Advances in Multilevel Modeling for Educational Research: Addressing Practical Issues Found in Real-World Applications

Day 1: Friday, November 14

Presenter: Sophia Rabe-Hesketh, Keynote Address (9:00am – 9:45am)

Title: Avoiding Omitted-Variable Bias in Multilevel Models

Abstract: Multilevel models are popular in education because the data tend to be clustered, with students (level 1) nested in classes or schools (clusters at level 2), or with occasions in a longitudinal study (level 1) nested in students (clusters at level 2). Multilevel models include covariates at both levels as well as cross-level interactions. Random intercepts and slopes represent any remaining between-cluster variability due to unobserved (or omitted) cluster-level covariates. If these unobserved covariates are correlated with covariates included in the model, then the random effects are correlated with the covariates and standard (restricted) maximum likelihood estimators of the regression coefficients suffer from omitted-variable bias. In that sense, inclusion of random-effects for clusters does not "control" for clusters. A popular solution to this problem is to use fixed-effects approaches, for example by including dummy variables for clusters. However, a problem with fixed-effects approaches is that they do not provide estimates of the coefficients of cluster-level covariates, and these coefficients are typically of central interest in education. In this talk, I will discuss simple methods for estimating all regression coefficients of multilevel models when random intercepts are correlated with level-1 covariates. I will then extend these approaches to the case of random slopes that are correlated with level-1 covariates, emphasizing the important additional problem and its solution: correlations of random slopes with variances of level-1 covariates. This is joint work with Anders Skrondal, Katherine Castellano, and Michael Bates.

Presenters: Ann O'Connell, D. Betsy McCoach & Gloria Yeomans-Maldonado (9:55am – 10:40am)

Title: Best Practices in Residual Diagnostics and Model Assessment in a Multilevel Framework

Abstract: Diagnostic approaches and assessment methods helpful in evaluating the prediction quality of single-level regression models are widely used, but corresponding residual techniques are not as developed or consistently applied to multilevel samples. This is problematic, since the quality of research conclusions is impacted by the quality of the models used to generate those conclusions. An informal review of published articles in top-tier educational research journals in the last year revealed that none of the multilevel articles presented residual diagnostics, and only one reported having looked at raw residuals. Our goal is to contribute to researcher understanding of the application of diagnostic methods useful for improving the derived model and in identifying unusual cases within multilevel studies. We begin with a review of basic multilevel model assumptions, including the interpretation and estimation of residuals. Next, we summarize the literature on multilevel model diagnostics, clarifying software-specific differences in available residual diagnostic methods, drawing parallels with its single-level counterparts. We propose best practice guidelines, and two multilevel applications are provided.

Presenters: Hong Jiao, Akihito Kamata, & Chao Xie (10:50am – 11:35am)

Title: Multilevel Cross-Classified Testlet Model for Complex Item and Person Clustering in Item Response Modeling

Abstract: Local independence is one of the assumptions for item response theory (IRT) models. Essentially, the applications of IRT models assume that independence holds among items after controlling for the assessed latent trait; the same is true of person independence. IRT models are not robust to violation of the independence assumption, however. This study introduces a multilevel cross-classified testlet model in accounting for complex item and person clustering effects in item response modeling. The complexity of item and person clustering lies in that items are not only clustered in testlets where local item dependence is likely to be present, but also clustered due to the assessment of different content areas. Thus, items are cross-classified by two types of grouping variables: testlets and content areas. Further, persons are clustered in such scenarios as students nested within schools. This study explores the model parameter estimation of the proposed model. It uses simulation data to demonstrate the effects of ignoring such complex clustering structure related to both items and persons in comparison with other competing models.

Presenters: Ji Seung Yang & Michael Seltzer (11:45am – 12:30pm)

Title: Handling Measurement Error in Predictors with a Multilevel Latent Variable Plausible Values Approach

Abstract: To handle measurement error and sampling error in predictors more properly in multilevel modeling settings in which latent predictors are measured by observed categorical variables, we propose a multilevel latent variable plausible values (MLVPV) approach. This approach draws substantially from the work of Mislevy, Beaton, Kaplan and Sheehan (1993), along with key ideas in the multiple imputation literature more generally, and in the Bayesian statistics literature, notably the concept of exchangeability. In this presentation, we outline the MLVPV approach and discuss its use in analyzing the relation between teacher reading instruction practices and student reading achievement at the end of first grade using data from the ECLS-K 1998-1999 cohort. MLVPV consists of two stages. The first stage entails specifying and imputing sets of values from a multilevel measurement model for the key practice measures that we wish to treat as latent variable predictors in our analyses of reading achievement outcome scores. The second stage consists of fitting multilevel models to the outcome data employing the imputed practice values as predictor variables. Two different approaches of drawing plausible values for the latent variables are discussed and compared through a simulation study.

Presenters: Bengt Muthén & Tihomir Asparouhov (2:00pm – 2:45pm)

Title: Bayesian Analysis of Multiple Indicator Growth Modeling Using Random Measurement Parameters Varying Across Time and Person

Abstract: Growth modeling for a factor measured by multiple indicators needs to account for possible measurement non-invariance. One classic growth model uses a wide-format, single-level approach but this leads to analyzing a large number of variables, with this number being the product of the number of factor indicators per time point and the number of time points. This enables the study of measurement invariance, although it is cumbersome due to the large number of measurement parameters. Another classic approach is to use a long-format; two-level approach where time is level 1 and subject is level 2. The number of analysis variables is now only the number of factor indicators per time point, but full measurement invariance has to be assumed. A third approach is to view item as level 1, time as level 2 and subject as level 3, but this enables measurement non-invariance with respect to only the intercepts. A new approach using Bayes Theorem allows measurement parameters to be random, varying across time and/or subjects. The Bayesian approach uses a cross-classified model where each observation is classified by the random modes of time and subject. This makes it possible to let the measurement intercepts and factor loadings be random parameters where their variation captures non-invariance across both time and subjects.

Presenters: Paras Mehta & Yaacov Petscher (2:55pm – 3:40pm)

Title: Everything I Know I Learned in Kindergarten: Potential Teacher Effects in Longitudinal Student Outcome Data

Abstract: Modeling growth in student achievement across years invariably involves a complex set of dependencies due to changing classroom nesting structure and student mobility. Traditional multilevel cross-classified models for such data make restrictive assumptions, particularly regarding persistent classroom/teacher effects across grades. Furthermore, the specification of the model itself tends to be unwieldy with many levels and complex dependencies. A general n-Level Structural Equation Modeling (NL-SEM) approach for complex dependent data is introduced. NL-SEM allows models with arbitrary number of levels. Each level may include a complete SEM model with observed and latent variables. Regression among observed and latent variables is allowed across any two levels that share a parent-child relationship. In effect, a full NL-SEM model is a DAG of SEM models. An empirical example illustrating alternative NL-SEM specifications of 'persistent teacher-effects' is presented using a large dataset of students' reading outcomes from grades 1 through 3. Interestingly, the data includes ID variables for students' kindergarten teachers even though student outcome data were not collected in kindergarten. The results indicate that kindergarten classrooms/teachers continue to have strong and persistent effects on student outcomes. Classrooms in later grades have a relatively smaller effect. Implications for "value added models" of teacher effects are discussed.

Presenters: Jeffrey Harring, S. Natasha Beretvas & Anita Israni (3:50pm – 4:35pm)

Title: A Model for Cross-Classified Nested Repeated Measures Data

Abstract: It is becoming increasingly common for educational researchers to measure students' change in attributes, skills or achievement over time while concurrently collecting data about those students' teachers. Multilevel modeling has been effectively utilized in this scenario to adequately summarize student growth and investigate determinants of this growth originating at both the individual and teacher levels. Embedded in this scenario are two methodological nuances that taken together have yet to be addressed in the methodological literature. The first is that measuring student growth over time typically necessitates that students are changing classrooms, and likely changing teachers. Cross-classified models are required to account for the variation among teachers. Secondly, teachers may be measured at multiple times each year and these times may or may not coincide with when student measurements are gathered. Thus, a model that takes into account the multilevel cross-classification aspects of the longitudinal design and permits the modeling of nested change processes must be employed. This presentation is devoted to the modeling of data in this scenario. Two real examples will be used to motivate the methodology and demonstrate its utility in answering interesting substantive questions.

Reception & Poster Session (5:00pm – 7:00pm)

SMEP & Pearson Travel Award Winners
Poster Presenters

Presenter: Brenden Bishop, Ohio State University
Title: Comparison of Bayesian and Frequentist Estimation for Multilevel Models
Abstract: There are compelling reasons why even the most seasoned quantitative specialists should take note of modern Bayesian modeling methods, in particular within the context of fitting complex multilevel models. Using a real dataset from a learning experiment, a comparison will be made between maximum likelihood (ML) estimation and Bayesian estimation using a simple change-point model. It will be shown that population average versions of such models are easily fit within both frameworks, but subject-specific models become difficult to fit in ML beginning with three random effects. An attempt will made using the method of Du Toit & Cudeck (2009) to accommodate a four dimensional multilevel mixed-model into being fit via ML. This reparameterized model will also be fit within a Bayesian framework. Finally, a change-point model with a smooth transition will be attempted, since continuous parameter gradients may aid ML estimation.

Presenter: Bruce Austin, Washington State University
Title: Use of Standard Deviations as Predictor Variables in Multilevel Models
Abstract: This study explored the use of standard deviations as predictors in multilevel models of teacher and student data for providing significant new sources of explanatory information. Variables included within-school standard deviations of attitudes, educational experience, and background of students from eight countries. Teacher outcomes were used as dependent variables. The standard deviations for student socio-economic status and mathematics achievement were significant predictors. Two-way interactions involving standard deviations with other measures were significant across outcomes and countries. The contribution of standard deviation to variance explained produced similar results to that of the mean aggregated variables for the same constructs.

Presenter: Chunhua Cao, University of South Florida
Title: Estimating Cross Level Covariate Moderation Effect in Multilevel MIMIC Models
Abstract: For multiple group analysis under the context of multilevel structural equation modeling (SEM), multilevel multiple-indicator multiple-cause (MIMIC) model has also been studied because of its flexibility in modeling the observed and latent variables, and in modeling multiple covariates and the interaction of those covariates. When one of the covariates is in the between level (e.g., private vs public school), and the other one in the within level (e.g., boy vs girl), and the effect of type of school on a latent variable varies for boys and girls, there is cross-level covariate interaction. This poster illustrates the performance of multilevel MIMIC in estimating cross level covariate interaction effect.

Presenter: Veronica Cole, University of North Carolina, Chapel Hill
Title: Recovering Highly Correlated Time Trends in Complex Multilevel Growth Models
Abstract: In longitudinal applications of multilevel models, relationships between growth parameters are often of interest. However, if growth parameters are highly collinear, these relationships may not be estimable. Fortunately, a number of estimation methods allow for the covariance matrix of growth parameters to be approximated in the event that an unstructured estimate would not be positive definite. The current simulation study compares several factor analysis-based methods of approximating the covariance matrix of the growth parameters. The approximation methods are compared in terms of bias of covariance parameter estimates, and recommendations are made for which method to use in different settings.

Presenter: Pega Davoudzadeh, University of California, Davis
Title: Early School Readiness Predictors of Grade Retention from Kindergarten through Eighth Grade: A Multilevel Discrete-time Survival Analysis Approach
Abstract: A crucial indicator of academic success is whether a child repeats a grade. Studies have found grade retention to be predicted by test scores, behavior, and demographic factors such as gender, age, ethnicity, and socioeconomic status. In the current study, using a multilevel, discrete-time survival analysis approach, we examined if and when a child is retained and its association with school readiness indicators using data from the Early Childhood Longitudinal Study – Kindergarten cohort. Grade retention was generally a rare occurrence, but found to be associated with child characteristics and early academic skills.

Presenter: Matthew Gaertner, Pearson
Title: Rescue Mission: Using Multilevel Models to Estimate the Effects of School Closure on Displaced Students
Abstract: In early 2006, the Riverside School Board voted to close Jefferson High School – an institution that had served the urban district's African-American and Latino populations for more than a century. Jefferson's 9th, 10th, and 11th grade students were instructed to transfer to other schools in the district to complete high school. Jefferson students had been struggling academically for years, so the school district touted the closure as a "rescue mission." Displaced students would have the opportunity to relocate to higher performing schools and would therefore stand a better chance of progressing academically and graduating on time. This research, based on multilevel models of standardized achievement scores, graduation, and dropout, suggests the closure had the opposite effect. Displaced students' test scores – rising prior to the closure – declined in the years that followed it, widening the already sizeable achievement gap between Jefferson students and their district peers. Graduation and dropout trends followed suit. Displaced students were more than twice as likely to exit early and 31% less likely to graduate on time, compared to similar students who had attended Jefferson prior to its closure. In isolation, declines in test scores could be cast as students simply protesting the test-based accountability policies that contributed to Jefferson's closure. In concert, however, negative trends across academic outcomes suggest that for the students immediately impacted, this policy decision intensified academic challenges rather than easing them.

Presenter: Megan Kuhfeld, University of California, Los Angeles
Title: Multilevel Item Factor Analysis of Student Ratings of Instructional Practice
Abstract: Student surveys of teacher practice and classroom environment are growing in popularity as a part of teacher evaluation systems. These surveys are traditionally scored by using sum or mean score methods that do not account for background covariates or possible survey response styles. This study examines a multilevel multidimensional item factor model with latent regression approach to scoring teachers based on students' ratings, compared with a traditional sum score approach, to estimate the level of bias in estimated scores from not properly accounting for key background characteristics in simulated data where there are true covariate effects.

Presenter: Mark Lai, Texas A & M University
Title: Bootstrap Confidence Intervals for Multilevel Effect Size
Abstract: Although many methodologists have urged the use of effect size measures accompanying tests of statistical significance, discussions on obtaining confidence intervals (CIs) multilevel effect sizes have been rare. In this presentation I explore the bootstrap as a viable and accessible alternative for obtaining CIs for multilevel standardized mean differences. A simulation is carried out to compare 10 procedures for constructing CIs in terms of empirical coverage probability, and results showed that, across all simulation conditions, the residual bootstrap with the bias-corrected and accelerated CI performed the best, followed by the asymptotic symmetric CI with the model-based method.

Presenter: Tyler Matta, University of Oregon
Title: Modeling English Proficiency Growth when Reclassification is Informative
Abstract: A challenge in estimating language attainment trends for limited English proficient students is that reclassification to full English proficiency results in non-ignorable attrition. This poster presents an approach for producing unbiased growth estimates by modeling the joint distribution of longitudinal English proficiency scores and time-to-reclassification simultaneously using a shared-parameter model.

Presenter: Melissa McTernan, University of California, Davis
Title: The Effect of the Assumed Error Distribution on Random Effect Estimates in a Multilevel Model Fit to Positive and Positively Skewed Data
Abstract: The gamma distribution contains properties that we often see in the educational and behavioral sciences; it is non-zero, positive, and right skewed. In this study I use multilevel data simulated from a gamma distribution with normal random effects to examine the effect of the assumed error distribution (gamma/normal), estimation method (MLE/REML), and the sampling distribution (normal/t) on the random effect estimates for varying sample sizes and degrees of skew. I also evaluate the tests of the fixed effects and the behavior of the likelihood ratio test (LRT) under the different assumed error distributions.

Presenter: Daniel Murphy, Pearson
Title: Cross-classified Random Effects Models for Assessing Rater Severity and Differential Rater Functioning
Abstract: This study demonstrates how the conventional multilevel model, the cross-classified multiple membership random effects model (*CCMMrem*) and the cross-classified random effects model (*CCrem*) can be used as raters' effects multilevel measurement models (MMMs). The three versions of these MMMs are presented along with extensions permitting assessment of impact, differential item and differential rater functioning. Distinctions are also made that emphasize when the more complex parameterizations might be needed and the types of differential rater functioning that can be assessed under each of the three models. Using a simulated dataset, the equivalence of item difficulty, person ability and rater severity estimation is empirically demonstrated for the *CCrem* and *CCMMrem* versions of the raters' effects MMM.

Presenter: Diep T. Nguyen, University of South Florida
Title: Investigation of Misspecifications in Level-1 Error Structure and Level-2 Residuals Normality in Single-Case Studies: A Simulation Study and Reanalysis of Published Data
Abstract: There are concerns with the application of multilevel modeling (MLM) to multiple-baseline data about whether an estimated model is correctly specified. This study examined the degree to which MLM estimates and inferences were impacted when a multiple-baseline single-case study ignored both dependencies in the level-1 error structure and non-normality in the level-2 residuals. In addition, a reanalysis of a published dataset was also performed to illustrate the use of multilevel models with multiple-baseline data, and demonstrate methods that can be used to probe for potential dependence in the level-1 error structure and non-normality in the level-2 residuals.

Presenter: Lorena Ortega, University of Oxford
Title: School and Teacher Effects on Chilean Children's Achievement Growth: A Cross-Classified Accelerated Growth Model
Abstract: The present study investigates school and teacher effects on student achievement growth in Chile. The data feature a cohort-sequential design comprising participants in 4 overlapping cohorts, together spanning grades 3 through 8. In order to address the research aims of the study, and appropriately account for the complex structure of the data, accelerated longitudinal designs, growth curve approaches, and cross-classified models are applied in combination. Results confirm that school effects are larger than those found in industrialized countries, teacher effects exceed school effects, and educational effects are larger when achievement progress over time, rather than achievement status, is studied.

Presenter: Bess Rose, Johns Hopkins University
Title: Using Cross-Classified Multiple Membership Growth Curve Modeling in Non-Hierarchical Multilevel Data Structures: The Effect of School Mobility and Concurrent Changes on Students' Academic Achievement
Abstract: This study examines the differential effects of school transfers on achievement, depending on whether the student also experienced changes in educational, neighborhood, or family environments when transferring. Because students usually belong to multiple schools over their educational history (and sometimes multiple schools within a school year), CCMM-GCM is used to model academic achievement over time (first-grade GPA is a function of the school(s) attended during the first year of school; annual change in GPA is a function of the school(s) attended subsequently). Results indicate that transfers are more harmful when concurrent changes occur in students' educational, neighborhood, or family environments.

Presenter: Cathy Zhang, University of British Columbia
Title: Examining the Method Effects Caused by Reverse Worded Items in Nested Data
Abstract: Previous studies have shown that reverse worded (RW) items often contaminate the covariance structure of the scale by creating method effects, thus affecting interpretations of data that rely on the accuracy of the scale. However, these method effects have never been demonstrated in nested data. Our study examined whether the method effects caused by RW items were present in different levels of a nested dataset that examines the relationship between extraversion and openness-to-experience. We found that method effects were present in both Level 1 and 2 units, thus they affected both within-person and between-person relationships between extraversion and openness-to-experience.

Advances in Multilevel Modeling for Educational Research: Addressing Practical Issues Found in Real-World Applications

Day 2: Saturday, November 15

Presenter: Li Cai & Kilchan Choi (9:00am – 9:45am)

Title: On the Importance of Advanced Psychometrics in Multi-Level Impact Evaluation Studies

Abstract: We propose a generalizable solution to measurement error and multilevel modeling issues in multisite randomized trials. Traditional measurement approaches, e.g., sum scores, do not enjoy the considerable benefits afforded by more modern model-based approaches to measurement. On the other hand, routine applications of model-based approaches, e.g., "off-the-shelf" IRT-based scaled scores, tend to ignore the lack of plausibility of certain inherent exchangeability assumptions in the experimental evaluation context. One example is the lack of full exchangeability of the post-treatment outcome measure for participants between treatment and control conditions. We propose a multilevel extension of the two-tier item factor model with latent gain specification (Cai, 2010b) to yield scaled scores for the outcome measure that are more sensitive to the intervention than standard measurement approaches. We use simulated data to illustrate our theoretical developments. This new methodology is also demonstrated with empirical data from a large-scale randomized controlled trial (RCT) conducted to assess the instructional effects of video games on rational number and fraction skills. Comparison of results from the hierarchical linear modeling of the IRT scaled outcome scores versus the observed summed scores revealed that the treatment condition showed significantly higher gains in learning than the comparison condition if the scaled scores are used.

Presenters: Jee-Seon Kim, Peter Steiner & Wen-Chiang Lim (9:55am – 10:40am)

Title: Mixture Modeling Strategies for Causal Inference with Multilevel Data

Abstract: Whereas the standard use of propensity score methods is rather well-established as a tool for making causal inferences in observational research, many aspects of the methods are less well-understood in the context of clustered or nested data structures. This study addresses fundamental issues and challenges in evaluating treatment effects from multilevel observational studies and suggests strategies for making proper inferences. Specifically, we demonstrate how to deal with a lack of overlap between treated and non-treated groups within clusters and also how to handle a heterogeneous population. It is shown that although one can exploit the multilevel structure of the data and borrow cases from other clusters, it is important to evaluate the heterogeneity of the population first and then pull cases across the clusters but within homogeneous subpopulations. This talk describes multilevel latent class models to account for unobserved heterogeneity in clustered data, and also demonstrates consequences when the heterogeneity in the population, such as different types of schools or regions (manifest subpopulations), or different selection and outcome processes (latent subpopulations), are not properly handled in the data analysis.

Presenters: Tracy Sweet & Qiwen Zheng (10:50am – 11:35am)

Title: Hierarchical Mixed Membership Stochastic Block Model with Network-Level Covariates

Abstract: Quantitative methods for analyzing social networks primarily focus on single networks but research in social sciences often covers a sample of networks, and only just recently have multilevel social network models been used in education research. We introduce the Hierarchical Mixed Membership Stochastic Block model with Network-level Covariates, a new multilevel social network model for networks with some form of subgroup or block structure that simultaneously related network-level covariates with subgroup mixing structure. We use this model to illustrate the effects of teacher classroom management style on student friendship networks.

Presenters: Daniel Bauer & Patrick Curran (1:00pm – 1:45pm)

Title: The Disjoint Between Measurement and Modeling in Longitudinal Research

Abstract: Much theoretical and applied work involving multilevel models has centered on model specification and estimation. A few salient examples include the importance of accounting for cross-classifications or multiple memberships, how to estimate indirect effects in mediation models, and determining an appropriate error structure for level-1 residuals. Less commonly addressed are the properties of the criterion variables for which these models are being specified. Here we draw attention to how criterion measurement impacts the results obtained from multilevel models. Many multilevel analyses are conducted using scale scores, for example, growth analyses of achievement test scores. In some cases, the score for an individual may be generated using a simple procedure, such as taking the total or average of an individual's item responses. In other cases, however, the score used is the expected or modal value of the posterior distribution of the latent trait given the individual's item responses (EAP or MAP, respectively). This latter approach is especially common in educational research, and has the advantage that it provides a principled way to generate scores when different individuals receive different item sets, as occurs when there is missing data, when using developmental scales or computerized adaptive tests, or when pooling data across independent studies to permit an integrative data analysis. Regardless of the approach, however, it is well-appreciated that scoring should be informed by the results of careful psychometric analyses. Typically, psychometric modeling is done with single-age or cross-sectional samples. The results obtained from these samples are used to establish a scoring algorithm, and it is this algorithm that is used in subsequent longitudinal research. Herein lies the disjoint: scores are based on psychometric analyses of (at most) cross-sectional data which contain no information on intra-individual variation, but it is precisely this variation that is of central interest when conducting longitudinal analyses of scores. In this presentation we show how this issue has manifested in our own applications of multilevel growth models, we use simulated data to evaluate how well the underlying growth process can be recovered when scores are based on cross-sectional psychometrics, and we conclude with future directions for research.

Presenters: Joop Hox, Stef van Buuren, & Shahab Jolani (1:55pm – 2:40pm)

Title: Incomplete Multilevel Data: Problems and Solutions

Abstract: Incomplete data are common in empirical research. The default solutions in software packages are very simplistic; the default is generally listwise deletion where a case with any variable missing is completely removed from the analysis. In multilevel data, missing values at the group level can be a serious problem. For example, when a teacher has no data on a single variable, listwise deletion means that the teacher plus the corresponding class is completely removed. Listwise deletion is clearly very inefficient. More importantly, any deletion scheme assumes that the remaining cases are representative for the entire original sample, meaning that it assumes that the missingness is completely random. This is a very strong assumption, unlikely to be true in real world data. Modern solutions to incomplete data are Full Information Maximum Likelihood (FIML) estimation, which includes the incomplete cases in the estimation, and Multiple Imputation (MI). The problem with FIML is that most available multilevel analysis software does not have it. The problem with MI is that one must use a multilevel procedure to generate the imputations. This presentation discusses missingness mechanisms, introduces the FIML and MI approaches, and shows how these can be used with currently available software.

Presenters: Laura Stapleton, Jeffrey Harring & Daniel Lee (2:50pm – 3:35pm)

Title: Sampling Weight Considerations for Multilevel Modeling of Panel Data

Abstract: National longitudinal panel surveys in education, such as the Early Childhood Longitudinal Study, include the problem of respondent attrition or nonsystematic missingness. On public-release data files for these studies, sets of panel weights are provided that allow the analyst to appropriately weight the observations in the study to account for non-response. Several sets of weights are provided depending on which waves of data are intended to be used in the analysis. However, these weight adjustments are based only on sampling information from the base year; information obtained from early waves in the survey program is not utilized. Other methods to address potential bias from non-response include full information maximum likelihood and multiple imputation; these methods could make use of responses obtained during data collection waves. In this presentation, we first describe common approaches to sampling weight adjustment for non-response (including cell adjustment and response propensity models), present arguments for using FIML and MI with auxiliary variables when undertaking longitudinal analyses, and demonstrate the differences in estimates from analyses when non-response adjusted panel weights are used as compared with FIML and MI utilizing the base year sampling weights only but with auxiliary information. These analyses utilize data from the National Center for Education Statistics.