MICRO-COMPUTER BASED TIME SERIES ANALYSIS

Paul Kershaw  (Principal Lecturer in Property)
Peter Rossini  (Lecturer in Property)
Rob Kooymans  (Lecturer in Property)

School of Real Estate,  Faculty of Social Sciences,
University of South Australia,  Adelaide Campus

PREFACE:

In a paper delivered to the 1992 Australasian Real Estate Educators Conference (Micro-Computer Based Real Estate Decision Making and Information Management - an Integrated Approach), the authors outlined their progress in developing a complete Urban Property Market Real Estate Analysis Package whose components can be used either together or separately. The aim has been to take concepts which are known and well-tried in both the academic world and in fields other than real estate, and to provide means by which they can be easily used by real estate practitioners who can afford neither the time nor trouble to become statisticians and computer programmers.

Development of Real Estate data storage and access facilities was completed in early 1991, with subsequent refinements taking place in development of a Windows format and a system for updating data. The Data Analysis Interface has demonstrated its potential for use in statistical analysis by the average real estate practitioner in the field. In particular, experiments with Multiple Regression Analysis in conjunction with a comprehensive residential real estate data base have shown real possibilities.

The third component of the Package was to be the Time Series Data Interface, and it is this component that the current paper addresses.

INTRODUCTION:

Time Series Analysis is a concept whose results are of vital concern to real estate practitioners in all fields, whether or not they realise that what they are dealing with is Time Series Analysis. Put simply, Time Series Analysis is the analysis of a sequential array of the values of an economic variable (Pappas & Brigham). It is usually carried out with the aim of ascertaining some pattern which may enable prediction of the behaviour of that data in the future. It facilitates an understanding of the past and enables the investigation of possible cause and effect relationships. The lay person is confronted with the results of time series analysis on a daily basis, in the form of seasonally adjusted unemployment statistics, CPI indices, stock market performance charts etc. Time Series Analysis is the basic tool of the Economist who is concerned with policy.

In the Real Estate field, both among practitioners and their clients, there is a fascination with the performance of "the Real Estate Market", whether it is an individual who wants to know the best time to buy an investment property, or an agent seeking confirmation of his or her assertion that real estate prices always rise. There is a strong interest in, and market for, studies which show the performance of real estate prices over time and which purport to show future trends. These are frequently shown in such publications as the REI Bulletin, in private newsletters and, in fairly crude form, in the daily press which sometimes sees fit to publish it as front page news.
Almost inevitably, much reliance is placed on so-called trends, cycles and seasonally adjusted figures, without the reader and sometimes, one suspects, the author, being remotely aware of what these mean and of the limitations of predictions based on time series data. The old saw, “Markets have no memory” (Fama) seems lost on many in the real estate market.

Current Time Series studies of the type mentioned above seem to suffer from three main problems:

1. They are expensive to produce, in that they require a lot of hours to collect, collate and analyse data.

2. They are rarely very frequent or timely, because of the costs of data and the effort involved to update the studies.

3. The more frequent and timely the studies are, the more they are likely to suffer from five fundamental problems of analysis.

The first analysis problem is the reliability of the data collected. In a volatile real estate market, a few extra weeks of data can make difference to interpretation. Commentators who want to appear authoritative about trends will sometimes use unverified data. For example, trends in residential sales volumes may be gleaned from transactions reported by agents, which can suffer from an "Optimism Factor".

The second problem lies in the interpretation of data over time. There have not been useable raw data whose consistency can be verified over time, resulting in comparisons over time of dubious consistency. For instance, the mean price of residential sales in a given area may be $150,000 now, whereas 10 years ago it may have been $80,000. But did the old mean include sales of units? Were sales of doubtful utility discarded? If the current researcher was not involved in obtaining the old figure, which is usually the case, there is a large element of doubt. The only way this can be resolved is to go back over the original data, which is currently very time-consuming, even if the original data is available.

The third problem, which is also data-related, is the scale and utility of the studies currently undertaken. Real estate practitioners of any experience are very wary of applying, for example, the increase in the average price of a house in Metropolitan Adelaide to house prices in the Adelaide suburbs in which they operate. They know that there are many sub-markets, and that any patterns which are apparent in aggregate figures may not come close to applying to the sub-markets in which they are interested. Most time series analysis of real estate data tends, in our experience, to be conducted at Local Government Area level at best, and only split sub-markets into house and unit sales, if that. This level of analysis is interesting and a nice talking point, but not a lot of use to either the serious researcher of real estate prices and sales volumes, or to the real estate practitioner trying to give advice to a client on what has been happening or is likely to happen in the client’s particular sub-market.

The fourth problem, loosely related to the third but occurring mainly because of ignorance of its potential existence, is that of the heterogeneity of housing and its effect on mean prices. Mean or median prices can be quite meaningless as a measure unless the analyst can be sure that the type of housing being sold is consistent over time.
As Wachter (1991) has observed, "In markets with rising per-capita incomes, the average and median values of units sold rise over time simply because higher priced units - larger units with more amenities - make up an increasing share of the total number of units sold."

The improvement in the quality of housing sold can cause average prices to rise without any increase in price for a housing unit with a given quality. Nevertheless, increases in average prices signifying improvements in quality are mistaken as increases in the prices of each unit sold. This problem can be addressed by hedonic or resale approaches, which will be addressed later in this paper.

The fifth problem relates to the time period for which data are available. Time Series Analysis, particularly the more complex methods, is hungry for data. This can be because some of the techniques waste data, but mainly occurs because cycles and trends in real estate occur over many years and not much can be gleaned from one or two years' data. Raw data on prices are available over quite a number of years in most parts of Australia, but the data needed to overcome the other problems of Time Series Analysis are not. In South Australia, we are fortunate in having such data going back over 20 years as a public record. Most other States have only commenced a systematic record of sale property attributes in the last two years, meaning that there is future potential but limited present ability to overcome the problems mentioned.

**CURRENT ACHIEVEMENTS:**

It became clear to the authors that a viable Time Series Analysis facility would rely on the ability to obtain good data over a lengthy time period, to extract and manipulate the large quantities of data that this entails, and to freely and rapidly select sub-markets on chosen parameters. Thus the ability to carry out good Time Series Analysis was one of the main original criteria for the development of UPmarket™. As mentioned in the Introduction, the data base objectives have been largely achieved, so it remained to experiment with Time Series Analysis methods themselves.

The main point of carrying out Time Series Analysis and associated techniques is to identify seasonal factors, cycles and trends which will help the real estate practitioner face the uncertain future with at least some idea of what may happen, to investigate the stability, efficiency and error terms of simple statistics and to validate or otherwise some of the assumptions such as the shape of the distribution of sales per period and the stability of that distribution. Historical records - a description of what has happened - are of academic and some popular interest, but the measure of a method's effectiveness is the plausibility of the forecasts which may be made from the seasonal factors, cycles and trends. Against this must be weighed the cost and convenience of use of the methods.

Time Series Analysis methods vary enormously in complexity, from simple naive models which assume that what happened last month will therefore happen this month (the model used informally by many real estate practitioners) to extremely complex econometric models which require a great deal of mainframe computing power and expertise in mathematics to construct and run. As Wheelwright and Makridakis have observed, "Consequently, knowing that the Box-Jenkins methodology (Auto-regressive Moving Average) does at least as well as large econometric models may lead one to wonder whether the use of econometric models is ever justified" - certainly the marginal costs far outweigh the marginal benefits.

For UPmarket™, the initial experiments have been in the use of a relatively simple Classical Decomposition Method, which is used to identify Seasonality and to provide Seasonal Indices, to identify Cycles and to identify a general Time Trend.
The UPmarket™ system lends itself to the use of many different methods of analysis - the authors needed to choose one to begin with. The Classical Decomposition method is a respectable one which has stood the test of usage in a wide range of economic forecasting situations. It is certainly more respectable than the naive methods which are consciously or unconsciously used in the Real Estate Industry at present. Importantly, UPmarket™ offers the advantage of scientific validity, in that results obtained by its use will be reproducible by other researchers. This is the primary advantage of going back to the raw data rather than relying on the results of other research without being certain of how that research was carried out.

As with all UPmarket™ analysis, the primary concern has been to construct a system which is easy to use and whose results can be readily interpreted by the average intelligent Real Estate professional. The ease of use problem has been overcome by placing the system in a standard Windows™ environment with links to the UPmarket™ data base, enabling the user to select the data required at the level required and to analyse it in the desired manner. UPmarket™ facilitates linkages to Excel™, Lotus 1-2-3™, Personal Computer based SPSSX™ packages, or Mainframe Computers for further analysis.

The output screens are a combination of numerical and graphical information of the type which is of most interest to Real Estate professionals.

Figure #1 provides a summary of the statistics on a per-period basis. In the example used, quarterly statistics at LGA level are given, but there is no reason why a user could not select, for example, annual statistics at a suburb level or monthly statistics for a group of LGA’s.
While Number of Sales, Minimum, Maximum, Mean and Median are frequently given in published statistics, Upmarket allows for the means of data which has been trimmed at the lower and higher ends by 5% and 10% (or any limits the user may care to define) which may have the effect of eliminating distorting sales. In the same vein, UPmarket allows the removal of those sales which the South Australian Lands Department has identified as involving other land, and which are therefore potentially distorting. Standard Deviation provides an indication of the spread of sales around the mean and therefore of the homogeneity of sales. Number of Land > 0 shows the number of sales which include data for land area, not the number of land sales. From these sales is calculated the mean price per square metre of (improved) land area. Number of Building > 0 shows the number of sales for which a building area is given, allowing the price per square metre of improvements to be shown. This figure is a good indicator of price movements, which goes some way to countering the problems of heterogeneity in the real estate markets. These simple statistics are graphed in Figure #2.

Figure #3 shows the results of Time Series Analysis performed on the same data, with Seasonal Indices enabling calculation of Seasonally Adjusted prices (not surprisingly, there is little seasonal fluctuation of prices, unlike numbers of sales). Trend analysis shows a definite upward trend in prices over time, while the Cycle Ratio indicates quite clearly why data are needed for a considerable period of time to identify cycles.
Figure 3 - Upmarket™ Direct Analysis of Time Series Data
UPmarket Time Series Analysis
Burnside LGA 1986-1990
Number of Sales

Seasonal factors

<table>
<thead>
<tr>
<th></th>
<th>96</th>
<th>97</th>
<th>99</th>
<th>99</th>
<th>Mean</th>
<th>INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>1.029086</td>
<td>0.993844</td>
<td>1.12967</td>
<td>1.117096</td>
<td>1.067424</td>
<td>1.083529</td>
</tr>
<tr>
<td>Apr</td>
<td>0.876513</td>
<td>0.985667</td>
<td>0.910683</td>
<td>1.015152</td>
<td>0.947254</td>
<td>0.957902</td>
</tr>
<tr>
<td>Jul</td>
<td>1.024356</td>
<td>1.193381</td>
<td>0.9728</td>
<td>0.955224</td>
<td>1.036515</td>
<td>1.008290</td>
</tr>
<tr>
<td>Oct</td>
<td>0.943729</td>
<td>0.938522</td>
<td>1.053630</td>
<td>0.896203</td>
<td>0.958023</td>
<td>0.950279</td>
</tr>
</tbody>
</table>

Trend Analysis

Type: Linear Trend

R Squared: 0.017361
F Ratio: 0.319014

Coeffs
- Constant: 196.2834
- 14.02315
- 13.93602

Growth: 0.60015
1.17063
0.563927

Standard Error of Estimate: 30.18774

Cycle Analysis

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Cycle Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-86</td>
<td>-0.24078</td>
</tr>
<tr>
<td>Apr-86</td>
<td>-0.04930</td>
</tr>
<tr>
<td>Jul-86</td>
<td>-0.07951</td>
</tr>
<tr>
<td>Oct-86</td>
<td>-0.07946</td>
</tr>
<tr>
<td>Jan-87</td>
<td>-0.17972</td>
</tr>
<tr>
<td>Apr-87</td>
<td>-0.05631</td>
</tr>
<tr>
<td>Jul-87</td>
<td>0.256018</td>
</tr>
<tr>
<td>Oct-87</td>
<td>0.086001</td>
</tr>
<tr>
<td>Jan-88</td>
<td>0.013236</td>
</tr>
<tr>
<td>Apr-88</td>
<td>0.14239</td>
</tr>
<tr>
<td>Jul-88</td>
<td>0.111017</td>
</tr>
<tr>
<td>Oct-88</td>
<td>0.276102</td>
</tr>
<tr>
<td>Jan-89</td>
<td>0.157306</td>
</tr>
<tr>
<td>Apr-89</td>
<td>-0.02944</td>
</tr>
<tr>
<td>Jul-89</td>
<td>-0.07039</td>
</tr>
<tr>
<td>Oct-89</td>
<td>-0.00945</td>
</tr>
<tr>
<td>Jan-90</td>
<td>-0.00811</td>
</tr>
<tr>
<td>Apr-90</td>
<td>0.006033</td>
</tr>
<tr>
<td>Jul-90</td>
<td>-0.02592</td>
</tr>
<tr>
<td>Oct-90</td>
<td>-0.25137</td>
</tr>
</tbody>
</table>

Figure 5 - UPmarket™ Direct Analysis of Time Series Data
Figure #4 illustrates information which is inevitably brought up when discussing real estate prices - the performance relative to inflation. The first graph shows actual prices compared with the first quarter price indexed for inflation. This shows real estate prices in the Burnside LGA over that period as falling behind and then outstripping inflation. The second graph uses the final quarter price as the starting point and deflates it by the CPI to give the indexed figure. The third and fourth graph compare the trends from Time Series Analysis with indexed trends. These give a much clearer picture of performance in relation to inflation. The point that needs to be made in examining these figures is that these trends vary across location and property type within the same time period, illustrating the folly of generalisation and the benefits of an accessible and flexible data base.

Figure #5 illustrates Time Series Analysis of Numbers of Sales. Comparison of the Cycle Ratio with that for Prices shows them to be a mirror image. This accords with economic theory, in which quantity sold moves in inverse relationship to prices, but not with popular perception which has Real Estate agents cheering increases in average prices. Perhaps "talking up the market" does not work to the agents' advantage!

FUTURE DEVELOPMENTS:

Earlier, the problems of heterogeneity of Real Estate markets were referred to, and it is to these problems that our developments of the system are to be addressed. To construct a constant quality index of prices and sales volumes, two approaches can be used.

Firstly, the hedonic approach requires detailed data on attributes that give value. These are already available in South Australia and will increasingly become so in other States. These data can be used in a regression equation to adjust house prices for changes in housing characteristics. For example, in a society where per capita income is rising, it is likely that more people are buying larger and better appointed houses. This could result in the mean or median price rising significantly over time, without the price of a house of a given size and quality rising at all. If the data base contains information on size and quality, as does the S.A. Lands Department's information, it ought to be possible to adjust individual sale prices to reflect a constant size and quality. The challenging part will be to incorporate this well-known and long-used idea (the U.S. Census Bureau uses it to construct its "Price Index of New One Family Houses Sold" Series) into the friendly environment of UPmarket™, then to ascertain the reliability of the information held on the data base concerning the size and quality of properties.

The second method of overcoming the problem is by use of repeat sales of the same property. UPmarket™ can identify South Australian Sales by Valuation Number, Certificate of Title and Address, or a combination of these. It is therefore possible to overcome the practical problem of tracking resales through the years. What needs to be addressed is the potential problem of distortion caused by individual units physically depreciating with age or, more commonly, being renovated and improved. Previous attempts at addressing this problem have been very expensive in resources and limited to large research organisations (Wachter; Case & Quigley; Abraham & Schauman; Hendershott & Thibodeau; Quan & Quigley), thus restricting availability to the majority of people who would benefit from being able to quickly carry out such research. The extent of the problem and the authors' ability to solve it, can only be revealed by experimentation.
REFERENCES:


