Paper 199-2009

Parsimony vs. Complexity: A Monte Carlo Investigation of Hierarchical and Cross-Classified Modeling Using SAS® PROC MIXED

Bethany A. Bell, University of South Carolina, Columbia, SC

Corina M. Owens, University of South Florida, Tampa, FL

Jeffrey D. Kromrey, University of South Florida, Tampa, FL

John M. Ferron, University of South Florida, Tampa, FL

ABSTRACT

Across a variety of disciplines, multilevel modeling continues to be a popular analytic approach among methodological and applied researchers. As such, a variety of reference and instructional materials are becoming increasingly available. However, such research primarily has been limited to purely hierarchical models. Far less research has been conducted on cross-classified random effects models (CCREMs). This Monte Carlo study focuses on the consequences of hierarchical vs. cross-classified modeling of cross-classified data on the estimation of fixed and random effects in terms of model convergence and both point and interval estimates as a function of level-2 sample size, degree of cross-classification, correlation of level-2 residuals, and model complexity. SAS/IML was used to simulate 1000 data sets for each of 648 conditions. Simulated data were then analyzed using different PROC MIXED model structures (e.g., 2-level models, 3-level models, and CCREMs). Results are presented in terms of statistical bias, confidence interval coverage, and rates of model non-convergence as a function of design factors.

Keywords: MONTE CARLO, MULTILEVEL MODELS, CROSS-CLASSIFIED DATA STRUCTURES, SAS/IML, SAS/STAT

INTRODUCTION

Multilevel models are being increasingly used across the social sciences to analyze nested or hierarchically structured data. There are many types of multilevel models, which differ in terms of the number of levels (e.g., 2, 3), type of design (e.g., cross-sectional, longitudinal with repeated measures, cross-classified), scale of the outcome variable (e.g., continuous, categorical), and number of outcomes (e.g., univariate, multivariate). These models have been used to address a variety of research questions involving model parameters that include fixed effects (e.g., average student socioeconomic status-mathematics achievement slope across schools), random level-1 coefficients (e.g., student socioeconomic status-mathematics achievement slope at a particular school), and variance-covariance components (e.g., amount of variation in the student socioeconomic status-mathematics achievement slope across schools). As the popularity of this analytical approach grows with both methodological and applied researchers, reference and instructional materials are becoming increasingly available (see, for example, Brown & Prescott, 2006; Bingenheimer & Raudenbush, 2004; Raudenbush & Bryk 2002; McCulloch & Searle, 2001). However, as these methods become better known and more widely used in research, issues that must be addressed when employing these techniques are also being identified.

For instance, although current research into multilevel modeling methodological issues has focused on a variety of issues including sample size (Giesbrecht & Burns, 1985, Hess, Ferron, Bell-Ellison, Dedrick, & Lewis, 2006; Kenward & Roger, 1997; Maas & Hox, 2002), model misspecification (Donoghue & Jenkins, 1992; Ferron, Dailey, & Yi, 2002, Schwartz, 1978), measurement error (Kromrey, Coraggio, Phan, Romano, Hess, Lee, Hines, & Luther, 2007), and missing data (Basilevsky, Sabourin, Hum, & Anderson, 1985; Raymond & Roberts, 1987; Roy & Lin, 2002), such research primarily has been limited to purely hierarchical models. Far less research has been conducted on cross-classified random effects models (CCREMs; Bell, Owens, Ferron, & Kromrey, 2008; Fielding, 2002; Raudenbush & Bryk, 2002; Myers & Beretvas, 2007). Furthermore, most of the studies that have investigated CCREMs have done so by comparing estimates from hierarchical models to those from CCREMs using real data sets that contain cross-classified data structures

(Bell et al., 2008; Fielding, 2002; Raudenbush & Bryk, 2002). For example, using data from the National Longitudinal Study of Adolescent Health (2005) and the Adolescent Health and Academic Achievement study (n.d.), Bell et al. (2008) examined and compared five different model structures: (a) a two-level model with students nested in schools and neighborhood variables included as contextual variables at Level 1, (b) a two-level model with students nested in neighborhoods and school variables included as contextual variables at Level 1, (c) a three-level model with students nested in schools nested in neighborhoods, (d) a three-level model with students nested in neighborhoods nested in schools, and (e) a cross-classified random effects model with students cross-nested in schools and neighborhoods. Although such comparisons are helpful, more controlled empirical investigations into CCREMs are warranted.

Unlike purely hierarchical data structures in which all level-1 units are nested in some level-2 unit, which may or may not be nested in some level-3 unit, cross-classified data structures occur when level-1 units are nested within two higher units, but the higher units are not nested within each other. For example, when thinking of students, schools, and neighborhoods, one scenario of purely hierarchical data would be students nested in schools nested in neighborhoods, thus all students in a particular school would live in the same neighborhood (see Figure 1). Or, students could be nested in neighborhoods nested in schools such that all students in a particular neighborhood attend the same school. However, neither of these three-level hierarchical structures is likely to exist. Instead, students who live in the same neighborhood often attend several different schools and students at a particular school often come from multiple neighborhoods (see Figure 2). That is, students are nested within schools and within neighborhoods, but the schools and neighborhoods are not nested within each other. Instead, students are cross-classified by school and neighborhood. This study focuses on the consequences of inappropriate modeling of cross-classified data on the estimation of fixed and random effects coefficients in terms of point estimates (statistical bias) and interval estimates (confidence interval accuracy and precision), and Type I error control and statistical power of tests associated with the fixed and random effects) as a function of the number of level-2 units, degree of cross-classification, correlation of level-2 residuals, and model complexity. By examining more complex models (i.e., models with various numbers of predictors and variance structures) and different degrees of cross-classification, this study contributes to our understanding of the behavior of hierarchical and cross-classified models under a variety of conditions.

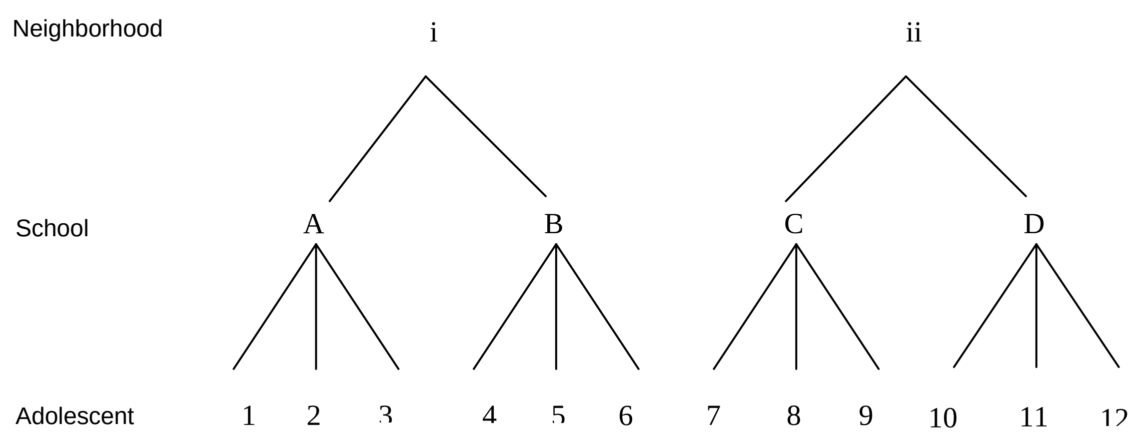

Figure 1. Example schematic of purely hierarchical data with adolecents nested in schools nested in neighborhoods.

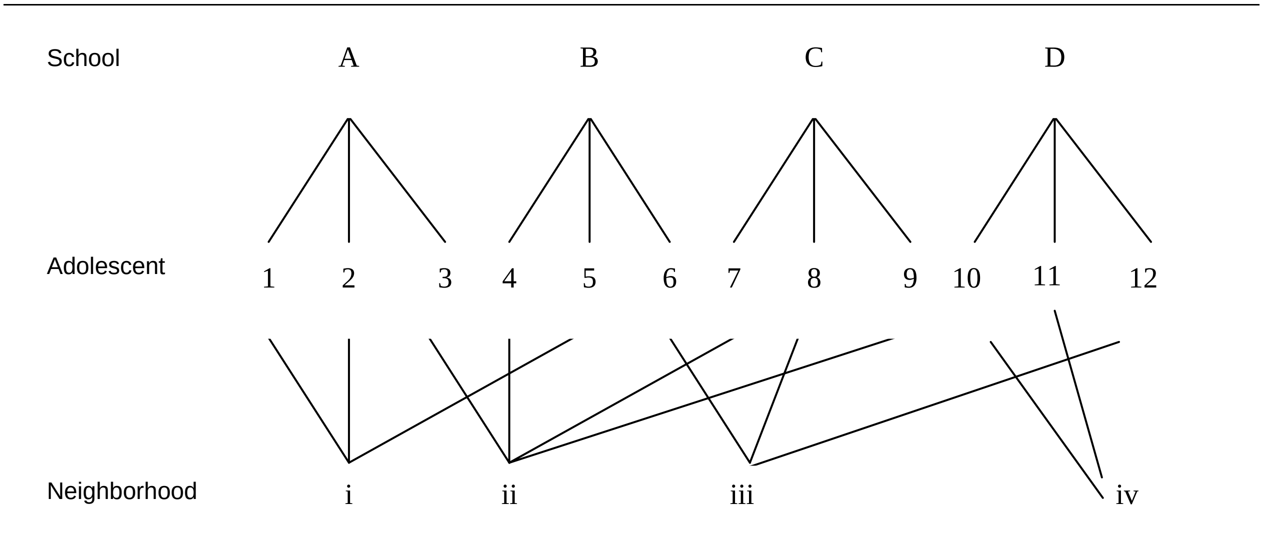

Figure 2. Example schematic of cross-classified data with adolecents nested within schools and neighborhoods.

METHOD

Although various forms of cross-classification can occur, for this Monte Carlo study, the data were generated to mirror the common theoretical cross-classification of students in schools and neighborhoods. Specific design factors and conditions examined include: (a) level-2 sample sizes (numbers of schools and neighborhoods of 50,50; 50,100; 50,200; 100,100; 100,200; and 200,200), (b) degree of cross-classification (Cramer's *V* values of .3, .5, .7, and .9, with lower Cramer's *V* values representing higher levels of cross-classification), (c) correlation of school and neighborhood intercept residuals (.0, .2, and .4), and (d) model complexity [9 different models varying in the number of level-1 predictors (1 or 3), level-2 predictors (1 or 2), whether interactions were included, and variance structure (random intercept only and random intercept and all slopes); see Table 1 for more details on model complexity]. These factors in the Monte Carlo study were crossed, yielding 648 conditions.

Table 1
Summary of model conditions examined

k_x	k_w	k_z	Random	Interactions
1	1	1	Intercept	0
1	1	1	Intercept, All X	0
1	1	1	Intercept, All X	XZ, WZ
3	1	1	Intercept	0
3	1	1	Intercept, All X	0
3	1	1	Intercept, All X	XZ, WZ
3	2	2	Intercept	0
3	2	2	Intercept, All X	0
3	2	2	Intercept, All X	XZ, WZ

Data were generated based on a cross-classified random effects model in which observations (n = 10,000) were cross-nested within two separate level-2 units (e.g., schools and neighborhoods). The level-1 errors were generated from a normal distribution with a variance of 1.0 using the RANNOR random number generator in SAS version 9.1.3 (SAS, 2004). The level-2 intercept errors were also generated from a normal distribution but with variance of .12 for schools and .06 for neighborhoods to produce the target cross-classified ICC = .15. Similarly, the slope error variances were set to the intercept variance multiplied by $1/(2*k_x)$. Intercept and slope fixed effects were set at .12 and .085, respectively, to yield statistical power of approximately .80 in the cross-classified model structure. The data were simulated such that the outcome variable and all predictors were normally distributed continuous variables, one level-1 predictor variable (x1) and one neighborhood level-2 predictor variable (w1) had no effect (for estimation of Type I error rate), and all other predictors had non-null effects (for estimating statistical power).

For each of the 648 conditions, an average of 1000 and no less than 400 data sets were simulated using SAS IML (SAS, 2004). The data simulation program was checked by examining the matrices produced at each stage of data generation. After each data set was generated, the simulated sample was analyzed following the same model structures as Bell et al. (2008): (a) a two-level model with observations nested in schools and neighborhood variables included at level-1, (b) a two-level model with observations nested in neighborhoods and school variables included at level-1, (c) a three-level model with observations nested in schools nested in neighborhoods, (d) a three-level model with observations nested in neighborhoods nested in schools, and (e) a cross-classified random effects model with observations cross-nested in schools and neighborhoods. All models were estimated using restricted maximum likelihood estimation and the residual degrees of freedom method via the MIXED procedure in SAS/STAT (SAS, 2004). The residual method was used to compute degrees of freedom to increase the speed of execution for the simulation. See Figure 3 for example code from the MIXED procedure for the most complex model examined (k_x = 3, k_w = 2, k_z = 2, interactions, and random intercepts and slopes). In each of the five model structures examined using PROC MIXED, Xs represent level-1 variables, Zs represent school-level variables, and Ws represent neighborhood-level variables.

In all analyses the covariance matrix of the level-2 errors, T, was modeled as variance components, and the covariance matrix of the level-1 errors was modeled as $\Sigma = \sigma^2\mathbf{I}$. Outcomes examined in this Monte Carlo study include: rate of model convergence and non-positive definite G-matrices, bias in the estimates of the fixed and random effects, confidence interval coverage for each effect, and average confidence interval width for each effect. In addition, Type I error rates and statistical power estimates were also examined. However, due to space limitation, limited results are presented below. Please contact us if you would like more information.

```
title 'Students nested in Schools';
proc mixed data=ccrem covtest noclprint CL;
by N_SCH N_NEIGH  kx kw kz ccresid tauc00 CramerV gamma0 gamma1 Replication;
class School;
model y3 = x1 x2 x3 w1 w2 z1 z2 x1*z1 w1*z1/s cl DDFM=RES;
random intercept x1 x2 x3/ sub= school type = VC;
run;

title 'Students nested in Neighborhoods';
proc mixed data=ccrem covtest noclprint CL;
by N_SCH N_NEIGH  kx kw kz ccresid tauc00 CramerV gamma0 gamma1 Replication;
class neighborhood;
model y3 = x1 x2 x3 w1 w2 z1 z2 x1*z1 w1*z1/s cl DDFM=RES;
random intercept x1 x2 x3/ sub= neighborhood type = VC;
run;

title 'Students nested in Schools nested in Neighborhoods';
proc mixed data=ccrem covtest noclprint CL;
by N_SCH N_NEIGH  kx kw kz ccresid tauc00 CramerV gamma0 gamma1 Replication;
class school neighborhood;
model y3 = x1 x2 x3 w1 w2 z1 z2 x1*z1 w1*z1/s cl DDFM=RES;
random intercept x1 x2 x3/ sub= neighborhood type = VC;
random intercept x1 x2 x3/sub = school (neighborhood)type = VC;
run;

title 'Students nested in Neighborhoods nested in Schools';
proc mixed data=ccrem covtest noclprint CL;
by N_SCH N_NEIGH  kx kw kz ccresid tauc00 CramerV gamma0 gamma1 Replication;
class neighborhood school;
model y3 = x1 x2 x3 w1 w2 z1 z2 x1*z1 w1*z1/s cl DDFM=RES;
random intercept x1 x2 x3/ sub= school type = VC;
random intercept x1 x2 x3/sub = neighborhood (school) type = VC;
run;
```

```
title 'Students Cross-Nested in Neighborhoods and Schools';
proc mixed data=ccrem covtest noclprint CL;
by N_SCH N_NEIGH  kx kw kz ccresid tauc00 CramerV gamma0 gamma1 Replication;
class neighborhood school;
model y3 = x1 x2 x3 w1 w2 z1 z2 x1*z1 w1*z1/s cl DDFM=RES;
random intercept x1 x2 x3/ sub= neighborhood type = VC;
random intercept x1 x2 x3/sub = school type = VC;
run;
```

Figure 3. Example PROC MIXED Code for 2-Level, 3-Level, and Cross-Classified Random Effects Models

COMPARISON OF STATISTICAL RESULTS

Model convergence was not a substantial problem with any of the model structures examined in the study. Convergence problems were noted in 1.5% of the cross-classified models and less than 1% for all other model structures. Similarly, low levels of statistical bias were evident for all parameter estimates. However, unlike model convergence and statistical bias, the rate of non-positive definite G-matrices varied across model structures (see Figure 4). Specifically, whereas average rates of non-positive G-matrices were relatively low (.0 for both 2-level models, < .001 for the cross-classified model and the 3-level school model, and .15 for the 3-level neighborhood model), the range of these rates varied. The 3-level neighborhood model had the most variability (range = .21), followed by the 3-level school model (range = .14), the cross-classified model (range = .07), and the 2-level neighborhood model (range = .001).

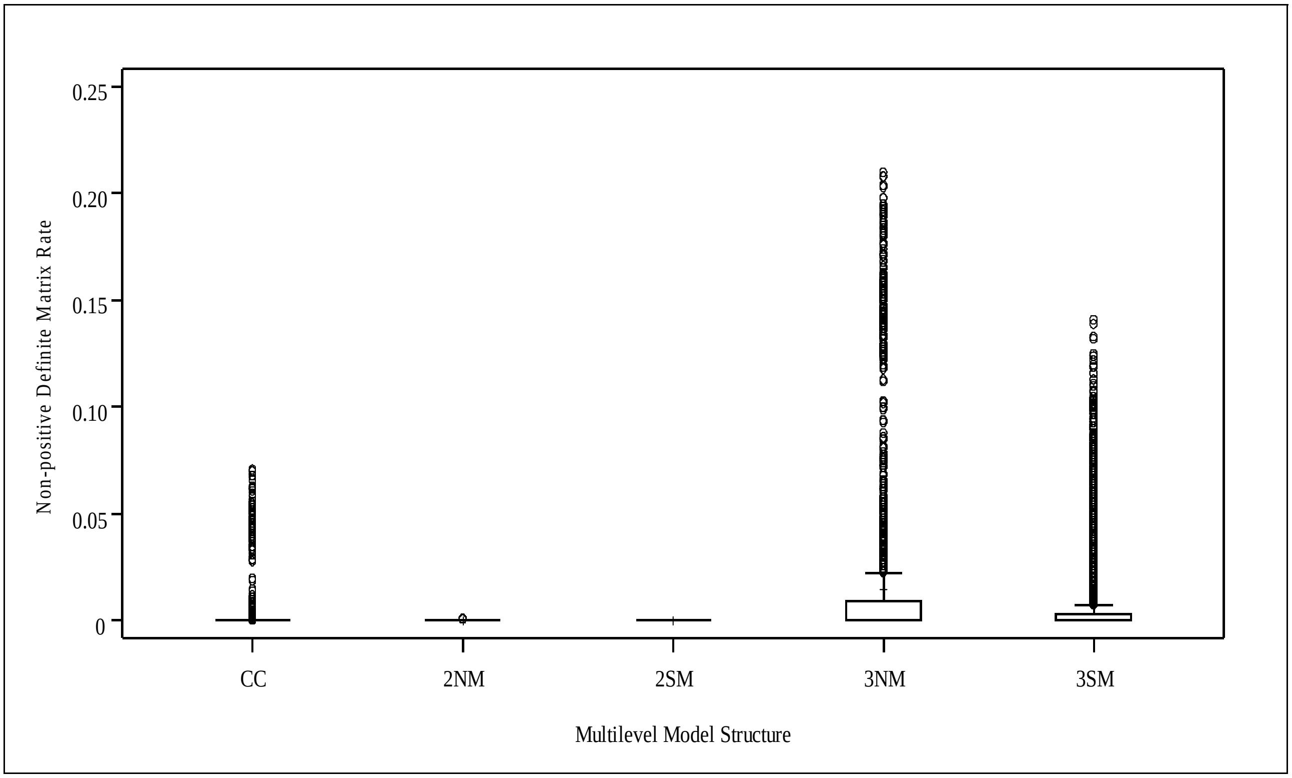

Figure 4. Rates of Non-positive Definite G-matrices in Model Estimation

Estimated Type I error rates for tests of the regression parameters for the level-1 variables and neighborhood variables are presented in Figures 5 and 6, respectively. For level-1 variables, on average, the cross-classified model and the two school models (2-level and 3-level) maintained rates close to the nominal alpha level. However, with the two neighborhood models, the average rates appear more liberal (e.g., >.10). For neighborhood variables, the cross-classified, 2-level neighborhood, and 3-level neighborhood models exhibited average Type I error rates close to .05. However, Type I error control within the two school models (2-level and 3-level) was notably greater than the desired alpha level of .05. In particular, the 3-level school model revealed slightly liberal average rates near .10 whereas the 2-level school model produced average rates greater than .30.

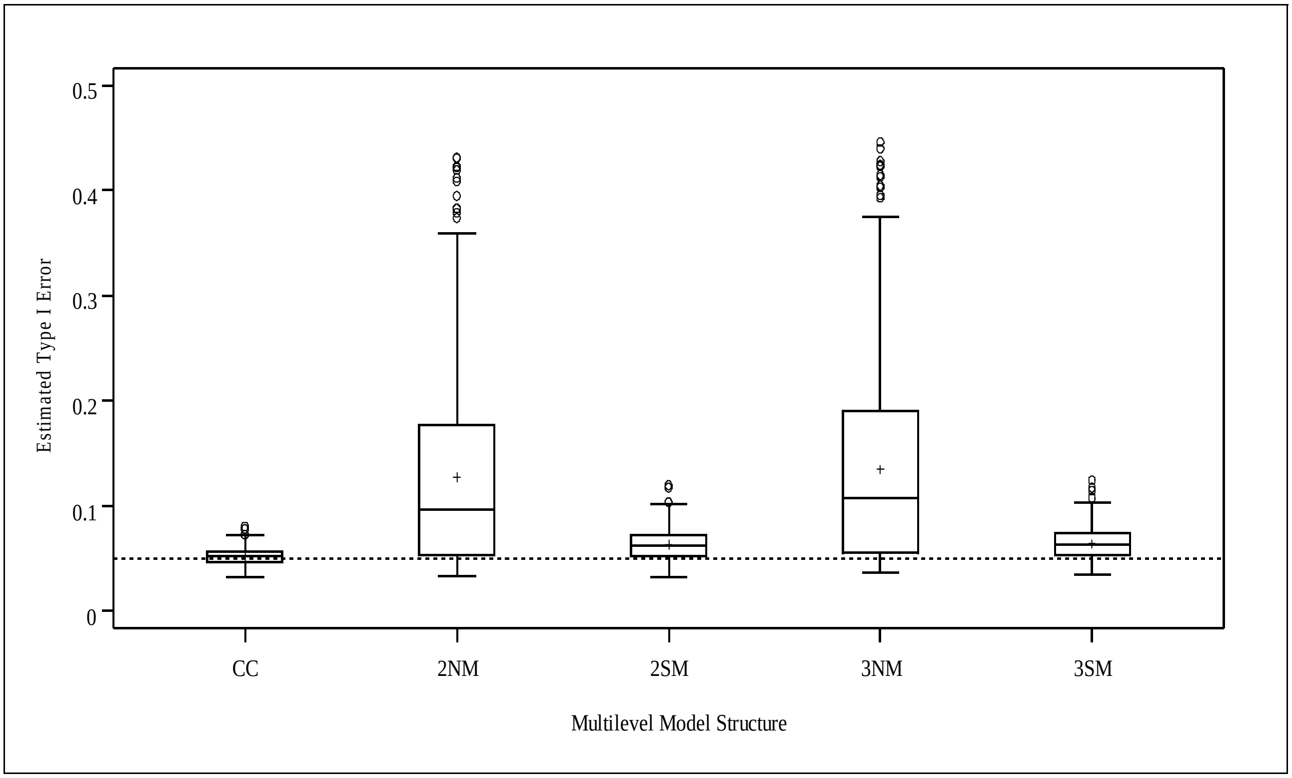

Figure 5. Estimated Type I Error Rates for Tests of Level-I Variables by Multilevel Model Structure (Fixed Effects)

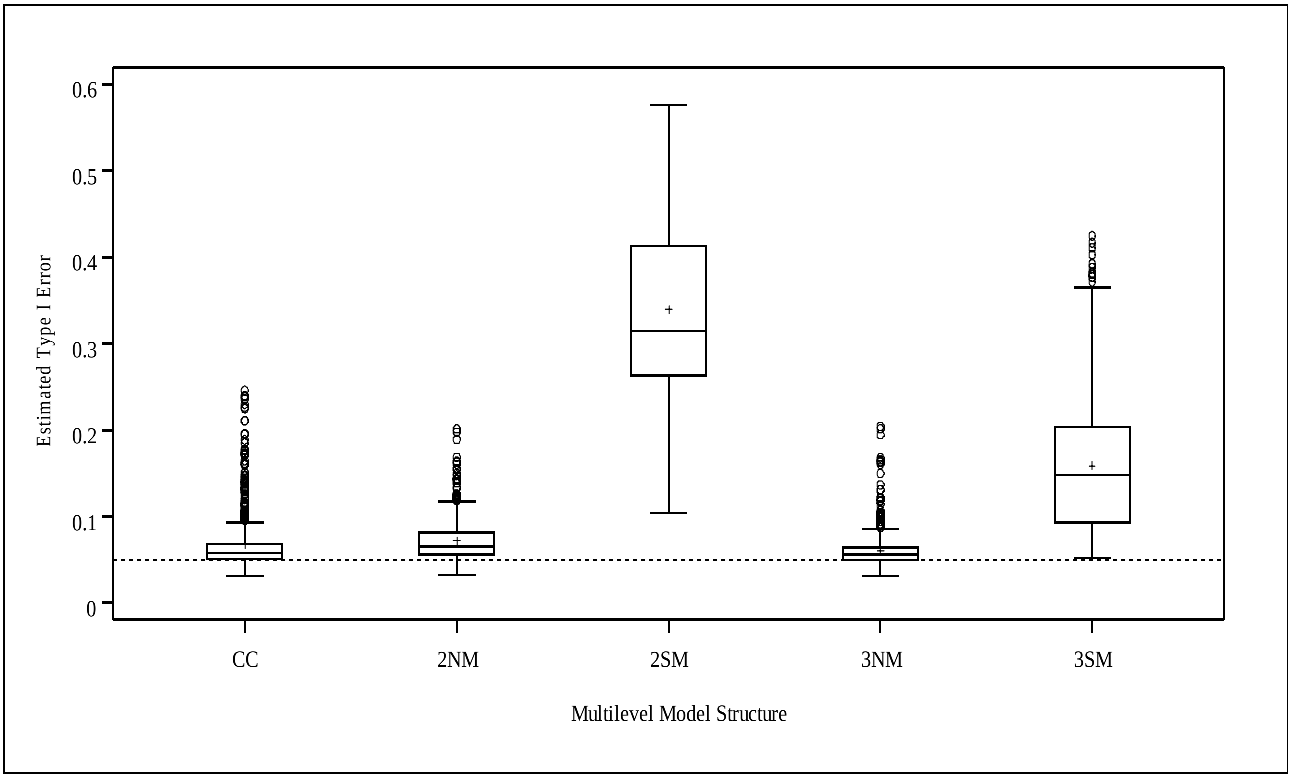

Figure 6. Estimated Type I Error Rates for Tests of Neighborhood Variables by Multilevel Model Structure (Fixed Effects)

Varying levels of estimated power for the tests of fixed effects parameters of model intercepts, level-1 variables, school variables, and neighborhood variables were evident across model structures. Across the five model structures, tests of fixed effects parameters for level-1 variables had the highest and most consistent power, with all model structures exhibiting averages greater than .95 and no conditions less than .80. Conversely, average power for the tests of fixed effects for model intercepts was less than .80 for all model structures. Similarly, except for the 2-level neighborhood model, tests of fixed effects parameters for school variables were underpowered, with the cross-classified and two school models (2-level and 3-level) exhibiting average power less than .50 (Figure 7). For the neighborhood variables, although estimated power for the tests of fixed effects parameters did vary, all model structures maintained average power greater than .60, with the cross-classified and two school models demonstrating average power greater than .80 (Figure 8).

The cross-classified models and the two school model structures provided confidence interval coverage at near nominal level for the level-1 fixed effects, while the two neighborhood model structures provided average coverage of less than 90%. Similarly, for the confidence intervals for the intercept, the cross-classified models and the two school model structures provided coverage closer to the nominal level than that of the neighborhood model structures. For the intercept parameter, however, the average coverage was only approximately 90% for the former model structures and approximately 70% for the latter.

The distributions of estimated confidence interval coverage for the fixed effects parameters of the school variables and the neighborhood variables are presented in Figures 9 and 10, respectively. For the school variables, the 95% confidence intervals provided nearly nominal level coverage across all conditions in the two school level model structures (2-level and 3-level) and in the cross-classified models. In contrast, the intervals in the neighborhood model structures provided undercoverage of the parameters. For the 2-level neighborhood models, the average interval coverage was less than 50%. The coverage improved for the 3-level neighborhood models, but average coverage of the 95% intervals remained less than 80%.

For coverage of the fixed effects parameters of the neighborhood variables (Figure 10), both neighborhood model structures and the cross-classified models provided near nominal level coverage in the majority of conditions. For these parameters, the school level model structures provided notable undercoverage. Although confidence interval coverage of the 3-level school models was superior to that of the 2-level school models, the average coverage remained less than 85%.

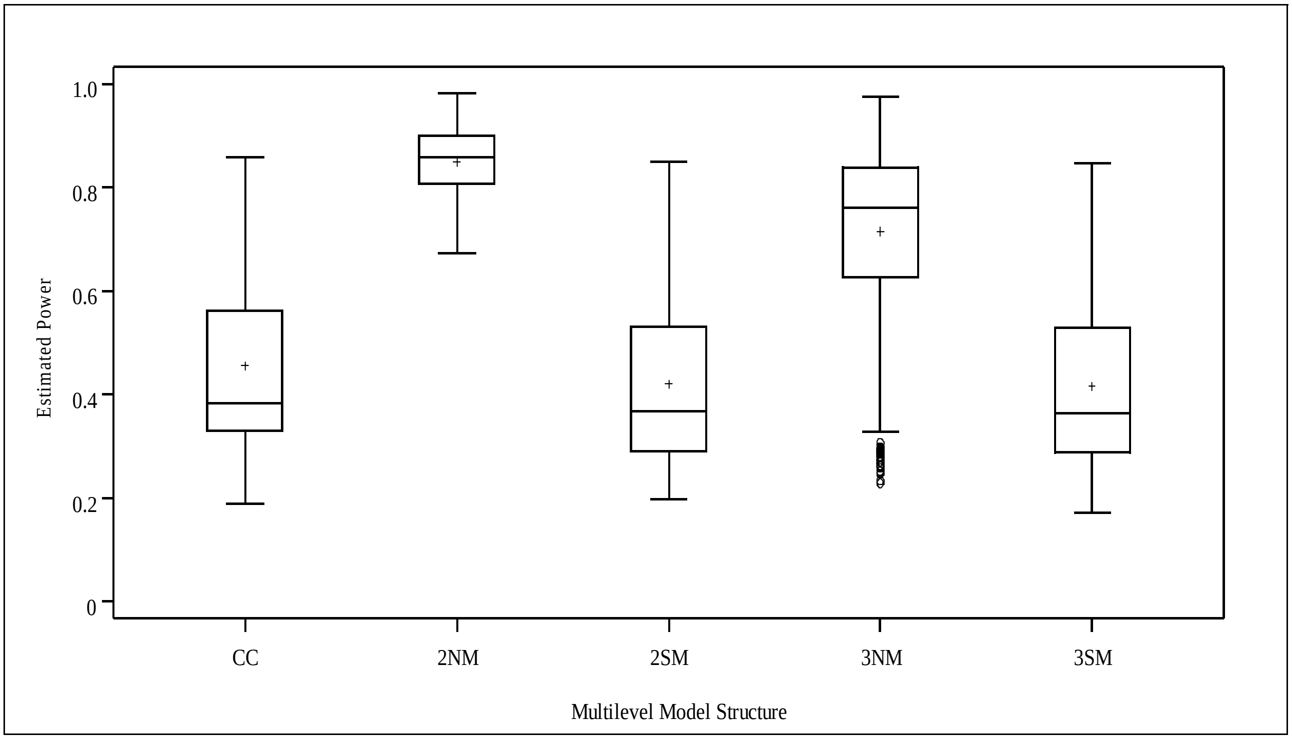

Figure 7. Estimated Power for Tests of School Variables by Multilevel Model Structure (Fixed Effects)

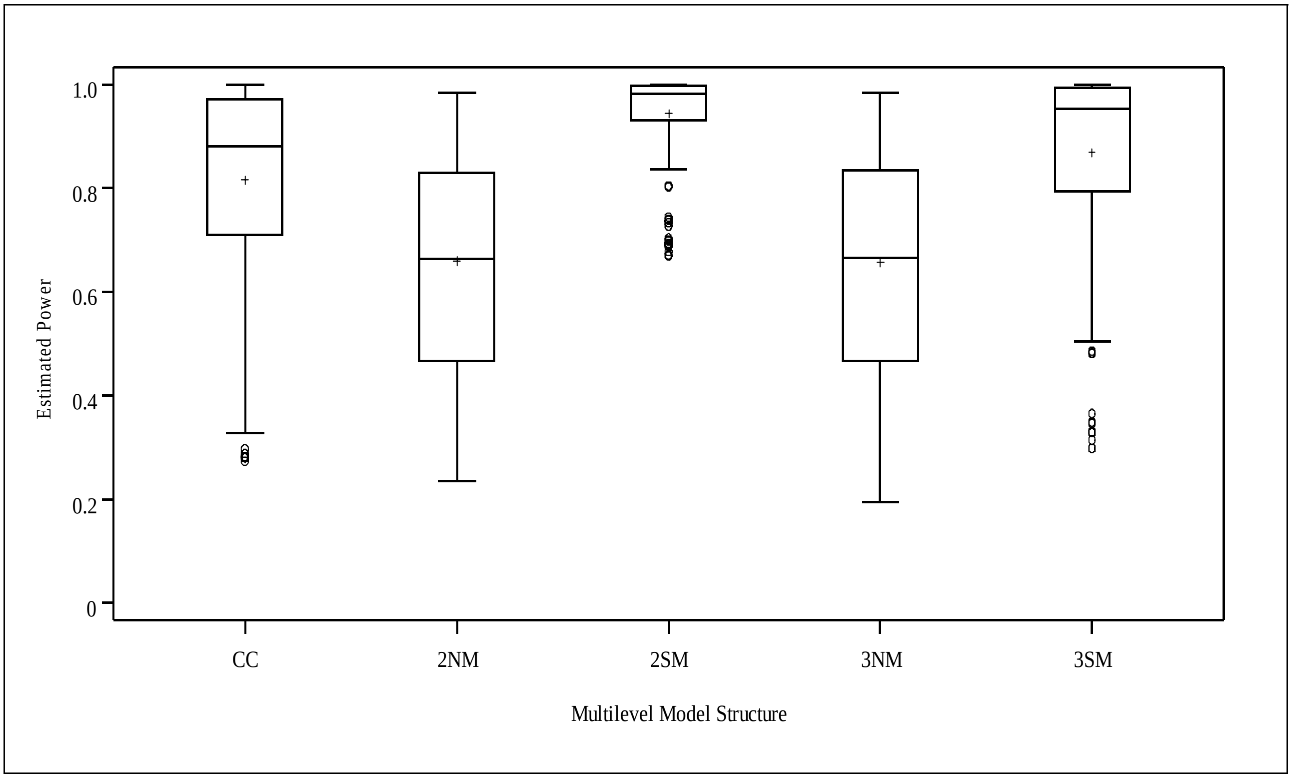

Figure 8. Estimated Power for Tests of Neighborhood Variables by Multilevel Model Structure (Fixed Effects)

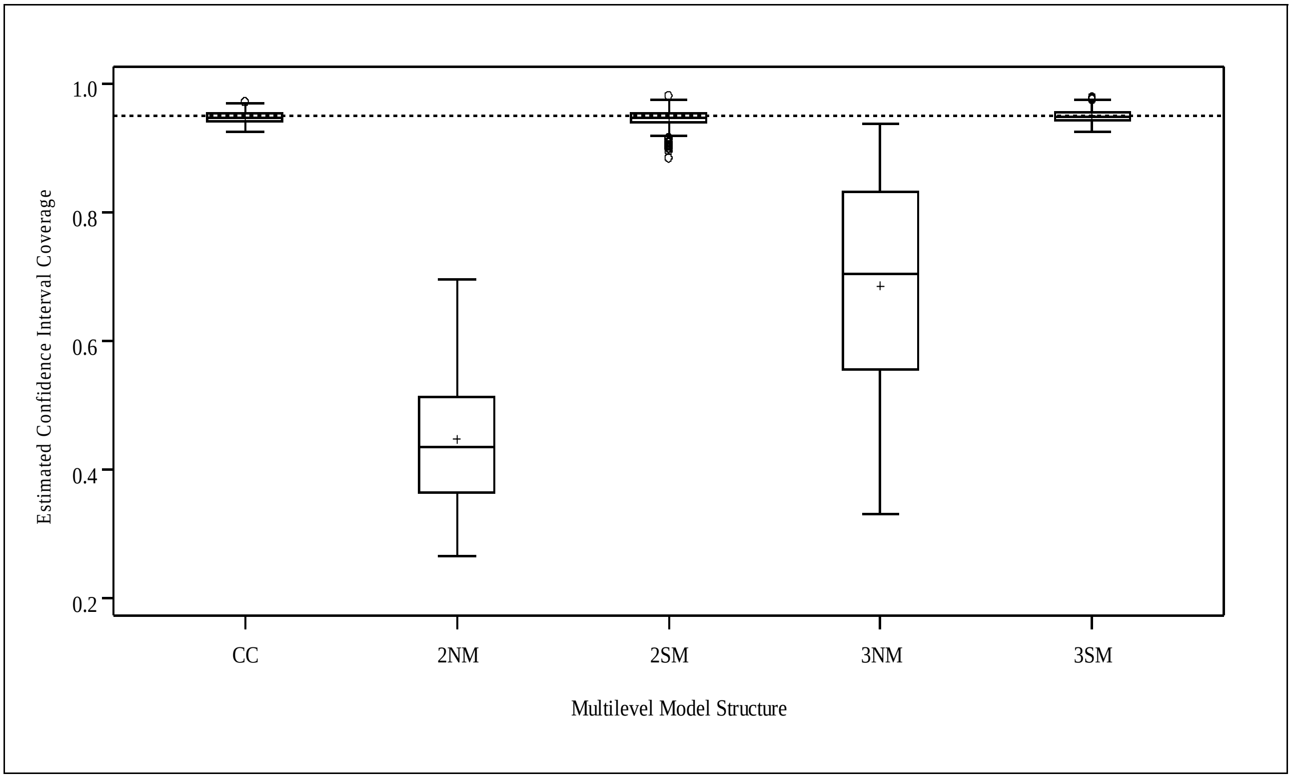

Figure 9. Confidence Interval Coverage for School Variables by Multilevel Model Structure (Fixed Effects)

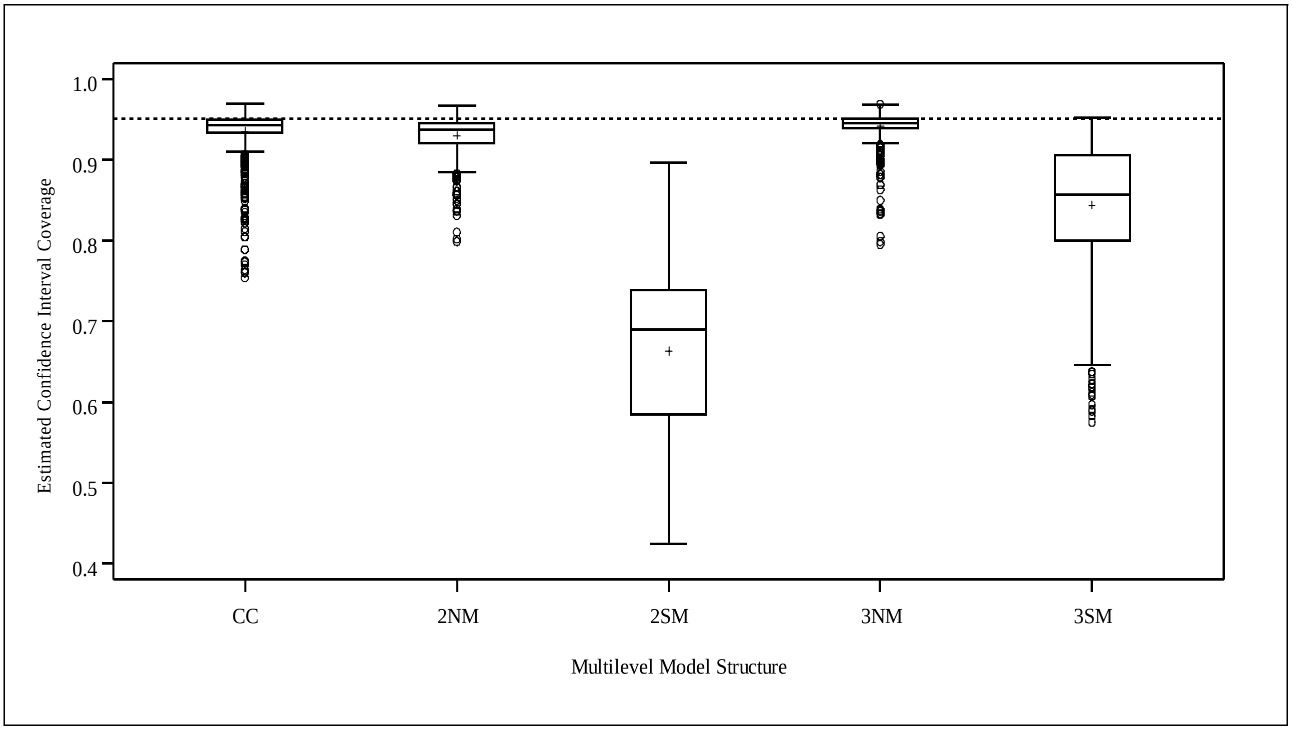

Figure 10. Confidence Interval Coverage for Neighborhood Variables by Multilevel Model Structure (Fixed Effects)

The distributions of the estimated interval coverage for the variance of the level-1 coefficients are presented in Figure 11. The cross-classified models provided coverage near the nominal level in the majority of conditions examined. Both of the 3-level model structures evidenced slight undercoverage in these confidence intervals and the undercoverage was more severe with the 2-level model structures. The most severe undercoverage of the intervals for the variance of the level-1 coefficients was provided by the 2-level neighborhood models in which the average coverage was less than 50%.

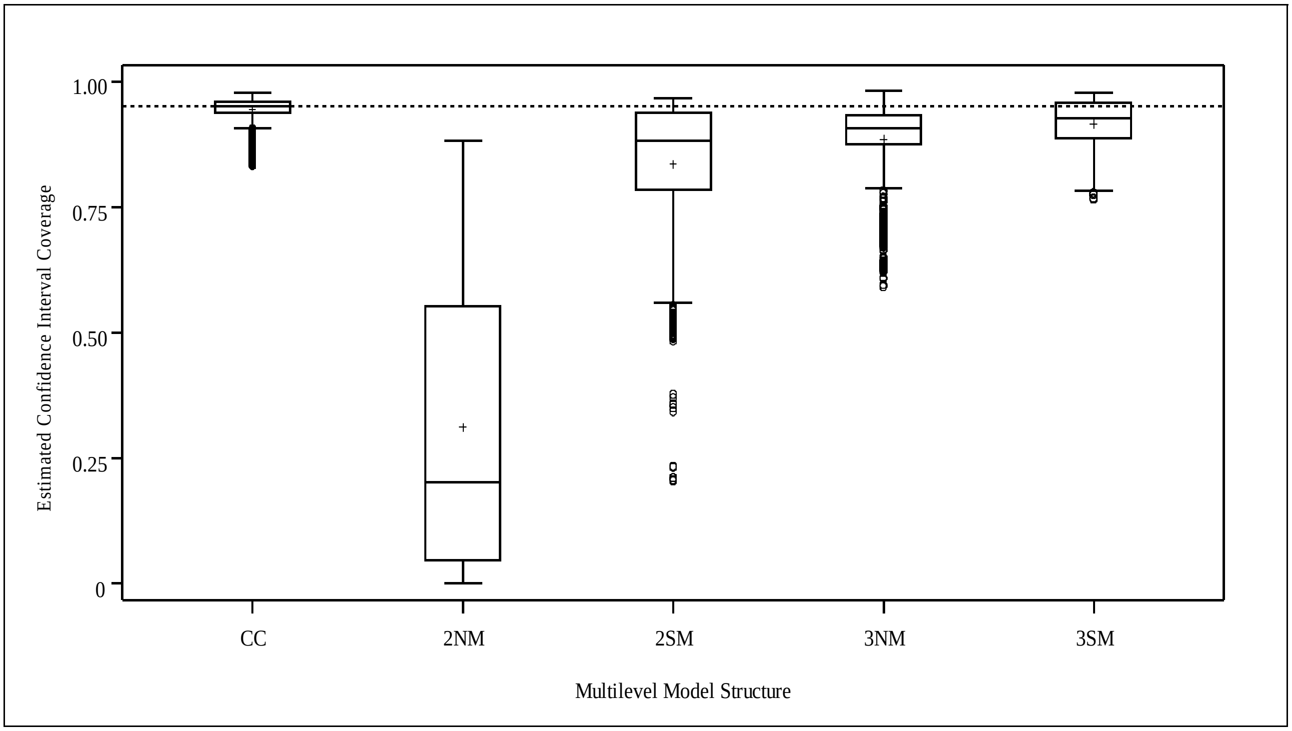

Figure 11. Confidence Interval Coverage for Variance of Level-1 Coefficients by Multilevel Model Structure (Random Effects)

The distributions of the confidence interval coverage estimates for the variance of the model intercepts are provided in Figure 12. The interval coverage of these variance parameters was notably lower than those of the level-1 coefficients. The best coverage was provided by the cross-classified models, with average coverage slightly below the nominal 95% level. The two 3-level model structures provided average interval coverage below 75% and the two 2-level model structures provide average coverage below 50%.

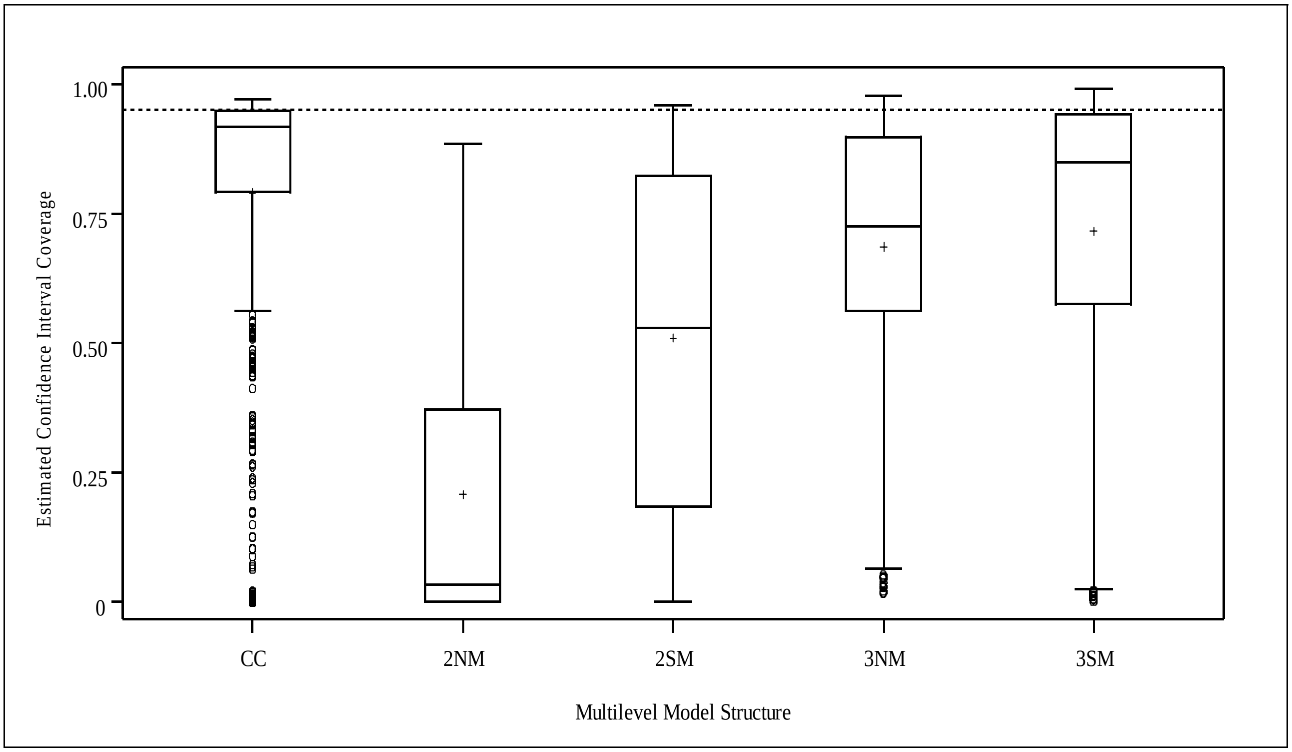

Figure 12. Confidence Interval Coverage for Variance of Intercepts by Multilevel Model Structure (Random Effects)

CONCLUSIONS

Findings from the current study support the notion that inappropriate modeling of cross-classified data matters. More specifically, estimates of statistical power, Type I error rates, and confidence interval coverage vary based on model specification. However, consistent with findings reported by Myers and Beretvas (2007), in terms of statistical bias, fixed effects from the current study did not demonstrate notable bias in any of the model structures. The inflated Type I error rates of the misspecified 2-level and 3-level models found in the current study are also consistent with findings presented by Myers and Beretvas (2007). Given the notable differences in the data generation and design factors between the current study and the study conducted by Myers and Beretvas (2007), these consistencies are note worthy.

General patterns in the findings from the current study are also consistent with the previous research of Bell, Owens, Ferron, and Kromrey (2008). More specifically, using data from the National Longitudinal Study of Adolescent Health (2005) and the Adolescent Health and Academic Achievement study (n.d.), Bell et al. (2008) noted that parameter estimates and model fit indices were similar across level-2 contexts (i.e., schools and neighborhoods) and not across model hierarchy (i.e., 2-level models and 3-level models). They also found statistical findings from the CCREMs were most similar to the two school models (Bell et al., 2008). In the current study, estimated Type I error rates for tests of the regression parameters for the level-1 variables were similar for the two neighborhood models and two school models. Furthermore, in terms of the estimated statistical power for the tests of fixed effects parameters for the school and neighborhood variables and the distributions of estimated confidence interval coverage for the fixed effects parameters of the school variables, findings from the CCREMs were most similar to the two school models (2-level and 3-level).

The only area where the CCREMs were clearly distinct from other model structures was with the estimated confidence interval coverage for the variance of level-1 coefficients and variance of model intercepts. For these random effects, not only was the CCREM the only model structure in which the 95% confidence intervals provided nearly nominal level coverage, but the coverage of the CCREM for both types of variance components was notably different than the 2-level and 3-level model structures. That the CCREM provided the best confidence interval coverage of level-1 variances is also consistent with Myers and Beretvas' (2007) findings.

In terms of model misspecification, the most grossly misspecified models were the 2-level school and 2-level neighborhood models. In these models, level-2 variables were included at level-1. For example, in the 2-level school model, neighborhood variables were included as level-1 variables and vice versa for the 2-level neighborhood model. Not surprising, the impact of these extreme levels of model misspecification were apparent in terms of estimated Type I error rates and estimated confidence interval coverage. The estimated Type I error rate for tests of the regression parameters for the neighborhood variables was greatest in the 2-level school model. Similarly, the worst estimated confidence interval coverage for school variables was found in the 2-level neighborhood model and the worse coverage of neighborhood variables was found in the 2-level school model.

In summary, findings from the current and previous research underscore the importance of appropriately modeling cross-classified data. Given how easy it is to estimate CCREMs (i.e., simply using two random lines in the PROC MIXED command in SAS), there is no reason for researchers to not estimate CCREMs when data contain a cross-classified structure. Furthermore, although findings from the current study add to our understanding of the consequences of model misspecification in relation to cross-classified data, there is still much to learn. Not only are future studies warranted, but further examination of the data obtained from the current study is forthcoming. Specifically, we will be conducting a closer examination of model structure performance by design factors, including level-2 sample size, degree of cross-classification, and correlation of level-2 residuals.

REFERENCES

Adolescent Health and Academic Achievement Study. (n.d.). *Adolescent Health and Academic Achievement Study*. Retrieved October 30, 2006, from http://www.prc.utexas.edu/ahaa/index.html

Basilevsky, A., Sabourin, D., Hum, D., & Anderson, A. (1985). Missing data estimators in the general linear model: An evaluation of simulated data as an experimental design. C*ommunications in Statistics*, *14*, 371-394.

Bell, B. A., Owens, C. M., Ferron, J. M., & Kromrey, J.D. (2008).Parsimony vs. complexity: A comparison of two-level, three-level, and cross-classified models using Add Health and AHAA data. In press, *Proceedings from the 16th Annual Southeast SAS Users Group.*

Bingenheimer, J. B. & Raudenbush, S. W. (2004). Statistical and substantive inferences in public health: Issues in the application of multilevel models. *Annual Review of Public Health*, *25*, 53–77.

Brown, H. & Prescott, R. (2006). *Applied Mixed Models in Medicine* (2nd ed.). John Wiley & Sons, LTD: New York.

Donoghue, J.R., & Jenkins, F. (1992). *A Monte Carlo study of the effects of model misspecification on HLM estimates*. Technical report, Educational Testing Service, Princeton, NJ. November 1992.

Ferron, J., Dailey, R., & Yi, Q. (2002). Effects of misspecifying the first-level error structure in two-level models of change. *Multivariate Behavioral Research*, *37*, 379-403.

Giesbrecht, F., & Burns, J. (1985). Two-stage analysis based on a mixed model: Large-sample asymptotic theory and small-sample simulation results. *Biometrics*, *41*, 477-486.

Hess, M.R., Ferron, J.M, Bell-Ellison, B.A., Dedrick, R.F., & Lewis, S.E. (2006, April). *The effects of sample size on fixed effect inferences*. Paper presented at the American Educational Association's Annual Meeting, San Francisco, CA.

Kenward, M., & Roger, J. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, *53*, 983-997.

Kromrey, J. D., Coraggio, J. T., Phan, H., Romano, J. T., Hess, M. R., Lee, R. S., Hines, C. V., and Luther, S. L. (2007, April). *Fallible regressors in multilevel models: Impact of measurement error on parameter estimates and hypothesis tests*. Paper presented at the American Educational Association's Annual Meeting, Chicago.

Maas, C.J.M. & Hox, J.J. (2002). Sample sizes for multilevel modeling. In: J. Blasius, J. Hox , E. de Leeuw & P. Schmidt (eds.) (2002). *Social Science Methodology in the New Millennium. Proceedings of the Fifth International Conference on Logic and Methodology*. Second expanded edition. Opladen, RG: Leske + Budrich Verlag (CD-ROM).

McCulloch, C. E. & Searle, S. R. (2001). *Generalized, linear, and mixed models*. New York: Wiley

Myers, J. L. & Beretvas, S. N. (2007). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research, 41,* 473-497.

National Longitudinal Study of Adolescent Health. (2005). *National Longitudinal Study of Adolescent Health*. Retrieved October 30, 2006, from http://www.cpc.unc.edu/addhealth

Raudenbush S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Raymond, M. R., & Roberts, D. M. (1987). A comparison of methods for treating incomplete data in selection research. *Educational and Psychological Measurement*, *47*, 13-26.

Roy, J., & Lin, X. (2002). Analysis of multivariate longitudinal outcomes with nonignorable dropouts and missing covariates: Changes in methadone treatment practices. *Journal of the American Statistical Association, 97*, 40-52.

SAS Institute Inc. (2004). *SAS, release 9.12* [computer program]. Carey, NC: SAS Institute Inc.

Schwartz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics*, *6*, 461-464.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please contact Bethany A. Bell at:
University of South Carolina
College of Education
Wardlaw 133
Columbia, SC 29208
phone: 803-777-2387
email: babell@sc.edu