

Dissertation topic: An investigation of causal inferences of Structural Equation Modeling in terms of intervention, counterfactual and realness of constructs

Chong Ho Yu

Many statisticians and social scientists subscribe to the view introduced by Cook and Campbell that causal inferences can be supported by experimental data, while only descriptions can be supported by non-experimental data. Clark Glymour has challenged this view with his use of Structural Equation Modeling (SEM). Glymour, Pearl, Woodward, and a few other scholars have argued that manipulation and intervention do not necessarily happen at the experimental level; in SEM the researcher can still conduct interventions by including/excluding variables, and by manipulating the structure of the model that represents the data. Glymour and his associates developed an algorithm named TETRAD, a plug-in to standard SEM software programs, that is capable of exhausting all possible combinations of paths in an SEM model. This approach is based on Fisherian counterfactual reasoning, which strives to yield the best model after exploring all possible scenarios.

It is noteworthy that persons who debate this approach have paid most attention to the “intervention” aspect, but have overlooked another important one. A SEM is composed of two models, namely, the factor model and the path model. The former is concerned with how observed variables can be combined to form theoretical constructs (factors), while the latter is about ordering the paths of these constructs in order to develop the structure of the model. The nature of theoretical constructs has been a battleground between realists and anti-realists. Larry Laudan bluntly proclaimed that factor analysis developed by early psychometricians had no ontology; factors are grouped only for mathematical convenience. If constructs (factors) are not real entities, then is it meaningful to talk about the cause and effect relationships among factors? Before the 16th century, supporters of the geocentric model proposed epicycles and equants to explain the behavior of heavenly bodies. Today we know that these constructs were used for saving the phenomena, and that none of them has any corresponding mapping to the actual movements.

It may be strange to compare epicycles and equants to psychological constructs such as “depression,” “anxiety,” “temporary insanity,” and “learned helplessness,” because these psychological constructs seem to be real. As a matter of fact, modern scholars have questioned

the reality of certain constructs that have been taken for granted for years. For example, Furedi points out that in Rwanda, people do not have words corresponding to “trauma” and “depression” even though Western scholars assume that some standard psychological model can be employed to explain the 1994 massacre.

Critics of Glymour’s approach also object to its imposition of strong assumptions on modeling, such as the Markov Chain condition and the faithfulness assumption. The justification offered by Glymour is that the stronger the assumptions the researcher makes, the more he/she can learn. As a matter of fact, many standard statistical procedures also require certain strong assumptions. For example, Analysis of Variance requires homogeneity of variance, normal distribution, and independence of observations. Regression analysis requires random residuals and the absence of multi-collinearity...etc. However, when some or all of those assumptions are violated, there are always other alternate procedures, such as non-parametric tests, which do not require any parametric assumptions, and robust procedures, which are immune to the violation of those assumptions. The weakness of Glymour’s methodology is obvious—it does not offer any alternatives when assumptions may not be met.

This project has four objectives:

1. Using constructs in the social sciences as examples, to discuss how the ontological status of constructs affects our understanding of an alleged causal model such as SEM. To be specific, when the ontological status of constructs is undetermined, to what extent can causal inferences be made in SEM?

2. To examine the usefulness and limitations of intervention or manipulation in SEM to uncover causal relationships.

3. To examine the applicability and limitations of counterfactual reasoning as the foundation of discovering the best model by exhausting all possible combinations.

4. To evaluate whether the faithfulness assumption and Markov Chain condition, as required in Glymour’s SEM, are reasonable assumptions, and how they may affect the causal inferences when the assumptions are violated.