Market modelling · 6
Path Analytic Models - 3

Chuck Chakrapani

Ideal models and the real world
As marketing researchers, we would all like to work with recursive models in which:

- a variable such as income clearly is linked to another variable such as sales; and
- there is an assumed hierarchy of relationship - such as age influencing net worth and net worth influencing the purchase of an expensive durable item.

Such models are intrinsically appealing, productive and (we hope) ultimately profitable.

In reality, model building is seldom that simple. Models are not always built to identify hierarchical influences. Cause and effect hypotheses are always open to question. Is affluence the 'cause' of good education or is good education the cause of affluence? Does higher advertising expenditure 'cause' higher sales or does higher sales 'cause' higher advertising? Our current state of knowledge does not allow us to answer such questions. It is also possible that the relationships are two-way processes: higher advertising can result in higher sales and higher sales can lead to higher advertising expenditures.

Time lagged models
The problem posed above can be handled through a new approach known as time lagged models. Let us consider a situation in which we are not sure how two variables are related. For instance in a model that is intended to predict voting behaviour we may not be certain whether Education influenced Income which in turn influenced Voting or Income influenced Education which in turn influenced Voting. How do we deal with such situations? There are many ways we can deal with them but first we need to know how important causal direction is to our final model.

When causality is not a major issue
In this situation the early paths of the model are usually of less significance. This implies that our results are not likely to be affected in any significant way by the direction of early arrows. If a model builder is uncomfortable with such an assumption, s/he can run the model both ways to make sure that the outcome is not affected if the relationships are reversed.

This approach, which is usually accepted in social sciences, makes even more sense in an applied discipline like marketing. After all, our main aim is to predict the outcome.

The usefulness of our prediction rests on the accuracy of the overall prediction rather than on the theoretical purity of the structural relationships.

When causality is a major issue
As we discussed in the previous section, when we are not sure of causal relationships in the early paths of a path diagram, we can still arrive at a model that is useful from a marketing point of view. But what if the assumption of causality is a crucial part of the model? For instance, if we have a path model which states that advertising influences sales, causality (ie. advertising influencing sales) is critical to the model and should be dealt with accordingly. In situations like these we can use the Cross-Lagged design. The Cross-Lagged design is used extensively in cases where it is necessary for us to understand the nature of causality.

Cross-Lagged designs
The Cross-Lagged design, often called Cross-Lagged Panel Correlation or CLPC is relatively straightforward

and simple to analyse-despite its imposing name.

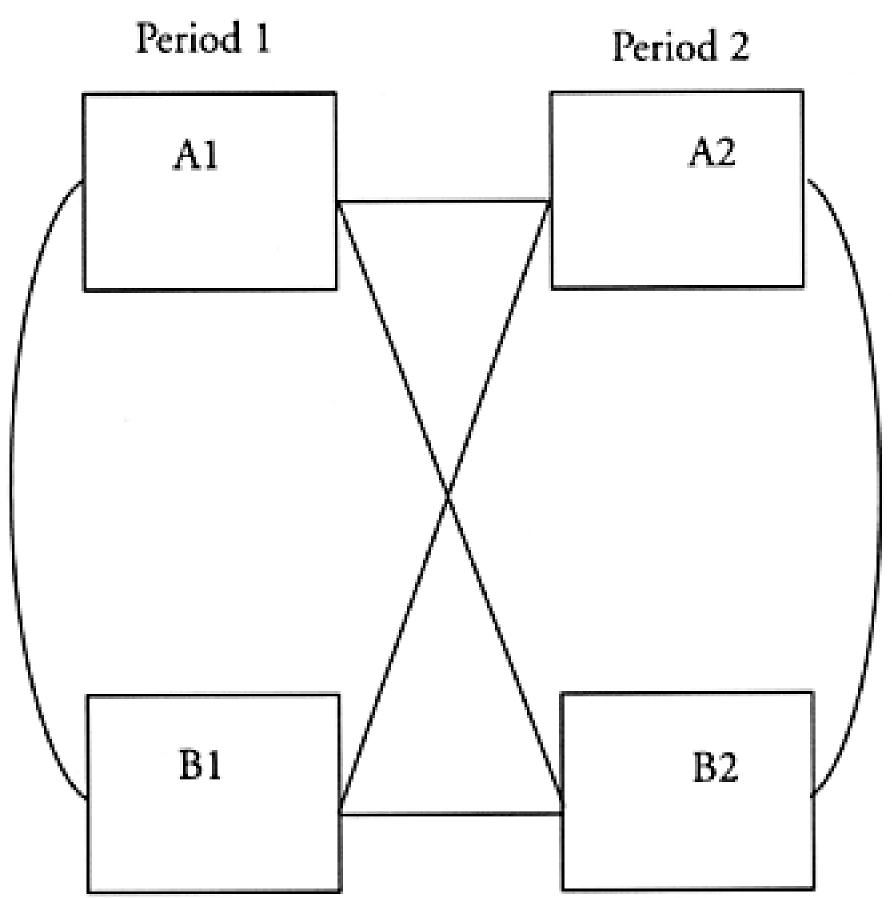
An example is shown in Exhibit 1. Two variables, A and B, are measured at two points in time, denoted as 1 and 2. This generates 6 correlations:

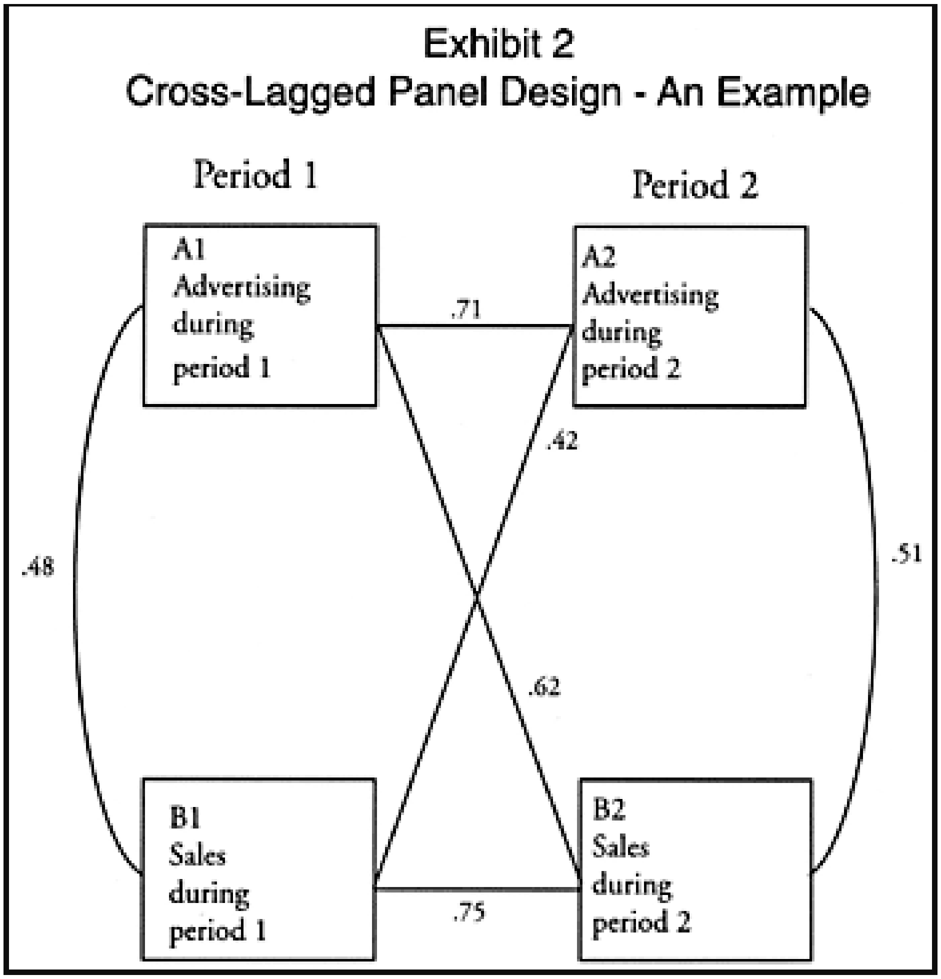
1. Correlation of A1 with A2
2. Correlation of A1 with B1
3. Correlation of A1 with B2
4. Correlation of B1 with A2
5. Correlation of B1 with B2
6. Correlation of B2 with A2

These correlations are grouped into sets of 2:

- **Synchronous correlations.** These are correlations between A1 & B1 and A2 & B2. These correlations are synchronous because the two variables in each pair is measured at the same time.
- **Autocorrelations.** These are correlations between A1 & A2 and B1 & B2. They are autocorrelations because the correlations relate to the same variables.
- **Cross-Lagged Correlations.** These are correlations between A1 & B2 and B1 & A2. They are Cross Lagged because the correlations are between different variables at different points in time.

Exhibit 1
Cross-Lagged Panel Design





A note on the diagram

You may also note that the diagram is different from diagrams which appeared in earlier articles. The lines connecting variables do not have any arrows since we are not sure of the direction of the relationships. Curved lines are drawn for synchronous relations because we do not know the direction of such relationships. (This is not the only way to draw the diagram. In fact different researchers tend to draw the diagram differently.)

The method of simple difference

One way to analyse the data is to use the method of simple difference. Let us consider an example: which path is more prominent -advertising influencing sales or sales influencing advertising?

To answer this question we set up the model (see Exhibit 2) and compute the coefficients as we described in earlier articles. The question to be answered is "Does advertising influence sales more than sales influence advertising?". To answer this question, we go through the following steps.

1. First, we compare the correlation between A1 (advertising in period 1) and B2 (sales in period 2) with the correlation between A2 (advertising in period 2) and B1 (sales in period 1). Suppose we find that the correlation between A1 and B2 is higher than the correlation between A2 and B1.

2. Secondly, we subject the difference to statistical significance testing. If the A1/B2 correlation turns out be significantly higher than A2/B1, then we conclude that advertising influences sales more than sales influences advertising.

The validity of the method of simple difference

Although the interpretation of the simple difference method as described above is generally accepted, questions have been raised about its validity in recent years. For instance, if the reliability of B2 is higher, all correlations with B2 will be higher. This would include the correlation between A1 and B2 (advertising in Period 1 with sales in Period 2). However, it is questionable whether we can infer from this that the direction of causality is from A1 and B2.

We would also expect that a good index of causality will remain stable over a period of time. This condition of *stationarity* is also difficult to prove. Marketing and research data, although relatively stable, seldom show the stability required to establish causality with certainty.

This brings us back to the model-builder. If we cannot infer causality with any great degree of certainty using the techniques we discussed so far, the next best alternative is to create them on the basis of theory. This means that the model-builder simply sets up the direction of causality and states explicitly the reasons for doing so. The reasons can be theoretical, empirical or both. This approach enables others to accept, reject or modify the assumptions of the model builder.

Why bother?
By now it should be obvious that setting up the method of simple difference to assess the direction of causality does not necessarily provide an unequivocal solution. In many cases, it is no more than another piece of evidence (in some cases not a strong one) in deciding the direction of causality. The model builder takes the ultimate responsibility for deciding the casual direction.

Many marketers who may expect the model to provide a definitive answer to marketing problems may be disappointed by the fact that the final resolution rests with the model builder (or the marketer).

So why bother? Why go through the complex model building process if decisions about causality finally rests on your assumptions rather than on the 'real' underlying relationships? There are 3 good reasons for building a model in this circumstance. First, an organizing principle makes us understand the data better. Second, such organizing principles enable us to predict with greater accuracy the effect of our marketing efforts. Third, it enables us to test alternative assumptions and move towards a more stable and defensible model.

A simple causal diagram, even when it is based on crude assumptions, can clarify our thinking, sharpen our perspective, point to data that need to be collected and raise some issues. To this extent, any modelling activity (if done right) forces us to organize our thoughts and examine our assumptions. It is better to make our assumptions explicit then leave them implicit.

More complex models
It is also possible to construct more complex models. Such models can involve techniques like factor analysis to groups variables that tend to correlate highly among themselves. Structural Equation Modelling is another possibility. (We will discuss Structural Equation Modelling later on in this series.)

However, even simple models force us to state our assumptions explicitly. While the techniques for testing our assumptions may not be perfect, they are a good start in understanding the complex structural relationships that underlie marketing dynamics.

Dr Chuck Chakrapani is a Toronto-based consultant, author and seminar leader. He works internationally.