Property Management

Emerald Article: Weight regression model from the sales comparison approach

Shih-MingYou, Chin-oh Chang

Weight regression model from the sales comparison approach

Shih-MingYou
Department of Land Administration, Ministry of the Interior, Taipei, Taiwan, and
Chin-oh Chang
National Chengchi University, Taipei, Taiwan

Abstract

Purpose – The purpose of this paper is to construct a weight model from the sales comparison approach.

Design/methodology/approach – Although weighted average of comparables into sales comparison value is commonly applied in the past, most papers only focus mathematical calculation. This paper examines the correlation between weight and attributes of 6,345 sales comparable properties adopting the multiple regression model.

Findings – This paper finds the price type, proximity of transaction date, inside the neighborhood area or not, total gross adjustment as percent, numbers of adjustments and the attributes of other comparables considered in one appraisal are significant on the weight of comparables. The expected MAPE and Hit rate criterions are passed after forecasting 10 percent validation samples modeled by 90 percent samples randomly surveyed.

Practical implications – The weighted average to determine the sales comparison value is reasonable since the value conclusion will "correlate" to indication of value derived by different comparables.

Originality/value – This paper discusses the weight model and forecasts weights directly instead of only forecasting value. By elaborating on the core question of weights, this paper hopes to assist the degree of science and objectivity of appraisal.

Keywords Process analysis, Sales forecasting, Correlation analysis, Real estate

Paper type Research paper

Introduction

Real estate appraisal comprizes the sales comparison, income, and cost approaches to value in general. Pagourtzi *et al.* (2003) states the sales comparison approach is the most widely used approach. When data are available, the sales comparison approach is the most direct and systematic approach to estimating value (International Valuation Standards Committee, is hereafter called, IVSC, 2007). The sales comparison approach is probably the most commonly used for developing an indication of value in the appraisal. The sales comparison approach is based on the principle of substitution. A knowledgeable, prudent purchaser will not pay more for a particular property than the cost of acquiring an equally desirable substitute property without delay. The principle of substitution is applicable for individual value-influencing differences between a comparable property and a subject property. The collective adjustments for the differences are applied to the comparable to make it equal to the subject reflecting the date of value, thereby estimating an indication of value for the subject property. When several similar or commensurate commodities, goods, or services are available, the one

Property Management
Vol. 27 No. 5, 2009
pp. 302-318

0263-7472
DOI 10.1108/02637470910998465

with the lowest price will attract the greatest demand and the widest distribution (Appraisal Institute, 2002).

Ventolo and William (2001) propose that, as a general rule, the sales comparison approach is the most reliable approach for single-family residences. The sales comparison approach has the widest applicability; it is useful in valuing residential, commercial, industrial, or agricultural property, and useful in applications involving both the type of property that would typically be occupied by its owner and the type that is rented to tenants. It is most commonly used when data on property attributes is available.

The application of sales comparison approach could be broken down into two phases. The first one is to select several comparables and estimate the adjusted sales price through adjustments on valuation date, situation, location and physical characteristics, etc. However, adjustments based on decades of experience, claimed by some appraisers themselves, are often criticized to be subjective and lacking econometric and quantitative support (Colwell *et al.*, 1983; Isakson, 1986; Galleshaw, 1992). Subsequently, there are some papers constructing coefficients from regression analysis using multiple regression model to be the reference for adjustments and improve the shortcoming of subjective judgment (Kang and Reichert, 1991; Todora and Whiterell, 2002).

After the adjusted sales price is estimated, the second phase is to determine the sales comparison value. To determine the sales comparison value, weights are usually assigned to each sales comparable. For example, when assigning weights using weighted average, "near neighbors criteria" is often proposed and comparable selection is based on nearness, such as Mahalanobis distance or Minkowski metric (Tshira, 1979; Isakson, 1986; Vandell, 1991; Isakson, 2002; Todora and Whiterell, 2002; Pagourtzi *et al.*, 2003). Although mathematical formula is commonly applied calculating weights of comparables, the hedonic utility mode is seldom applied to gauge the relationship between the weight of comparable and corresponding attributes in terms of market value perspective. The hedonic model has been widely applied to improve the degree of science and objectivity in adjustments for the first phase. However, the hedonic model is rarely used in the calculation of weights for the second phase, which leave the room for further study. Appraisal Institute (2008) states that in reconsiling value indications in the sales comparison approach, the appraiser evaluates the number and magnitude of adjustments and the importance of the individual elements of comparison in the market to judge the relative weight a particular comparable sale should have in the comparative analysis. Even though the researches illustrated above might calculate weight in compliance with the weighting principle. Some papers might demonstrate the feasibility and accuracy of weighted average based on the forecasting performance by the degree the predicted "value" reflect the transaction price. In comparison, this paper will discuss the weight model and forecasts "weight" directly instead of "value". By elaborating on the core question of weights, this paper hopes to increase the degree of science and objectivity of appraisal.

Moreover, using mathematical calculation to measure the weight after comparing the nearness of comparable like Isakson (1986), Kang and Reichert (1991), Todora and Whiterell (2002), and Pagourtzi *et al.* (2003) research might focus on the "target" comparable and seem treat each comparable independently. However, since several sales comparables are required in the sales comparison approach appraising one subject and those comparables would determine the sales comparison value jointly, the weight of one comparable not only decided by its absolute nearness, but also the relative nearness of other comparables in the same "group". By doing so, it will better reflect the concept of "correlation". To sum up, without sufficiently reasonable explanation to analyze the correlation of the weights from different value indications, even if the adjustments of comparable are performed cautiously, it might overshadow the elaborate process that preceded it. Regardless of the quantity of evidence available, the responsibility of the appraiser goes beyond the manipulation of numbers.

For above, the research problems of this paper are as follows:

RP.1 Analyzing the relationship between the weight and the attributes of each comparable.

This paper is going to adopt the multiple regression model to examine the relationship between weight, reflecting the market value, and internal factors or attributes of each comparable in the sales comparison approach.

RP.2 Comparing the forecasting performance whether considering relative attributes of other comparables within the same group or not.

Although weighed average is often applied in the sales comparison approach, the nearness concept – the greater the distance (difference of attributes), the smaller the weights – by arithmetic calculation is mainly adopted. If incorporating other comparables within the same group, will the forecasting performance increase? What the forecasting performance of weight model will become considering this factor?

After additional discussion of the motivations for this research, a review of the literature on hedonic price model and sales comparison approach are provided. The data, methods and empirical result in this research are then described. The article concludes with a summary of the results and their implications.

Literature review

Multiple regression analysis

Regression analysis (hedonics) represents the first major property valuation technology transferred from academics to practitioners (Colwell and Dilmore, 1999). Lancaster (1965) proposes that consumers purchase products to increase satisfaction and are served with the offer of commodity comprising a variety of characteristic. Rosen (1974) also states the product is composed of lots of characteristic and the price should be determined by those characteristic. The regression analysis had been used by academics in the 1920s. The name hedonics was given by an auto industry analyst in the 1930s, to convey the idea that value relates to features that provide the user with utility or pleasure, borrowed from a psychological term representing pleasant states of mind. Besides the price is always used as the dependent variable studied by a bunch of papers, rental is also often used as dependent variable. Those papers explain the quantitative degree of dependent variable by means of hedonic price theory. Similar to price and rental, the weight might be the dependent variable showing the utility the attributes of weights support. This paper tries to construct the hedonic weight model, based on the premium or discount of weight reflecting the attributes of comparables and is shown as Equation (1):

$$W_i = \alpha_i + \sum_{i=1}^{n} \beta_{ij} X_{ij} + \phi_i$$

where:

$W_i$is the weight of comparable i;

α_i is the intercept of comparable i;

β_{ij} is the coefficient of the attribute j of the comparable i;

X_{ij} is the attribute j of the comparable i; and

ϕ_i is the error term of comparable i.

Adelman and Griliches (1961) might be the earliest paper studying the relationship between the value and attributes of the real estate. Kinnard and Boyce (1978) apply the hedonic price theory to the sales comparison approach. Their article use coefficients estimated from the regression to be the reference for adjustment and link the hedonic price to the sales comparison approach. Colwell *et al.* (1983) demonstrate how to derive adjustment factors using the ordinary least squares (OLS) method rather than appraiser judgment. Kang and Reichert (1991) establish the Grid Adjustment Method (GAM, AGM is also dubbed) applying the hedonic price model and compare the grid adjustment method with the regression method. They find the average error of the AGM smaller than the multiple regression analysis and the error by the percentage adjustment method is the least one in the AGM.

Weighted average of the sales comparison approach

Isakson (1986) presents a technique called the Nearest Neighbors Appraisal Technique (NNAT). He proposes the final value estimate is calculated as a weighted average of the actual selling prices of the comparable properties. The NNAT uses the Mahalanobis distance to select comparables and assign weights to the actual selling prices of the comparable properties. The Mahalanobis distance is calculated according to Equation (2):

$$D_{ij}^2 = (X_i - X_j)E^{-1}(X_i - X_j)'$$

where:

D_{ij} = Mahalanobis distance between property i and j;

X = a vector of the factor-coordinates of the property; and

E = the factor-coordinate covariate covariance matrix of all of the properties.

After the distance D is arrived, the next step is to calculate the weights W_{ij} of adjusted sales comparables according to the equation:

$$W_{ij} = (1/D_{ij}^2)/\sum_{i=1}^{k}(1/D_{ij}^2)$$

The NNAT eliminates the need to calculate the adjusted selling prices of the comparable properties. Vandell (1991) presents a minimum variance (among the adjusted values of the comparable sales) approach for selecting and weighting comparable properties. Further, Gau *et al.* (1992) present a variation of Vandell's techniques in which the coefficient of variation replaces Vandell's variation as the measure to be minimized. Todora and Whiterell (2002) demonstrate how results of the sales comparison approach, developed using mass appraisal techniques, can be explained in a single-property appraisal. In their paper, value indicant is revealed form each comparable, and a final value is correlated with weights calculated by Minkowski metric. Minkowski metric is shown as:

$$\sum w_i[abs(x_{si} - x_{ci})/x_{si}]$$

Where w_i = the weight assigned to the attribute i, x_{si} = the value of the subject property i, and x_{ci} = the value of the comparable property i. Pagourtzi *et al.* (2003) also suggest the process of finding comparables utilizing "*distance*" to establish a measure of comparability between the subject and the comparable under consideration. It is computed by weighting the differences in characteristic between the subject and the comparable. The distance, *D,* is calculated in advance to estimate W_i as follows (McCluskey and Borst, 1997):

$$W_i = \frac{1}{(D/2)^2 + D_i^2 + \left[2D(|ASP_i - SP_i\ |SP_i)\right]^2}$$

$$W = \sum_{i=1}^{n} W_i,\ \text{Weighted estimate} = \sum_{i=1}^{n} \frac{W_i}{W} ASP_i$$

where:

ASP_i = adjusted sale price for comparable *i*.

SP_i = sale price of comparable *i*.

D_i = distance for comparable *i*.

D = max of D_i.

After discussion on the review of the literature, research design and data source are provided in next section.

Research design and data source

Attributes of weight

This paper analyzes the relationship between the weight and attributes of sales comparable using 2,115 residential and commercial land benchmark value. With three comparables are selected in one land benchmark appraisal, 6,345 comparables are collected from 25 cities and counties in Republic of China in 2007 and 2008. We use the weights decided by valuers on each adjusted sales price from comparable to be the dependent variable. Kinnard (1971) states correlation is applied throughout the entire appraisal framework, in each of the approaches to value estimate. It is not restricted to the section of the appraisal report titled "correlation". Hentschel and Tosh (1982) propose the appraiser can check his reasoning by testing the various components within each approach against the results of the other approaches and the final value conclusion. Appraisal Institute (2008) states that in the reconciliation process, the appraiser often asks several questions about the data and techniques used in the sales comparison approach: Is the comparable property similar in terms of physical characteristics and location? Are the characteristics of the transaction similar to those expected for the subject property? Hence, this paper tries to discover the relationship between the weight and the factors calculated during the process of each appraisal approach as Kinnard (1971), Hentschel and Tosh (1982) have suggested.

Price type. Uncertainty is always embedded in the price since market error would make the price not reflecting the true market value due to the bargaining power, the influence of real estate brokers, the difference of information obtained, and the heterogeneity of real property. Hence, the transaction price would become a random distribution according to the nature of each transaction price affected by those factors abovementioned. Clapp (1990) finds the market value could be estimated from transaction price plotted in the normal price distribution. The uncertainty of the transaction price will influence the accuracy of market value pursued. From the seller's perspective, if the list price is substantially higher than the expected list price, then a prospective buyer is less likely to visit the house (Anglin *et al.*, 2003). If a prospective buyer is less likely to visit the house, the transaction chance will be slim and won't become a value indication. When the listing price set too high, a prospective buyer will have more negotiation room to bargain (Asabere and Huffman, 1993; Springer, 1996). Therefore, this paper infers that the asking or listing price will make the discovery of market value more uncertain. Besides, since the asking or listing price communicate the message that the bargain has not been completed yet, the inspection of transaction condition might not be executed and the date of transaction could not be confirmed. It will exert negative effect on the estimation of adjusted selling price. In addition, for transaction price, even valuers have collected transaction price, the authenticity of the price might not be easily verified. This paper demands valuers to confirm the nature of each sales price collected. The price type will be composed of three types: "verified transaction price", "transaction price not verified", and "asking or listing price". These three types will be the dummy variables and the asking or listing price is the base. Opposed to the "asking or listing price", the expected sign of "verified transaction price", "transaction price not verified" are positive.

Proximity of transaction date (months). An appraiser first selects several similar properties from among all the properties that have recently been sold (Pagourtzi *et al.*, 2003), the "recently" sold properties better offer insights to observe participants from the market. If comparable properties that occurred under market conditions differ from those applicable to the subject property on the effective date of valuation, adjustment is required for any difference that affect their values. The adjustment for market conditions is often referred to as a time adjustment. In general, appraisers select comparables that are recently transacted to better reflecting the date of valuation of the subject. For the reason that the market condition is changing, the earlier the transaction date of sales comparable was, the less applicability the reference for observing the market behavior of current situation. This paper calculates the month difference between transaction date of the comparable and valuation date of the subject. Proximity of transaction (months) date is a continuous variable. The expected sign is negative.

Inside the neighborhood area or not. The area valuers will search for comparables, for a broader concept, would be the primary market area. It is the geographic area within which the subject property and comparable properties are substitutable and affect prices each other. For a narrower and specific range, it will be the neighborhood area, which is the geographic area with a high level of homogeneity within which a number of properties are surrounded by subject property or comparable properties, which possess same or similar use purposes with those properties. If the comparables are collected from the neighborhood area, they might share similar location characteristic with the subject and have a similar value level. Inherently, the difference of location influence between subject and comparable might be slight and could even be omitted. Besides, location is always an important variable on property price for multiple regression analysis (Sirmans *et al.*, 2005). Hence, if the comparable is located inside the neighborhood area, the dummy value will take 1 and 0, instead. The expected sign is negative.

Numbers of adjustments. The most common property components, or elements of comparison, considered for their influence on value are as follows (IVSC, 2007):

- Real property rights conveyed.
- Financing terms.
- Conditions of sale.
- Expenditures made immediately after purchase.
- Market conditions (time).
- Location.
- Physical characteristics.
- Economic characteristics.
- Use (zoning).
- Non-realty components of value.

If the sales are similar otherwise, less accuracy may be attributable to the comparable property that required the larger adjustments (Appraisal Institute, 2008). The comparables with the smallest adjustments are better in compliance with the principle of substitution. Featherston (1968) suggests valuers to consider that which approach required fewer adjustments to allow for differences in the subject approach. Numbers of adjustments of the comparable is a continuous variable. The expected sign is negative.

Total gross adjustment as percent. If one or two comparable transactions require fewer total adjustments than the other comparable transactions, an appraiser may attribute greater accuracy and give more weight to the value indications obtained from these transactions, particularly if the magnitude of the adjustments is proximately the same (Appraisal Institute, 2008). The best sales comparable to be chosen is that result in the smallest errors in predicting the value of the subject property. To minimize adjustment errors, one would want to pick sales that have very small adjustments (Kummerow, 2003). If only calculating the adjustment by totaling the positive and negative adjustments and subtracting the smaller amount from the larger amount, a net adjustment figure may be misleading because one cannot assume that any inaccuracy in the positive and negative adjustments will cancel each other out. The magnitude of net adjustments would not be the only indicator measuring accuracy. To supplement the "numbers of adjustments" variable, the total gross adjustment would be a good indication measuring the similarity of comparable. Considering above, the total gross adjustment as percent of the comparable is one of the explanatory variables. The expected sign of this continuous variable is negative.

The distance of other comparables within the same group. As Figure 1 presents, one appraisal of subject property (Δ) for sales comparison approach is determined by three round-shape comparables. Among 6,345 comparables, every three comparables constitute one group and there are 2,115 groups circled with dotted line. Since the weight of one target comparable (•) correlate with other two comparables (◦). The distances of three round-shape comparables are different within the same group. The group circled with larger dotted line implies less prosperous the primary market area is and less comparables could be selected. For the calculation of distance, the five variables listed above should be incorporated borrowed from the concept of papers mentioned above, i.e. Isakson (1986, 2002), Vandell (1991), Todora and Whiterell (2002), and Pagourtzi *et al.* (2003). In order to balance the magnitude of those continuous and dummy variables, average numbers of 6,345 samples on five variables are calculated in advance. Followed by summation of five figures, which are estimated with each value of variable divided by its corresponding whole sample average value, then the distance of 6,345 samples will be completed as the below equation shows:

$$D_i = \sum_{j=1}^{m} \left(X_{ij} / \sum_{i=1}^{n} \frac{X_{ij}}{n} \right)$$

Where D_i = distance for comparable i or the difference of attributes between subject property (Δ) and comparable property i (◦), X_{ij} is the attribute j of the comparable i; n is the number of total samples; m is the number of attributes. For instance, five attributes and distance of comparable F are calculated as Table I presents.

According to Table I, the distance of comparable F is 7.97, which is greater than the average of whole samples[1]. For example, if subject α is located in a metropolitan city with bunch of comparables, the similarities of three comparables, i.e. A, B and C, to the subject are greater and the distances of them are smaller, say, 3, 4 and 5. The weights of

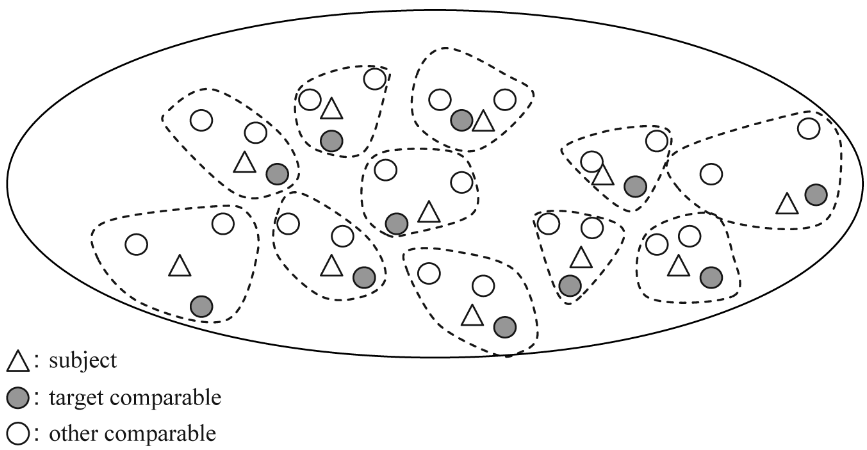

Figure 1.
The relationship between the subject and the comparable

A, B and C are assumed to be 50 percent, 30 percent, 20 percent according to the nearness concept- the greater the distance (difference of attributes), the smaller the weights. On the other hand, subject β is located in a suburban area with few comparables, the similarities of three comparables, i.e. D, E and F, to the subject are smaller and the distances are greater, say, 6, 7 and 7.97. The weights of D, E and F are assumed to be 40 percent, 35 percent, 30 percent. All the comparables consisting A to F are included in whole samples for hedonic model. When comparable C is compared with D, although the distance of comparable C is 5, while the D is 6.5; however, the weight of comparable C, 20 percent, is smaller than 40 percent, the weight of comparable D, would become unreasonable. Hence, besides the attributes of target comparable (•), considering the distances of other comparables (◦) as well within the same group will better explaining how each weight on comparable is formed. Within each group, several comparables are selected and the sum of weights is 100 percent. The greater the distance of other comparable (◦) within one group stands for the similarity of other comparable (◦) to the subject is lower. Thus the weight of target comparable (•) might be greater comparatively. The expected sign on the weight of target comparable (•) is negative compared to the distances, the continuous variables, of other comparables (◦) within the same group.

Although this paper utilizes sales comparables covering different cities and counties, the weights assigned by valuers won't be affected by region factors; therefore dummy variable on location shall be omitted. In terms of the valuation date of each adjusted selling price, 3,027 comparables from 1,009 land benchmark value are on March 31, 2007, and 3,318 comparables from 1,106 land benchmark value are on March 31, 2008. Year factor as a dummy variable is required for that the date of valuation for samples from 2007 and 2008 is different. In overall, attributes of weights are summarized in Table II.

Weight model using the OLS

For the specification of multiple regression model, Follain and Malpezzi (1980) propose the coefficients of semi-log model could be the percentage of the dependent variable influenced by the variation of unit of attribute and suggest semi-log is superior than linear regression model. This paper establishes the weight model with the explanatory variables from Table I as follows:

$$\log(SW_i) = \alpha_i + \sum_{i=1}^{n} \beta_{ij} X_{ij} + \sum_{i=1}^{n} \sum_{k=1}^{m} \beta_{ik} X_{ik} + \phi_i$$

Table I.
Calculation of the distance of comparable F

Value of attribute	Price type	Proximity of transaction date	Inside the neighborhood area or not	Total gross adjustment as %	Numbers of adjustments
Comparable F	2	6	1	0.185	9
Average	1.8593	4.8327	0.4162	0.1029	6.1917
Distance	2/1.8593 + 6/4.8327 + 1/0.4162 + 0.185/0.1029 + 9/6.1917 = 7.97				

Attribute/variable	Symbol	Description	Variable type
Price type			
Verified price	SA1	"Verified price", "price not verified", and "asking or listing price"(base)	Dummy
Price not verified	SA2		
Proximity of transaction date (months)	SB	Month difference between transaction date and valuation date	Continuous
Inside the neighborhood area or not	SC	Comparable located inside the neighborhood area: 1; comparable located outside the neighborhood area: 0(base)	Dummy
Total gross adjustment as a percentage	SD	Total gross adjustment percentage of comparable	Continuous
Numbers of adjustments	SE	Numbers of adjustments of comparable	Continuous
Distance of other comparable 1	Scon1	The summation of the ratio, with the figure of attribute to the average figure of whole sample attribute, on 5 attributes of comparable 1 in the same group	Continuous
Distance of other comparable 2	Scon2	The summation of the ratio, with the figure of attribute to the average figure of whole sample attribute, on five attributes of comparable 2 in the same group	Continuous

Table II.
Variable explanation on the attributes of weights

where:

$\log(SW_i)$ is the log of weight of comparable i;

α_i is the intercept of comparable i;

β_{ij} is the coefficient of the attribute j of the comparable i;

n is number of total samples; m is number of samples within each group;

X_{ij} is the attribute j of the comparable i;

X_{ik} is the attribute k within the same group of the comparable i; and

ϕ_i is the error term of comparable i.

Namely, the function of log of weight is shown as Equation (9):

$$\log(SW_i) = \text{f(SA1, SA2, SB, SC, SD, SE, Scon1, Scon2, Year)}$$

The measurement of forecasting performance

Pace and Gilley (1993) suggest the effective criteria measuring the accuracy of forecasting performance should focus on the hold-out samples. This paper adopts cross validation method by randomly surveying 90 percent from total observations to pave the way for forecasting 10 percent hold-out samples. The MAPE and Hit-rate are employed for the measurement of forecasting performance. By the reference of relevant literature (Drives Jonas/IPD, 1990; Matysiak and Wang, 1995; Calhoun, 2001), if the Hit-rate hit the mark of 30 percent within 10 percent range and 70 percent within 20 percent range, the forecasting performance would arrive at the standard being good.

The reason we use land benchmark value to be our empirical samples is that: the land value benchmark is the representative land in a neighborhood area. It is appraised to reflect the market value and serve as a transaction reference for the public. Since all of the land benchmark values have been confirmed by the land value review committee composed of scholars, officials, representative of Real Estate Appraiser Guilds, etc. after appraised by valuers, we can say the land benchmark value would reflect the market value, and weight of each comparable determined by valuers would become the basis supporting the market value. Therefore, these 6,345 weights would be worthy constructing the weight model and serve future application need.

Empirical results

This paper analyzes the relationship between the weight and attributes of 6,345 comparables. For sample selection, we deletes the outliers using Dffits criteria, since it is better than R-student, Covratio, Cook'D criteria proven in the past. After randomly surveys 5,711 samples, 90 percent of whole 6,345 samples, and deletes outliers from 5,711 sales comparables, the remaining 5,350 sales comparables are used for empirical analysis. The descriptive analysis is shown as Table III.

The least weight of the sales comparison approach is 5 percent, while the most one is 80 percent (Table III). The multiple regressions analysis is shown as Table IV using SAS software:

In order to highlight the importance and significance of inclusion of distance of other comparables, model 1 and model 2 in Table IV are compared. Model 1 excludes the distance of other comparables, while model 2 is inclusive of distance of other comparables. As Table IV presents, although all explanatory variables are significant at 1 percent level and are consistent with the expected sign. With the F-statistic is 223.51 but Adj R-Sq is only 24.34 percent, which imply some important independent variables might not be incorporated in model 1 yet. For model 2, with CI-statistic at 4.44, no significant collinearity is found. Besides all explanatory variables are significant at 1 percent level except year dummy and are consistent with the expected sign, the F-statistic increase to 1013.68 and Adj R-Sq stands at 63.01 percent opposed to model 1 (Table IV). It shows the model fit is good and the degree independent variables explain the variation of weights is high. Consequently, the coefficients of explanatory variables of model 2 are elaborated further.

Variable	Mean	Std. dev.	Min	Max
Weight of sales comparable (%)	33.9	6.0	5	80
Price type	1.8593	0.4610	0	3
Proximity of transaction date (months)	4.8327	5.0013	0	48
Inside the neighborhood area or not	0.4162	0.4930	0	1
Total gross adjustment (%)	10.29	7.42	0	47
Numbers of adjustments	6.1917	3.9030	0	23
Distance of other comparable 1	5.1574	1.9693	1.3457	17.219
Distance of other comparable 2	5.1486	1.9856	1.3457	17.219

Table III.
Descriptive analysis

Table IV. Results of regression analysis by the OLS

Variable	Expected sign	Model 1: exclusion of distance of other comparables					Model 2: inclusion of distance of other comparables				
		Coefficient	*t*-statistic	*p*-statistic	Standardized β	Rank	Coefficient	*t*-statistic	*p*-statistic	Standardized β	Rank
Intercept		−0.97431	−79.77	<0.0001	0		−1.24768	−134.78	<0.0001	0	
Price type (verified price)	+	0.10038	8.59	$<0.0001^{**}$	0.21578	2	0.12548	15.59	$<0.0001^{**}$	0.26542	7
Price type (price not verified)	+	0.05482	5.05	$<0.0001^{**}$	0.12783	5	0.05603	7.53	$<0.0001^{**}$	0.12913	8
Proximity of transaction date (months)	−	−0.00754	−17.26	$<0.0001^{**}$	−0.20852	3	−0.01368	−43.53	$<0.0001^{**}$	−0.3805	3
Inside the neighborhood area or not	−	−0.03782	−7.44	$<0.0001^{**}$	−0.10301	6	−0.06556	−17.9	$<0.0001^{**}$	−0.17397	6
Total gross adjustment as percent	−	−0.38223	−7.55	$<0.0001^{**}$	−0.15681	4	−0.9191	−25.31	$<0.0001^{**}$	−0.36838	5
Numbers of adjustments	−	−0.01001	−10.15	$<0.0001^{**}$	−0.21602	1	−0.01936	−27.16	$<0.0001^{**}$	−0.40625	4
Distance of other comparable 1	+						0.045	43.47	$<0.0001^{**}$	0.44106	2
Distance of other comparable 2	+						0.04525	44.34	$<0.0001^{***}$	0.44775	1
Year		0.01089	2.5	0.0124^{*}	0.03009		−0.01002	−3.23	0.0013^{*}	−0.02699	
		F − statistic:223.51		Adj − R^2:0.2434			F − statistic:1013.68		Adj − R^2:0.6301		
		CI − statistic:4.96					CI − statistic:4.44				

Note: *Significant at 95 percent; **significant at 99 percent

Analysis of weights of the sales comparison approach

The weight of comparable will increase 13.37 percent[2] and 5.76 percent respectively when the price type is "verified transaction price" and "transaction price not verified" as opposed to the base, asking or listing price. With one more month for proximity of transaction date, the weight will decrease 1.37 percent. The weight will decrease 6.35 percent if the comparable is outside the neighborhood area compared with the base if the comparable is inside the neighborhood area. With 1 percent increase for the total gross adjustment as percent, the weight will decrease 0.91 percent. With one unit increase for the numbers of adjustments, the weight will decrease 1.9 percent. The abovementioned independent variables are all significant at 99 percent confidence level. The greater the distance of other comparables, representing the similarity degree is lower for the price formation as opposed to the subject property, the closer the target comparable is. With one unit increases for the distance of other comparable 1 and comparable 2, the weight (of target comparable) will increase 4.5 percent and 4.52 percent respectively (Table IV).

For the rank of standardized, the distance of comparable 2 and 1 are the first and second place (Table IV). Apparently, it is quite important for the inclusion of the attributes of other comparables when constructing the weight model. In other words, the logic for considering the combined attributes of other comparables is that, since at least three comparables have to be collected, the weight of the target comparable will not only determined by its attributes, but also be dependent on relative relationship of other two comparables within the same group. The rank on standardized beta supports the idea. Hence, it is necessary to incorporate the combined attributes of other two comparables into the model to express the weight formation of concerned factors. In addition, for individual attribute of each sales comparable, proximity of transaction date affect the weight the most except to the distance of other comparables.

Forecasting performance of the weight

After inputting the coefficients of regression model, which is established with 90 percent samples randomly surveyed from weights of all the sales comparable, and the attributes of 10 percent validation samples to the Equation (9), the predicted weights are produced in Equation (10):

$$\log(SWi) = -1.24768 + 0.12548 \times \mathrm{SA1} + 0.05603 \times \mathrm{SA2} - 0.01368 \times \mathrm{SB} - 0.06556 \times \mathrm{SC} - 0.9191 \times \mathrm{SD} - 0.01936 \times \mathrm{SE} + 0.045 \times \mathrm{Scon1} + 0.04525 \times \mathrm{Scon2} - 0.01002 \times \mathrm{Year}$$

On the other hand, in order to compare with model 1, the MAPE and Hit rate are calculated for model 1 using the same procedure reaching the Equation (10). The function is presented as Equation (11):

$$= -0.97431 + 0.10038 \times \mathrm{SA1} + 0.05482 \times \mathrm{SA2} - 0.00754 \times \mathrm{SB} - 0.03782 \times \mathrm{SC} - 0.38223 \times \mathrm{SD} - 0.01001 \times \mathrm{SE} + 0.01089 \times \mathrm{Year}$$

The MAPE and Hit rate are shown in Table V.

As the Table V presents, the MAPE is only 13.11 percent and the Hit rate within 20 percent range is 83.73 percent for model 2: inclusion of Distance of other comparables. Both figures are better than model 1: exclusion of Distance of other comparables with the MAPE at 17.48 percent and the Hit rate within 20 percent range is 76.08 percent (Table V). For model 2, the MAPE and Hit rate criterion are passed after forecasting 10 percent validation samples modeled by 90 percent samples randomly surveyed. Clearly, the weight of one adjusted selling price is not only decided by absolute distance of the target comparable, but also relative distance of other comparables within one group. By doing so, the meaning of correlation is better fulfilled. The higher forecasting performance could also be a proof.

Conclusion

Sales comparison approach is one of the three real estate appraisal approaches to value. Since it is always required to collect several comparables, weighted average of those comparables into sales comparison value is commonly applied using mathematical formula to calculate weights in past literature. However, the weights produced won't necessarily be supported by market perspective only by means of arithmetical calculation. This paper examines the correlation between weight and attributes of 6,345 sales comparables adopting the multiple regression model. This paper finds the price type, proximity of transaction date, inside the neighborhood area or not, total gross adjustment as percent, numbers of adjustments and the attributes of other comparables considered in one appraisal are significant on the weight of comparables. The expected MAPE and Hit rate criterions are passed after forecasting 10 percent validation samples modeled by 90 percent samples randomly surveyed.

The performance of model 2 is superior to model 1, which communicates the idea: since at least three comparables need to be collected, the weight of the subject comparable will not only be determined by its absolute attributes, but also be dependent on relative relationship of other two comparables within the same group.

We have not found any papers constructing the weight model using hedonic regression analysis. One of the explainable reasons might be the difficulty gathering weight outcomes from appraisal reports. It is not easy to see empirical analysis emphasizing the weight regressed on weights affecting attributes. This paper discusses the weight model and forecasts weights directly instead of only forecasting value. By elaborating on the core question of weights, we hope to assist the degree of science and objectivity of appraisal.

	Forecasting performance (%)	Model 1 (%)	Model 2 (%)
MAPE		17.48	13.11
Hit Rate	10	45.77	64.59
	20	76.08	83.73

Table V. Forecasting performance of the weight of the sales comparison approach

Notes

1. The logic leads us to get the distance of the average of whole sample will be 5, with the calculation of $1.8593/1.8593 + 4.8327/4.8327 + 0.4162/0.4162 + 0.1029/0.1029 + 6.1917/6.1917 = 5$.
2. Because the price type is a dummy variable, for the semi-log model, the explanation on coefficients should be: $\exp(0.12548)-1 = 0.133692$. Subsequent dummy variables are calculated accordingly.

References

Adelman, I. and Griliches, Z. (1961), "On an index of quality change", *Journal of the American Statistical Association*, Vol. 56, pp. 535-46.

Anglin, P.M., Rutherford, R.C. and Springer, T.M. (2003), "The trade-off between the selling price of residential properties and time-on-the-market: the impact of price setting", *Journal of Real Estate Finance and Economics*, Vol. 26 No. 1, pp. 95-111.

Appraisal Institute (2002), *The Dictionary of Real Estate Appraisal*, 4th ed., Chicago, IL, p. 281.

Appraisal Institute (2008), *Appraisal Institute The Appraisal of Real Estate*, 13th ed., Chicago, IL.

Asabere, P.K. and Huffman, F.E. (1993), "Price concessions, time on market, and the actual sale price of homes", *Journal of Real Estate Finance and Economics*, Vol. 6 No. 2, pp. 167-74.

Calhoun, C.A. (2001), *Property Valuation Methods and Data in the United States*, Housing Finance International, London, pp. 12-23.

Clapp, J.M. (1990), "A methodology for constructing vacant land price indices", *Journal of American Real Estate and Urban Economics Association*, Vol. 18 No. 3, pp. 274-93.

Colwell, P.F. and Dilmore, G. (1999), "Who was first? An examination of an early hedonic study", *Land Economics*, Vol. 75 No. 4, pp. 620-6.

Colwell, P.F., Cannaday, R.E. and Wu, C.C. (1983), "The analytical foundations of adjustment grid methods", *Real Estate Economics*, Vol. 11 No. 1, pp. 11-29.

Drivers Jonas/IPD (1990), *The Variance in Valuations*, Drivers Jonas/Investment Property Databank, London.

Featherston, J.B. (1968), "Correlation and final value estimate", *The Real Estate Appraiser*, Vol. 34, pp. 40-6.

Follain, J.R. and Malpezzi, S. (1980), *Dissecting Housing Value and Rent*, The Urban Institute, Washington, DC.

Galleshaw, M.W. (1992), "Appropriate uses of economic characteristics in the sales comparison approach", *Appraisal Journal*, Vol. 60 No. 1, pp. 91-8.

Gau, G., Lai, T.-Y. and Wang, K. (1992), "Optimal comparable selection and weighting in real property valuation: an extension", *AREUEA Journal*, Vol. 20 No. 1, pp. 107-23.

Hentschel, J.J. and Tosh, D.S. (1982), "The appraiser's conscience: final correlation", *The Real Estate Appraiser and Analyst*, Summer, pp. 55-9.

International Valuation Standard Committee (2007), *International Valuation Standard*, 8th ed., London.

Isakson, H.R. (1986), "The nearest neighbors appraisal technique: an alternative to the adjustment grid methods", *Journal of the American Real Estate and Urban Economics Association*, Vol. 14 No. 2, pp. 274-86.

Isakson, H.R. (2002), "The Linear Algebra of the sales comparison approach", *Journal of Real Estate Research*, Vol. 24 No. 2, pp. 117-28.

Kang, H.B. and Reichert, A.K. (1991), "An empirical analysis of hedonic regression and grid-adjustment techniques in real estate appraisal", *AREUEA Journal*, Vol. 19, pp. 70-91.

Kinnard, W.N. (1971), *Income Property Valuation: Principles and Techniques of Appraising Income-Producing Real Estate*, Heath Lexington Books, Lexington, MA, pp. 442-3.

Kinnard, W.N. and Boyce, B.N. (1978), *An Introduction to Appraising Real Property*, Society of Real Estate Appraisers, Chicago, IL.

Kummerow, M. (2003), "Theory of real estate valuation: an alternative way to teach real estate price estimation methods", research monograph, Department of Land Economy and Valuation, Curtin University, Perth, p. 29.

Lancaster, K. (1965), "The theory of qualitative linear systems", *Econometrica*, Vol. 33 No. 2, pp. 395-409.

McCluskey, W. and Borst, R. (1997), "An evaluation of MRA, comparable sales analysis and ANNs for the mass appraisal of residential properties in Northern Ireland", *Assessment Journal*, Vol. 4 No. 1, pp. 47-55.

Matysiak, G. and Wang, P. (1995), "Commercial property market prices and valuation: analyzing the correspondence", *Journal of Property Research*, Vol. 12, pp. 181-202.

Pace, R.K. and Gilley, O.W. (1993), "Translating prior information across specifications to improve predictive accuracy", *Journal of Business & Economic Statistics*, Vol. 11 No. 3, pp. 301-9.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T. and French, N. (2003), "Real estate appraisal: a review of valuation methods", *Journal of Property Investment and Finance*, Vol. 21 No. 4, pp. 383-401.

Rosen, S. (1974), "Hedonic price and implicit market: product differentiation in pure competition", *Journal of Political Economy*, Vol. 32, pp. 34-55.

Sirmans, G.S., Macpherson, D.A. and Zietz, E. (2005), "The composition of hedonic pricing models", *Journal of Real Estate Literature*, Vol. 13 No. 1, pp. 1-44.

Springer, T.M. (1996), "Single-family housing transactions: seller motivation, price, and. marketing time", *Journal of Real Estate Finance and Economics*, Vol. 13, pp. 237-54.

Tchira, A. (1979), "Comparable sales selection: a computer approach", *The Appraisal Journal*, Vol. 48 No. 1, pp. 86-98.

Todora, J. and Whiterell, D. (2002), "Automating the sales comparison approach", *Assessment Journal*, Vol. 9 No. 1, pp. 25-33.

Vandell, K.D. (1991), "Optimal comparable selection and weighting in real property valuation", *Real Estate Economics*, Vol. 19 No. 2, p. 213.

Ventolo, W.L. and William, M.R. (2001), *Fundamental of Real Estate Appraisal*, 8th ed., DearBorn Real Estate Education, Chicago, IL, p. 71.

Further reading

Appraisal Institute (1967), *Appraisal Terminology and Handbook*, 5th ed., Chicago, IL.

Hagood, W.D. (1978), "Reconciliation analysis (formerly called Correlation)", *The Real Estate Appraiser*, Vol. 44, pp. 42-4.

Jud, G.D., Winkler, D.T. and Kissling, G.E. (1995), "Price spreads and residential housing market liquidity", *Journal of Real Estate Finance and Economics*, Vol. 11 No. 3, pp. 251-60.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T. and French, N. (2003), "Real estate appraisal: a review of valuation methods", *Journal of Property Investment and Finance*, Vol. 21 No. 4, pp. 383-401.

Todora, J. and Whiterell, D. (2002), "Automating the sales comparison approach", *Assessment Journal*, Vol. 9 No. 1, pp. 25-33.

About the authors

Shih-Ming You is a Certified Real Estate Appraiser and Section Chief in the Department of Land Administration, Ministry of Interior in Republic of China. He is in charge of land right regulation and is currently completing his PhD in the Department of Land Economics, National Chengchi University. He is the "Young Person's Prize Winner" presenting a topic about appraisal reconciliation in 23rd Pan Pacific Congress of Appraisers, Valuers and Counselors in San Francisco, California, 2006. Shih-Ming You is the corresponding author and can be contacted at: shiminyou@gmail.com

Chin-oh Chang is a Professor of Department of Land Economics, National Chengchi University and teaches real estate investment, housing policy and economics. He completed his PhD degree at Department of City and Regional Planning, University of Pennsylvania. His research areas include land and housing policy, real estate investment and financial analysis, urban and housing economics. His research papers have been published at *Urban Studies, Housing Studies, International Real Estate Review, Journal of Property Research, Taiwan Economic Review, Academia Economic Papers, Journal of Management*, and *Management Review*.