Propensity Score Analysis in Structural Equation Modeling: A New Approach Dealing with Selection Bias in Quasi-Experimental Studies

Yonghwan Chang, University of Florida
Yong Jae Ko (Advisor), University of Florida
Walter Leite, University of Florida

Research/statistical methodology Thursday, May 29, 2014 20-minute oral presentation (including questions) (Allegheny)

Abstract 2014-014 10:35 AM

In social science including sport management, structural equation modeling (SEM) has been an extremely popular tool especially for the examinations of latent variables and multisampling models (MacCallum & Austin, 2000). However, one common concern scholars have when using SEM is a selection bias problem that occurs due to potential covariates. For example, in examining the influence of team identification on consumption behavior, scholars may face serious problems associated with such important covariates as past consumption experience, socio-demographic variables, cognitive and affective involvement level, personality traits, need states, and current emotional states. As such, most researchers include several most viable factors based on parsimony principle in their structural model because it may not feasible to include all covariates. However, the selection of covariates may lead to biased interpretation of the result due to selection bias.

Propensity score analysis

As an effective method of controlling selection bias, we would suggest to apply the propensity score analysis (PSA; Rosenbaum & Rubin, 1983). PSA is a set of statistical techniques that allows the researcher to perform quasi-experimental analyses on survey data (Austin, 2011). This approach enables scholars to significantly reduce selection bias that occurs due to observed covariates (Thoemmes & Kim, 2011). Since most observational studies have the same or similar restraints (i.e., selection bias), PSA receives tremendous attention from scholars in many scientific areas, including the social sciences where randomization is not feasible (Thoemmes & Kim, 2011).

The propensity score refers to the conditional probability of treatment, which expresses how likely a participant is to be assigned or to be selected for the treatment condition, given certain observed baseline covariates. Propensity scores are estimated using logistic regression, probit regression, discriminant analysis, or boosted regression trees that use multiple background information to predict the probability of receiving the treatment (McCaffrey et al., 2004; Rosenbaum & Rubin, 1983). Once the propensity score is estimated, the data is conditioned using various types of PSA, such as matching, weighting, and stratification (Thoemmes & Kim, 2011).

The successful implementation of PSA is determined by evaluating the overlap of propensity scores and covariate balance. Overlap refers to the degree to which the distributions of propensity scores for treated and untreated are similar (Stuart, 2010). The area of the distribution of propensity scores where there is adequate overlap is known as the ‘area of common support.’ If both overlap and covariate balance are achieved, and there are no unobserved covariates, then the treatment assignment is strongly ignorable (Rosenbaum, 1984) and causal inference is possible. Therefore, the assumption of strong ignorability requires that the treatment assignment is independent of the potential outcome distributions, given observed covariates. The strong ignorability of treatment assignment assumption also requires that no values of any covariate are associated with a treatment probability of zero or one (i.e., 0 < P [T = 1 | Z] < 1).

Finally, after the balancing process carried out using various types of PSA, one can estimate treatment effects, and the specific estimator used may vary widely. There are two types of estimation of treatment effects including (1) average treatment effect (ATE) and (2) average treatment effect for the treated (ATT). ATE is the average effect, at the population level, of moving an entire population from untreated to treated. ATT is the difference between the mean outcomes of the treated and the mean potential outcomes of the treated if they had not been exposed to treatment.

An application of structural equation modeling with propensity scores
Based on this understanding, we suggest a new method to deal with selection bias, an integrative analysis of SEM with marginal mean weighting through stratification (MMW-S). MMW-S is a newly emerging PSA that combines key elements of propensity score: weighting and stratification. MMW-S generates more robust analytic results when compared to other widely used PSA, such as matching, stratification and weighting (Hong, 2010). Hong (2011) recommended specific steps for this integrative analysis: (1) sorting all the units in the analytic sample in an ascending order on the logit propensity score, (2) dividing the sample into five strata with equal proportions of units, as Cochran (1968) suggests, five strata typically remove about 90% of the selection bias, (3) checking common support by comparing the range of the estimated logit propensity score across the treatment groups within each stratum, and (4) computing the marginal mean weight for either ATE or ATT through the following formulas:

For ATE, \( MMW-S = \frac{ns \cdot \text{pr}(Z=0)}{nz,0} \)  
For ATT, \( MMW-S = \frac{nz=1,s \cdot \text{pr}(Z=0)}{nz=0,s \cdot \text{pr}(Z=1)} \) if \( Z = 0 \); \( MMW-S = 1 \) if \( Z = 1 \)

The authors will present this innovative method by using empirical study with data gathered from golf club customers (N = 333). Briefly, we used MMW-S to adjust 16 observed covariates including personality traits, socio-demographic factors, and patron status in the examination of the moderating effect of product value orientation (i.e., private club: n = 161 vs. public club: n = 172) in the relationship between four brand leadership factors (i.e., quality, value, innovativeness, and popularity) and word-of-mouth recommendation. Then, we estimated both the ATE and moderated effects using multiple-group SEM (Kaplan, 2009). We implemented this newly developed method using R 2.15.2 (R Development Core Team, 2012).

The results showed significantly different patterns given that a series of hierarchically nested multiple-group CFA models allowed for a valid comparison between means of latent variables across groups. SEM model fits were also acceptable. For example, before performing MMW-S, quality level of golf course was significant (\( \beta = .33, p = .00 \)) for public club consumers when they recommend the golf course to others, while after performing MMW-S quality level of golf course was not significant (\( \beta = .13, p = .20 \)) for public club consumers when they recommend the golf course to others.

Methodological implications

From a methodological point of view, the current research first introduces an integrative analysis with MMW-S and SEM in the social science area; in fact, no study applies the same type of analysis presented in the current study in any fields except for few similar applications (e.g., Hoshino et al., 2006; Kaplan, 1999; Leite et al., 2012). It is believed that the SEM results estimated after MMW-S should be more robust than the SEM results estimated before MMW-S because MMW-S shows strong reduction in selection bias, meaning that MMW-S increased the internal validity of the results. This result may have important implications for the issue of selection bias in future observational research since existing research mostly neglect this issue, even when they have either the same or similar limitations; this is an important matter that must be addressed because such neglect generates a critically biased interpretation. As Thoemmes and Kim (2011) suggest, propensity score methodology and associated procedures are most likely going to remain as statistical tools in the area of social science, including sport marketing and management.