variables (Cont'd)

8. Independent Variable -Orientation (NE, SE, NW, SW)

Orientation of flats is also an important variable. The importance of the orientation of the flat is associated with the climate in Hong Kong. Flats with openings facing south are most popular while those with openings facing north are least popular in Hong Kong. Winds mostly come from the South and North in summer and winter respectively. In summer wind is welcomed since it can be made use of to regulate the indoor temperature and ventilation. On the contrary, in winter wind is always unwelcomed since it lowers the indoor temperature. Besides that, flats with openings facing the West are also not desirable since indoor temperature is expected to greatly increase during sunset. There is no specific disadvantage for flat located on east. In short, flats with openings facing South or East is more popular in the market, which results in higher price.

9. Independent Variable – Balcony

Balcony is often seen as a green feature in residential buildings and is thus favourable by the building occupants. This was proved in some researches that balcony can function as an "environmental filter" in reducing noise level from outside, providing extra leisure space and enhancing energy efficiency. As a result, the balcony is able to have favourable effect towards the property price regardless of the view outside.

2.5 Data collection and method

Most of the data such as flat size, GFA, SFA, block and price information are from the excel file given in the course. While additional items including orientation, view, number of rooms in each flat, provisions of balcony and roof are found from maps and floor plan on the internet and property agencies such as Centaldata (2015) and GoHome.com.hk (2015).

2.6 Equation and expected result of the Models

2.6.1 Functional Form 1

$$\ln(NP)=\alpha_0 + \alpha_1 GFA + \alpha_2 AGE + \alpha_3 FL + b_0 D_t + b_1 D_{t+1} + \cdots + b_n D_{t+n} + \varepsilon$$

Where-

NP is the nominal price

GFA is the gross floor age;

AGE is the building age;

FL is the floor level;

$\boldsymbol{\alpha_0}$ is a constant term;

$\boldsymbol{\alpha_i}$ is the coefficient of the i^{th} attritube, where i = 1,2,3…,11;

D_t is a time dummy variable which is equal to 1 at the time t. Otherwise, it is 0;

$\boldsymbol{b_0}$ is the coefficient of the responding time dummies D_t;

$\boldsymbol{\varepsilon}$ is the error term.

Expected sign of coefficients

Variables	Expected sign of coefficients
GFA	+
FL	+
AGE	-

Table 1 Expected sign of coefficients of Model 1

2.6.2 Functional Form 2

$$\ln(\text{NP}) = \alpha_0 + \alpha_1 GFA + \alpha_2 AGE + \alpha_3 FL + \alpha_4 FL^{1.1} + \alpha_5 BAL + \alpha_6 GREEN + \alpha_7 OVIEW + \alpha_8 ROOF + \alpha_9 SW + \alpha_{10} NE + \alpha_{11} SE + b_0 D_t + b_1 D_{t+1} + \cdots + b_n D_{t+n} + \varepsilon$$

Where-

NP is the nominal price

GFA is the gross floor age;

AGE is the building age;

FL is the floor level;

BAL is a dummy variable which is equal to when there is a balcony. Otherwise, it is 0.

GREEN is a dummy variable which is equal to 1 when there is a green view. Otherwise, it is 0;

OVIEW is a dummy variable which is equal to 1 when there is an open view. Otherwise, it is 0;

ROOF is a dummy variable which is equal to 1 when there is a roof. Otherwise, it is 0.

SW is a dummy variable which is equal to 1 when it is facing south west. Otherwise, it is 0;

NE is a dummy variable which is equal to 1 when it is facing north east. Otherwise, it is 0;

SE is a dummy variable which is equal to 1 when it is facing north west. Otherwise, it is 0;

α_0 is a constant term;

α_i is the coefficient of the i^{th} attritube, where i = 1,2,3…,11;

D_t is a time dummy variable which is equal to 1 at the time t. Otherwise, it is 0;

b_0 is the coefficient of the responding time dummies D_t;

ε is the error term.

2.6.2 Functional Form 2 (Cont'd)

Expected sign of coefficients

Variables	Expected sign of coefficients
GFA	+
FL	+
$FL^{1.1}$	+
AGE	-
ROOF	+
OVIEW	+
GREEN	+
BAL	+
SW	Unknown
NE	Unknown
SE	Unknown

Table 2 Expected sign of coefficients of Model 2

2.6.3 Functional Form 3

$$\ln(NP_t) = \alpha_0 + \alpha_1 GFA + \alpha_2 AGE + \alpha_3 FL + \alpha_4 FL^{1.1} + \alpha_5 BAL + \alpha_6 \text{OVIEW} * GREEN + \alpha_7 OVIEW * GREEN * BAL + \alpha_8 ROOF - \alpha_9 NE + \alpha_{10} \text{SE} + b_t D_t + b_{t+1} D_{t+1} + \cdots + b_{t+n} D_{t+n} + \varepsilon$$

Where-

NP is the nominal price

GFA is the gross floor age;

AGE is the building age;

FL is the floor level;

BAL is a dummy variable which is equal to when there is a balcony. Otherwise, it is 0.

GREEN* OVIEW is a dummy variable which is equal to 1 when there is a green view and open view. Otherwise, it is 0;

GREEN* OVIEW *BAL is a dummy variable which is equal to 1 when there is an open view, green view and balcony. Otherwise, it is 0;

ROOF is a dummy variable which is equal to 1 when there is a roof. Otherwise, it is 0.

2.6.3 Functional Form 3 (Cont'd)

NE is a dummy variable which is equal to 1 when it is facing north east. Otherwise, it is 0;

SE is a dummy variable which is equal to 1 when it is facing north west. Otherwise, it is 0;

$\boldsymbol{\alpha_0}$ is a constant term;

$\boldsymbol{\alpha_i}$ is the coefficient of the i^{th} attritube, where i = 1,2,3…,11;

D_t is a time dummy variable which is equal to 1 at the time t. Otherwise, it is 0;

$\boldsymbol{b_0}$ is the coefficient of the responding time dummies D_t;

$\boldsymbol{\varepsilon}$ is the error term.

<u>Expected sign of coefficients</u>

Variables	Expected sign of coefficients
GFA	+
FL	+
$FL^{1.1}$	+
AGE	-
ROOF	+
BAL	+
OVIEW* GREEN	+
OVIEW*GREEN*BAL	+
NE	Unknown
SE	Unknown

Table 3 Expected sign of coefficients of Model 3

CHAPTER 3

EMPIRICAL RESULTS AND INTERPRETATION

3.1 Regression Result

Model 1

Dependent Variable: ln(NP)
Method: Least Squares
Included observations: 6888

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GFA	0.001780	1.77E-05	100.5642	0.0000
AGE	-0.002287	7.67E-05	-29.83087	0.0000
FL	0.003526	9.66E-05	36.47937	0.0000
C	-0.335437	0.081579	-4.111826	0.0000
YR=1991,MO=4	0.090553	0.087625	1.033408	0.3015
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
YR=2015,MO=8	1.764463	0.080485	21.92297	0.0000

R-squared	0.964760	Mean dependent var	0.234329
Adjusted R-squared	0.963178	S.D. dependent var	0.395343
S.E. of regression	0.075863	Akaike info criterion	-2.277615
Sum squared resid	37.93246	Schwarz criterion	-1.982791
Log likelihood	8141.105	Hannan-Quinn criter.	-2.175941
F-statistic	609.6007	Durbin-Watson stat	1.156199
Prob(F-statistic)	0.000000		

Table 4 Regression result of Model 1
Notes: *represent significant at 5% level.

3.1 Regression Result (Cont'd)

Model 1

Based on the regression result, the equation should be rewritten as below with coefficient substituted.

$$\ln(NP_t) = -0.33544 + 0.00178GFA - 0.00229AGE + 0.003526FL + b_tD_t + b_{t+1}D_{t+1} + \cdots + b_{t+n}D_{t+n} + \varepsilon$$

Therefore,

$$\boldsymbol{NP_t} = \mathbf{E}\ (-0.33544 + 0.00178GFA - 0.00229AGE + 0.003526FL + \varepsilon) * e^{b_t}$$

Where e^{b_t} is the time dummy at time t.

According to the p value of the variables, all of them are statically significant at 1% level. This can show that the variables are explanatory variables to the nominal price. Furthermore, the adjusted R^2 is 0.96476. This means that the regression model can fit most of the actual data. Therefore, the regression model is satisfactory.

Comparison of expected sign and actual sign of coefficients.

Variables	Expected sign of coefficients	Actual sign of coefficients
GFA	+	+
FL	+	+
AGE	-	-

Table 5 Comparison of expected sign and actual sign of coefficients of Model 1

3.1 Regression Result (Cont'd)

<u>Model 2</u>

Dependent Variable: ln(NP)
Method: Least Squares
Included observations: 6888

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.700619	0.081281	-8.619763	0.0000
GFA	0.001735	1.90E-05	91.38488	0.0000
AGE	-0.001387	9.63E-05	-14.40372	0.0000
FL	0.027220	0.001557	17.48647	0.0000
OVIEW	0.017692	0.002635	6.714413	0.0000
GREEN	0.019248	0.003795	5.071733	0.0000
BAL	0.026837	0.003226	8.320075	0.0000
FL^1.1	-0.016160	0.001058	-15.26809	0.0000
NE	-0.021227	0.002855	-7.435162	0.0000
SW	-0.012806	0.002952	-4.338644	0.0000
SE	-0.005807	0.002512	-2.311483	*0.0208
ROOF	0.019425	0.008391	2.314961	*0.0206
YR=1991,MO=4	0.053968	0.083030	0.649982	0.5157
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
YR=2015,MO=8	1.734259	0.076251	22.74399	0.0000

R-squared	0.968432	Mean dependent var	0.234329
Adjusted R-squared	0.966974	S.D. dependent var	0.395343
S.E. of regression	0.071846	Akaike info criterion	-2.385310
Sum squared resid	33.98057	Schwarz criterion	-2.082545
Log likelihood	8520.008	Hannan-Quinn criter.	-2.280897
F-statistic	664.3027	Durbin-Watson stat	1.277309
Prob(F-statistic)	0.000000		

Table 6 Regression result of Model 2

Notes: *represent significant at 5% level.

3.1 Regression Result (Cont'd)

Model 2

Based on the regression result, the equation should be rewritten as below with coefficient substituted.

$$\ln(NP_t) = -0.700619 + 0.001735GFA \pm 0.001387AGE + 0.027220FL - 0.016160FL^{1.1} + 0.026837BAL + 0.019248GREEN + 0.017692OVIEW + 0.019425ROOF - 0.012806SW \pm 0.021227NE \pm 0.005807\text{SE} + b_tD_t + b_{t+1}D_{t+1} + \cdots + b_{t+n}D_{t+n} + \varepsilon$$

Therefore,

$$\boldsymbol{NP_t} = \mathbf{E}\ (-0.700619 + 0.001735GFA - 0.001387AGE + 0.027220FL - 0.016160FL^{1.1} + 0.026837BAL + 0.019248GREEN + 0.017692OVIEW + 0.019425ROOF - 0.012806SW - 0.021227NE - 0.005807\text{SE} + \varepsilon) * e^{b_t}$$

Where e^{b_t} is the time dummy at time t.

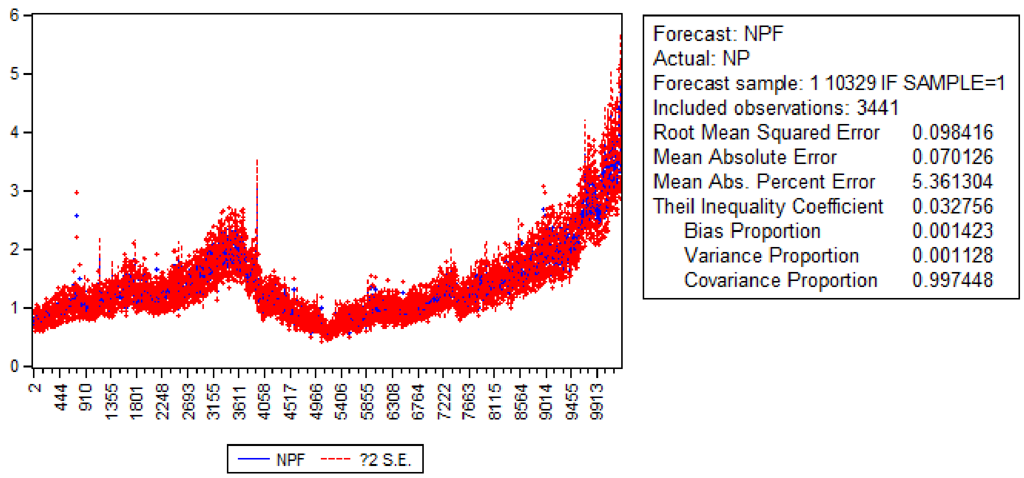

Figure 4 The out of sample forecast of Model 2

3.1 Regression Result (Cont'd)

Model 2

The out of sample forecast shows that the regression model attain a 5.36% mean absolute percentage error.

Variables	Expected sign of coefficients	Actual sign of coefficients
GFA	+	+
FL	+	+
$FL^{1.1}$	+	(-)
AGE	-	-
ROOF	+	+
OVIEW	+	+
GREEN	+	+
BAL	+	+
SW	Unknown	-
NE	Unknown	-
SE	Unknown	-

Table 7 Comparison of expected sign and actual sign of coefficients of Model 2

3.1 Regression Result (Cont'd)

Model 3

Dependent Variable: ln(NP)
Method: Least Squares
Included observations: 6886

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GFA	0.001773	1.84E-05	96.14695	0.0000
AGE	-0.001279	9.36E-05	-13.65790	0.0000
FL	0.050563	0.003316	15.24963	0.0000
C	-0.803355	0.080394	-9.992749	0.0000
FL^1.1	-0.032350	0.002291	-14.11993	0.0000
NE	-0.010382	0.002805	-3.701637	0.0002
SE	-0.001603	0.002255	-0.710709	0.4773
OVIEW*GREEN*BAL	-0.025087	0.005594	-4.484824	0.0000
ROOF	0.031342	0.008487	3.693105	0.0002
BAL	0.033880	0.003935	8.609794	0.0000
OVIEW*GREEN	0.048590	0.003038	15.99249	0.0000
YR=1991,MO=4	0.052833	0.082931	0.637075	0.5241
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
YR=2015,MO=8	1.739932	0.076159	22.84616	0.0000

R-squared	0.968508	Mean dependent var	0.234392
Adjusted R-squared	0.967058	S.D. dependent var	0.395369
S.E. of regression	0.071759	Akaike info criterion	-2.387853
Sum squared resid	33.89341	Schwarz criterion	-2.086005
Log likelihood	8525.378	Hannan-Quinn criter.	-2.283755
F-statistic	668.0562	Durbin-Watson stat	1.281566
Prob(F-statistic)	0.000000		

Table 8 Regression result of Model 3

3.1 Regression Result (Cont'd)

Model 3

Heteroskedasticity Test: White

F-statistic	37.07741	Prob. F(304,6581)	0.0000
Obs*R-squared	4347.606	Prob. Chi-Square(304)	0.0000
Scaled explained SS	154527.2	Prob. Chi-Square(304)	0.0000

Test Equation:
Dependent Variable: WGT_RESID^2
Method: Least Squares
Date: 11/19/15 Time: 16:00
Sample: 1 10329 IF SAMPLE=0 AND FL<36
Included observations: 6886

Table 9 Heteroskedasticity Test of Model 3

After conducting the Heteroscedasticity Test, a probability of 0.0000 is observed. This means that the hypothesis of no heteroscedasticity is rejected. In other words, heteroscedasticity is observed

Model 3 after heteroscedasticity correction

Dependent Variable: ln(NP)
Method: Least Squares
Included observations: 6886
Weighting series: FL^-2
Weight type: Inverse standard deviation (EViews default scaling)
White heteroskedasticity-consistent standard errors & covariance

3.1 Regression Result (Cont'd)

Model 3 after heteroscedasticity correction

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GFA	0.001538	0.000111	13.83955	0.0000
AGE	-0.001134	0.000453	-2.503612	*0.0123
FL	0.206874	0.033337	6.205612	0.0000
FL^1.1	-0.145141	0.025003	-5.804985	0.0000
NE	-0.032443	0.015525	-2.089731	*0.0367
SE	0.037525	0.010844	3.460592	0.0005
OVIEW*GREEN*BAL	0.042142	0.023881	1.764633	**0.0777
ROOF	0.232447	0.078385	2.965459	0.0030
BAL	0.031151	0.016254	1.916502	**0.0553
OVIEW*GREEN	0.028529	0.016007	1.782317	**0.0747
C	-0.885912	0.171000	-5.180763	0.0000
YR=1991,MO=4	0.034793	0.030931	1.124868	0.2607
YR=1991,MO=5	0.127123	0.019272	6.596130	0.0000
YR=2015,MO=7	1.721961	0.024678	69.77634	0.0000
YR=2015,MO=8	1.760056	0.024707	71.23638	0.0000

Weighted Statistics

R-squared	0.991716	Mean dependent var	0.168243
Adjusted R-squared	0.991335	S.D. dependent var	1.515803
S.E. of regression	0.134281	Akaike info criterion	-1.134618
Sum squared resid	118.6830	Schwarz criterion	-0.832770
Log likelihood	4210.488	Hannan-Quinn criter.	-1.030520
F-statistic	2600.550	Durbin-Watson stat	1.842062
Prob(F-statistic)	0.000000	Weighted mean dep.	0.132000

Unweighted Statistics

R-squared	0.892094	Mean dependent var	0.234392
Adjusted R-squared	0.887126	S.D. dependent var	0.395369
S.E. of regression	0.132831	Sum squared resid	116.1332
Durbin-Watson stat	0.933246		

Table 10 Regression result of Model 3 after heteroscedasticity correction

Notes: *represent significant at 5% level and **represent significant at 10% level

3.1 Regression Result (Cont'd)

Model 3

According to the Durbin-Waston Stat, the value is 1.842062, which is smaller than 2, it represents that there may be a positive first order autocorrelation.

Breusch-Godfrey Serial Correlation LM Test:

Obs*R-squared	0.000000	Prob. Chi-Square(2)1.0000

Table 11 LM test of Model 3 after correction

From the result of LM test, the probability of 1.0000 is observed and it rejected that there is an autocorrelation.

Based on the regression result, the equation should be rewritten as below with coefficient substituted.

$$\begin{aligned}\ln(NP_t) = &- 0.885912 + 0.001538GFA - 0.001134AGE + 0.0206874FL \\ &- 0.145141FL^{1.1} + 0.0.031151BAL + 0.028529\text{OVIEW} * GREEN \\ &+ 0.042142OVIEW * GREEN * BAL + 0.232447ROOF \\ &- 0.032443NE + 0.037525S\text{E} + b_t D_t + b_{t+1} D_{t+1} + \cdots + b_{t+n} D_{t+n} \\ &+ \varepsilon\end{aligned}$$

Therefore,

$$\boldsymbol{NP_t} = \mathbf{E}\,(- 0.885912 + 0.001538GFA - 0.001134AGE + 0.0206874FL - 0.145141FL^{1.1} + 0.0.031151BAL + 0.028529\text{OVIEW} * GREEN + 0.042142OVIEW * GREEN * BAL + 0.232447ROOF - 0.032443NE + 0.037525S\text{E} + \varepsilon) * e^{b_t}$$

Where e^{b_t} is the time dummy at time t.

3.1 Regression Result (Cont'd)

<u>Model 3</u>

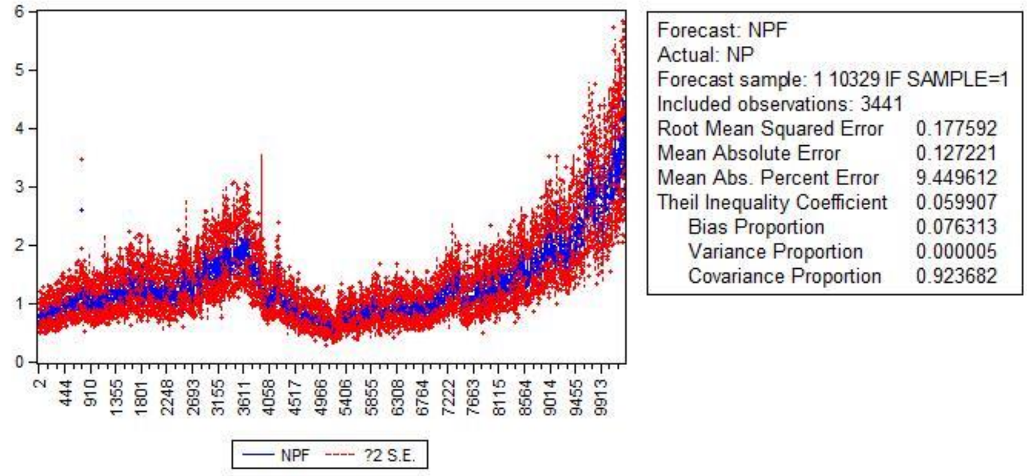

Figure 5 The out of sample forecast of Model 3

The out of sample forecast shows that the regression model attain a 9.44% mean absolute percentage error.

According to the p value of the variables, most of them are statically significant at 1% level and all of them are significant at 10% significant level. This can show that the variables are explanatory variables to the nominal price. Furthermore, the adjusted R^2 is 0.991716. This means that the regression model can fit most of the actual data. Therefore, the regression model is good. Since model 3 has a greater R^2, model 3 is adopted.

3.1 Regression Result (Cont'd)

Model 3

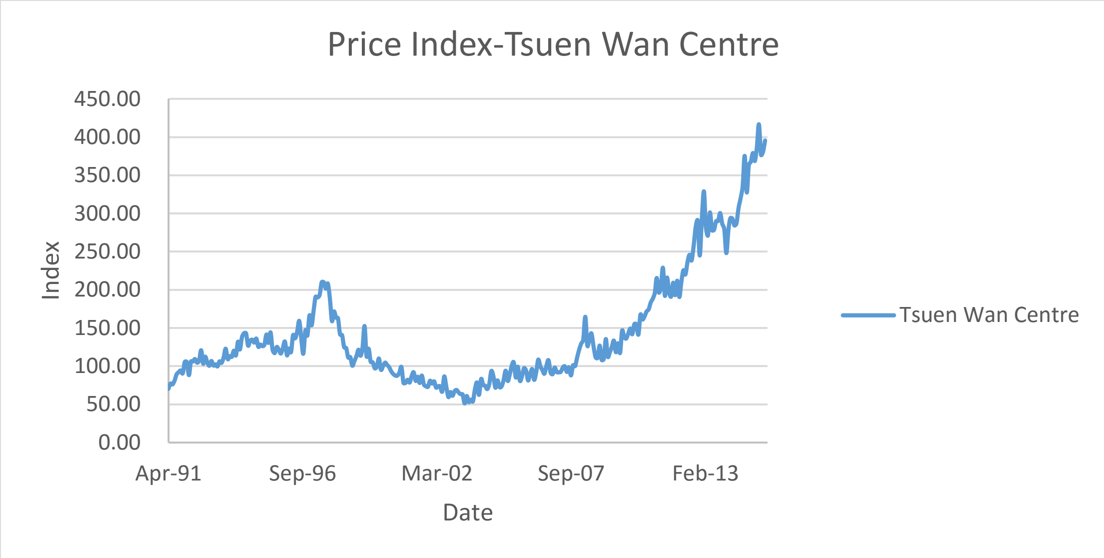

Figure 6 Price Index of Tsuen Wan Centre

Generally, the Price Index of Tsuen Wan Centre have increased from around 70 units to nearly 400 units. The price index increased gradually from around 70 units to around 200 units in 1997. A peak of about 200 units was reached in 1997. After that, it fell gradually from around 200 units in 1997 to nearly 50 units in 2003. In 2003, there was a trough in price index, which was also the lowest point from 1991 to 2015. Starting from 2003, it has started to rise gradually and jumped in 2008. From then, it has rose steeply and reached a plateau in 2012. Then, it has increased steeply and reached another plateau in 2013. After that, there was a dramatic increase in price index in 2014, arriving at nearly 400 units.

3.1 Regression Result (Cont'd)

Model 3

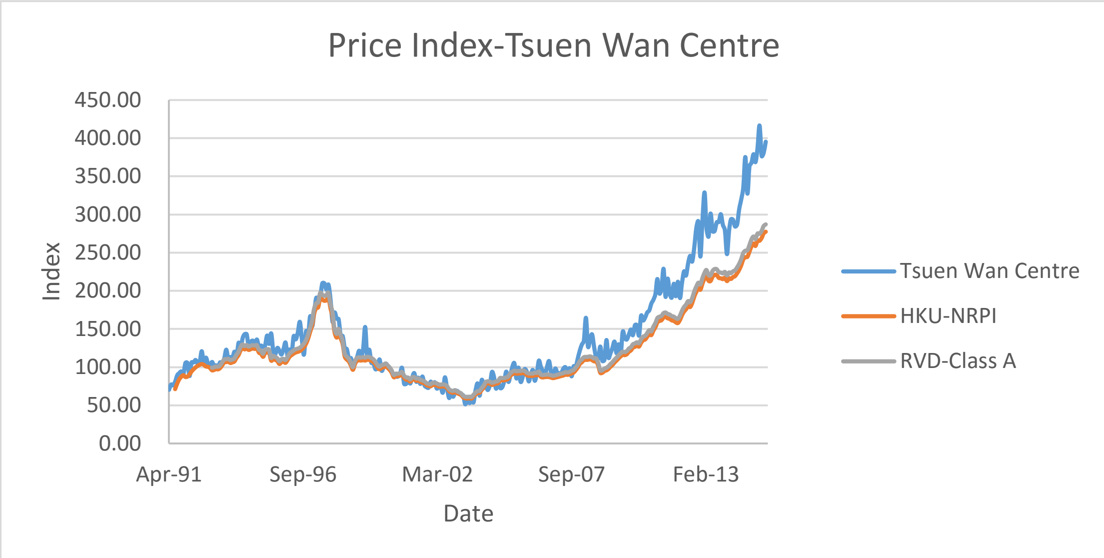

Figure 7

Comparison of price index of Tsuen Wan Centre, HKU-NRPI and RVD-Class A[1]

As a whole, the Price Index of Tsuen Wan Centre has demonstrated a similar trend with the HKU-NRPI and the price index of Class A flats according to RVD. The gradual increase from 1991 to 1996 was nearly the same. They all reached a peak at 1997 and then decreased to a trough in 2003. They all increased gradually from 2003 to 2008. However, the gradient of increase of HKU-NRPI and RVD-Class A were less steeply than that of Price Index of Tsuen Wan Centre. The HKU-NRPI and RVD-Class A arrived at around 275 units at the end, while at the same time, the Price Index of Tsuen Wan Centre reached nearly 400 units.

[1] RVD index is manipulated to let the index at Jan 2000 be 100

3.1 Regression Result (Cont'd)

Model 3

Expected Sign of Coefficients

Variables	Expected sign of coefficients	Actual sign of coefficients
GFA	+	+
FL	+	+
$FL^{1.1}$	+	(-)
AGE	-	-
ROOF	+	+
BAL	+	+
OVIEW* GREEN	+	+
OVIEW*GREEN*BAL	+	+
NE	Unknown	-
SE	Unknown	+

Table 12 Comparison of expected sign and actual sign of coefficients of Model 3

3.2 Implication

According to the regression model 3, positive effect of Floor level, GFA, Roof, Open View, Green View and Balcony are observed. Except the dummy variables, GFA gives the strongest effect to the Nominal price since it has the largest value. The time dummies, which is a representation of price index, tell us that there is a trend of increasing property price over the years by more than 500% from April, 1991 to August, 2015

3.3 The Interactive Model

The valuation of flats can be done by directly inputting the floor and GFA, as well as selecting the desired building as there are 19 buildings in total. Additional requirements like the provision of green or open view, wind direction, provision of roof and balcony can also be selected as criteria for valuation.

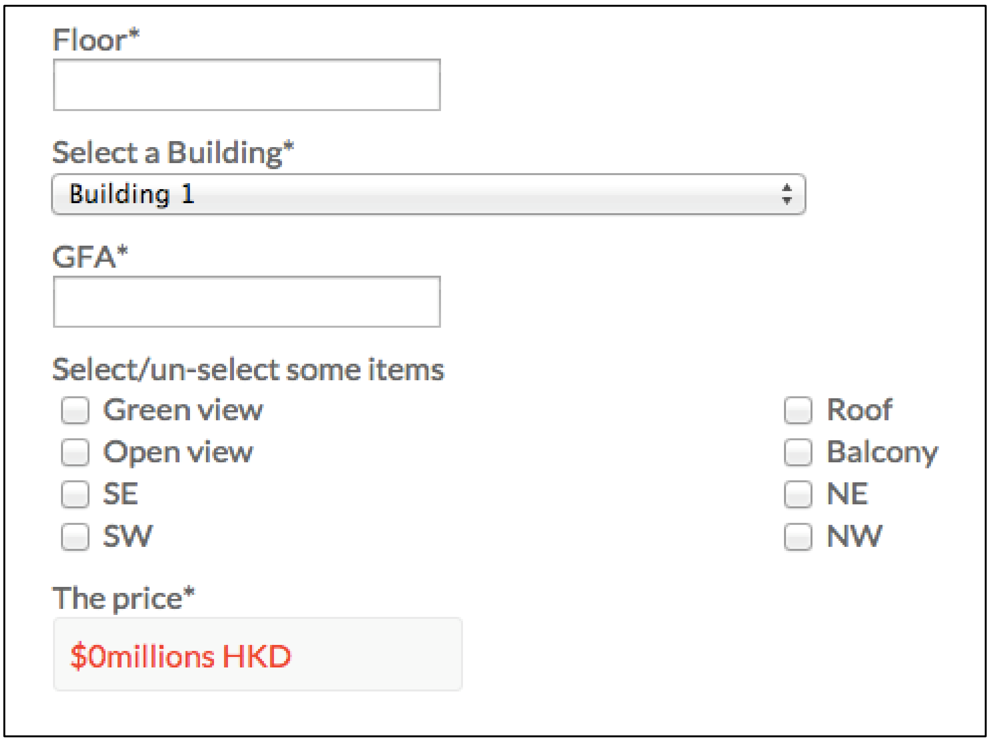

Figure 8 The Interactive Model

As the latest EPRC data used for constructing the model dates back to August 2015, newly available data should be continuously input in the future to keep the valuation and price index up-to-date.

3.4 Plans for Updating and Maintenance

Problems

Time Lag

Time lag exists when collecting and applying the property prices of Tsuen Wan Centre. As it is not possible to obtain the transaction record exactly right after the transaction, the most updated transaction records are mostly from the last month. As a result, the Hedonic Pricing Model can only apply the last month data to produce the result of property price in this current month. Therefore, slight errors and inaccuracy of the estimated property may occur.

Possible Difficulty in Collection of Property Price Data

There may be problems in the collection of property price data used to update the pricing model. This is due to the fact that the data of property prices in Tsuen Wan Centre is not public. Extra cost and time may be incurred in collecting the new data.

Possible Changes in Regression Model

Changes may take place in the influence of variables to the pricing model. Conditions such as views blocked by new built structures or redevelopment of the district will affect the coefficients of the variable in regression model, leading to the inaccuracy of the model if it is not properly updated.

3.4 Plans for Updating and Maintenance (Cont'd)

<u>Method</u>

Updating the pricing model with newly available transaction records in a regular manner.

Data

Data regarding transaction records of flats in the Tsuen Wan Centre will be obtained in the EPRC. Data concerning accuracy and factors of the model may be acquired through the feedback from the users.

Frequency

The Hedonic Pricing Model will be updated every month by adding the data in the last month; while the regression model will be reviewed for its accuracy every 4 months.

Updating Procedures

Data of transaction records and conditions of flats will be collected from the EPRC and then they will be organized and analyzed in relevant software. After the analysis, new regression model will be constructed and after final edit as well as adjustment, the interactive model will be updated for public use.

CHAPTER 4

DISCUSSION

Illustration will be given below to explain the sign of coefficients of some of the variables above.

Balcony

Balcony is often seen as a green feature in residential buildings and is thus favourable by the building occupants. This was proved in a research that balcony does function as an "environmental filter" in reducing noise level from outside, providing extra leisure space and enhancing energy efficiency. As a result, the balcony is able to have favourable effect towards the property price regardless of the view outside (Chau, K.W., S.K. Wong, 2004). This justifies the positive coefficient of balcony which indicates the increase of property price with the presence of balcony.

Open View and Greenery View

The different views of the residential buildings are associated with the comfortableness of building occupants. Open view tends to provide the occupants a less dense feeling and unblocked scene. Meanwhile, greenery view may give refreshing feeling and provide natural space instead of concrete jungle.

Both of the views therefore have a positive effect to the property value.

Open View, Greenery View and Balcony

This dummy variable indicates that "1" refers to flats with all three of the factors: open view, greenery view and balcony; while "0" refers to flats not possessing all three of the factors. As explained above, the open view, greenery view and balcony apply positive effect to the property value. Further to it, the presence of balcony can even maximize the value of the open and greenery view as more space is provided to enjoy the advantages of open and greenery view. As a result, the presence of all of the three factors can further raise the property value.

Orientation – Northeast

The importance of orientation of building is often related to the climate in Hong Kong. The main reason is the heating and cooling effect. As wind often comes from North in winter, flats facing northeast thus encounter wind in winter which further reduces the indoor temperature (B. Andersson, 1985).This is unwelcome by the occupants as more cost may be incurred in heating. Hence, flats with orientation to northeast tends to lower the property value.

Orientation – Southeast

Further to the above explanation, flats facing southeast is exactly contrary to the flats facing northeast. Due to the cooling effect brought by the wind come from South in summer, this helps to reduce the cost in air conditioning (B. Andersson, 1985).The property value can therefore be enhanced when the flat locating on southeast.

CHAPTER 5

CONCLUSION

This study aims at looking for the attributes that affecting the housing price of Tsuen Wan Centre. To find out the correlation between the housing price and pre-defined attributes, a hedonic price model is adopted for the empirical study. The transaction record from EPRC is extracted and analyzed by hedonic price model. The empirical results show that the physical feature of a flat contributed to the price of the flat mostly. However, some neighborhood attributes such as green view and open view are shown to be significant to the housing price. This may due to the psychological benefit to the occupant (Need source). Therefore, people are more willing to pay for these attributes.

CHAPTER 6

LIMITATION

Choices of attributes

1. The index may be subjective and affected by the attributes adopted, since the hedonic price index is adopted. According to Chau (2006), this method controls quality constant by including all the qualities variables. However, it is not possible to include all variables because there is always a lack of information.

2. Since there is a lack of information, not all the attributes can be adopted for the hedonic price model. Therefore, some important attributed may be ignored and this will affect the regression result.

CHAPTER 7

REFERENCE

1. Brandt Andersson. (1985). “The Impact of Building Orientation on Residential Heating and Cooling”. Retrieved November 13, 2015from http://www.sciencedirect.com/science/article/pii/0378778885900052

2. Chau, K. W. (2006). Index Construction Method for The University of Hong Kong All Residential Price Index (HKU-APRI). Retrieved November 10, 2015 from The University of Hong Kong, Department of Real Estate & Construction Web site: http://hkureis.versitech.hku.hk/

3. Chau, K. W., Wong, S. K., & Chung, Y. Y. (2004). The value of the provision of a balcony in apartments in hong kong. *Property Management*, *22*(3), 250-264. Retrieved from http://search.proquest.com/docview/213400937?accountid=14548

4. Centadata. (2015). “Tsuen Wan Centre”. Retrieved October 3, 2015, from http://www1.centadata.com/eptest.aspx?type=2&code=DMXSZHLXHD&info=tr&page=0

5. EPRC LTD. (2015).”About Us”. Retrieved November 9, 2015, from http://eprc.com.hk/

6. GoHome.hk.com. (2015). “Tsuen Wan Centre - Floor Plan”. Retrieved October 3, 2015, from http://community.gohome.com.hk/Tsuen-Wan/Tsuen-Wan-Centre/floorplan/

7. G, S. S., Macpherson, D. A., & Zietz, E. N. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature,13*(1), 3-43. Retrieved from http://search.proquest.com/docview/200098809?accountid=14548

8. HKIS. (1999). “Code of Measuring Practice”. Retrieved November 11, 2015, from http://www.hkis.org.hk/ufiles/MPracticeCode-199903.pdf

9. Hutcheson, G. D. (2011). Ordinary Least-Squares Regression. In L. Moutinho and G. D

10. Hutcheson, The SAGE Dictionary of Quantitative Management Research. Pages 224-228. Retrieved from https://datajobs.com/data-science-repo/OLS-Regression-[GD-Hutcheson].pdf

11. Ling Leng1, Tianyi Zhang1, Lawrence Kleinman 2, Wei Zhu1 (2007) Ordinary Least Square Regression, Orthogonal Regression, Geometric Mean Regression and their Applications in Aerosol Science. Journal of Physics: Conference Series. Vol 78.. Retrieved from http://iopscience.iop.org/article/10.1088/1742-6596/78/1/012084/pdf

12. Mok H.M.K. Chan P.P.K. & Cho Y.S. (1995). "A Hedonice Price Model for Private Properties in Hong Kong." The Journal of Real Estate Finance and Economics 10(1):37-48

13. Raymond Y.C. Tse, & Peter E.D. Love. (2000). Measuring residential property values in hong kong. *Property Management, 18*(5), 366. Retrieved from http://search.proquest.com/docview/213397845?accountid=14548