An Empirical Investigation of Four Market-Derived Adjustment Methods

*Joseph B. Lipscomb**
*J. Brian Gray***

Abstract. **This study uses published data on 422 market sales of FHA/VA insured/guaranteed houses to examine and compare four methods of estimating market-derived adjustment values to be employed in the sales comparison appraisal approach. These four adjustment methods are variations and combinations of matched pair and multiple regression analysis. Two major conclusions drawn from the results are: (1) regression on a matched pair data set is equivalent to matched pair analysis using regression coefficients as secondary adjustments and produces the same primary adjustment estimate for the feature of interest, and (2) even under relatively ideal circumstances, market-derived adjustments contain a high degree of uncertainty.**

Introduction

The sales comparison approach to real estate appraisal is the method most heavily relied upon by the appraisal and lending industries to estimate the market value of a single-family residence. In the sales comparison approach, the appraiser is required to find at least three similar single-family residences located near the subject property that have sold recently. These houses are referred to as *comparable sales*. Where there are differences between a comparable sale and the subject property, adjustments must be made to the comparable sale to estimate what price that property would have sold for had it been identical to the subject. The method of determining the size of these adjustments is the topic of this research.

Matched pair or *paired sales analysis* is the adjustment method currently recommended by the major appraisal organizations. The objective in matched pair analysis is to isolate two properties whose only difference is the property feature in question. Theoretically, the difference in the sale prices of these two properties can be attributed solely to that one feature and represents the adjustment amount for the presence (or absence) of that feature. The reality is that sale prices are affected by factors other than the physical characteristics reported in an FHA appraisal. For example, the sale price for a single-family residence is dependent to some degree on the relative negotiating skills of and the levels of information held by the buyer and seller.

*Texas Christian University, P.O. Box 32868, Fort Worth, Texas 76129.
**University of Alabama, P.O. Box 870226, Tuscaloosa, Alabama 35487-0226.
Date Revised—January 11, 1990; Accepted—March 30, 1990.

Appraisers often rely heavily on their experience and subjective judgement in making adjustments. Pressure is being put on the appraisal industry to use adjustments that are derived from market data, based on factual information, and supportable in court if necessary [3]. This issue has been addressed recently by Grissom, Robinson and Wang [4]. Other recent proposals include those made by Cronan, Epley and Perry [2] and Kroll and Smith [6]. However, the standard matched pair technique and regression analysis currently remain the primary methods utilized by the appraisal industry.

Our study examines four existing approaches to estimating adjustment amounts from published data on actual sales. These adjustment methods are variations and combinations of matched pair and multiple regression analysis. In the next section, we discuss matching in general and the particular form of matching we use in conjunction with three of the four adjustment techniques described in section three. In section four, we apply these techniques to a data set of 422 sales to individually estimate adjustment values for four property features. We conclude by drawing comparisons among the four adjustment methods and discussing the precision of their estimates.

Matching

Three of the four adjustment methods described in the next section require a matching phase. In this section, we discuss the concept of matching and the particular features of the matching approach used in our analysis of the data in section four. The reader should note that "matching" in this study refers to the selection of pairs to estimate adjustment values as opposed to finding the best comparable sales for an FHA appraisal.

For any collection of property sales, a binary feature of interest identifies two subgroups, those sales with the feature of interest and those sales without the feature. The smaller group of n sales (usually without the feature of interest) is called the *subject sales* set. The larger group of m sales is referred to as the set of *potential matching sales* or *potential matches*. The goal in matching is to find the most similar match among the potential matches for each subject sale based on important *physical characteristics* (*explanatory variables*).

In the event of a "perfect" match, the difference in sale prices between the subject sale and its match can be attributed to the feature of interest and to variation not accounted for by the explanatory variables (i.e., noise), assuming that no important explanatory variables have been left out. Given that perfect matches are rare, a measure of the difference (dissimilarity) between a subject sale and a potential match is needed to judge the closeness of a match.

Since physical characteristics such as square footage and number of bedrooms are expressed in very different units of measurement, a common first step before measuring dissimilarity is to standardize the values of the explanatory variables [11] as follows:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_i}{S_i}, \quad (1)$$

where Z_{ij} is the standardized value of the jth observation on the ith explanatory variable (zero mean and unit standard deviation), X_{ij} is the jth observation on the ith explanatory variable, $\overline{X}_i$ is the average of the ith explanatory variable across all $N=n+m$ observations, and S_i is the standard deviation of the ith explanatory variable across all N observations $(i=1, 2, \ldots, p; j=1, 2, \ldots, N)$.

Secondly, since some explanatory variables are more important than others, their contributions to the dissimilarity measure should be weighted accordingly. Instead of using subjective belief to arrive at importance weights [11], we perform a forward regression of sale price on the explanatory variables (not including the feature of interest) for the entire data set and use each variable's sequential contribution to R^2 as its importance weight in the dissimilarity calculation. Others have proposed using the regression coefficients as importance weights in matching [10].

The measure of dissimilarity between the jth subject sale $(j=1, 2, \ldots, n)$ and the kth potential match $(k=1, 2, \ldots, m)$ used in this study is given by

$$d_{jk}=\sum_{i=1}^{p} w_i\,(Z_{ij}-Z_{ik})^2, \tag{2}$$

where w_i is the importance weight given to the ith variable, Z_{ij} is the standardized value of the ith explanatory variable for the jth subject sale, and Z_{ik} is the standardized value of the ith explanatory variable for the kth potential match. This dissimilarity measure is the squared weighted euclidean distance between the subject sale and its potential match based on their standardized explanatory variable values.

An optimal matching of subject sales to potential matches that minimizes the total dissimilarity of the matching can be found using the *assignment method* [7]. The assignment method is needed because each sale in the set of potential matches must be used no more than once to preserve independence in the paired price differences. However, when the number of potential matches m is large compared to the number of subjects n, the so-called "greedy" approach (sequentially choosing the closest remaining match at each step until each subject sale is matched) usually comes very close to the optimal solution found by the assignment method. Rosenbaum [8] and Rubin [9] provide further discussion on matching units in an observational study.

Adjustment Methods

The four adjustment methods under investigation in this study are:

1. *Matched Pair Mean Adjusted Difference (MP–*$\overline{\textit{Adj D}}$*)*
2. *Matched Pair Regression (MP–MRA)*
3. *Matched Pair Differences Regression (MP–D–MRA)*
4. *Regression on All Sales (MRA)*

All four methods involve regression, and the first three also involve a matching phase. Both regression and matching are methods of conditioning (i.e., adjusting) for the effects of sale price determinants of secondary interest so that the effect of the primary feature of interest is revealed. Each approach has its merits and disadvantages. Regression assumes a particular form of the relationship (usually linear) between sale price and its determinants and may be subject to the bias of model misspecification. On the other hand, matching is a model-free approach, but its results can be compromised by poor matches. Thus, matching often requires a larger data set than that needed for regression when the correct form of the regression model is known.

To make the exposition of the four methods clearer, it is necessary to introduce some additional notation. Let Y represent sale price and let X_i represent the ith explanatory variable (physical characteristic), $i = 1, 2, \ldots, p$. After defining a feature of interest for which an adjustment is to be estimated, matching is used to generate n pairs of subject and matched sales. The notation given below is useful in denoting quantities related to a given matched pair. (The subscript identifying the matched pair is unnecessary and has been left out to avoid clutter.)

$$W = \text{dummy variable for feature of interest} = \begin{cases} 0, & \text{for the subject sale} \\ 1, & \text{for the matched sale,} \end{cases}$$

Y_0 = sale price of the subject sale,
Y_1 = sale price of the matched sale,
D = $Y_1 - Y_0$ = price difference for the matched pair,
X_{i0} = value of the ith explanatory variable for the subject sale,
X_{i1} = value of the ith explanatory variable for the matched sale,
X'_i = $X_{i1} - X_{i0}$ = ith explanatory variable difference for the matched pair.

We now discuss each of the four adjustment methods in detail.

Matched Pair Mean Adjusted Difference ($MP\text{–}Adj\,\overline{D}$): This is the standard matched pair analysis approach (traditional appraisal grid approach) in which the average price difference between perfectly matched pairs is used to estimate the adjustment for the feature of interest (i.e., the *primary adjustment*). If the two sales of a matched pair are not perfectly matched in every feature except the feature of interest, then *secondary adjustments* are made to compensate for these unwanted differences. Since there were no perfect matches in this study, all pairs required secondary adjustments to construct the equivalent of perfect matches. After these secondary adjustments, the remaining differences in matched pair sale prices are averaged to produce the primary adjustment estimate.

One source of secondary adjustment values is the set of coefficients from a regression of sale price on the explanatory variables calculated for the $2n$ sales in the matched pair data set [1]. Sequential analysis [4, 5] is an alternative, but is not considered here because it does not make full use of the available data in deriving secondary adjustments. See Colwell et al. [1] for a further discussion of adjustment grid methods.

More specifically, once the matching process has been carried out to obtain the n pairs of subject and matched sales, the $MP\text{–}Adj\,\overline{D}$ computation is performed in the following series of steps. First, the regression for the $2n$ sales in the matched pairs data set is computed:

$$\hat{Y} = \hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \ldots + \hat{\alpha}_p X_p + \hat{\alpha}_w W. \tag{3}$$

The dummy variable W is included in the regression to condition on the primary feature of interest so that the other parameter estimates (secondary adjustments) are not biased by its effect. The coefficient $\hat{\alpha}_i$ can be interpreted as a per unit secondary price adjustment for the ith explanatory variable. Second, a difference grid is created by differencing the values of each explanatory variable for the subject sale and its match in each pair. The primary adjustment feature is excluded from this grid since its value is represented by the final adjusted price difference. Third, multiplying the differences $X'_i = X_{i1} - X_{i0}$ in the grid by the corresponding regression coefficients α_i and summing the results provides the total net secondary adjustment due to differences in the p explanatory variables, given by:

$$\hat{\alpha}_1 X'_1 + \hat{\alpha}_2 X'_2 + \ldots + \hat{\alpha}_p X'_p. \tag{4}$$

Fourth, this total net secondary adjustment is subtracted from the price difference D which results in the adjusted price difference for that pair:

$$Adj\ D = D - [\hat{\alpha}_1 X'_1 + \hat{\alpha}_2 X'_2 + \ldots + \hat{\alpha}_p X'_p]. \tag{5}$$

Finally, these adjusted price differences are averaged over the n matched pairs to produce a primary adjustment estimate for the feature of interest, which is denoted by $\overline{Adj\ D}$.

A major deficiency with the MP–$\overline{Adj\ D}$ approach is the difficulty involved in estimating the standard error of the resulting adjustment estimate. Although the price differences D are independent of each other, the adjusted price differences $Adj\ D$ are not independent because the secondary adjustments used to compute $Adj\ D$ involve all the paired data. As such, the standard error of the primary adjustment $\overline{Adj\ D}$ cannot be estimated directly from the sample standard deviation of the $Adj\ D$ values.

Matched Pair Regression (MP-MRA): The second adjustment approach is based on the multiple regression model in equation (3) with sale price as the dependent variable, the physical characteristics as the independent variables, and a dummy variable W for the feature of interest. As before, the regression is computed from the $2n$ individual sales contained in the matched pair data set. In the *MP–MRA* approach, however, the dummy variable coefficient $\hat{\alpha}_w$ serves as the estimate of the primary adjustment for the feature of interest. The standard error of this estimate is provided in the output of most regression packages.

It is important to note that the dummy variable regression coefficient from *MP–MRA* is always equal to the adjustment derived from the traditional MP–$\overline{Adj\ D}$ approach using regression coefficients for secondary adjustments, i.e., $\hat{\alpha}_w = \overline{Adj\ D}$. This fact is significant and useful for two reasons. First, it shows that multiple regression with a dummy variable, which is comparatively simple to perform, provides the same adjustments as the more cumbersome grid process with regression coefficient adjustments. Secondly, the standard error of the primary adjustment estimate, a very valuable piece of information for judging the accuracy of an estimate, is readily available from the output of the *MP–MRA* approach but is inaccessible from the MP–$\overline{Adj\ D}$ method.

Matched Pair Differences Regression (MP–D–MRA): The third approach is a multiple regression of the sale price difference within a matched pair on the differences in the explanatory variables for the pair. This approach has conceptual appeal and has been studied by Rubin [9] and employed by Isakson [5] in an appraisal adjustment study. For the n matched pairs, the regression of differences is computed as follows:

$$D' = \hat{\gamma}_0 + \hat{\gamma}_1 X'_1 + \ldots + \hat{\gamma}_p X'_p. \tag{6}$$

The intercept term $\hat{\gamma}_0$ estimates the adjustment for the feature in question. Since the difference in the dummy variable W is equal to 1 for all pairs of sales in the matched sample, the intercept term represents the primary adjustment. The standard error of this estimate can be found in the output of most regression packages.

Regression on All Sales (MRA): The fourth approach uses a regression of sale price on the explanatory variables (including W) to estimate the effect of an adjustment factor on sales price. The regression is computed for all $N=n+m$ sales as follows:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_p X_p + \hat{\beta}_W W. \tag{7}$$

The quantity $\hat{\beta}_W$ estimates the sale price adjustment for the feature of interest.

In the next section, these four adjustment methods are applied to a collection of 422 sales to investigate the abilities of the techniques to estimate sale price adjustments for four features.

Data Analysis

The data were taken from the Durham Report, a local reporting service with the sole purpose of providing FHA and VA comparable sales data in Tarrant County, Texas. This report includes most of the FHA and VA sales in the county and contains every piece of information necessary to use any of these sales as a comparable sale in an FHA appraisal.

Standard appraisal practice requires adjusting for the effect on sale prices of (1) favorable financing, (2) condition of sale, (3) market condition, and (4) location and physical characteristics. All sales in this study were financed with FHA or VA mortgage loans provided by third-party lenders, and points were within the customary range Therefore, no financing adjustments were made. It is the practice of the Durham Report to report only sales that meet standard arms-length sale conditions. Therefore, no adjustments are needed for condition of sale. The sales occurred between July 1985 and March 1986. An analysis of the price per square foot over time indicated that date of sale had no discernible effect on sale price during this period. As a result, no market condition (time) adjustment was required. The effects of location were controlled by limiting the sales data to selected MAPSCO coordinates in southwest Arlington, Texas. These contiguous coordinates were carefully chosen for similarity of the houses and neighborhood characteristics and to avoid pockets where houses deviate greatly from the average FHA-insured house in the area. Location was further controlled by heavily emphasizing closeness in physical distance during the matching process. The differences in physical characteristics are dealt with directly through the adjustment methods described in the previous section. A summary of the sale prices and physical characteristics of the houses in the sales data set are given in Exhibit 1.

In this study, four physical features (brick exterior, garage type, fireplace, and half-bath) were selected as examples for adjustment. Since our objective is to investigate the precision of the four adjustment methods, we chose four binary features that had the minimum number of subjects (n) required for model fitting. Continuous variables were not included because the added dimension of incremental size introduces additional subjectivity into the adjustment process. For a discussion of continuous variable adjustment, see Colwell et al. [1].

Exhibit 1
Descriptions of Variables

		SUMMARY STATISTICS		
		Mean	Median	Std. Dev.
DEPENDENT VARIABLE				
PRICE	= sale price	79,061	79,000	10,522
EXPLANATORY VARIABLES				
Continuous:				
SIZE	= living area (sq. ft.)	1436	1419	222
AGE	= age in years	3.47	3.00	1.69
		DESCRIPTION (1st Row) NUMBER OF SALES (2nd Row)		
Discrete:				
		2	3	4
BED	= number of bedrooms	29	372	21
		2	2.5	Other
BATH	= number of bathrooms	368	33	21
		1-Car	2-Car	Other
CAR	= garage type	32	386	4
		Excellent	Good	
COND	= condition of property	125	297	
Binary:				
		Yes	No	
BRICK	= brick exterior	407	15	
HALFSTRY	= one-and-a-half story	41	381	
TWOSTRY	= two-story house	30	392	
INSWDS	= insulated windows	207	215	
SPSYS	= sprinkler system	10	412	
MW	= microwave oven	68	354	
FIRE	= wood-burning fireplace	410	12	
FENCE	= fence	158	264	
OVEN	= oven	420	2	
DISH	= dishwasher	421	1	

Exhibit 2 reports the adjustment estimates for each of these features that were individually determined by the four adjustment methods described in the previous section. For the three estimates derived directly from regression coefficients, the associated t-statistics, degrees of freedom, and adjusted R^2 values are also reported.

Although four features are considered for adjustment, we present the analysis for one feature (brick exterior) in detail and then summarize the results for the other three features at the end of this section. In our data set of 422 sales, there are $n = 15$ non-brick (i.e., siding) houses in the subject set and $m = 407$ brick veneer houses in the set of potential matches. The first block of rows in Exhibit 2 contains the estimated adjustments for *BRICK* for each of the four adjustment methods. We now describe their calculations in detail.

First, the matching technique described in section two was used to find the most similar matches to the subject sales. The sequential R^2 contributions were calculated from the full data set with a forward selection regression of sale price on all explanatory variables in Exhibit 1, except *BRICK*. These values were then used as importance weights in calculating the dissimilarity measure in equation (2) for all $n \times m$ possible matches. Although location

Exhibit 2
Estimated Adjustment Value for Four Methods

	(1) Matched pair mean adjusted difference (*MP–Adj D*)	(2) Matched pair regression (*MP–MRA*)	(3) Matched pair differences regression (*MP–D–MRA*)	(4) Regression on all sales (*MRA*)
BRICK				
Estimate	6,764	6,764	6,980	6,213
Std. Error	N/A*	1,407	2,044	1,797
Sample Size	$n=15$	$2n=30$	$n=15$	$N=422$
Variables	$p=11$	$p=12$	$p=11$	$p=12$
t-statistic		4.809	3.414	3.457
Deg. Freedom		DF=17	DF=3	DF=409
Adjusted R^2		0.87	−0.36	0.65
GARAGE				
Estimate	4,723	4,723	3,967	6,657
Std. Error	N/A*	1,732	2,825	1,662
Sample Size	$n=32$	$2n=64$	$n=32$	$N=418$
Variables	$p=12$	$p=13$	$p=12$	$p=13$
t-statistic		2.726	1.404	3.231
Deg. Freedom		DF=50	DF=19	DF=404
Adjusted R^2		0.69	−0.055	0.65
FIREPLACE				
Estimate	2,657	2,657	3,170	179
Std. Error	N/A*	2,928	N/A**	2,072
Sample Size	$n=12$	$2n=24$	$n=12$	$N=422$
Variables	$p=11$	$p=12$	$p=11$	$p=12$
t-statistic		−0.907	N/A**	0.086
Deg. Freedom		DF=11	DF=0	DF=409
Adjusted R^2		0.75	N/A**	0.65
HALFBATH				
Estimate	1,099	1,099	1,460	231
Std. Error	N/A*	2,447	3,723	1,389
Sample Size	$n=33$	$2n=66$	$n=33$	$N=401$
Variables	$p=15$	$p=16$	$p=15$	$p=16$
t-statistic		0.449	0.862	0.116
Deg. Freedom		DF=49	DF=17	DF=384
Adjusted R^2		0.41	−0.09	0.59

*The standard error of the matched pair mean adjusted difference cannot be calculated for *MP–Adj D*.
**No degrees of freedom remaining for error estimate after fitting model [12].

was not included as a variable in the forward selection regression to determine importance weights, a proximity variable (i.e., physical distance between houses) based on MAPSCO coordinates was given equal weight with the highest weighted variable (*SIZE*) for matching purposes. Since $m=407$ was much larger than $n=15$, the greedy approach was employed to find the best matches. The majority of subjects were assigned matches in their MAPSCO cell (a 3000 ft. by 3000 ft. area). Thus, the effects of location were minimized.

After the $n=15$ matches were identified, the regression in equation (3) was calculated based on the thirty sales in the matched data set to estimate the secondary adjustments for the grid (*MP–Adj D*) technique. All of the explanatory variables in Exhibit 1 were used in this regression except *FENCE, DISH, SPSYS* and *OVEN*, which were constant for the thirty homes in the matched data set. Exhibit 3 shows the grid adjustment calculation for *BRICK* based on this regression.

Along the top of Exhibit 3 are the explanatory variables and their per unit secondary adjustments computed from the regression. Although most of the physical characteristic coefficients are reasonable, the coefficient for *MW* (Microwave) is −$5,283. Since only one house in the thirty houses of the *BRICK* matched set has a microwave, the *MW* coefficient is obviously not reliably estimated. The fifteen rows in the top half of Exhibit 3 show how each of the fifteen subject sales and their matches differ in sale price and in the explanatory variables. The variation in the fifteen price differences is quite large even though the matches are reasonably close as seen in Exhibit 3. Secondary adjustments should reduce this variability. Multiplying the explanatory variable differences by the per unit secondary adjustment values produces the fifteen rows at the bottom of Exhibit 3. Within each row, the secondary adjustments are subtracted from the unadjusted price difference D shown in the leftmost column to produce the adjusted difference $\overline{Adj\ D}$ shown in the rightmost column. The average of the adjusted differences, $\overline{Adj\ D}$ = $6,764, is shown in the lower right-hand corner of Exhibit 3 and the first row of Exhibit 2. The standard deviation of the fifteen adjusted price differences ($3,881) is smaller than for the unadjusted differences ($5,447), but still large for practical purposes. As explained in section three, an estimate of the standard error of the $MP\text{–}\overline{Adj\ D}$ estimate cannot be directly computed from the adjusted price differences since these differences are not independent. We now discuss the calculation of the *BRICK* adjustment using the other three methods.

The matched pair regression approach (*MP–MRA*) uses the same regression calculated in the previous method for estimating secondary adjustments. The estimated coefficient for *BRICK* is $6,764 as reported in Exhibit 2. (The other regression coefficients were previously given at the top of Exhibit 3.) Note that the primary adjustment estimates for the first two methods are equal as discussed in section three.

In the matched pair differences regression (*MP–D–MRA*) method, the fifteen unadjusted price differences of the *BRICK* matched pair data set are regressed on the corresponding differences in the explanatory variable values. The estimate of $6,980 reported in Exhibit 2 is the intercept of this fitted model.

The final method considered, regression on all sales (*MRA*), involves a regression of all $N=n+m$ sale prices on the same twelve physical characteristics used in the *MP–MRA* approach, including the dummy variable for *BRICK*. The same variables were used in both methods for comparability reasons. The adjustment estimate of $6,213 and its estimated standard error do not differ greatly from those of the other techniques.

The four adjustment methods were also applied to three other features of interest: one-car garage ($n=32$) vs. two-car garage ($m=388$), two and one-half baths ($n=33$) vs. two baths ($m=368$), and no fireplace ($n=12$) vs. fireplace ($m=410$). The adjustment estimates for these features derived from the four methods are presented in Exhibit 2. Note that n and m do not always add to 422 (e.g., to account for the one-bath houses not included in the half-bath analysis). Also note that the number of variables, p, used in the regressions varies across the four features. In each *MP–MRA* regression, all explanatory variables listed in Exhibit 1 were used with the exceptions of the feature of interest and any variables that were constant across the matched data set. Each *MRA* regression used the same variables as the corresponding *MP–MRA* regression.

The regression on all sales method (*MRA*) appears to provide more precise estimates (as

Exhibit 3
Matched Pair Analysis for Brick Adjustment

		CAR	SIZE	AGE	BED	BATH	COND	HALFSTRY	TWOSTRY	INSWDS	MW	FIRE
UNIT SECONDARY ADJUSTMENTS→		2799	26.88	−2557	3532	1040	−3177	−5940	−7456	−869	−5283	2652
		MATCHED PAIR DIFFERENCES GRID										
PAIR	PRICE DIFFERENCE											
1	2450	1	30	0	0	0	0	0	0	0	0	0
2	4050	0	−66	1	0	−0.5	0	0	0	0	0	−1
3	9300	0	73	0	0	1	−1	0	0	0	0	0
4	12050	1	−4	−1	0	1	0	0	0	1	0	0
5	4000	0	110	2	1	0	0	0	0	−1	0	−1
6	2450	0	−27	0	0	−1	0	0	0	1	0	0
7	7755	1	−79	−1	0	0.5	1	−1	1	0	0	0
8	9000	0	24	−4	0	1	1	0	0	1	0	−1
9	11200	−2	88	0	0	0	0	0	0	0	0	1
10	17300	1	−60	0	1	0	0	0	0	0	0	0
11	9240	1	24	0	1	0	0	0	0	0	1	0
12	7050	1	29	−1	−1	0.5	0	−1	1	0	0	0
13	18050	1	129	0	0	1	0	0	0	0	0	0
14	8250	0	−41	1	−1	0	0	0	−1	0	0	0
15	−2135	0	−12	0	0	0	0	1	0	−1	0	0

Exhibit 3 (Continued)

	UNADJUSTED PRICE DIFFERENCE (D)	ADJUSTMENT GRID — SECONDARY ADJUSTMENTS											ADJUSTED PRICE DIFFERENCE (*Adj D*)
1	2450	2799	806	0	0	0	0	0	0	0	0	0	−1155
2	4050	0	−1774	−2557	0	−520	0	0	0	0	0	−2652	11553
3	9300	0	1962	0	0	1040	3177	0	0	0	0	0	3121
4	12050	2799	−108	2557	0	1040	0	0	0	−869	0	0	6631
5	4000	0	2957	−5114	3532	0	0	0	0	869	0	−2652	4408
6	2450	0	−726	0	0	−1040	0	0	0	−869	0	0	5085
7	7755	2799	−2124	2557	0	520	−3177	5940	−7456	0	0	0	8696
8	9000	0	645	10228	0	1040	−3177	0	0	−869	0	−2652	3785
9	11200	−5598	2365	0	0	0	0	0	0	0	0	2652	11781
10	17300	2799	−1613	0	3532	0	0	0	0	0	0	0	12582
11	9240	2799	645	0	3532	0	0	0	0	0	−5283	0	7547
12	7050	2799	780	2557	−3532	520	0	5940	−7456	0	0	0	5442
13	18050	2799	3468	0	0	1040	0	0	0	0	0	0	10743
14	8250	0	−1102	−2557	−3532	0	0	0	7456	0	0	0	7985
15	−2135	0	−323	0	0	0	0	−5940	0	869	0	0	3259
												$\overline{Adj\ D}$	6764

judged by estimated standard errors) than the other three methods. However, the *MRA* estimates deviate substantially from those of the first three adjustment methods and are unreasonable in some cases (e.g., the *FIREPLACE* adjustment of $179 and the *HALF-BATH* adjustment of $231). The first three adjustment methods produce reasonable estimates that are fairly close to one another. While using regression on all sales results in coefficients with smaller standard errors and larger *t*-statistics, the adjustments provided are potentially biased. The complete data set contains many observations that are significantly different from those in the matched pairs set, i.e., the complete data set covers a much larger range of values for the explanatory variables. While the linear regression models (3) and (6) are probably reasonable approximations over the range of physical characteristics covered by the matched data set, the validity of the linearity assumption in model (7) over the much larger region represented by the entire data set is in doubt. The resulting bias due to possible model misspecification in the full data regression may offset any gains due to increased sample size. In other words, with *MRA* we may have accurate estimates of the parameters of a potentially incorrect model. For these reasons, we find *MRA* to be a potentially inferior technique for estimating adjustment amounts.

Exhibit 2 shows that the standard errors of the *MP–D–MRA* estimates are 45–63% larger than those of the *MP–MRA* method for the four adjustments examined in this study. This was contrary to our prior opinion that this method would yield the most accurate estimate among the four adjustment methods. The near-zero adjusted R^2 values show that the differences in the explanatory variables remaining after matching do not explain any further variation in the price differences.

The two remaining methods, *MP–Adj D* and *MP–MRA*, are equivalent and produce the same adjustment estimate in all cases. However, the *MP–MRA* method is superior since it provides a standard error estimate and is easier to calculate.

Finally, we should point out that although some of the *t*-statistics are highly significant (especially for *BRICK* and *GARAGE*), the associated standard errors of the estimates may be large for practical use. For example, using the *MP–MRA* approach, the *t*-statistic for *BRICK* was a highly significant 4.809 (p-value = 0.0001), yet a 95% confidence interval for its adjustment runs from $3,795 to $9,732. Some of the variability in the *BRICK* estimate can be attributed to the fact that the amount of brick used varies with the size of the house. Therefore, the adjustment estimate for *BRICK* should be interpreted as an average value.

Conclusion

This study used published data on 422 markets sales of FHA/VA-insured/guaranteed houses to examine and compare four methods of extracting market-derived adjustment values to be employed in the sales comparison appraisal approach. The four adjustment methods are variations and combinations of matched pair analysis and multiple regression analysis. Four physical features (brick exterior, garage type, fireplace, and half-bath) were chosen as adjustment value examples.

With a small risk of over-generalizing from the evidence presented here, we make the following observations:

- Perfectly matched pairs for adjustment purposes, i.e., those that require no secondary adjustments, seldom exist.

- The matched pair mean adjusted difference or traditional grid approach (*MP–Adj* $\overline{D}$) and the matched pair regression approach (*MP–MRA*) provide more reasonable estimates than regression on all sales (*MRA*) and more precise estimates (as judged by estimated standard error) than matched pair differences regression (*MP–D–MRA*) based on the statistical evidence presented in Exhibit 2.
- The *MP–Adj* $\overline{D}$ and *MP–MRA* methods are equivalent and produce the same adjustment estimate. Nevertheless, *MP–MRA* is more desirable than *MP–Adj* $\overline{D}$ because it is simpler to perform and provides a readily available measure of the standard error of the adjustment estimate.
- There are some primary adjustment features for which market data may provide a fairly clear indication of their effect on market values (e.g., *GARAGE* and *BRICK* in this study). Other features can be highly inconclusive (e.g., *HALFBATH* and *FIREPLACE* in this study).
- Although the statistical evidence provided by the *t*-statistics is impressive in some of the examples given here, the standard errors of the adjustment estimates are still quite large from a practical point of view.

In conclusion, among the methods considered here, regression coefficients computed from sufficiently large matched pair data sets appear to be the best source of market-derived adjustment values. However, due to the multicollinearity that typically accompanies sale price data and to the many sources of unexplained variation in sale prices, adjustment estimates computed from regression model coefficients or from matched pair analysis are subject to a great deal of uncertainty, as measured by their estimated standard errors.

References

[1] Peter F. Colwell, Roger E. Cannaday and Chunchi Wu. The Analytical Foundations of Adjustment Grid Methods. *AREUEA Journal* 11 (Spring 1983), 11–29.

[2] Timothy P. Cronan, Donald R. Epley and Larry G. Perry. The Use of Rank Transformation and Multiple Regression Analysis in Estimating Residential Property Values with a Small Sample. *Journal of Real Estate Research* 1 (Fall 1986), 19–31.

[3] Federal Home Loan Bank Board, Office of Examinations and Supervision. Memorandum R41c. September 11, 1986.

[4] Terry V. Grissom, Rudy R. Robinson and Ko Wang. A Matched Pairs Analysis Program in Compliance with FHLBB Memorandum R41b/c. *The Appraisal Journal* 55 (January 1987), 42–68.

[5] Hans R. Isakson. Analysis of Price Adjustment Amounts Derived From Paired Sales of Single-Family Properties. Unpublished paper presented at the Annual Conference of the American Real Estate Society, 1989.

[6] Mark J. Kroll and Charles Smith. Buyer's Response Technique—A Framework for Improving Comparable Selection and Adjustment in Single-Family Appraising. *The Journal of Real Estate Research* 3 (Spring 1988), 27–35.

[7] Harold Kuhn. The Hungarian Method for the Assignment Problem. *Naval Research Logistics Quarterly* 2 (Spring 1955), 83–98.

[8] Paul R. Rosenbaum. Optimal Matching for Observational Studies. *Journal of the American Statistical Association* 84 (December 1989), 1024–32.

[9] Donald B. Rubin. Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association* 74 (June 1979), 318–28.
[10] Hollis A. Swett and Jo Ann Whalen. The Use of Multiple Regression Analysis to Select and Adjust Comparable Sales. *Assessors Journal* (March 1977), 17–33.
[11] Arnold Tchira. Comparable Sales Selection—A Computer Approach. *The Appraisal Journal* 47 (January 1979), 86–98.
[12] R. Zerbst. A Caution on the Adjustments of Comparable Sales. *Real Estate Appraiser* 43 (January/February 1977), 28–29.