

# Survey Data Modeling With Mplus

Tihomir Asparouhov  
tihomir@statmodel.com  
*MUTHEN & MUTHEN*  
*3463 Stoner Ave.*  
*Los Angeles, CA 90066*  
*USA*

Bengt Muthén  
bmuthen@ucla.edu  
*Graduate School of Education & Information Studies*  
*UCLA*  
*Social Research Methodology Division*  
*Los Angeles, CA 90095-1521*  
*USA*

## **Abstract**

This talk presents new Mplus features for latent variable analysis with complex survey data. A first part of the talk presents the general latent variable modeling framework underlying Mplus. This framework allows unparalleled modeling flexibility with a variety of observed dependent variable types, both continuous and categorical (mixture) latent variables, and several ways of handling complex survey data. The framework includes analyses as varied as SEM, IRT, growth modeling, latent class analysis, and multilevel modeling. A second part of the talk discusses technical issues related to handling multivariate modeling with complex survey data that features stratification, unequal probability sampling (weights), and clustering. Estimation techniques implemented in Mplus for survey data analysis are described, such as pseudo maximum-likelihood, pseudo weighted least

squares and multilevel pseudo maximum-likelihood. The likelihood ratio test and tests of model fit available in Mplus for complex survey data are presented. The effects of different design features on parameter estimation and chi-square testing are discussed. An overview of weighted multilevel estimation in Mplus is given. Various factors that affect the quality of the estimation are discussed and evaluated through simulation studies, such as size of clusters, level of informativeness, intraclass correlation and effective sample size. Comparison between different latent variable estimation techniques for complex designs with small number of PSUs is presented as well. A third part presents latent variable applications to a national survey on alcohol dependence and abuse. Applications focus on new types of latent class and factor mixture models allowing for complex survey data, including multilevel latent class factor analysis. Nonignorable missing data modeling is also illustrated. Comparisons are made with conventional techniques ignoring various features of the complex survey design.