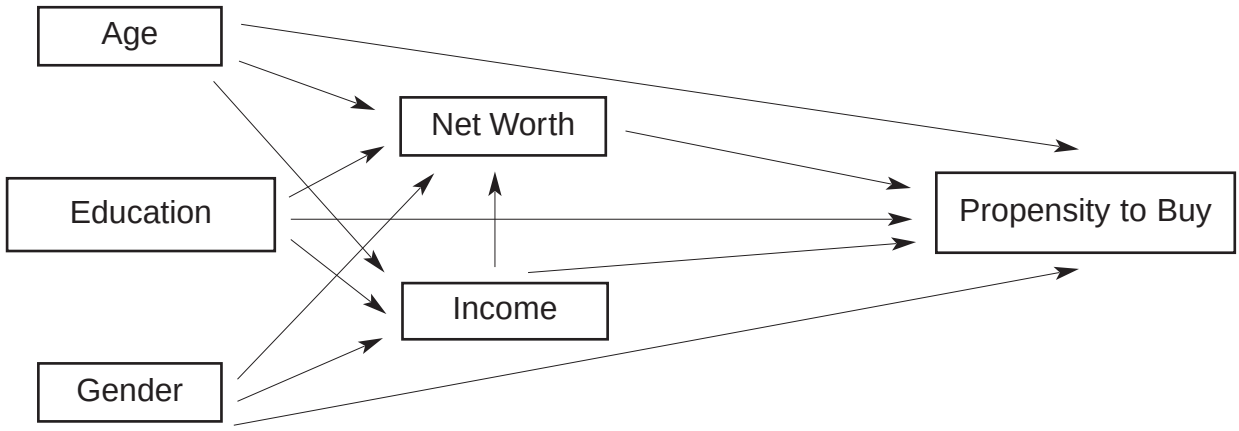
Marketing modelling
Path Analytic Models - 2

By Chuck Chakrapani

It is of great interest to market researchers and social scientists to identify how a set of selected variables influence attitudes and opinions. Path analytic models can be of interest in such situations. As an example, consider a study in which the researcher is interested in understanding how demographic characteristics such as age, income and gender influence a person's propensity to buy a financial product.

Exhibit 1
Just Identified Model



Just Identified Model

To solve the above problem, the researcher may set up a model as shown in Exhibit 1. The model connects every exogenous variable to every possible endogenous variable. This is called a *just identified model*.

Preliminary analysis of the model

The first step in analysing this model is to regress each dependent variable on all preceding independent variables. Thus a variable can both be an independent and a dependent variable depending on the context and where it occurs in the path model. The results are shown in Exhibit 2.

Identifying non-causal paths

Let us first consider the direct effects given in Exhibit 2.

• Propensity to Buy was regressed on five variables (Age, Education, Gender, Net Worth and Income). Of these, Income showed no relationship to Propensity to Buy (Coefficient = 0).

- Net Worth was regressed on four variables (Age, Education, Gender and Income). Education showed no relationship to Net Worth (Coefficient = 0).
- Income was regressed on three variables (Age, Education and Gender).

Exhibit 2
Effects of Independent Variables on Propensity to Buy, Net worth and Income

	Effect			
	Causal			
Variables	Direct	Indirect	Total	Noncausal
		Propensity to buy		
Net Worth	.21	.00	.21	.00
Income	.00	.11	.11	.08
Age	-.17	-.03	-.20	.11
Education	.18	.02	.20	.00
Gender	-.33	.02	0.31	.01
		Net Worth		
Income	.54	.00	.54	-.05
Age	-.15	.00	-.15	-.03
Education	.00	.11	.11	.00
Gender	.21	-.11	-.10	.05
		Income		
Education	.20	.00	.20	.00
Age	.00	.00	.00	.07
Gender	-.21	.00	-.21	.00

The analysis show that the following coefficients are 0:
- Income and propensity to Buy
- Education and Net Worth
- Age and Income.

Now we can respecify the model eliminating these relationships since they contribute nothing to our understanding of a person's propensity to buy, either directly or indirectly.

Overidentified Model

Given that the above three direct effects are 0, we can eliminate the paths from our Just Identified Model and respecify the model. Dropping paths from a Just Identified Model results in what is known as an *Overidentified Model*.

Advantages of an overidentified model

An Overidentified Model has several advantages over the Just Identified Model.
- The Just Identified Model cannot be used for testing the significance of variables. When we include all parameters, the correlations will always be perfectly reproduced.

• An Overidentified Model is parsimonious since it excludes attributes that contribute nothing to our understanding of the dependent variable (in our example, Propensity to Buy).

Further refinement of the model

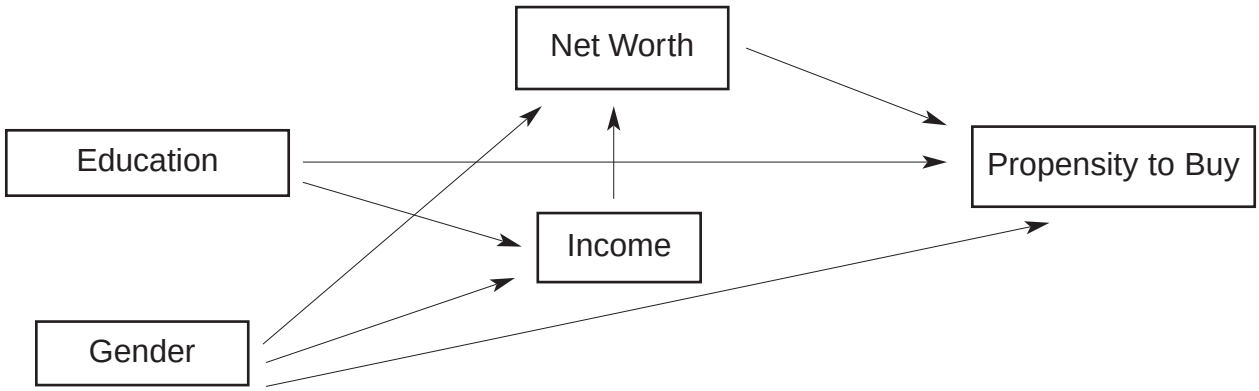
So far we eliminated only the obvious non-relationships. A non-zero coefficient does not necessarily mean that the relationship between two variables is significant from a statistical point of view.

We can now subject the beta coefficients to statistical significance testing. Suppose we decide to keep only those attributes that are significant at the 95% level of confidence. We apply significance tests to all coefficients and find the following two coefficients are not statistically significant at the specified level.

• Propensity to Buy and Age
• Age and Net Worth.

the revised model is shown in Exhibit 3.

Exhibit 3
The Overidentified Model



The logic of the paths

At this stage, the researcher is not simply concerned with statistical significance. The meaning is important as well. (For instance, it may well be that there is no logical connection between Education and Propensity to Buy the product under consideration. Education might have been included in the Just Identified Model only because the researcher might have wanted to assess the indirect importance of Education on Propensity to Buy. If the researcher has reason to believe that Education has no direct effect on Propensity to Buy, s/he can remove the direct path from Education to Propensity to Buy, notwithstanding the statistical criterion. The reason for this is that causality is the working hypothesis of the researcher and is not decided by the model itself.

In essence, the question asked of the model is not “Which of these paths are causally related?” Rather, the question posed by the researcher is “What evidence does the data offer to support my model of causality?” Statistical procedures do not specify the model. They only test it.

The revised model

Assuming that our model does not have any logical inconsistencies, we can simply revise the model based on the statistical significance of the coefficients. After eliminating the paths with zero or non-

significant coefficients, we arrive at a revised model that is considerably simpler than the Just Identified Model. The model clearly shows how easily the Just Identified Model can be modified using simple statistical analysis.

Calculating the revised coefficients

Once we eliminate certain paths from the model, we cannot use the coefficients attached to the remaining paths. The reason is that excluding some paths may change the beta value of the remaining paths. Accordingly, we need to regress the remaining variables as follows:

- Propensity to Buy on three variables (Education, Gender, and Net Worth).
- Net Worth on two variables (Gender and Income).
- Income on two variables (Education and Gender).

Testing the overidentified model

The analysis described will give us revised coefficients for each of the paths. But is the model really valid? To assess whether the model really fits or not, we can subject the model to a chi-square statistical test. If the obtained chi-square value is not significant, it would mean that the model does not fit, even though the individual coefficients themselves were significant.

It is important to distinguish between the two types of significance tests. In the earlier significance test of the Just Identified Model, we were testing whether the individual coefficients themselves were significant or not. When we test the Overidentified model, we are not testing the individual coefficients, but whether the model, as specified by the researcher, is itself valid or not.

Problems with model testing

One major problem with model testing relates to sample sizes. When our sample sizes are small, we may conclude that the model does not fit even when it does, simply because the sample size may not be large enough to detect statistical significance.

One way to overcome this problem (as some statisticians suggest) is to use the Q test which is not affected by the sample size.

The model may fit and yet be incomplete

Even when the coefficients are significant and even when the overall model shows a reasonable fit, it does not necessarily follow that the model is complete. For instance, in our model the exogenous variable, Education, may itself be influenced by variables not in the model. If those variables are taken into account, Education, by itself, may have little impact on Propensity to Buy.

It cannot be emphasized enough that when a model fits, it does not mean that we have established how the phenomenon works. What we have established is as follows:

- We have a formal hypothesis about what variables influence a phenomenon such as Propensity to Buy.
- Statistical evidence offers no evidence that the model is inappropriate.
- Because we have no reason to believe that the model is incorrect, we assume that the model is correct.
- The final coefficients show the exact nature of the relationships among variables given our assumption that we have specified the model correctly and have subjected it to statistical testing.

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