Regression Modeling in Calgary—A Practical Approach

The City of Calgary Assessment Department Abstract

This case study shows how the Calgary Assessment Department converted from mass appraisal based on the cost approach to mass appraisal based on the sales comparison approach and multiple regression analysis (MRA) for an in-house reappraisal of residential property. Issues of computer technology and modeling are discussed, and models are developed and applied to a small city.

Introduction

In August 1996, a team of appraisers was formed to examine and implement the reappraisal of approximately 255,000 residential single-family and condominium properties using the sales comparison approach and multiple regression analysis (MRA). The new values will be implemented for taxation in 1999 using 1998 as the base year. The assessed values on the current tax roll are based on replacement cost of the structure and an estimated market value of land. The formidable challenge was, therefore, to develop and implement a process that would move our jurisdiction from one method of valuation to another.

The potential efficiencies of implementing an MRA-based valuation system are numerous. Some notable improvements would result, such as fewer appeals over time, values that are easier to explain, and labor efficiency allowing more time to improve data quality (verify sales, confirm property detail, and review final assessed values).

Theoretically, MRA can be used to value all classes of property within our jurisdiction, but the first area of investigation was single-family homes and duplexes, for which data were available and in a usable form. In Calgary, there are about 200,000 single-family residences, a volume that makes MRA beneficial.

MRA has been applied extensively to single-family homes and condominiums throughout the United States since the 1970s, but in Canada is not generally used as the primary valuation tool. Several issues needed to be resolved to ensure MRA efforts in Calgary were successful:
- What type of database were we going to use (and would it reside on a mainframe or on the PC)?
- What model form would be used?
- What and how many variables would be used?
- How could the variables be standardized?
- How would we deal with different coefficient values for the same variables in similar neighborhoods?
- What would we do with variables that occur only occasionally?
- How would we decide on market areas or property groups?
- Where would we perform the calculations on the population?
- What statistical software would we use?
- How would we present and defend the values?
- What organizational structure would best support mass appraisal?

To answer some of these questions and to investigate processes we sent a team to Colorado, Arizona, and Florida. These site visits allowed us to question and to experience MRA analysis, value presentation, and value defense in a variety of jurisdictions. We also contracted with Robert Gloudemans to analyze and assist us in our work, once our first models were complete, and he has since returned to teach twenty-four of our staff the SPSS modeling course. The Mass Appraisal Team, as of November 1996, consisted of a project manager, project consultant, project leader, and six modelers.

This paper will be presented in three parts:

- an overview of multiple regression and how it has been used in Calgary
- development and use of computer technology and adaptation of the existing technology
- an example of the direct application of MRA in an assessment jurisdiction

**Overview of Mass Appraisal Using MRA Establishing Market Areas**

The city of Calgary was initially divided into five regions, determined by geographic boundaries and economic zones. Smaller market areas were then established within each region. Models were developed for each market area to estimate property values. However, the regional models were developed first to serve as overall benchmarks for the market area models. The regional models also provided an overall view of how the markets worked across the city. From the regional models significant key variables were chosen, where applicable, to be used as a starting point in the specification of the market area models. Other variables particular to the individual market areas were then added. This approach provided for a consistency in model development between appraisers but still allowed the use of variables that might be significant at the market area level. For example, variables relating to size and quality were significant in both regional and market area models, but the significance of some variables, such as view, varied between market areas and regions.

Large regional areas were also modeled to explore the relationships between variables (to provide guidelines in estimating values for variables with few sales or occurrences in the smaller market areas). These variables or relationships could be used in a market area model or constrained and included in the final value if appropriate. (Constrained coefficients should be used carefully and sparingly to deal with influences that cannot be accurately determined by the model. The result of a regression model is the best linear unbiased estimator [BLUE], and the more constraining of variables [that may or may not actually influence value] the less likely is the statistical result to be optimal. In terms of defense of the final values, it becomes increasingly difficult to justify effects on value that have little empirical or statistical basis.)
The city was divided into twenty-four market areas. Along with natural and manmade geographic boundaries, we looked at the many similarities among adjacent communities. Selling price per square meter of gross building area was also used to determine market area boundaries. No discontinuous market areas were used, and all market areas were confined to legal community boundaries for this general assessment. The twenty-four market areas provided adequate sales for each model. There were also enough observations for property attributes so as to minimize the need to constrain variables.

**MRA**

MRA measures the effect of various characteristics or influences on a single variable. In the case of the mass appraisal of single-family residences, an MRA model measures the contribution, or effect, of living area, lot size, number of fireplaces, and other property characteristics on sale price.

For each property that sold over a given period, we know the sale price and property characteristics. Given enough sales, a model can be built based on sold properties. Once the model is refined and has attained satisfactory assessment and statistical results, it can be applied to every property in the area of study. It is possible to estimate the assessments because we know the main characteristics of every property in Calgary. The reasonableness of the final estimate would have to be verified using appraisal theory, experience, and comparable sales.

**Model Forms**

Three functional forms are commonly applied when valuing real properties: linear (additive), nonlinear (log-log or multiplicative), or a combination of these two forms (a hybrid model). In the area of assessment, the linear functional form is the dominant choice (as reported by assessment offices we consulted and by Robert Gloudemans), given its advantages of explainability and the fact that residential market value generally has an additive nature. We plan to explore the option of a nonlinear model form and make our choice based on the model results.

An example of a linear, or additive, MRA model with two variables is:

\[
\text{Sale price} = b_0 + b_1 \times \text{living area} + b_2 \times \text{lot size} + \text{error term}. \tag{1}
\]

If the coefficient \(b_1\) is found to be \$400, then an increase in living area of one square meter adds \$400 to the market value of a home in the market area being studied.

An example of a multiplicative, or nonlinear, model is:

\[
\ln(\text{Sale price}) = a_0 + a_1 \times \ln(\text{living area}) + a_2 \times \ln(\text{lot size}) + \text{error term}, \tag{2}
\]

where \(\ln\) denotes the calculation of the natural logarithm of each term. An estimated model might look like the following:

\[
\ln(\text{Sale price}) = 10.00 + 0.25 \times \ln(\text{living area}) + 0.10 \times \ln(\text{lot size}).
\]
In this case, the coefficient of 0.25 on living area is interpreted differently than in the additive case. An increase of living area of 10 percent leads to an increase of market value of 2.5 percent. This different type of relationship has the important characteristic that bigger houses are valued at less per square meter than smaller ones—a common finding in real estate valuation. This principle is also useful for age (on average, new homes depreciate more in absolute terms than older ones), lot size, and many other characteristics of a property. The nonlinear form is also useful when analyzing a region or market area with a wide range of living area, lot size, and other characteristics.

If the nonlinear models perform significantly better than linear models, in terms of assessment and statistical measures, it remains to decide whether the greater challenge of explaining nonlinear model results outweighs the benefits of a better model that produces better values. Nonlinear coefficients can be converted to dollar amounts, and most importantly, nonlinear models emulate two principles of appraisal theory: nonlinear depreciation (improvements do not generally depreciate a fixed amount per year) and diminishing returns (the value of a home does not increase by the same amount for each additional square foot).

Furthermore, as we found in several jurisdictions in the United States, presenting comparable sales is the primary line of defense and explanation. It would be unusual to choose to discuss the detail of functional forms and coefficients with the taxpayer. It is preferable to concentrate on the final values and the comparable sales used in the analysis. This makes the choice of functional form much easier to make—choose whatever model type works significantly better in each market area or groups of market areas. If the performance is indistinguishable, then linear models should be used given their advantage of explainability.

With this in mind, both forms of model will be estimated and their results compared. A decision matrix will be completed for each market area to help choose the best model. Clearly, this process involves some subjectivity. For instance, you have to decide which is the most important test result. If the R-squared is relatively high and the standard error low, is there a need to be concerned with any other statistic? What if the model isn't stable? When there is a difference in statistics such as the coefficients of variation (COVs) and coefficients of dispersion (CODs), what amount of difference should be considered significant enough to consider changing model forms? The decision matrix can essentially be broken down into two categories: indications of statistical quality and indications of assessment performance.

**Parsimony: Careful and Economical Use of Variables**

The two-variable MRA models used above were for illustrative purposes only and would not work in practice, given the numerous factors or variables that enter into determining property values. The modeling of residential properties, and real estate in general, usually requires much more complex models. Nonetheless, one must be careful not to make the models so large and complex that they become difficult to explain and manage. Careful and economical use of only the key variables that theoretically and empirically influence value will be used. This is the concept of parsimony.

The goal of modeling is to explain a lot with a little. In other words, the building of regression models is best applied if you can adequately explain some phenomenon with a set of key variables, as opposed to many variables. It is conceivable in our case to estimate models with several hundred variables, but the benefits of a model that can adequately explain sales prices with a few key variables are apparent. From the basic model types, we added variables that were thought to influence value in a particular market area. Key variables included such items as total
size by quality type, effective age, lot size, basement area, garages, fireplaces, and location attributes.

**How to Evaluate Models?**

The error term is one of the most important outputs of a regression model. The error term, the difference between the estimate produced by the regression equation and the actual sales prices of the properties that were used in the MRA model, is important in determining the validity of the model, and hence the values or assessments produced.

In a good regression model, the error term has several key characteristics: the error term is normally distributed with a mean of zero and has constant variance, which implies that it is uncorrelated with sale price (see, for example, Gujarati [1988,279] and IAAO [1990a, 386-87]. Furthermore, the error term should show no identifiable pattern with respect to sale price or any of the explanatory variables. If it does, the model may be misspecified, indicating that an incorrect functional form has been used or that key variables have not been included or transformed properly in the current model.

The practical implication of any of these characteristics of the error term is that the other statistics produced (such as the t-tests, F-test, and Chow tests) are not reliable. For example, a Chow test result might suggest that the model is stable over a period of years (that is, the coefficients are the same for the model in 1994, 1995, and 1996). However, if there is heteroscedasticity (non-constant variance) present (see IAAO 1990a, 386-87), the coefficients may actually be unstable over time. Another example would be the F-test, which tests that all of the coefficients in the model are not equal to zero; again, the coefficients may in fact all be statistically equal to zero, although the test result suggests the opposite.

There are several ways to test explicitly for violations of the above conditions. One of the most current tests is called the White test, which checks for any patterns left in the error term. The White test is run after the valuation model has been estimated and is a regression in which the dependent variable is the square of the error term and the independent variables are basically the variables used (as well as cross products and squared versions of the independent variables). This formulation tests not only for heteroscedasticity, but also for correct model specification and form (see White [1980]). The practical implication of this violation is that many of the other statistics produced to verify the validity of an MRA model are not reliable.

If heteroscedasticity is evident, more modeling work is necessary. This is not to say that the model should be discarded or that the values produced by the model are not valid--rather, the estimated values of the sold and unsold properties must be verified by several people knowledgeable about appraisal, the data inventory must be checked for errors, and comparable sales must be used to ensure that the extent of this statistical problem is not so serious as to damage the reasonableness and uniformity of assessments or the median or mean level of assessment.

Another way to analyze the residuals or error term is to plot the errors against any number of variables in the regression model. This method is not as efficient as the White test, but it provides for a visual review and is more intuitive. As standard modeling practice, both methods will be used.

The Chow test determines whether or not the estimated model is stable over the sample period. In our case, we are using thirty months of sales data, so it is important to determine if the model
changes significantly over the period. That is, if we estimate three models--one each for 1994, 1995, and the first half of 1996--are the models (the estimated coefficients) the same? If the three models are not significantly different, we should be confident that any values estimated by the thirty-month model will be reliable. After all, if the market can be adequately modeled over the past thirty months as opposed to the past six, we can be more confident in our resulting assessments. Again, more information is better, but a compromise is necessary. Some jurisdictions use up to sixty months of sales data. We think our compromise of thirty months is reasonable.

Standard statistics produced by all MRA modeling software are the R-squared, t-statistics, and the standard error. The R-squared tells us what percentage of the variance in the dependent variable, in our case sale price, has been explained by the variables being used on the right-hand side of the equation (in the above example, living area and lot size). The closer the R-squared is to one, the higher the percentage of the variance that has been explained by the model; the closer to zero, the lower the percentage of the variance that has been explained by the model. An R-squared of 0.5 indicates that 50 percent of the variance has been explained by the variables used in the model. Careful interpretation of the R-squared is required. It is a common misconception that a low R-squared is indicative of a poor model; this may or may not be true. The R-squared is only one small part of the picture and must be used only in combination with a set of tests used to determine the validity of a model.

The t-statistics are indications of whether a particular coefficient (such as $b_1$ in equation 1) of a variable (living area) is statistically significant; that is, is the estimated coefficient statistically equal to zero? A rule of thumb for interpreting a t-statistic is that if it is higher than 2.0 in absolute value, the estimated coefficient is statistically significant or, equivalently, the particular variable has a statistically significant effect on sale price. Having a t-statistic lower than the absolute value of 2.0 does not preclude keeping the variable in the model. The t-statistic is only one measure to consider. If the variable makes good valuation sense (gives a reasonable value for that variable) and should logically be included as an indicator of value (for example, attached garages in a model that already includes detached garages), it should remain in the model even with t-statistics less than the absolute value of 2.0.

The standard error is determined from the error term after the MRA model has been estimated. It tells the modeler the probability distribution of the error term. The average error of almost all MRA models is zero, and a good result is that the error term is normally distributed. If the standard error of a model is 20,000, then 67 percent of the errors made by the model in estimating the actual sale prices are between -20,000 and 20,000; 95 percent of the errors are between -40,000 and 40,000; and 99 percent are between -60,000 and 60,000.

Is this amount substantial? If the average price of the sales used in the model is $75,000 then the model is poor. However, if the average price is $150,000, the model is predicting sales prices better. Given this problem, the standard error is commonly divided by the average sale price of the properties used in the model to find out what percentage the standard error is of the average price. In homogeneous areas, this percentage should be less than 10 percent; in more heterogeneous areas, the percentage may be as high as 20 percent.

An extremely useful method to test the validity of a model (after the MRA model building process has been completed) is to conduct an extensive ratio study of the sales and the estimated values of those properties from the MRA model. The mean and median levels of assessment (the assessment-sale price [A/S] ratio), the COV (with respect to the mean), or COD (with respect to the median) are all measures of appraisal level and uniformity. These, along with checks on the outlier ratios and good analytical skills, will tell us whether the statistical problems indicated by
the above tests have severe effects on the assessments or uncover any other valuation problems that might exist.

What will be the primary criterion to accept a model? This is a difficult question to answer, but, primarily, the estimated coefficients must be reasonable and defensible in the opinion of our team and other assessment professionals. Unfortunately, there is no standard statistical test to measure reasonableness; therefore we must rely on our experience and judgment.

If the standard of reasonableness is met, then the above set of tests helps us to have even more confidence in our results. If the reasonable model fails to pass some of the formal statistical tests, it should not be discarded, but more modeling work is necessary. Market areas or communities may not be adequately defined (that is, within a neighborhood, there is a property detail that cannot be adequately explained in the current model). Again, more modeling work should be done, which might include a redefinition of market areas, smaller stratifications, or the recalibration of the model using another functional form for some or all of the variables involved.

We have found with regard to tests and statistics that more information is better than less. Although few of the jurisdictions we visited that rely on MRA use the White or Chow tests or extensive ratio studies, we believe that the additional information they make available can be an advantage. This preference also applies to the use of additive and multiplicative model forms. The preference for more information must, however, be balanced against the resources available.

Assessment Performance

There are two main criteria to evaluate model performance (MRA or otherwise). For mass appraisal these should be the level of appraisal (specifically, the median level of the A/S ratio) and the uniformity of values (dispersion around the median A/S ratio) in relation to market value. In a good model, on average, the assessed values should be equivalent to the sales prices in the market under study, and the COD should be 10 or less for homogeneous communities and 15 or less for heterogeneous ones (IAAO 1990b). (The Province of Alberta's standards for the citywide median A/S ratio must be between 0.90 and 1.10, with a COD of less than 15 for residential properties.)

Using these two criteria, we can then attempt to address the following critical issues:

- What MRA model gives us better results in terms of levels of appraisal and uniformity of assessments?
- Does MRA produce better results based on level and uniformity than the cost-based/market value for land values used in the 1994 General Assessment?

Typically the MRA models, regardless of the functional form used, perform either as well as or much better than the cost-based/market value for land values used in the 1994 General Assessment. This finding alone justifies the move to an MRA-based approach for valuing Calgary's residential properties.

Field Checks and Audits

Final values are calculated on the population, once the model is accepted as complete. The final test is then to conduct field reviews of values produced by the model. A simple spreadsheet with
key information can be used to compare properties in the same neighborhood, street, and market area. Values should be compared between market areas, and some citywide adjustments may be needed based on the regional models. The final field reviews, if the models perform well, may bring forward data problems that can be resolved before the values go to the roll. Outliers, both high and low, should be investigated and documented throughout the process.

The entire process will be documented throughout each stage, and the same base program and standards for each area analyzed will help us to construct detailed audit trails which will make the defense of our work credible and supportable.

Transferring the Values to the Roll

Once a model is accepted as complete, the values estimated by the model are directly applied to all properties and sent to the roll.

Embracing Available Computer Technology

In developing the process for MRA, the Mass Appraisal Team had to incorporate available computer technology to accomplish the mass appraisal of property values.

With the introduction of Pentium personal computers (PCs) and powerful software applications, assessors have been given more flexibility in the use of computer technology. The City of Calgary Assessment Department's computer systems are in a state of major transition, moving from a rather rigid mainframe environment to a more flexible client-server, PC-based environment. The department is taking advantage of available computer technology to do more with fewer assessors, and to do annual mass appraisals on approximately 255,000 residential single-family and condominium properties.

From a systems support perspective, revisions were required. The existing system supported the cost approach to value, which resided on a mainframe. With a pending reappraisal of residential properties, an alternative system and new processes evolved to support the sales comparison approach using MRA. Initially, there was a preference for the mass calculation of assessments to continue to reside on the mainframe. However, the time estimates given for program development on the mainframe made this a prohibitive option, and the variable transformations required for a mass appraisal using MRA reside in a Microsoft Access software application.

Multiple Regression Access Application

The City of Calgary Assessment Department's MRA application uses Microsoft Access to manage two key processes: the processing of data and the mass appraisal calculation on the entire residential inventory. Other supporting functions were required such as a comparable selection facility for defense and explanation of assessments.

The development and use of the multiple regression Access application and SPSS software was considered to be an interim solution to support the mass appraisal process. The Access application is a working prototype. The department has determined that its systems needs for the future will be met through a geographical information system/computer-assisted mass appraisal (GIS/CAMA) system. The Access application functions will be duplicated and enhanced in the GIS/ CAMA system. Extensive research and several site visits to other jurisdictions led to the
conclusion that the department's GIS/CAMA systems needs could not be requisitioned from an external source. Development of an in-house GIS/CAMA system began on January 1, 1997.

**Rapid Application Development and Security**

The Microsoft Access application was scoped and constructed during a six-month period from late 1996 to early 1997 using Rapid Application Development (RAD) methods. The application and database are kept in separate Access files for ease of promoting new application versions without upsetting the data tables. The programming language "Access Basic" is used extensively in the application, which is supported by a programmer analyst with intermediate or advanced Access or Visual Basic skills.

The security option built into Access was used for this application. Before an appraiser can use the application, his or her ID must be added to one or more defined ID user groups. All the objects in the application and database files have permissions assigned at the user group level.

**Selection of a Statistical Analysis Tool**

SAS, NCSS, Eviews, and SPSS were evaluated based primarily on performance, ease of use, and cost. SPSS was chosen for a number of reasons:

- SPSS appears to be the most commonly used modeling/analysis software tool among other jurisdictions.
- SPSS can handle very large data sets, as found in major cities.
- Operating within the Windows environment, SPSS has a user-friendly point and click interface.
- SPSS provides graphical features and statistical summaries that supplement the MRA application.
- From an audit or quality control perspective, SPSS has a journal file that can keep a record of procedures which have been run.
- SPSS has programming functions that enable the creation of standard procedures in a syntax file. This ensures a consistent approach in the valuation modeling process, yet provides the modeler with the option to explore additional model functions or procedures.
- There is good technical and training support available for SPSS.

**Main Functions of Application**

The multiple regression Access application imports data from the mainframe, transforms the data, and then exports the transformed data to a statistical analysis tool. The adjustments from the valuation models are copied into the application, which facilitates the mass calculation of assessments. The assessments generated are then passed from Access to the mainframe, where the assessment account creation and maintenance facilities reside and from which information is pulled to generate assessment notices. The main functions of the multiple regression Access application are accessible from the applications main menu. The main functions are as follows:

- **Start Extra** starts a window to connect to the mainframe.
- **Import Database Extract** allows the appraiser to define what type of data extract is to be imported, specify the path to the file, and have it imported. Options are provided to replace null numeric values with zeros and trim blanks from text fields. The filename is used as a "data set" name to track the set of records through the application.
- **Run Transformations** runs mathematical transformations on data with the click of a button. The transformations can create new variables or convert old cost data elements into variables required for MRA. For example, the transformation process can convert the number of plumbing fixtures to
the number of bathrooms. The programmer analysts involved in the development of the Access application, with some ingenuity, have continually improved on the speed of this process.

- Create Analysis File writes out a selected set of records to a file that can be imported into a statistical analysis software tool, such as SPSS.
- Start SPSS starts an SPSS session.
- Mass Calculation permits the replication of the valuation model derived through the MRA application. Once the valuation model is duplicated, the appraiser can trigger the mass appraisal calculation on a specific population data set.
- Create Upload File writes the mass calculated assessments for a data set to an external file in preparation for an upload of assessments back to the mainframe environment.
- Delete Data Set is provided to enable the appraiser to remove data sets from the database in a controlled environment.

There are additional functions as follows:

- Browse Data Set or Table is a dialogue that provides a flexible interface to view data sets or systems tables.
- Ratio Studies Report provides standard ratio studies that incorporate all typical measures of uniformity and level of assessment (median, mean, and range of MS ratios; and COV, COD, and price-related differential [PRD]).
- Comparable Selection Report is a reporting facility with built-in flexibility for the appraiser. Comparables can be selected based on the criteria set by the appraiser. The appraiser can run the report for any given number of comparables, which can be selected from either the sales data set or the entire population data set. Comparables are ranked through the built-in computation of the Euclidean Distance Metric, a dissimilarity index.
- Shift Analysis Report provides a summary of the shifts in value that have occurred between individual properties or groupings of properties due to the revaluation.
- Representativeness Report provides critical information on whether the sales data set is representative of all property types in the population data set.

**The City of Airdrie: A Practical Application of Multiple Regression Models**

The Access application was used successfully to process data required for multiple regression models developed as part of the assessment services provided to the City of Airdrie, a small city north of Calgary. Performing the mass appraisal calculation on the entire residential inventory of Airdrie proved to be fast, efficient, and accurate and helped develop a prototype that could be applied to Calgary. It also provided clear direction for development of the upcoming GIS/CAMA project. The reappraisal of all residential properties in Airdrie for the 1997 assessment roll provided an excellent opportunity to develop and test the use of MRA to calibrate sales comparison models. MRA-generated values could then be applied directly to the assessment roll.

The City of Airdrie, located about thirty kilometers (eighteen miles) north of Calgary, has a population of 14,507 (1994). The majority of employed Airdrie adults (58.3 percent) work in Calgary. Residential housing units in Airdrie are distributed as follows:

- Single-family and semi-detached 4,700
- Townhouse 90
- Manufactured homes 700

Three multiple regression models were developed: a single-family model (which included semi-detached units), a condominium townhouse model, and a manufactured homes model. The properties were valued on market value as of July 1996.
Sales and Property Data

Sales used for analysis were for the period January 1, 1994, to June 30, 1996. All sales were screened so as to provide valid, arm's-length transactions for improved properties. The sales were time adjusted from date of sale to the valuation date of July 1996 based on an analysis of time trends within neighborhoods and market areas. The final number of valid sales used was 991.

Property characteristics and influences for residential properties in Airdrie are stored in the same residential database that exists for Calgary properties. A recent reinspection of homes provided property data that were confirmed and up-to-date. Sales and property data were downloaded to the multiple regression Access application using a residential extract program developed in-house. Standard variable transformations were done, and the data set was then exported to SPSS for analysis and development of the multiple regression models.

Considerable detail is stored for both structure and land characteristics. However, information collected for structures was collected ostensibly for the cost approach to value. Cost information was then converted for the sales comparison approach through standard variable transformations. For example, a transformation was done to convert the number of plumbing fixtures into the number of bathrooms. For consistency and a minimum of effort for the modeler, most transformations were done using the multiple regression Access application.

Multiple Regression Models The Single-Family Model

Variables used in the regression model for single-family and semi-detached homes related to

- Size by quality type
- Structure type
- Effective age
- Structure attributes
- Land area
- Land attributes and property influences
- Location variables for neighborhoods

The single-family regression results with adjustments are shown in table 1.

Condominium Townhouse Model

There are five condominium townhouse projects in Airdrie. Variables used in the regression model related to

- Size
- Structure attributes
- Effective age
- Binary variables for each project

Confirm Values and Property Data

After the preliminary regression models were developed, estimated values were calculated for all residential properties. The modeler and appraiser responsible conducted field reviews by driving
by the properties to confirm values and determine changes required for the model or property data. The regression model was then rerun and adjusted to account for documented changes.

Upload of Assessed Values

A direct application of the multiple regression models was used to calculate assessed values based on the constant and coefficients provided by the models. Assessed values were uploaded through the same Access application to the mainframe, which ultimately produced the assessment roll.

Assessment Quality

The final multiple regression models produced extremely satisfactory results. The models were statistically sound, with coefficients that were logical and explainable from a valuation perspective.

Although the previous assessed values met accepted standards, the level and uniformity of the new values improved dramatically. The level of assessment, which measures the ratio of assessed values to market values, improved for all groups of properties. The legislated level of assessment in Alberta is a range of 0.90 to 1.10 for residential properties. The median level of assessment using values from the regression models derived for Airdrie improved from 0.93 to 1.00 for single-family homes. The level of assessment for condominiums was 0.99.

The COD measures the average percentage deviation from the median ratio. The Alberta Standards of Assessment Regulation requires the COD for residential properties to be 15 percent or less. The uniformity of assessed values derived from the regression models improved for both residential and condominium properties. The COD for single-family properties improved from 6.13 percent to 5.03 percent. The COD for condominiums was a satisfactory 5.10 percent. Table 2 summarizes levels of assessment and measures of uniformity.

The notices of assessment for Airdrie were mailed May 27, 1997. Very few appeals were received (approximately fifty residential appeals, less than 1 percent of the properties). This indicates the assessed values determined through the use of multiple regression models were well received by property owners.

Presentation and Defense of Assessed Values

The use of multiple regression analysis to calibrate a sales comparison model produced excellent results. The resulting values should be presented and defended within the context of comparable properties that have sold. Our intention is to demonstrate that the assessed values are reasonable in relation to market value and in comparison to similar properties through the use of a comparable selection facility which will present the subject and comparable properties. Also, ratio studies will include enough detail to demonstrate that groups of properties by such characteristics as quality type and structure are valued at similar levels in relation to market value.
Summary

The City of Calgary Assessment Department has developed an integrated, practical approach using MRA to estimate accurate, reliable assessed values based on market value. We have combined the use of current PC technology with the application of sound valuation analysis and judgment to produce mass appraisal valuation models that meet or exceed accepted measures of uniformity. PC software applications developed in house facilitate the analysis of sales and property data. They also calculate and measure the quality of property assessments, based on a market value standard.

Our goal is to develop multiple regression valuation models that can be applied directly to the assessment roll. In the case of the city of Airdrie we have accomplished our goal. We are confident we can accomplish the same for the city of Calgary. Having completed a number of residential regional and market area models, our next step is to apply a similar approach to multifamily, commercial, and industrial properties. To date, we have developed preliminary sales comparison models for condominium warehouses and commercial retail properties. We are now exploring the development of income models for commercial and office properties.

Multiple regression valuation models and PC technology can form the core of a practical and effective approach to providing quality market value assessments. At the city of Calgary, we are satisfied that we have begun to translate theory into practical reality. The major portion of the assessed values implemented for Calgary's 1999 taxation year will be completed using a direct application of multiple regression valuation models.

Table 1:

Summary of Model Coefficients (Airdrie Single-Family)

<table>
<thead>
<tr>
<th>Size quality (m²)</th>
<th>Value t-statistic[*]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality fair area</td>
<td>314/m²   6.93</td>
</tr>
<tr>
<td>Quality average area</td>
<td>333/m²   12.32</td>
</tr>
<tr>
<td>Quality good area</td>
<td>347/m²   11.66</td>
</tr>
</tbody>
</table>

Structure type (m²)

| Bilevel area                      | 69/m²   2.96        |
| Bungalow area                     | 81/m²   3.51        |
| Split-level area                  | 99/m²   4.06        |
| Two-story area                    | 21/m²   0.95        |
| Two-story basement-less area      | -50/m²  -1.52       |

Effective age

| Effective age > 25 years         | -679/year  -6.93   |
| Effective age <= 25 years        | -1,106/year -11.71 |

Structure attributes

| Attached house size              | -91/m²  -8.17     |
| Total basement finished area     | 64/m²    5.54      |
| Vaulted ceiling (binary)         | 3,845    4.35      |
| Fireplace (binary)               | 3,091    5.33      |
| Attached garage area             | 294/m²   12.51     |
| Detached garage area             | 177/m²   11.03     |
| Number of plumbing fixtures      | 1,109/fixture 5.29 |
Walkout basement (binary)                      3,108    2.68

Land

Actual land area                            23/m²    7.69

Land attributes (binary)

Traffic influence                          -2,743   -2.71
Commercial influence                      -4,919   -2.03
Multifamily influence                    -6,625   -2.39
Corner lot                                -1,339   -1.45
Golf course                               10,942    4.99
Direct river access                       7,504    3.22
Train influence                            -7,304   -2.18

Subneighbourhood codes (binary)

SNC1                                      1,048    0.67
SNC2                                     -3,200   -1.68
SNC4                                     1,556    1.14
SNC5                                     -2,798   -0.83
SNC7                                     -875    -0.64
SNC8                                     7,808    3.90
SNC10                                    -5,238   -4.19
SNC12                                     3,352    2.67
SNC13                                    -966    -0.90
SNC16                                    -3,172   -1.75

Constant/intercept                       54,816   20.47

Table 2: Level and Uniformity Measures: Airdrie
Legend for Chart:
A - Level of assessment Median
B - Measures of uniformity, Coefficient of variation
C - Measures of uniformity, Coefficient of dispersion

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>1.00</td>
<td>6.53</td>
<td>5.03</td>
</tr>
<tr>
<td>Condominiums</td>
<td>0.99</td>
<td>6.40</td>
<td>5.10</td>
</tr>
<tr>
<td>Manufactured homes</td>
<td>0.99</td>
<td>7.00</td>
<td>5.60</td>
</tr>
</tbody>
</table>

References


By Robert Dalgiesh and Lawrence Buchart