MACRO CONSULTING INC.

WHITE PAPER©

Structural Equation Modeling (SEM) is a powerful, comprehensive technique that combines confirmatory factor analysis and linear regression.

What are Structural Equation Models

A Brief Introduction

Paul Richard "Dick" McCullough

Structure Equation Models are confirmatory models that allow complex relationships to be modeled and statistically evaluated. SEMs simultaneously perform confirmatory factor analysis and linear regression analysis.

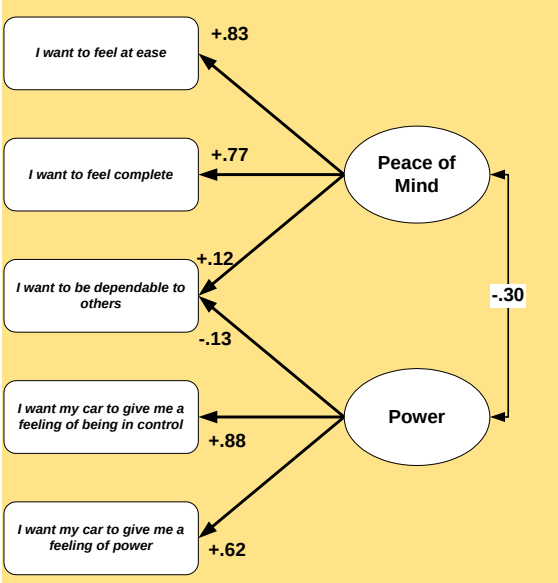
SEMs are often expressed visually in a schematic form called path diagrams. In path diagrams, rectangles represent observed variables and ellipses represent unobserved variables. Observed variables are variables for which we have direct data, such as questions in the new car buyer survey. Unobserved variables are latent factors that are assumed to exist but for which we do not have direct data, such as attitude factors or brand image halo effects. Arrows leading from latent factors (unobserved variables) to observed variables represent factor loadings or regression coefficients. Double-headed arrows connecting two unobserved variables represent covariance between the two latent factors.

The confirmatory factor analysis portion of the structural equation model is called a measurement model. A simple measurement model is illustrated next page in Figure 1.

The arrows leading away from the elliptical shapes represent the relationship between an individual statement (shown as a rectangle with rounded corners) and a latent factor (the ellipse). The number next to the arrow reflects the strength of relationship. The larger the number, the more strongly the statement loads on the factor. The double-headed arrows reflect covariance between factors. By closely examining the load factors and covariances, potentially rich and nuanced interpretations of the latent factors can be developed.

In Figure 1, for example, it can be seen from the covariance of -.30 between factors Peace of Mind and Power that the more a respondent feels a desire for peace of mind, the less likely they are to feel the desire to be powerful. Further, the factor loading of -.13 indicates that if they have a desire for power, they tend not to want to be dependable to others. Conversely, the load factor of +.12 indicates that if they feel a need for peace of mind, they do want to be dependable to others. These findings are entirely fictitious and displayed to illustrate the concept only.

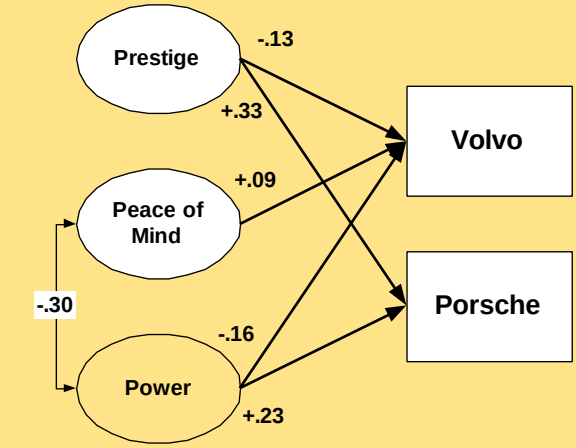
Figure 1



Factors derived in the confirmatory factor analysis can be included as independent variables in the regression models. There can be multiple regression models allowing for multiple dependent variables. The regression model portion of the structural equation model is called the structural model or the causal model. A structural model is illustrated below in Figure 2.

In the model displayed in Figure 2, it can be seen that Volvo purchases are driven by a desire for Peace of Mind (+.09) and negatively driven by desires for Prestige (-.13) and Power(-.16). Again, these findings are hypothetical.

Figure 2



Hypothetical models are specified based on current category understanding. Survey data are fitted to these models validating or rejecting a specific hypothetical model structure. The final model structure reflects valid causal relationships between attitudinal factors and product attributes, creating genuine insight into customer attitudes that motivate relevant behaviors.

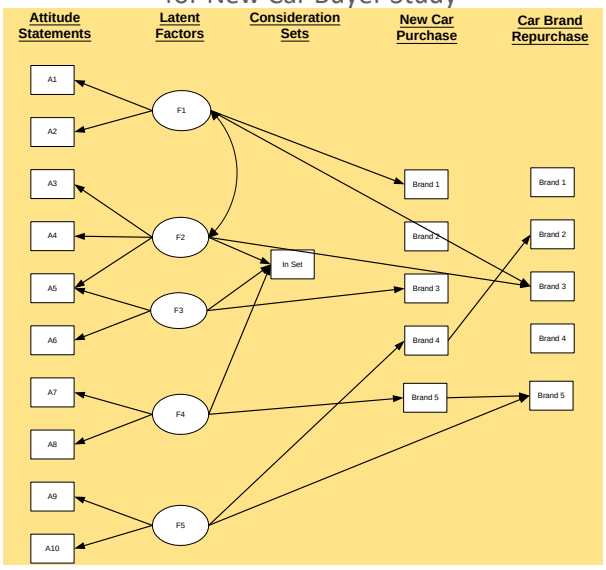
Attitudinal and belief factors can be derived from attitudinal statements in the survey instrument. By using multiple attitude statements to derive latent factors, reliability and validity of the attitudinal factors are greatly enhanced. Car purchase behavior can be used in the SEM as dependent variables as can stated consideration sets. The attitudinal factors can be regressed against the car buying behaviors in a series of simultaneous linear regression models.

A simple path diagram of a fictitious SEM combining the measurement model and the structural model is shown below in Figure 3 to illustrate the suggested model framework within the context of a new car buyer study.

Commercial models will be substantially more complex. One can see in Figure 4 an SEM example based on a commercial dataset. This SEM displays a complex web of variable relationships; observed to unobserved, observed to observed and unobserved to unobserved. Each line in the path diagram in Figure 4 represents a statistically significant relationship between two variables. It is important to note that the analyst must also test each potential model parameter for interpretability. Given the large number of potential parameters that could be included in the model, the likelihood of spurious relationships can be very high. SEMs must be built with an eye toward reasonableness as well as statistical significance.

Figure 3

Hypothetical SEM
for New Car Buyer Study



Notice in Figure 3 there are rectangles for being in the consideration set (In Set) and also for various brands. These serve as dependent variables in the model. The ellipses represent the attitudinal factors that predict the outcomes or behaviors of the dependent variables. Notice also that brands occur twice, once for most recent purchase and once for next purchase. The purchase of certain brands of cars will predict the next purchase. Those brands will have strong brand loyalty. This will be indicated by an arrow from the brand in the first purchase column to the same brand in the next purchase column.

Figure 4

Actual SEM
based on Commercial Survey Data

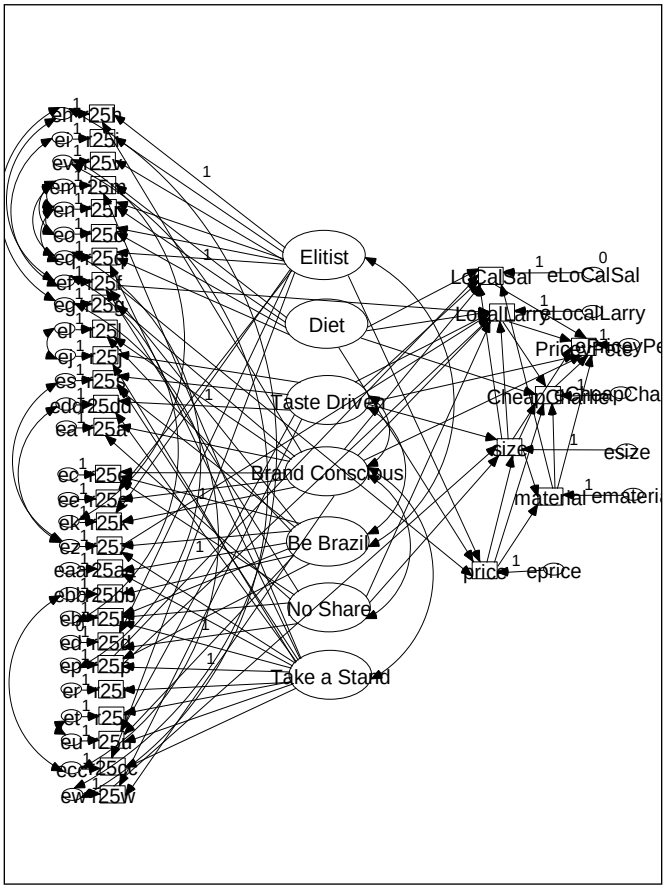


Figure 4 above demonstrates that SEMs have the potential to identify and quantify extremely complex relationships between attitudes and behaviors. Understanding these relationships can create rich interpretations of market dynamics that could have significant impact on business strategies.

Because SEMs are potentially complex models with a large number of parameters to estimate, sample sizes are typically at least 400 for simpler models up to one thousand or more for more advanced models.

MACRO CONSULTING, INC.
Marketing Research / Advanced Analytics

Telephone: 650-823-3042

General Inquiries:
info@macroinc.com

Advanced Analysis Inquiries:
analysis@macroinc.com

richard@macroinc.com

www.macroinc.com