The Linear Algebra of the Sales Comparison Approach

Author Hans R. Isakson

Abstract

This study presents the sales comparison approach to value as a system of linear algebra equations. The difference between appraiser and academician views of this system of equations is shown to pivot on the one-price assumption. The need of appraisers to subjectively derive adjustment factors is shown to be an artifact of the common practice of appraisers to assume multiple indicated values. The use of multiple regression analysis and other statistical techniques by academicians is shown to be possible only given the one-price assumption. Suggestions are presented for the improvement of both views.

The sales comparison approach is known by several different names in the appraisal literature. In some of the older literature it is called the market data approach, while elsewhere it is referred to as the grid adjustment technique. In this study, the sales comparison approach is equated to the technique employed by appraisers in which the selling prices of comparable properties are adjusted to arrive at a set of indicated values of a subject property. The adjustments are made to reflect differences in the elements of comparison (hereafter referred to as property characteristics) between the subject and comparable properties. Finally, the set of indicated values are boiled down to a final estimate of value through a process called reconciliation.

The sales comparison approach is of interest because it is widely regarded by most appraisers as the approach that produces the most reliable estimate of the value of a subject property, especially when there are many recently sold properties comparable to the subject property. But, appraisers do not make use of a particularly large number of these comparable properties. Usually, appraisers combine their expert judgment with a relatively small number of comparable sales to arrive at a final estimate of value.

The sales comparison approach is also of interest to academicians. Some academicians like to study the sales comparison approach in order to offer methods for reducing or eliminating the subjective judgments used by appraisers. Much of this literature can be found in the Lentz and Wang (1998) review of 141 academic books and articles that deal, in one way or another, with the sales comparison approach. For example, one of the landmark articles that focuses on the sales comparison approach is Colwell, Cannaday and Wu (1983), which demonstrates how to derive adjustment factors using the ordinary least squares (OLS) method rather than appraiser judgment. Indeed, the need for traditional appraiser estimates of adjustment factors can be completely eliminated by multiple regression analysis. Yet, most appraisers do not make use of the Colwell, Cannaday and Wu techniques.

Some academicians look at other aspects of the sales comparison approach. Diaz (1990) studies the comparable property selection processes of students and experts, and reports that experts examine less data than students and follow different search processes than students. In a similar study, Spence and Thorson (1998) find that estimates of value from the sales comparison approach derived by experts were more accurate than those derived by students. Manaster (1991) reports that if the sales comparison approach is followed correctly, different appraisers using slightly different methods can and do arrive at final estimates of value that are very similar. Crookham (1995) examines the quality of commercially-generated comparable sales databases commonly used by appraisers who apply the sales comparison approach, while Boronico and Moliver (1997) explore the contradictory estimates obtained using different units of comparison.

A series of interesting academic studies look at the sales comparison approach with greater detail than others. Isakson (1986, 1988) presents a technique dubbed the Nearest Neighbors Appraisal Technique (NNAT) in which the final estimate of value is calculated as a weighted average of the actual selling prices of the comparable properties. The weights in NNAT are determined by a multi-dimensional measure of similarity; comparable properties more similar to the subject are given greater weight. 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, while Gau, Lai and Wang (1992, 1994) present a variation of Vandell's techniques in which the coefficient of variation replaces Vandell's variation as the measure to be minimized. In both techniques, the adjusted values of the comparable properties are calculated using the dollar additive technique and OLS-derived adjustment factors. Isakson's NNAT, Vandell's minimum variance, and Gau, et al.'s minimum coefficient of variation are computationally complex. In all of these studies, certain calculations are made for all possible combinations of the comparable properties taken 2, 3, 4, . . . *n* at a time, where *n* is the total number of comparable properties. Green (1994) looks closely at the Vandell and Gau, et al. techniques and shows that under classical OLS assumptions, Vandell's approach is preferable. But, in the presence of omitted variables or heteroscedasticity, Green finds that the Gau, et al. approach is better. The Isakson, Vandell and Gau, et al. techniques have not been given much attention in the literature since their initial appearances.

Usually, when academicians wish to estimate or analyze real estate values, they rarely use the sales comparison approach. Instead, academicians prefer to make use of a wide variety of statistical techniques. Most of the time, the academicians are more interested in the parameter estimates (adjustment factors in appraiser lingo) than the estimates of value that come out of their statistical analyses. A closer look at how appraisers and academicians view the valuation process can shed light on why they view the process of valuation so differently.

One way to better understand the differences between appraisers and academicians is to cast the sales comparison approach in a manner that brings these differences to the surface. Mathematical expressions of calculation techniques often provide a better understanding of the techniques. Pace (1998) presents the sales comparison approach in a mathematical format (matrix algebra). In doing this, Pace is better able to focus on some of the more interesting aspects of this approach. Unfortunately, Pace's model does not completely capture the sales comparison approach as practiced by appraisers.

The purpose of this study is to examine the sales comparison approach within a linear algebra framework in order to shed some light on how and why practicing appraisers view the approach so differently than do academicians. First, the sales comparison approach, as used by appraisers, is expressed as a system of equations. Next, the change academicians (including Pace) typically make to the approach is introduced. Within the framework of the academic version of the approach, the various classes of solutions are examined in detail. Finally, some implications of the findings are discussed.

The Appraisers' Model

Traditionally, the appraisal literature does not express the sales comparison approach as a system of equations. Instead, appraisers are taught a series of steps to follow when employing the approach. A characteristic of the approach is the derivation of an intermediate estimate of value from each of the comparable properties. These intermediate estimates are referred to as the "indicated values" of the subject property. Appraisers then "reconcile" the set of indicated values into a single number called the "final estimate of value." As seen below, these multiple, intermediate indicated values explain why appraisers rely so heavily on subjective judgment when using the sales comparison approach.

The exploration of the sales comparison approach will begin by expressing it as a series of mathematical equations. Where:

S = A $(1 \times n)$ vector of n indicated values of a particular subject property;
X = A $(j \times 1)$ vector of the j property characteristics of the subject property;
P = A $(1 \times n)$ vector of the selling prices of n comparable properties;
Z = A $(j \times n)$ matrix of the j characteristics of the n comparable properties;
A = A $(1 \times j)$ vector of j adjustment factors; and
I = A standard (1 x n) unit vector (all elements of **I** are 1).

Define **V** as the (1×1) final estimate of value of the subject property, and **W** as a ($1 \times n$) vector of the weights attached to each of the n comparable property. In terms common to appraisal textbooks, **P** is measured in appropriate units of comparison (price per square foot, per acre, etc.), while **X** and **Z** contain appropriate elements of comparison (location, physical characteristics, financing terms, etc.).[1]

Now, the sales comparison approach can be expressed in linear algebra terms as a system of three equations:

$$\mathbf{S} = \mathbf{P} + \mathbf{A}(\mathbf{XI} - \mathbf{Z}), \tag{1}$$

$$\mathbf{V} = \mathbf{WS}', \tag{2}$$

$$\mathbf{WI}' = 1. \tag{3}$$

Equation (1) defines a vector containing the *n* indicated values of a subject property as the adjusted selling price of each of *n* comparable properties. The net adjustment for each of the *n* comparable properties in Equation (1) is the sum of the product of the difference (between the subject and comparable property) in each of *j* characteristics and the adjustment factor for that characteristic. The unit vector in the parenthesis in Equation (1) is multiplied by the vector of subject property characteristics in order to form a matrix of the appropriate dimensions for use in calculating the differences between the property characteristics of subject and comparable property.

In Equation (1), the indicated price derived from each of the *n* comparable properties is *not* forced to be the same. This practice is universal among appraisers who use the comparable sales approach. That is, appraisers obtain a different indicated value from each comparable property included in the approach. Indeed, appraisers are generally taught to expect to find a different indicated value from each comparable property.

Equation (2) defines the final indicated value of the subject property as the weighted average of the indicated values of *n* comparable sales. Appraisers subjectively determine the weights in **W**. Although, appraisers are taught that the weights they select should reflect the relative strength of each comparable property in contributing to the final estimate of value.

Equation (3) simply forces the weights used in Equation (2) to add up to one, so that the final indicated value lies somewhere between the highest and lowest indicated values derived from each of the n comparable sales. Equations (2) and (3) constitute the appraisal practice called reconciliation.

Equations (1) − (3) are all fully deterministic. That is, appraisers assume that none of the terms are random variables. An appraiser observes **P**, **X** and **Z** and has prior knowledge of **A** and **W**. Thus, **S** is fully determined by the know values of **P**, **X**, **Z** and **A**. The weights, **W**, are set by the appraiser after calculating **S**, and **V** is fully determined by the previously calculated value of **S** and the appraiser's expert judgment about the weights, **W**. Appraisers do not view any of the terms in Equations (1) − (3) as having some sort of underlying probability distribution. Instead, they view all of the terms as fully determined by a combination of market data and their expert opinions.

The Effect of Multiple Indicated Values

In general, a unique numerical solution for **S**, **A** and **V** is not possible in the above model without information beyond **P**, **Z** and **X**. Given observations of: (1) the selling prices of the comparable properties (**P**); (2) the characteristics of the comparable properties (**Z**); and (3) the characteristics of the subject property (**X**), a unique numerical solution for **V** *cannot* be found. The reason for this confounding result is very simple; there are *always* too many unknowns. Specifically, Equation (1) contains $n + j$ unknowns (the *n unique* indicated values derived from the n comparable properties plus the adjustment factors for the j characteristics) in n equations (one equation for each comparable property). So, as long as $j > 0$, an exact solution for **P**, **A** and **V** can never be found no matter how big n (the number of comparable properties) becomes. Every comparable property added to the model brings with it at least one characteristic ($j > 0$) that needs adjustment plus at least one *unique* indicated value. At least one property characteristic will always need adjustment, because no two properties are exactly alike. Two identical structures cannot occupy the same space at the same time; they will always be in different locations. Even repeat sales of the same property have differences in the time of sale for which adjustments must be made. Thus, as long as the size of **S** is greater than 1×1 (*i.e.*, multiple indicated values), a unique numerical solution cannot be found in the appraisers' model of the sales comparison approach.

Estimates of the Adjustment Factors

In the sales comparison approach, a unique solution (final estimate of value) requires additional information beyond the information contained within the recently sold comparable properties and the characteristics of the subject property. The appraisal textbooks deal with this problem by instructing appraisers to bring in additional information in the form of estimates of the adjustment factors from outside the realm of the comparable properties being used in the approach. Often, these estimates consist of an "educated guess" referred to as "judgment" or "expert opinion" in the appraisal literature. Some techniques are suggested for making these estimates, such as cost analysis, graphic analysis, sensitivity analysis, etc. One of these techniques, paired data analysis, is very popular among appraisers and deserves further attention.

Paired Data Analysis

Equation 1 can be modified to illustrate the mathematical equivalent of the paired data analysis technique. In paired data analysis, the appraiser calculates an adjustment factor from two recently sold properties that are alike in all characteristics except one. Assume two sales of properties that are identical in all respects except one, the characteristic for which an adjustment factor will be calculated. To keep things simple, the terms in Equation (1) are redefined as follows:

$\mathbf{S} = S_1$, the observed selling price of the first property;
$\mathbf{X} = X_1$, the observed amount of the characteristic for which an adjustment factor is being calculated for the first property;
$\mathbf{P} = S_2$, the observed selling price of the second property;
$\mathbf{Z} = X_2$, the observed amount of the characteristic for which an adjustment factor is being calculated for the first property; and
$\mathbf{A} = A$, the adjustment factor.

Now, solve Equation (1) for A and get:

$$A = (S_1 - S_2)(X_1 - X_2)^{-1}. \tag{4}$$

The solution to Equation (4) is scalar, because in the paired data analysis technique, adjustment factors are found one at a time. As long as $X_1 \neq X_2$, the solution for A exists. In other words, the paired data analysis technique is consistent with Equation (1).

The Academic Model

The major difference between how appraisers and academicians view the sales comparison approach pivots on something called the one-price assumption. Academicians typically assume that there is a single, scalar value for S and then build models in which the academicians make additional assumption about the nature of S to estimate it. Appraisers, on the other hand, use a model in which no explicit assumption is made regarding the distribution of S. Instead, appraisers calculate various values of S that are not constrained to be the same. Equation (1) can be easily alterd to fit the academicians view by constraining all of the indicated values to be the same. Equation (1) can be modified to represent the academic model by defining S as the indicated value to be estimated in the academic version of the model.[2] This produces the following single equation model:

$$S\mathbf{I} = \mathbf{P} + \mathbf{A}(\mathbf{XI} - \mathbf{Z}). \tag{5}$$

Equations (2) and (3) in the appraisers' model are not necessary in the academic model, because in the academic model S, the indicated value, is assumed to be scalar. Subsequently, there is no need to include the weights, $\mathbf{W}$, in the academic model. Moreover, so far, no assumptions are necessary about the distribution of any of the terms in Equation (5).

Equation (5) can also be examined as a system of simultaneous equations with $j + 1$ unknowns (the j adjustment factors and the indicated value of the subject property) in n equations (the n comparable properties). Consider also how to determine the values of $\mathbf{A}$ and S that must simultaneously arise from the observed values for $\mathbf{P}$, $\mathbf{X}$ and $\mathbf{Z}$. In general, there are three special cases in which something can be said about the numerical solutions for $\mathbf{A}$, and S:

1. When $(j + 1) > n$, no unique numerical solution for $\mathbf{A}$ and S exists and additional information is needed to find a unique numerical solution, thus Equation (5) is *overidentified.*
2. When $(j + 1) = n$, a unique numerical solution for $\mathbf{A}$ and S can exist, thus Equation (5) is *just identified*.[3]
3. When $(j + 1) < n$, multiple numerical solutions for $\mathbf{A}$ and S exist, thus Equation (5) is *underidentified*.

Each of these three cases is examined in greater detail below.

Case 1: Overidentification

In Case 1, the number of unknowns $(j + 1)$ exceeds the number of equations (the number of comparable sales). This situation also exists in the appraisers' model, although with greater severity. The only way a solution can be obtained is to introduce additional information from outside Equation (5). Because the indicated value of the subject property is the end objective of the model, the only sort of information that can give us a solution is some independent observations of the adjustment factors. This is the same situation the appraisers face. But, in the academic model less information is needed than in the appraisers' model.

In Equation (5), observations of only $(j - n + 1)$ adjustment factors are needed in order to be able to obtain a system of equations with a unique solution. Estimates (subjective or objective) of *all* of the adjustment factors are not needed in order to make Equation (5) just identified (*e.g.*, have a unique numerical solution). Estimates are only needed for enough of the adjustment factors to make $(j + 1) = n$. The magic number of adjustment factors needed to make Equation (5) just identified is $(j - n + 1)$. For example, there are three comparable properties that call for eight adjustment factors to obtain an indicated value of the subject property, we need estimates of only $(8 - 3 + 1)$ or six of the adjustment factors. In other words, $n - 1$ (two in the above example) of the adjustment factors can be estimated *simultaneously* along with the indicated value of the subject property using Equation (5).

There is no need to estimate all of the adjustment factors involved in the academic model in order to calculate a unique indicated value of the subject property. But, in the appraisers' model, estimates of *all* of the adjustment factors are necessary, which all but guarantees a different indicated value for each comparable property. When this happens, the appraiser must bring in Equations (2) and (3) (reconciliation) to arrive at a final estimate of value. Thus, the appraisers' use of multiple and different indicated values forces them to use Equations (2) and (3). If appraisers would adopt the one-price assumption, then their use of judgment could be reduced.

Case 2: Just Identification

Just identification is the result when Equation (5) meets the condition that $(j + 1) = n$. In other words, when there is *one more* comparable property than the number of characteristics, a unique numerical solution can be found for the adjustment factors as well as the indicated value of the subject property using Equation (5).

This result suggests that the more comparable properties the better. But, this may not always be the case. As more comparable properties are added to the mix, the number of characteristics needing adjustment often increases. Adding more comparable properties can do more harm than good, if doing so also increases the number of characteristics that need adjustment by more than one. Indeed, if the number of characteristics needing adjustment (j) increases faster than the number of comparable properties (n), the appraiser's work unnecessarily increases, because the number of adjustment factors that must estimate also increases.

Occasionally, appraisal textbooks and appraisal course workbooks contain just identified cases to illustrate the sales comparison approach. These illustrations are useful in teaching the sales comparison approach, because they can be solved without any additional information or appraiser judgment. But, as a practical mater, the just identification case is rarely seen in actual appraisal reports, because in practice, appraisers do not use the one-price assumption made to get Equation (5).

Case 3: Underidentification

In the event that $(j + 1) < n$, there are more than enough comparable properties to find a solution. Indeed, if all possible combinations of the comparable properties for which $(j + 1) = n$ are used, there would be many numerical solutions.[4] So, the challenge when there is underidenfication is to find some way to combine all of the many solutions into a single best solution. This is the case favored by academicians, who approach the challenge by using numerous statistical methods to extract an estimate of the unknowns (along with many other interesting statistics) from a set of real estate sales data.

Statistical Techniques

Academicians prefer the underidentified case, because it allows them to make use of various statistical techniques to estimate **A** and S. Equation (5) can be transformed into the equivalent of an OLS equation by assuming that the value of the subject property is equal to the sum of the products of the characteristics of the subject property (**X**) times the corresponding adjustment factor for each characteristic (**A**), or that $S = \mathbf{AX}$. Now, **AX** can be substituted for S on the left hand side of Equation (5). The **AXI** terms cancel, producing:

$$\mathbf{P} = \mathbf{AZ}. \tag{6}$$

Adding a well-behaved error term to the right hand side of Equation (6) and assuming the appropriate distributional forms for **P**, **A** and **Z,** yields a traditional OLS model, from which **A** can be estimated given observations of **P** and **Z**. After OLS estimates of **A** are obtained, the price (value) of the any property with characteristics **X** can be calculated using $S = \mathbf{AX}$.

Unfortunately, there are still many subjective judgments for academicians to make. Of course, the statistical technique that most academicians prefer is multiple regression analysis using the OLS technique. This technique is very popular with academicians because, given well-behaved data, OLS parameter estimates are best linear unbiased estimates (BLUE),[5] and BLUEs are highly prized by academicians.

However, even academicians must employ some judgment when selecting and applying OLS to well-behaved real estate sales data. For example, the best functional form for use in multiple regression analysis is difficult to determine. Exactly which property characteristics (elements of comparison) to include in a multiple regression model are not easy to determine, especially when dependent on sales data collected by other parties (*i.e.*, MLSs, tax officials, etc.). Whether or not to apply factor or principle component analysis to the property characteristics has also been hotly debated by academicians. Variants of multiple regression analysis, such as stepwise regression analysis, ridge regression, etc. have also been debated by academicians. Indeed, there appears to be as many (or more) judgments for academicians to ponder within the underidentified case as there are judgments for appraisers to make in their version of the overidentified case.

Some Insights

The mathematical expression of the sales comparison approach from the point of view of appraisers and academicians sheds considerable light on the practices of each, and why appraisers view the estimation of value very differently than do academicians. The key difference between appraisers and academicians can be found in the assumptions they make about the nature of the value of the subject property. Appraisers work within the world of multiple indicated values, thereby forcing them to bring in their expert judgments regarding the adjustment factors and weights in Equation (1). Most academicians embrace the one-price assumption, which enable them to apply their expert judgments in the use of various statistical techniques and methods. But, the mathematical expressions presented provide considerable insight into the ways in which both appraisers and academicians might improve their work.

First, appraisers could improve their view of the sales comparison approach by adopting the one-price assumption used by academicians. By adopting this assumption, appraisers would need to estimate fewer adjustment factors. Indeed, $n - 1$ of the adjustment factors need not be estimated in the just identified case. An appraiser could allow the $n - 1$ most difficult to estimate adjustment factors to be estimated simultaneously with an (unique) estimate of value. Moreover, this process would allow appraisers to completely remove the reconciliation steps from the sales comparison approach.

Academicians could also benefit from a better understanding of the simple linear algebra of the sales comparison approach. For example, academicians have yet to examine the just identification case when given multiple comparable properties. Although the number of potential exact solutions is very large, the distribution of these solutions could prove very interesting. Given the speed and versatility of computers, an exercise of this sort is certainly not beyond the realm of possibilities.

Conclusion

This study presents mathematical models of the sales comparison approach as viewed by appraisers and academicians. Examination of these two models highlights why appraisers and academicians view the estimation of value from real estate sales data so differently. Specifically, academicians adopt a key assumption (one price) that enable them to apply a wide variety of statistical estimation techniques to a set of real estate sales data. Appraisers do not adopt this assumption. Instead, they seem to prefer using very few comparable properties in order to sell their expert judgments regarding adjustment and weighting factors. In any event, both appraisers and academicians could benefit from a better understanding of the models presented in this study.

The task of appraisers could be made much easier if they were to adopt the academicians' one-price assumption. Adoption of this assumption would reduce the number of adjustment factors that the appraiser must estimate. The task of academicians could be improved by exploring techniques for the estimation of value other than multiple regression analysis. In any event, the fundamental real estate valuation problem can be understood better by making use of the findings of this study.

Endnotes

1 Academicians should note that appraisers place the characteristics of the properties in column vectors, rather than in row vectors.

2 The rows and columns could also be transposed, but doing so adds an additional layer of complexity. So, for the sake of simplicity, the rows and columns will remain the same as defined in the appraisers' model.

3 The just identified case also requires that the inverse of ($\mathbf{XI} - \mathbf{Z}$) exist. Throughout the discussion of the just identified case, it is assumed that this inverse exits.

4 The number of possible combinations of n objects taken $(j + 1)$ at a time is $n!/(j + 1)!(n - j - 1)!$

5 See Thiel (1971) or any other test on multiple regression analysis for a discussion and proofs of BLUE.

References

Boronico, J. S. and D. M. Moliver, Appraisal Reliability and the Sales Comparison Approach, *The Appraisal Journal*, 1997, 331–36.

Colwell, P. F., R. E. Cannaday and C. Wu, The Analytical Foundations of Adjustment Grid Methods, *Journal of the American Real Estate and Urban Economics Association*, 1983, 11:1, 11–29.

Crookham, J., Sales Comparison Approach: Revisited, *The Appraisal Journal*, 1995, 177–81.

Diaz, J. III., The Process of Selecting Comparable Sales, *The Appraisal Journal*, 1990, 533–40.

Gau, G. W., T-Y. Lai and K. Wang, Optimal Comparable Selection and Weighting in Real Property Valuation: An Extension, *Journal of the American Real Estate and Urban Economics Association*, 1992, 20:1, 107–24.

——., A Further Discussion of Optimal Comparable Selection and Weighting, and A Response to Green, *Journal of the American Real Estate and Urban Economics Association*, 1994, 22:4, 655–63.

Green, R. K., Optimal Comparable Weighting and Selection: A Comment, *Journal of the American Real Estate and Urban Economics Association*, 1994, 22:4, 647–54.

Isakson, H. R., The Nearest Neighbors Appraisal Technique: An Alternative to the Adjustment Grid Methods, *Journal of the American Real Estate and Urban Economics Association*, 1986, 14:2, 274–86.

——., Valuation Analysis of Commercial Real Estate Using the Nearest Neighbors Appraisal Technique, *Growth and Change*, 1988, 19:2, 11–24.

Lentz, G. and K. Wang, Residential Appraisal and the Lending Process: A Survey of Issues, *Journal of Real Estate Research*, 1998, 15, 11–39.

Manaster, M. S., Sales Comparison Approach: A Comparative Analysis of Three Appraisal Reports on the Same Property, *The Real Estate Appraiser*, 1991, May, 12–26.

Pace, R. K., Total grid estimation, *Journal of Real Estate Research*, 1998, 15:1/2, 101–14.

Spence, M. T. and J. Thorson, The Effect of Expertise on the Quality of Appraisal Services, *Journal of Real Estate Research*, 1998, 15:1/2, 205–15.

Thiel, H., *Principles of Econometrics*, New York, NY, 1971.

Vandell, K. D., Optimal Comparable Selection and Weighting in Real Property Valuation, *Journal of the American Real Estate and Urban Economics Association*, 1991, 19:2, 213–39.

Hans R. Isakson, University of Northern Iowa, Cedar Falls, IA 50614-0124 or Hans.Isakson@uni.edu.