9  A Multi-trait Multi-method Validity Test of Partworth Estimates

Wagner Kamakura and Muammer Ozer

9.1  Introduction

Conjoint analysis has already been widely accepted by marketing researchers as a popular instrument for the measurement of consumer preferences. Typical applications of conjoint analysis include new product design based on the relationship between product features and predicted choice behavior, benefit segmentation based on attribute preferences, etc. The popularity of conjoint analysis among marketing researchers hinges on the belief that it produces valid measurements of consumer preferences for the features of a product or service, and that it provides accurate predictions of choice behavior.

Given this importance, a vast literature has already emerged on the validity of conjoint analysis for the measurement of preferences. However, as we will discuss it in more details in our literature review section, these tests usually involve collecting data from two separate conjoint tasks (usually in the same interview), estimating partworths based on the first task, using the estimates to make predictions about the second task, and measuring predictive fit. We argue that these tests are more akin to test-retest reliability assessments than validation tests. Our literature review also indicates that comparative conjoint studies have yielded inconclusive results partly due to the different validity measures used as a basis for validation.

The purpose of this chapter is to compare various conjoint models across their partworth estimates based on actual behavior. Because we are comparing those methods across their partworth estimates, we use a Multitrait-Multimethod (MTMM) framework to assess the relationships among the methods and the partworth estimates. We test the relationships by using both the traditional MTMM analysis and a direct product methodology. The following section presents our literature review. After that, we provide a summary of the MTMM and direct product methodologies. Then, we present details about our research design and the conjoint models that we used to generate partworth estimates. Finally, we discuss our results and conclude the chapter with managerial and research implications.

9.2  Literature Review

Previous studies have proposed a number of methods to improve the validity and reliability of conjoint analysis. For example, Hagerty (1985) used a Q-type factor analysis to determine optimum weights that optimize the expected mean squared
error of prediction in a validation sample. He tested his methodology by using both synthetic and real data. He manipulated the amount of overlaps in the clusters of partworth estimates in two Monte Carlo simulations. He also asked student subjects to rank-order eighteen job descriptions based on five attributes. Two of the eighteen job descriptions were used as a holdout sample. He compared his results with those of non-overlapping clustering and individual-level clustering. Based on Mean Squared Error (MSE) and the first choice prediction criteria, he showed that his methodology yielded higher predictive accuracy.

In another study, Kamakura (1988) used an agglomerative hierarchical method for simultaneous segmentation and estimation of conjoint models. He used a least squares procedure to identify segments that maximize the predictive validity of the segment-level partworth estimates. He compared his results with those of the two-stage segmentation procedure (individual-level estimation and clustering of subjects based on partworth estimates) by using both synthetic and empirical data based on holdout samples. His synthetic data included simulated preference rankings whereas his real data consisted of preference ratings of twenty-seven full-profile descriptions of checking account services. The holdout sample included eight profiles with the same attributes. He concluded that his methodology and the two-stage segmentation procedure yielded similar partworth estimates, but his proposed methodology gave consistently more accurate results. Ogawa (1987) also suggested a similar procedure, but used a logit-based estimation methodology. By using both simulated data and a set of preference data for Japanese automobiles, he was able to show that his partworth estimates were internally valid. Similar to earlier studies, his validity assessment was also based on a holdout sample.

Wedel and Steenkamp (1989) proposed a fuzzy clusterwise regression algorithm to allow consumers to possess partial membership in multiple segments. Wedel and Steenkamp (1989) compared their procedure with a clusterwise regression and Hagerthy’s optimal weighting method. They first used a simulated data to validate the computational efficiency of their method. They later used a data set about customer satisfaction with respect to eight stock market scenarios and another data set for meat products. The results of the simulation showed that the methodology was computationally sound. In addition, judging by the percentage of first choices accurately predicted in a holdout sample with similar tasks, they concluded that their method and the clusterwise regression gave consistent results. They also concluded that their results were slightly better than those of Hagerthy’s method for well-defined clusters, but were worse for diffuse clusters. Similarly, DeSarbo et al. (1992) introduced a latent class methodology that allowed overlapping clusters for simultaneous segmentation and parameter estimation by using a mixture of multivariate conditional normal distributions. They successfully applied the methodology to a conjoint experiment on remote controls for cars, simultaneously identifying segments and generating partworth estimates within each segment. When they compared their methodology with various OLS procedures, they found that their latent class procedure outperformed them based on a likelihood-ratio test and a goodness of fit index (R-square).