The Impacts of Ignoring a Crossed Factor in Analyzing Cross-Classified Data

Wen Luo \(^a\), Oi-man Kwok \(^b\)

\(^a\) University of Wisconsin, Milwaukee
\(^b\) Texas A&M University

Online Publication Date: 01 March 2009

To cite this Article Luo, Wen and Kwok, Oi-man(2009)'The Impacts of Ignoring a Crossed Factor in Analyzing Cross-Classified Data', Multivariate Behavioral Research, 44:2, 182 — 212

To link to this Article: DOI: 10.1080/00273170902794214

URL: http://dx.doi.org/10.1080/00273170902794214

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
The Impacts of Ignoring a Crossed Factor in Analyzing Cross-Classified Data

Wen Luo
University of Wisconsin at Milwaukee

Oi-man Kwok
Texas A&M University

Cross-classified random-effects models (CCREMs) are used for modeling non-hierarchical multilevel data. Misspecifying CCREMs as hierarchical linear models (i.e., treating the cross-classified data as strictly hierarchical by ignoring one of the crossed factors) causes biases in the variance component estimates, which in turn, results in biased estimation in the standard errors of the regression coefficients. Analytical studies were conducted to provide closed-form expressions for the biases. With balanced design data structure, ignoring a crossed factor causes overestimation of the variance components of adjacent levels and underestimation of the variance component of the remaining crossed factor. Moreover, ignoring a crossed factor at the kth level causes underestimation of the standard error of the regression coefficient of the predictor associated with the ignored factor and overestimation of the standard error of the regression coefficient of the predictor at the (k – 1)th level. Simulation studies were also conducted to examine the effect of different structures of cross-classification on the biases. In general, the direction and magnitude of the biases depend on the level of the ignored crossed factor, the level with which the predictor is associated at, the magnitude of the variance component of the ignored crossed factor, the variance components of the
predictors, the sample sizes, and the structure of cross-classification. The results were further illustrated using the Early Childhood Longitudinal Study-Kindergarten Cohort data.

In social and behavioral sciences, multilevel data are very common. When the levels in multilevel data are strictly nested or hierarchical, the technique of hierarchical linear modeling (HLM) can be used to model the clustering effects. However, many multilevel data do not have pure hierarchical structures. For example, students are nested within schools and neighborhoods at the same time, whereas schools and neighborhoods are not nested within but crossed with each other (Raudenbush & Bryk, 2002). In longitudinal studies, participants may move to different clusters during the course of a study. The mobility causes a nonhierarchical data structure in which repeated measures are cross-classified by participants and clusters (Raudenbush & Bryk, 2002).

For cross-classified multilevel data, cross-classified random-effects models (CCREMs) are used to investigate the relationships among variables within a given level and across levels (Goldstein, 1986, 1995; Rasbash & Goldstein, 1994; Raudenbush, 1993). Despite the flexibility of CCREMs, they have not been commonly used in educational research.

Meyers and Beretvas (2006) conducted a simulation study assessing the impacts of misspecifying CCREMs due to the restrictions of the studies (e.g., only two-level cross-
classified data were examined in the Meyers and Beretvas 2006 study). This study aims to provide a more synthesized and complete picture of the impacts of misspecifying CCREMs. The study has two major focuses: (a) examining the more general three-level cross-classified model with cross-classification at either the top or the intermediate level and (b) examining the effect of different structures of cross-classification on the biases of parameter estimates.

First, the level at which a crossed factor is ignored could affect the biases in parameter estimates. Moerbeek (2004) investigated the consequences of ignoring a level of nesting in a three-level hierarchical linear model and found that the bias depends on which level is ignored (i.e., the top level or the intermediate level). Therefore, it is important to investigate the impacts of ignoring a crossed factor using the more general three-level CCREMs in which cross-classification could occur at either the top level or the intermediate level.

In social and behavioral sciences, three-level cross-classified empirical data are fairly common. For example, Hough (2006) examined business segment performance using three-level data with two crossed factors at the top level. In the data set, yearly business segment performance observations at Level 1 were nested within business segments at Level 2 that were cross-classified by corporations and industries at Level 3. This data structure is presented in Figure 1a using the classification diagram (Browne, Goldstein, & Rasbash, 2001).

Crossed random factors can also occur at the intermediate level in a three-level structure. Fielding’s (2002) study, as mentioned previously, was an example. Jayasinghe, Marsh, and Bond’s (2003) study was another example in which

![Crossed random factors at the top versus the intermediate level.](a) Crossed random factors at the top level. (b) Crossed random factors at the intermediate level.

**FIGURE 1** Crossed random factors at the top versus the intermediate level. (a) Crossed random factors at the top level. (b) Crossed random factors at the intermediate level.
ratings at Level 1 were cross-classified by assessors and proposals at Level 2 that were nested within fields of study at Level 3 (see Figure 1b for the data structure).

Cross-classified multilevel models were also applied to data with more than three levels. For example, Marsh, Martin, and Cheng (2008) used five-level cross-classified models to investigate gender effect in classroom motivation and climate. In their sample, motivation related outcomes were cross-classified by students and school subjects (i.e., math, English, and science), students and school subjects were both nested within classrooms, classrooms were nested within teachers, and teachers were nested within schools.

Second, previous studies did not thoroughly consider the impact of different structures of cross-classification on the biases of parameter estimates. In Meyers and Beretvas’s (2006) study, they only used the number of empty cells in cross-classified data as an indicator of the degree of cross-classification and found little effect of the number of empty cells on the observed biases. Nevertheless, the observed biases might be affected by the distribution of the empty cells (i.e., the structure of cross-classification) rather than the number of empty cells. Moreover, the contradictory findings between Fielding’s (2002) and Meyers and Beretvas’s studies may be due to the different structures of cross-classification in their data.

Cross-classified data structures can be generally categorized into two main types: *complete* vs. *partial* cross-classification. In completely cross-classified data, the probability of a lower level unit affiliating with any clusters of the crossed factors is approximately the same across all units. In other words, units in a cluster of one crossed factor could affiliate with any clusters of the other crossed factor and vice versa. Consider the example of students cross-classified by schools and neighborhoods. In a complete cross-classification condition, students from a specific school can live in any neighborhoods and students from a specific neighborhood can go to any schools in the sample. In reality, however, students living in certain neighborhoods only go to certain schools and students attending certain schools only live in certain neighborhoods. We call this kind of cross-classification a partial cross-classification because units in a cluster of one crossed factor can only affiliate with *part* (some but not all) of the clusters of the other crossed factor.

We conducted two main studies to examine the impact of misspecification of CCREMs on parameter estimates in a three-level model framework. Study 1 investigated cross-classified models in which the two random factors were crossed at the top level. Study 2 investigated models with cross-classification at the intermediate level. Within each study, analytical results were first presented, followed by a demonstration with simulated data and a simulation study focusing on structures of cross-classification. A short empirical demonstration using the Early Childhood Longitudinal Study–Kindergarten Cohort (ECLS-K) data was included to illustrate the analytical and simulation results.
STUDY 1: IGNORING A CROSSED FACTOR AT THE TOP LEVEL

Analytical Study

Consider the example of cross-classification at the top level as shown in Figure 1a. Suppose there are a total of $n_1$ corporations (F1) and $n_2$ industries (F2). Within each of the $n_1 \times n_2$ cells (i.e., unique combination of corporations and industries), there are a total of $n_3$ business segments (B) randomly distributed in these cells. Within each business segment, there are $n_4$ yearly performance observations (A). A random intercept cross-classified model with one predictor associated with each level and random factor is specified as follows:

\[
\begin{align*}
\text{Level 1:} & \quad Y_{ij(kl)} = \gamma_0(jkl) + \gamma_1 X_{ij(kl)} + \epsilon_{ij(kl)} \\
\text{Level 2:} & \quad \gamma_0(jkl) = \gamma_{00}(kl) + \gamma_2 W_{j(kl)} + \mu_{0j(kl)} \\
\text{Level 3:} & \quad \gamma_{00}(kl) = \gamma_{0000} + \gamma_3 Z_k + \gamma_4 S_l + \nu_{00k} + \omega_{00l},
\end{align*}
\]

where $i$ indexes Level 1 unit A ($i = 1, \ldots, n_4$), $j$ indexes clusters of Level 2 factor B ($j = 1, \ldots, n_3$), $k$ indexes clusters of Level 3 crossed factor F1 ($k = 1, \ldots, n_1$), and $l$ indexes clusters of Level 3 crossed factor F2 ($l = 1, \ldots, n_2$).

In the fixed part of the model, $\gamma_{0000}$ is the overall intercept, $\gamma_1$ is the effect of the Level 1 predictor $X_{ij(kl)}$, $\gamma_2$ is the effect of the Level 2 predictor $W_{j(kl)}$, $\gamma_3$ is the effect of the predictor $Z_k$ that is associated with F1, and $\gamma_4$ is the effect of the predictor $S_l$ that is associated with F2. It is assumed that the Level 1 predictor $X_{ij(kl)}$ has identical distribution within each Level 2 unit and the Level 2 predictor $W_{j(kl)}$ has identical distribution within each Level 3 cell. It is further assumed that predictors do not have variance components at other levels or factors. In the random part of the model, $\mu_{0j(kl)}$ is the random effect of the Level 2 factor B [$\mu_{0j(kl)} \sim N(0, \psi)$], $\nu_{00k}$ is the random effect associated with the Level 3 factor F1 [$\nu_{00k} \sim N(0, \tau)$], $\omega_{00l}$ is the random effect associated with the Level 3 factor F2 [$\omega_{00l} \sim N(0, \zeta)$], and $\epsilon_{ij(kl)}$ is the Level 1 residual [$\epsilon_{ij(kl)} \sim N(0, \theta)$]. It is assumed that all random effects and residuals are independent from each other (i.e., covariances between all random effects and residuals are zero).

The variance components (i.e., $\theta$, $\psi$, $\tau$, and $\zeta$) can be estimated from sample data. Using the iterative generalized least squares (IGLS) approach for cross-classified multilevel models (Goldstein, 1986; Rasbash & Goldstein, 1994), we obtained the estimated variance components in Model 1:
where $SS_R$ is the sum of squares of the Level 1 residuals, $SS_B$ is the sum of squares of the residuals of Level 2 clusters, and $SS_{F1}$ and $SS_{F2}$ are the sums of squares of the residuals of clusters in $F1$ and $F2$, respectively (see Appendix for the formulas). It is noted that IGLS is equivalent to maximum likelihood when the random effects are normally distributed (Goldstein, 1986).

When a crossed factor ($F2$) is ignored, the data is analyzed using a hierarchical linear model and the predictor associated with $F2$ is modeled as a Level 1 predictor. The estimated* variance components in the misspecified model are

$$
\hat{\theta} = \frac{SS_R}{(n_4 - 1)n_1n_2n_3}
$$

$$
\hat{\psi} = \frac{1}{n_4} \left[ \frac{SS_B}{n_1n_2n_3 - n_1 - n_2 + 1} - \frac{SS_R}{(n_4 - 1)n_1n_2n_3} \right]
$$

$$
\hat{\xi} = \frac{1}{n_2n_3n_4} \left( \frac{SS_{F1}}{n_1 - 1} - \frac{SS_B}{n_1n_2n_3 - n_1 - n_2 + 1} \right)
$$

$$
\hat{\xi} = \frac{1}{n_1n_3n_4} \left( \frac{SS_{F2}}{n_2 - 1} - \frac{SS_B}{n_1n_2n_3 - n_1 - n_2 + 1} \right).
$$

Comparing the variance component estimates of the correct model (CCREM) in Equation 2 and those of the misspecified model (HLM) in Equation 3, we obtained

$$
\hat{\theta}^* = \hat{\theta}
$$

$$
\hat{\psi}^* = \hat{\psi} + \frac{n_3(n_2 - 1)}{n_2n_3 - 1} \xi
$$

$$
\hat{\xi}^* = \hat{\xi} - \frac{n_2 - 1}{n_2(n_2n_3 - 1)} \xi.
$$
So, ignoring a crossed factor at the top level does not affect the estimated Level 1 residual variance but results in overestimating the variance component of Level 2 and underestimating the variance component of the remaining crossed factor.

The fixed effects are estimated using generalized least squares (GLS) estimators. In balanced design, the consistency of GLS estimator does not depend on the specification of the random effects of the model (Kreft & de Leeuw, 1998). Hence, we only examined the effect of ignoring a crossed factor on the efficiency of the estimator, that is, the variance of the regression coefficient. The variances of the regression coefficients in Model 1 are

\[
\begin{align*}
\text{Var}(\hat{\gamma}_1) &= \frac{\hat{\theta}}{S_X^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_2) &= \frac{\hat{\theta} + n_4 \hat{\psi}}{S_W^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_3) &= \frac{\hat{\theta} + n_4 \hat{\psi} + n_2 n_3 n_4 \hat{\tau}}{S_Z^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_4) &= \frac{\hat{\theta} + n_4 \hat{\psi} + n_1 n_3 n_4 \hat{\xi}}{S_S^2 n_1 n_2 n_3 n_4}
\end{align*}
\]

(5)

where $S_X^2$, $S_W^2$, $S_Z^2$, and $S_S^2$ are the variances of $X_{ij(kl)}$ (the Level 1 predictor), $W_{j(kl)}$ (the Level 2 predictor), $Z_k$ (the predictor associated with the crossed factor F1), and $S_l$ (the predictor associated with the crossed factor F2), respectively. When the crossed factor F2 is ignored, the variances of the regression coefficients become

\[
\begin{align*}
\text{Var}(\hat{\gamma}_1)^* &= \frac{\hat{\theta}^*}{S_X^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_2)^* &= \frac{\hat{\theta}^* + n_4 \hat{\psi}^*}{S_W^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_3)^* &= \frac{\hat{\theta}^* + n_4 \hat{\psi}^* + n_2 n_3 n_4 \hat{\tau}^*}{S_Z^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_4)^* &= \frac{\hat{\theta}^* + n_4 \hat{\psi}^*}{S_S^2 n_1 n_2 n_3 n_4}
\end{align*}
\]

(6)
Substituting Equation 4 into Equation 6 and comparing the results with those of the correct model in Equation 5, we obtained

\[
\begin{align*}
\text{Var}(\hat{\gamma}_1) &= \text{Var}(\hat{\gamma}_1) \\
\text{Var}(\hat{\gamma}_2) &= \text{Var}(\hat{\gamma}_1) + \frac{(n_2 - 1)}{S^2 \sigma^2 n_1 n_2 (n_2 n_3 - 1)} \\
\text{Var}(\hat{\gamma}_3) &= \text{Var}(\hat{\gamma}_1) \\
\text{Var}(\hat{\gamma}_4) &= \text{Var}(\hat{\gamma}_1) - \frac{n_1 n_2 (n_3 - 1) + (n_1 - 1)(n_2 - 1)}{S^2 \sigma^2 n_1 n_2 (n_2 n_3 - 1)}
\end{align*}
\]

Therefore, ignoring a crossed factor at the top level will not affect the variances of the regression coefficients of the predictors associated with Level 1 and the remaining crossed factor. However, the variance of the regression coefficient of the predictor at Level 2 will be overestimated and the variance of the coefficient of the predictor associated with the ignored crossed factor will be underestimated.

**Demonstration With Simulated Data**

A simulated data set was used to illustrate the analytical results. Using completely balanced design, we generated data with 50 corporations (F1) and 20 industries (F2). Within each of the 1,000 (20*50) cells there are four business segments. Within each business segment, there are eight repeated measures of yearly performance. The example data were generated based on Model 1 and all the predictor variables were generated with the standard normal distribution and the assumptions specified in Model 1. The generated data set was analyzed using two models: (a) the correctly specified model (CCREM), in which both crossed factors were included, and (b) the misspecified model (HLM), in which the crossed factor of industry was ignored. The data set was generated using SAS 9.1 and the models were then estimated using SAS PROC MIXED with restricted maximum likelihood estimation method (SAS Institute Inc., 2004).

As shown in Table 1, ignoring the crossed factor of industry resulted in lower estimated variance component of the corporation ($\hat{\psi}$), higher estimated variance component of business segment ($\hat{\psi}$), and unchanged estimated residual variance ($\hat{\theta}$). The regression coefficient estimators were not affected. However, the standard error of the regression coefficient of the predictor associated with business segment ($\hat{\gamma}_2$) was too high and the standard error of the regression coefficient of the predictor associated with industry ($\hat{\gamma}_4$) was too small in the misspecified model. The standard errors of the regression coefficients of the predictors associated with yearly performance ($\hat{\gamma}_1$) and corporation ($\hat{\gamma}_3$) were unaffected.
Partially Cross-Classified Data

In the analytical study and the simulated data example, the data are balanced and completely cross-classified. It is assumed that the probability of a business segment affiliating with any industry or corporation is the same. However, in reality it is more likely that a certain business segment in a corporation can only affiliate with some industries, that is, the data set is more likely to be partially cross-classified. To examine whether the analytical results can be generalized to partially cross-classified data, we conducted the following simulation study.

Hierarchical, Partially Cross-Classified, and Completely Cross-Classified Data Structure

Consider an example in the educational setting in which repeated measures are nested within students and students are cross-classified by neighborhoods (F1) and schools (F2). If all neighborhoods have designated schools (e.g., students from Neighborhoods 1, 2, and 3 can only attend School 1; students from Neighborhoods 4, 5, and 6 can only attend School 2, etc.), the data structure is strictly hierarchical with repeated measures nested within students, students nested within neighborhoods, and neighborhoods nested within schools. On the other hand, if only some but not all of the neighborhoods have designated schools and students in the nondesignated neighborhoods can choose any schools they would like to attend, the data become partially cross-classified. Furthermore, if none of the neighborhoods have designated schools and students in any

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ignoring a Crossed Factor at the Top Level</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Correct Model (CCREM)</th>
<th>Misspecified Model (HLM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression coefficients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeated measures: $\gamma_1$</td>
<td>.500</td>
<td>.005</td>
</tr>
<tr>
<td>Business segment: $\gamma_2$</td>
<td>.500</td>
<td>.009</td>
</tr>
<tr>
<td>Corporation: $\gamma_3$</td>
<td>.541</td>
<td>.034</td>
</tr>
<tr>
<td>Industry: $\gamma_4$</td>
<td>.514</td>
<td>.091</td>
</tr>
<tr>
<td>Variance components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corporation ($\alpha$)</td>
<td>.054</td>
<td>.052</td>
</tr>
<tr>
<td>Industry ($\xi$)</td>
<td>.148</td>
<td>—</td>
</tr>
<tr>
<td>Business segment ($\psi$)</td>
<td>.206</td>
<td>.348</td>
</tr>
<tr>
<td>Residual ($\theta$)</td>
<td>.599</td>
<td>.599</td>
</tr>
</tbody>
</table>

Note. CCREM = Cross-Classified Random-Effects Model; HLM = Hierarchical Linear Model.
neighborhoods can attend any schools, the data structure is completely cross-classified.

Data Generation

Using the mechanism described earlier, we generated different levels of partially cross-classified data from two extreme conditions: (a) crossed factor F1 nested within crossed factor F2 and (b) crossed factor F2 nested within crossed factor F1. The details of the data generation are presented in the following section.

Generating cross-classification from the hierarchy of F1 nested within F2. We first created the hierarchy of eight repeated measures nested within each student, 10 students nested within each neighborhood, and three neighborhoods nested within each of the 20 schools (see Figure 2a for the structure). The model for data generation is

\[
\begin{align*}
\text{Level 1:} & \quad Y_{ijkl} = \gamma_{0jkl} + \gamma_1 X_{ijkl} + \varepsilon_{ijkl} \\
\text{Level 2:} & \quad \gamma_{0jkl} = \gamma_{00kl} + \gamma_2 W_{jkl} + \nu_{0jkl} \\
\text{Level 3 (F1):} & \quad \gamma_{00kl} = \gamma_{000l} + \gamma_3 Z_{kl} + \nu_{00kl} \\
\text{Level 4 (F2):} & \quad \gamma_{000l} = \gamma_{0000} + \gamma_4 S_l + \omega_{0000}.
\end{align*}
\]

(8)

where \(i\) indexes repeated measures, \(j\) indexes students, \(k\) indexes neighborhoods, and \(l\) indexes schools. The assumptions of the distributions of the random effects and the predictors were the same as those in Model 1 in the analytical study. Following Meyers and Beretvas’s (2006) study, we used .50 as the generating value for the fixed regression coefficients (i.e., \(\gamma_1\) to \(\gamma_4\)). The overall intercept \(\gamma_{0000}\) was generated with the value of .10. The variances of \(\nu_{00kl}, \rho_{000l}, \nu_{0jkl}\) and \(\varepsilon_{ijkl}\) and were generated with values of .05, .15, .20 and .60, respectively.

To create the partially cross-classified structure as shown in Figure 2b, we randomly selected 5 neighborhoods (e.g., Neighborhoods 5, 12, 28, 33, and 57) and then randomly assigned half of the students from each of these 5 neighborhoods to the original designated schools (e.g., half of the students from Neighborhood 5 still attend School 2) but the other half of the students to a randomly selected non-designated school (e.g., the other half of students from Neighborhood 5 attend School 4). Using this procedure, we can create a more cross-classified data structure by selecting more neighborhoods and schools. Figure 2c illustrates a cross-classified structure in which 10 neighborhoods are randomly selected and the students in each of these 10 neighborhoods are then
randomly assigned to five different schools (including the original designated school). If we select all neighborhoods and randomly assign the students to any schools, we will have the completely cross-classified data structure.

To simplify the presentation, we call the selected neighborhoods with students attending both designated and nondesignated schools *feeders* and the corresponding assigned (designated and nondesignated) schools *receivers* hereafter. To create different degrees of partial cross-classification, we selected three levels for the number of feeders (5, 25, and 45 feeders) and two levels for the number of receivers (two and eight receivers). Combining the two factors, there were six conditions and in each condition 500 data sets were generated.

*Generating cross-classification from the hierarchy of F2 nested within F1.* We also examined the other extreme condition in which F2 (the ignored

---

![Figure 2](image_url)

**FIGURE 2** Different structures of cross-classification. (a) F1 (neighborhoods) nested in F2 (schools). (b) 5 feeders (neighborhoods) and 2 receivers (schools). (c) 10 feeders (neighborhoods) and 5 receivers (schools). (continued)
crossed factor) was originally nested with F1. We started with the same hierarchical structure of repeated measures nested within students, students nested within neighborhoods, and neighborhoods nested within schools and then generated different levels of partial cross-classification in the same way (i.e., with the same parameter settings). In this scenario, however, neighborhood was the ignored factor (F2) and school was the remaining factor (F1). The corresponding model for the data generation is shown as follows:

\[
\begin{align*}
\text{Level 1:} & \quad Y_{ijkl} = \gamma_{ijkl} + \gamma_1 X_{ijkl} + \epsilon_{ijkl} \\
\text{Level 2:} & \quad y_{ijkl} = \gamma_{00k} + \gamma_2 W_{ijkl} + \mu_{ijkl} \\
\text{Level 3 (F2):} & \quad y_{00kl} = \gamma_{000l} + \gamma_4 S_{kl} + \omega_{000} \\
\text{Level 4 (F1):} & \quad y_{000l} = \gamma_{0000} + \gamma_3 Z_l + v_{000}. 
\end{align*}
\]
Relative Biases of the Variance Component Estimates

The generated data were analyzed using the correctly specified CCREM and misspecified HLM (by ignoring the crossed factor F2) separately. The variance component estimates based on CCREM and HLM were summarized across the 500 replications. The relative bias for each variance component estimate was calculated using the following equation:

\[ B(\hat{\sigma}) = \frac{\hat{\sigma} - \sigma}{\sigma}, \]  

where \( \hat{\sigma} \) is the estimated variance component and \( \sigma \) is the true parameter value. A negative relative bias indicates an underestimation of the parameter (i.e., the estimated value is smaller than the true parameter value), whereas a positive
relative bias indicates an overestimation of the parameter (i.e., the estimated value is larger than the true parameter value).

Relative Biases of the Estimated Standard Errors of the Regression Coefficients

The relative biases of the estimated standard errors of the regression coefficients were computed using the following equation:

$$B(\hat{\sigma}_\gamma) = \frac{\hat{\sigma}_\gamma - \hat{\sigma}_{\gamma,\text{EMP}}}{\hat{\sigma}_{\gamma,\text{EMP}}}.$$  \hspace{1cm} (11)

where $\hat{\sigma}_\gamma$ is the estimated standard error of a regression coefficient and $\hat{\sigma}_{\gamma,\text{EMP}}$ is the empirical standard error, calculated as the standard deviation of the 500 estimates of in the true model.

Results

Table 2 shows that regardless of the position of F1 and F2 in the original hierarchy and the level of partial cross-classification, there was no bias in the estimated Level 1 residual variance (i.e., $\hat{\theta}$). When F1 was almost nested within F2 (i.e., five feeders and two receivers), ignoring F2 caused almost no bias in the estimated Level 2 variance component [$B(\hat{\psi}) = .04$] but large positive bias in the estimated variance component of the remaining crossed factor F1 [$B(\hat{\tau}) = 2.70$]. As the number of feeders and receivers increased, that is, as the data structure became more cross-classified, the bias in $\hat{\psi}$ became larger and the bias in $\hat{\tau}$ became smaller. When F2 was almost nested within F1, ignoring F2 caused positive bias in both estimated Level 2 and F1 variance components [$B(\hat{\psi}) = .45$ and $B(\hat{\tau}) = 1.15$]. As the data structure became more cross-classified, the bias in $\hat{\psi}$ became larger and the bias in $\hat{\tau}$ became smaller.

Regardless of the structure of cross-classification, the bias in the estimated standard error of the regression coefficient associated with the Level 1 predictor was acceptable according to the cutoff value of .10 recommended by Hoogland and Boomsma (1998). For the Level 2 predictor, there was almost no bias in the estimated standard error of the regression coefficient when F1 was almost nested within F2 [$B(\hat{S}_{\gamma,\text{F2}}) = -.02$]. A positive bias of this estimated standard error emerged and became larger as the data structure became more cross-classified. When F2 was almost nested within F1, there was small positive bias in the estimated standard error of the regression coefficient of the Level 2 predictor [$B(\hat{S}_{\gamma,\text{F1}}) = .11$] and the bias became slightly larger as the data structure became
TABLE 2
Ignoring a Crossed Factor at the Top Level With Partially Cross-Classified Data

<table>
<thead>
<tr>
<th>N of Feeders</th>
<th>N of Receivers</th>
<th>B(θ)</th>
<th>B(ψ)</th>
<th>B(ζ)</th>
<th>B(S_f1)</th>
<th>B(S_f2)</th>
<th>B(S_f3)</th>
<th>B(S_f4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
<td>.00</td>
<td>.04</td>
<td>2.70</td>
<td>.01</td>
<td>−.02</td>
<td>.35</td>
<td>−.40</td>
</tr>
<tr>
<td>8</td>
<td>.00</td>
<td>.06</td>
<td>2.55</td>
<td>.01</td>
<td>.05</td>
<td>.33</td>
<td>−.44</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>.00</td>
<td>.18</td>
<td>2.04</td>
<td>.00</td>
<td>.01</td>
<td>.32</td>
<td>−.58</td>
</tr>
<tr>
<td>8</td>
<td>.00</td>
<td>.31</td>
<td>1.42</td>
<td>−.01</td>
<td>.12</td>
<td>.29</td>
<td>−.65</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>.00</td>
<td>.33</td>
<td>1.38</td>
<td>.02</td>
<td>.06</td>
<td>.29</td>
<td>−.63</td>
</tr>
<tr>
<td>8</td>
<td>.00</td>
<td>.56</td>
<td>.36</td>
<td>.02</td>
<td>.22</td>
<td>.11</td>
<td>−.67</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N of Feeders</th>
<th>N of Receivers</th>
<th>B(θ)</th>
<th>B(ψ)</th>
<th>B(ζ)</th>
<th>B(S_f1)</th>
<th>B(S_f2)</th>
<th>B(S_f3)</th>
<th>B(S_f4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
<td>.00</td>
<td>.45</td>
<td>1.15</td>
<td>−.02</td>
<td>.11</td>
<td>.01</td>
<td>−.47</td>
</tr>
<tr>
<td>8</td>
<td>.00</td>
<td>.46</td>
<td>1.14</td>
<td>.00</td>
<td>.14</td>
<td>.03</td>
<td>−.48</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>.00</td>
<td>.52</td>
<td>.94</td>
<td>−.04</td>
<td>.12</td>
<td>.05</td>
<td>−.48</td>
</tr>
<tr>
<td>8</td>
<td>.00</td>
<td>.59</td>
<td>.50</td>
<td>.03</td>
<td>.15</td>
<td>.07</td>
<td>−.50</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>.00</td>
<td>.58</td>
<td>.66</td>
<td>.05</td>
<td>.17</td>
<td>.07</td>
<td>−.50</td>
</tr>
<tr>
<td>8</td>
<td>.00</td>
<td>.69</td>
<td>.15</td>
<td>.04</td>
<td>.17</td>
<td>.09</td>
<td>−.55</td>
<td></td>
</tr>
</tbody>
</table>

Note. θ is the estimated Level 1 residual variance, ψ is the estimated Level 2 variance component, ζ is the estimated variance component of the remaining crossed factor F1, S_f1 is the estimated standard error of the regression coefficient of Level 1 predictor, S_f2 is the estimated standard error of the regression coefficient of the Level 2 predictor, S_f3 is the estimated standard error of the regression coefficient of the predictor associated with the remaining crossed factor F1, S_f4 is the estimated standard error of the regression coefficient of the predictor associated with the ignored crossed factor F2.

For the predictor associated with F1 (i.e., the remaining crossed factor), there was a large positive bias in the standard error of the regression coefficient when F1 was almost nested within F2 [B(S_f1) = .35]. As the data became more cross-classified, this bias became smaller. On the other hand, a very small bias in the estimated standard error of the coefficient of the F1 predictor was found when F2 was almost nested within F1 [B(S_f2) = .01], and this bias became slightly larger as data became more cross-classified. For the standard error of the regression coefficient of the F2 predictor (i.e., the predictor of the ignored crossed factor), there was always a substantial negative bias and the bias increased as the data structure became more cross-classified.
Analytical Study

Consider the example of cross-classification at the intermediate level as shown in Figure 1b. Suppose there are \( n_1 \) fields of study (B). Within each field, there are \( n_2 \) assessors (F1) and \( n_3 \) proposals (F2). Within each unique combination of assessors and proposals, there are \( n_4 \) ratings (A). A random intercept cross-classified model with one predictor associated with each level and random factor is specified as follows:

\[
\begin{align*}
\text{Level 1:} & \quad Y_{i(jk)l} = \gamma_0(jk)l + \gamma_1 X_{i(jk)l} + \varepsilon_{i(jk)l} \\
\text{Level 2:} & \quad \gamma_0(jk)l = \gamma_{000l} + \gamma_2 Z_{j l} + \gamma_3 S_{kl} + \nu_{0jl} + \omega_{0kl} \\
\text{Level 3:} & \quad \gamma_{000l} = \gamma_{0000} + \gamma_4 W_l + \mu_{00l},
\end{align*}
\]  

where \( i \) indexes Level 1 unit A \((i = 1, \ldots, n_4)\), \( j \) indexes clusters of Level 2 crossed factor F1 \((j = 1, \ldots, n_2)\), \( k \) indexes clusters of Level 2 crossed factor F2 \((k = 1, \ldots, n_3)\), and \( l \) indexes clusters of Level 3 factor B \((l = 1, \ldots, n_1)\). The distributions of the random effects and residuals are \( \varepsilon_{i(jk)l} \sim N(0, \tau) \), \( \omega_{0kl} \sim N(0, \theta) \), and \( \mu_{00l} \sim N(0, \psi) \). The covariances between all random effects and residuals are zero.

The estimated variance components in Model 12 are as follows:

\[
\begin{align*}
\hat{\sigma}^2 &= \frac{SS_R}{(n_4 - 1)n_1n_2n_3 + n_1(n_2 - 1)(n_3 - 1)} \\
\hat{\tau} &= \frac{1}{n_2n_4} \left[ \frac{SS_{F1}}{(n_4 - 1)n_1} - \frac{SS_R}{(n_4 - 1)n_1n_2n_3 + n_1(n_2 - 1)(n_3 - 1)} \right] \\
\hat{\theta} &= \frac{1}{n_2n_4} \left[ \frac{SS_{F2}}{(n_3 - 1)n_1} - \frac{SS_R}{(n_4 - 1)n_1n_2n_3 + n_1(n_2 - 1)(n_3 - 1)} \right] \\
\hat{\psi} &= \frac{1}{n_2n_3n_4} \left[ \frac{SS_B}{n_1 - 1} - \frac{SS_{F1}}{(n_2 - 1)n_1} - \frac{SS_{F2}}{(n_3 - 1)n_1} \right] \\
&\quad - \frac{SS_R}{(n_4 - 1)n_1n_2n_3 + n_1(n_2 - 1)(n_3 - 1)},
\end{align*}
\]  

where \( SS_R \) is the sum of squares of the Level 1 residuals, \( SS_{F1} \) and \( SS_{F2} \) are the sums of squares of the residuals of clusters in F1 and F2, respectively, and
$SS_B$ is the sum of squares of the residuals of Level 3 clusters (see Appendix for the formulas).

Ignoring the Level 2 crossed factor $F2$ yields

$$
\begin{align*}
\hat{\theta}^* &= \frac{SS_R + SS_{F2}}{n_1n_2(n_3n_4 - 1)} \\
\hat{\xi}^* &= \frac{1}{n_3n_4} \left[ \frac{SS_{F1}}{(n_2 - 1)n_1} - \frac{SS_R + SS_{F2}}{n_1n_2(n_3n_4 - 1)} \right] \quad (14) \\
\hat{\psi}^* &= \frac{1}{n_2n_3n_4} \left[ \frac{SS_B}{n_1} - \frac{SS_{F1}}{(n_2 - 1)n_1} \right].
\end{align*}
$$

Comparing the estimated variance components in the misspecified model and the correct model, we obtained

$$
\begin{align*}
\hat{\theta}^* &= \hat{\theta} + \frac{n_4(n_3 - 1)}{n_3n_4 - 1} \hat{\xi} \\
\hat{\xi}^* &= \hat{\xi} - \frac{n_3 - 1}{n_3(n_3n_4 - 1)} \hat{\xi} \quad (15) \\
\hat{\psi}^* &= \hat{\psi} + \frac{1}{n_3} \hat{\xi}.
\end{align*}
$$

Hence ignoring a crossed factor at the intermediate level causes the Level 1 and Level 3 variance components to be overestimated and the variance component of the remaining crossed factor $F1$ to be underestimated.

The variances of the regression coefficients in Model 12 are

$$
\begin{align*}
\text{Var}(\hat{\gamma}_1) &= \frac{\hat{\theta}}{S^2_{\hat{\gamma}_1}n_1n_2n_3n_4} \\
\text{Var}(\hat{\gamma}_2) &= \frac{\hat{\theta} + n_3n_4\hat{\xi}}{S^2_{\hat{\gamma}_2}n_1n_2n_3n_4} \\
\text{Var}(\hat{\gamma}_3) &= \frac{\hat{\theta} + n_2n_4\hat{\xi}}{S^2_{\hat{\gamma}_3}n_1n_2n_3n_4} \\
\text{Var}(\hat{\gamma}_4) &= \frac{\hat{\theta} + n_3n_4\hat{\xi} + n_2n_4\hat{\psi} + n_2n_3n_4\hat{\psi}}{S^2_{\hat{\gamma}_4}n_1n_2n_3n_4}. \quad (16)
\end{align*}
$$
Ignoring the Level 2 crossed factor $F_2$ yields

\[
\begin{align*}
\text{Var}(\hat{\gamma}_1)^* &= \frac{\hat{\theta}^*}{S_X^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_2)^* &= \frac{\hat{\theta}^* + n_3 n_4 \hat{\xi}^*}{S_X^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_3)^* &= \frac{\hat{\theta}^* + n_2 n_4 \hat{\xi}^*}{S_X^2 n_1 n_2 n_3 n_4} \\
\text{Var}(\hat{\gamma}_4)^* &= \frac{\hat{\theta}^* + n_3 n_4 \hat{\xi}^* + n_4 n_5 \hat{\zeta}^*}{S_X^2 n_1 n_2 n_3 n_4}.
\end{align*}
\]

Substituting Equation 15 into 17 and comparing the results with those in Equation 16, we obtained

\[
\begin{align*}
\text{Var}(\hat{\gamma}_1)^* &= \text{Var}(\hat{\gamma}_1) + \frac{n_3 - 1}{S_X^2 n_1 n_2 n_3 (n_3 n_4 - 1)} \hat{\xi} \\
\text{Var}(\hat{\gamma}_2)^* &= \text{Var}(\hat{\gamma}_2) \\
\text{Var}(\hat{\gamma}_3)^* &= \text{Var}(\hat{\gamma}_3) - \frac{n_2 n_3 (n_4 - 1) + (n_2 - 1) (n_3 - 1)}{S_X^2 n_1 n_2 n_3 (n_3 n_4 - 1)} \hat{\xi} \\
\text{Var}(\hat{\gamma}_4)^* &= \text{Var}(\hat{\gamma}_4).
\end{align*}
\]

According to Equation 18, ignoring a crossed factor ($F_2$) at the intermediate level causes overestimation of the variance of the regression coefficient of the Level 1 predictor but underestimation of the variance of the regression coefficient of the predictor associated with the ignored crossed factor ($F_2$). The variances of the regression coefficients of the predictors associated with the remaining crossed factor ($F_1$) and Level 3 are not affected.

**Demonstration With Simulated Data**

A simulated data set was used to illustrate the analytical results. A total of 25 fields of study were generated. Within each field of study, there were 50 assessors ($F_1$) and 20 proposals ($F_2$). Within each of the 1,000 (50*20) cells, there were 4 ratings. Hence, there are 4,000 (4*50*20) ratings in each field of study. Model 12 was used to generate the example data.
TABLE 3
Ignoring a Crossed Factor at the Intermediate Level

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Correct Model (CCREM)</th>
<th>Misspecified Model (HLM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>Regression coefficients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rating: $\gamma_1$</td>
<td>.501</td>
<td>.002</td>
</tr>
<tr>
<td>Assessor: $\gamma_2$</td>
<td>.498</td>
<td>.009</td>
</tr>
<tr>
<td>Proposal: $\gamma_3$</td>
<td>.453</td>
<td>.025</td>
</tr>
<tr>
<td>Fields of study: $\gamma_4$</td>
<td>.457</td>
<td>.057</td>
</tr>
<tr>
<td>Variance components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fields of study ($\psi$)</td>
<td>.087</td>
<td></td>
</tr>
<tr>
<td>Assessor ($\tau$)</td>
<td>.100</td>
<td></td>
</tr>
<tr>
<td>Proposal ($\zeta$)</td>
<td>.305</td>
<td></td>
</tr>
<tr>
<td>Residual ($\theta$)</td>
<td>.497</td>
<td></td>
</tr>
</tbody>
</table>

Note. CCREM = Cross-Classified Random-Effects Model; HLM = Hierarchical Linear Model.

As shown in Table 3, ignoring the crossed factor of proposals resulted in lower estimated variance component of the assessors ($\tau$), higher estimated variance component of fields of study ($\psi$) and residual variance ($\theta$). Consistent with the analytical results, ignoring the crossed factor of proposals caused higher estimated standard error of the regression coefficient of the predictor at the rating level ($\gamma_1$) and lower estimated standard error of the regression coefficient of the predictor associated with proposals ($\gamma_3$). The standard errors of the regression coefficients of the predictors associated with assessors ($\gamma_2$) and fields of study ($\gamma_4$) were unaffected.

Partially Cross-Classified Data

Similar to Study 1, we conducted a simulation study to investigate the effect of ignoring a crossed factor at the intermediate level with partially cross-classified data.

Data Generation

Using similar procedures as presented in Study 1, we first generated the data from the two extreme conditions: (a) the hierarchy of F1 nested within F2 (i.e., 10 ratings nested within each assessor, three assessors nested within each proposal, and 20 proposals nested within each of the 25 fields of study) and (b) the hierarchy of F2 nested within F1. Then, we created different levels of partial cross-classification by manipulating the number of feeders and receivers.
The variance components of fields of study (ψ), assessors (τ), proposals (ζ), and residual (θ) were generated with values of .10, .10, .30, and .50, respectively. We doubled the magnitude of the variance components of the two crossed factors because the crossed factors were at the intermediate level and larger variance components are more likely to occur at the lower level. The other conditions for data generation were the same as those in Study 1.

Results

The results presented in Table 4 show that when F1 was almost nested within F2 (i.e., five feeders and two receivers), ignoring F2 caused almost no bias in the estimated Level 1 residual variance \( \hat{B}(\theta) = .03 \) but large positive bias in the estimated variance component of the remaining crossed factor F1 \( \hat{B}(\tau) = 2.73 \).

<table>
<thead>
<tr>
<th>N of Feeders</th>
<th>N of Receivers</th>
<th>( B(\hat{\theta}) )</th>
<th>( B(\hat{\tau}) )</th>
<th>( B(\hat{\psi}) )</th>
<th>( B(\hat{\zeta}_1) )</th>
<th>( B(\hat{\zeta}_2) )</th>
<th>( B(\hat{\zeta}_3) )</th>
<th>( B(\hat{\zeta}_4) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
<td>.03</td>
<td>2.73</td>
<td>.11</td>
<td>.01</td>
<td>.44</td>
<td>-.42</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>.05</td>
<td>2.61</td>
<td>.11</td>
<td>.04</td>
<td>.39</td>
<td>-.47</td>
<td>-.01</td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>.16</td>
<td>2.09</td>
<td>.11</td>
<td>.08</td>
<td>.41</td>
<td>-.56</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>.27</td>
<td>1.50</td>
<td>.15</td>
<td>.12</td>
<td>.30</td>
<td>-.63</td>
<td>.00</td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>.28</td>
<td>1.44</td>
<td>.15</td>
<td>.10</td>
<td>.32</td>
<td>-.64</td>
<td>-.03</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>.49</td>
<td>.38</td>
<td>.14</td>
<td>.24</td>
<td>.12</td>
<td>-.65</td>
<td>-.01</td>
</tr>
</tbody>
</table>

Note. \( \hat{\theta} \) is the estimated Level 1 residual variance, \( \hat{\tau} \) is the estimated variance component of the remaining crossed factor F1 at Level 2, \( \hat{\psi} \) is the estimated Level 3 variance component, \( \hat{\zeta}_1 \) is the estimated standard error of the regression coefficient of the Level 1 predictor, \( \hat{\zeta}_2 \) is the estimated standard error of the regression coefficient of the predictor associated with the remaining crossed factor F1, \( \hat{\zeta}_3 \) is the estimated standard error of the regression coefficient of the predictor associated with the ignored crossed factor F2, and \( \hat{\zeta}_4 \) is the estimated standard error of the regression coefficient of the Level 3 predictor.
As the data structure became more cross-classified, the bias of \( \hat{\theta} \) became larger and the bias of \( \hat{\tau} \) became smaller but still substantial. There was a consistent positive bias in the estimated Level 3 variance component (\( \hat{\psi} \)).

When \( F_2 \) was almost nested within \( F_1 \), ignoring \( F_2 \) caused positive bias in both the estimated Level 1 residual variance [\( B(\hat{\theta}) = .38 \)] and the estimated variance component of the remaining crossed factor \( F_1 \) [\( B(\hat{\tau}) = 1.16 \)] but not in the estimated Level 3 variance component [\( B(\hat{\psi}) = -.01 \)]. As the data structure became more cross-classified, the bias of \( \hat{\theta} \) became larger and the bias of \( \hat{\tau} \) became smaller.

For the Level 1 predictor, there was no bias in the estimated standard error of the regression coefficient when \( F_1 \) was almost nested within \( F_2 \) [\( B(\hat{S}_{\tau_1}) = .01 \)]. A positive bias emerged and became larger as the data structure became more cross-classified. On the other hand, there was a positive bias in the estimated standard error of the Level 1 regression coefficient when \( F_2 \) was almost nested within \( F_1 \) [\( B(\hat{S}_{\tau_2}) = .12 \)] and the bias also became larger as the data structure became more cross-classified.

For the predictor associated with the remaining crossed factor \( F_1 \), there was substantial positive bias in the estimated standard error of the regression coefficient when \( F_1 \) was almost nested within \( F_2 \) [\( B(\hat{S}_{\tau_1}) = .44 \)]. As the data became more cross-classified, the bias became smaller. On the other hand, there was little bias in the estimated standard error of the regression coefficient of the \( F_1 \) predictor when \( F_2 \) was almost nested within \( F_1 \) [\( B(\hat{S}_{\tau_2}) = -.02 \)]. Positive bias in this estimated standard error emerged as the data became more cross-classified.

For the predictor associated with \( F_2 \) (i.e., the ignored crossed factor), there was always a large negative bias in the estimated standard error of the regression coefficient and this bias increased as the data structure became more cross-classified. For the Level 3 predictor, there was almost no bias regardless of the structure of cross-classification.

**DEMONSTRATION WITH EMPIRICAL DATA**

We used data from the Early Childhood Longitudinal Study–Kindergarten Cohort (ECLS-K) for the purpose of a simple empirical demonstration of our analytical and simulation results.

**Cross-Classification at the Top Level**

In the ECLS-K data, schools and assessors were partially crossed because a school could have several designated assessors and an assessor could assess students in several designated schools (see Figure 3). In the example data, scores on different measures were nested within students and students were cross-classified by schools and assessors. Scores on different measures
FIGURE 3
Structure of cross-classification of schools and assessors in the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K)

Note: Numbers in the cells represent frequencies.
were modeled as the function of the type of the measures (i.e., \( \text{MATH} = 1 \) if a measure was related to math skills, whereas \( \text{MATH} = 0 \) if a measure was related to verbal skills), student’s social and economical status (\( \text{SES} \)), and school type (\( \text{PRIVATE} = 1 \) for private schools and \( \text{PRIVATE} = 0 \) for public schools):

\[
\begin{align*}
\text{Level 1: } & \quad \text{SCORE}_{ij(kl)} = \gamma_{0j(kl)} + \gamma_1 \text{MATH}_{ij(kl)} + \epsilon_{ij(kl)} \\
\text{Level 2: } & \quad \gamma_{0j(kl)} = \gamma_{00(kl)} + \gamma_2 \text{SES}_{j(kl)} + \mu_{0j(kl)} \\
\text{Level 3: } & \quad \gamma_{00(kl)} = \gamma_{0000} + \gamma_3 \text{PRIVATE}_k + \nu_{00k} + \omega_{00j/l}.
\end{align*}
\]

(19)

Predictors associated with the assessors (i.e., the ignored crossed factor) were not included in the model due to the lack of such information in the original data.

Table 5 presents the parameter estimates based on the correctly specified CCREM and the misspecified HLM in which the assessor crossed factor was ignored. The comparison shows that the variance component of the ignored assessor crossed factor was redistributed to the remaining school crossed factor and the student level but not to the score level. The standard error of the regression coefficient of school type was higher in the HLM than in the CCREM. The difference was very small because the magnitude of the variance component of the ignored assessor crossed factor was comparatively small. The standard error of the regression coefficients of the type of measures was unaffected. It is

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Correct Model (CCREM)</th>
<th>Misspecified Model (HLM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>Regression coefficients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1: MATH</td>
<td>-.1383</td>
<td>.0104</td>
</tr>
<tr>
<td>Level 2: SES</td>
<td>.1014</td>
<td>.0092</td>
</tr>
<tr>
<td>Level 3 F1: PRIVATE</td>
<td>-.0137</td>
<td>.0222</td>
</tr>
<tr>
<td>Variance components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School</td>
<td>.0025</td>
<td>.0026</td>
</tr>
<tr>
<td>Assessor</td>
<td>.0003</td>
<td>—</td>
</tr>
<tr>
<td>Student</td>
<td>.0063</td>
<td>.0065</td>
</tr>
<tr>
<td>Residual</td>
<td>.1512</td>
<td>.1512</td>
</tr>
</tbody>
</table>

*Note.* ECLS-K = Early Childhood Longitudinal Study-Kindergarten Cohort; CCREM = Cross-Classified Random-Effects Model; HLM = Hierarchical Linear Model; MATH = 1 for measures related to math skills and MATH = 0 for measures related to verbal skills; SES = Students’ social and economic status; PRIVATE = 1 for private schools and PRIVATE = 0 for public schools.
noted that the regression coefficient associated with students’ SES and school type were slightly different between the HLM and the CCREM results. This may be due to the facts that (a) the data were unbalanced and (b) the predictor variables had variance component at other levels.

Cross-Classification at the Intermediate Level

In the case of cross-classification at the intermediate level, students were cross-classified by schools and assessors that were nested within regions. The student’s verbal composite score was modeled as the function of the student’s social and economical status (SES) and school type (PRIVATE):

\[
\begin{align*}
\text{Level 1:} & \quad \text{VERBAL}_{i(jk)l} = \gamma_0(jk)l + \gamma_1\text{SES}_{i(jk)l} + \epsilon_{i(jk)l} \\
\text{Level 2:} & \quad \gamma_0(jk)l = \gamma_{000}l + \gamma_2\text{PRIVATE}_{j,l} + v_{0jl} + \mu_{0jl} \\
\text{Level 3:} & \quad \gamma_{000}l = \gamma_{0000} + \mu_{0jl}.
\end{align*}
\]

Due to the lack of information in the original data, predictors associated with assessors and regions were not included in the model. Table 6 shows that the variance component of the ignored assessor crossed factor was redistributed to the remaining crossed factor of school, the top level (i.e., regions), and the lower level (i.e., students). The standard errors of the regression coefficients of students’ SES and school type were slightly higher in the HLM than in the CCREM.

<table>
<thead>
<tr>
<th>TABLE 6</th>
<th>Empirical Demonstration With ECLS-K Data: Cross-Classification at the Intermediate Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
<td>Correct Model (CCREM)</td>
</tr>
<tr>
<td>Regression coefficients</td>
<td></td>
</tr>
<tr>
<td>Level 1: SES</td>
<td>.1025</td>
</tr>
<tr>
<td>Level 2 F1: PRIVATE</td>
<td>.0504</td>
</tr>
<tr>
<td>Variance components</td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td>.0000</td>
</tr>
<tr>
<td>School</td>
<td>.0034</td>
</tr>
<tr>
<td>Assessor</td>
<td>.0020</td>
</tr>
<tr>
<td>Residual</td>
<td>.0376</td>
</tr>
</tbody>
</table>

Note. ECLS-K = Early Childhood Longitudinal Study-Kindergarten Cohort; CCREM = Cross-Classified Random-Effects Model; HLM = Hierarchical Linear Model; SES = Students’ social and economic status; PRIVATE = 1 for private schools and PRIVATE = 0 for public schools.
DISCUSSION

Cross-classified multilevel data are fairly common in social and behavioral sciences. Cross-classified random-effects models (CCREMs) were developed to analyze such data with great flexibility. However, due to some reasons (e.g., missing higher level cluster IDs\(^1\)), researchers may not be able to take the cross-classified structure into account and have to treat the data as strictly hierarchical in the analyses. We investigated the impacts of ignoring a crossed random factor at either the top level or the intermediate level on the estimation of the variance components and the variances of the regression coefficients. Through both analytical and simulation studies, we found that (a) the effect of ignoring a random crossed factor is similar to that of ignoring a level of nesting and (b) the direction and magnitude of the biases depend not only on factors such as magnitude of the variance component of the ignored crossed factor, sample size, and the level with which the predictor variable is associated but also on the structure of cross-classification.

Similarity of Ignoring a Crossed Factor and Ignoring a Level of Nesting

A few recent studies have investigated the consequences of ignoring a level of nesting in multilevel models with strictly hierarchical structures (Moerbeek, 2004; Tranmer, 2001; Van den Noortgate, Opdenakker, & Onghena, 2005; Van Landeghem, De Fraine, & Van Damme, 2005). Within the framework of a balanced design and random intercept model (i.e., all regression coefficients are fixed instead of random), analytical results showed that when the highest level is ignored, its variance component is added to the variance component at the level right below it. If an intermediate level is ignored, its variance component is split into two parts and redistributed to the variance components at the adjacent levels (i.e., the level above and the level below). When the \(k\)th level is ignored, the standard errors of the regression coefficients of the predictor variables at the \(k\)th level are generally underestimated and those at the \((k - 1)\)th level are overestimated.

The results from our analytical studies showed that the impacts of ignoring a crossed factor are similar to those of ignoring a level of nesting in hierarchical models. When the data structure is completely cross-classified and a crossed factor at the top level is ignored, the variance component of the ignored factor is added to the variance component at the level right below it. When a crossed

\(^1\)For example, students within the same school can change classrooms over time. However, researchers may only be able to access the information related to students’ ID and their corresponding school IDs but not any information related to the change of classroom (or classroom IDs).
factor at the intermediate level is ignored, part of its variance component is added to the variance component at the level above it and part added to the variance component at the level below it.

Interestingly, according to Equations 4 and 15, ignoring a crossed factor will cause the estimated variance component of the remaining crossed factor to decrease. However, from the equations it is easy to see that the amount of reduction will be trivial if the number of the lower level units in the remaining crossed factor is large.

Similar to ignoring a level of nesting, ignoring a crossed factor at the $k$th level causes overestimation of the standard errors of the regression coefficients of the predictor variables at the $(k - 1)$th level and underestimation of the standard errors of the regression coefficients of the predictor variables associated with the ignored factor. The standard errors of the regression coefficients of the predictor variables associated with the remaining crossed factor and at the $(k + 1)$th level are generally unaffected.

Structure of Cross-Classification

The simulation studies with partially cross-classified data showed that the degree of cross-classification can dramatically affect the direction and magnitude of the observed bias of parameter estimates. Consider the structure of the cross-classification as a continuum with the remaining crossed factor (e.g., $F_1$ in our simulation studies) nested within the ignored crossed factor (e.g., $F_2$ in our simulation studies) as one end of the continuum and the ignored crossed factor nested within the remaining crossed factor as the other end of the continuum. The condition with the two factors completely crossed is at the middle of the continuum. When the data structure is closer to the two extreme conditions, the two crossed factors are more like two nested levels and ignoring a crossed factor is more similar to ignoring a level of nesting in hierarchical models. Therefore, the direction and magnitude of the biases depend on the relative position of the ignored factor in the hierarchy.

Figure 4 shows the redistribution of the variance component of the ignored top-level crossed factor under different structures of cross-classification. When the remaining crossed factor ($F_1$) is almost nested within the ignored crossed factor ($F_2$), a very small part of $F_2$ variance component is added to the variance component at the level below (represented by the dashed arrow in Figure 4a), and a large part of the variance component is added to the variance component of the remaining Level 3 factor $F_1$ (represented by the bold arrow). As the data become more cross-classified, the part that is added to Level 2 variance component increases. When the data is completely cross-classified, all the variance component of $F_2$ is added to the Level 2 variance component (represented by the bold arrow in Figure 4b). On the other hand, when $F_2$ is almost nested within
FIGURE 4  Distribution of the variance component of the ignored crossed factor (F2) at the top level. Arrows point to the recipients of the variance component of the ignored crossed factor F2. Solid bold arrows indicate that a large proportion of the variance component of F2 is added to the recipient, solid thin arrows indicate a small proportion, and dashed arrows indicate that little of the variance component of F2 is added to the recipient. (a) F1 almost nested within F2. (b) F1 and F2 completely cross-classified. (c) F2 almost nested within F1.

F1, a large part of the variance component of F2 is redistributed to the variance component at the level below and a small part of it is redistributed to the variance component of the remaining crossed factor F1 (see Figure 4c). Additionally, the variance of F2 is not redistributed to the Level 1 residual variance regardless of the structure of cross-classification.

Figure 5 shows the redistribution of the variance component of the ignored intermediate-level crossed factor under different structures of cross-classification. When the remaining crossed factor (F1) is almost nested within the ignored crossed factor (F2), a very small part of F2 variance component is added to the variance component at the level below (represented by the dashed arrow in Figure 5a), a large part of the variance component is added to the variance component of F1 (represented by the bold arrow), and a small part is added to the variance component at Level 3 (represented by the thin arrow). On the other hand, when F2 is almost nested within F1, little of the variance component of F2 is redistributed to Level 3 variance component (represented by the dashed arrow in Figure 5c). A large part of the F2 variance component is added to Level 1 variance component (represented by the bold arrow) and a small part of it is added to F1 variance component (represented by the thin arrow). When the data structure is completely cross-classified, a large part of F2 variance component is added to the Level 1 variance component and a small part of it is added to Level 3 variance component (Figure 5b).

This finding also explained the inconsistent results from studies in the literature. The estimated variance component of the remaining crossed factor was found to be substantially overestimated in Meyers & Beretvas’s (2006) study.
IGNORING A CROSSED FACTOR IN CROSS-CLASSIFIED MODELS

FIGURE 5 Distribution of the variance component of the ignored crossed factor (F2) at the intermediate level. Arrows point to the recipients of the variance component of the ignored crossed factor F2. Solid bold arrows indicate that a large proportion of the variance component of F2 is added to the recipient, solid thin arrows indicate a small proportion, and dashed arrows indicate that little of the variance component of F2 is added to the recipient. (a) F1 almost nested within F2. (b) F1 and F2 completely cross-classified. (c) F2 almost nested within F1.

but underestimated in Fielding’s (2002) study. One of the possible reasons is that the data structure in the latter study is more cross-classified than that in the former study.

LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Study results need to be interpreted in light of certain limitations. Both the analytical and simulation studies are within a restrictive framework, such as the fixed-effect predictors (i.e., no random effect associated with any regression coefficient), the correct specification of the fixed part of the model, and the assumptions of the normal distribution of the response variable and residuals. Thus, extrapolating the current findings to models beyond these conditions should be done with caution. As Van Landeghem et al. (2005) pointed out, the direction and magnitude of the bias in the estimated standard error of the regression coefficient of a predictor variable could depend on the variance components of the predictor variable itself at the other levels. In addition, random coefficient models are also common for multilevel data. Some analytical results have been obtained for the strictly hierarchical data (Berkhof & Kampen, 2004; Lange & Laird, 1989), but results for cross-classified random coefficient models are still unknown. Further investigation on these issues, including the
impact of ignoring a crossed factor in random coefficient models, models with misspecification in either the fixed part (Raudenbush & Bryk, 2002) or the random part (Kwok, West, & Green, 2007), or both is needed.

In this study we provided one possible way to control for different levels of partial cross-classification structures by generating data from the extreme condition in which one crossed factor was completely nested within the other. We manipulated the number of feeders and receivers to create different degrees of cross-classification. More efforts would be needed to examine other potential ways to represent the structure of cross-classification and to examine the mechanisms that cause different types of structures.

**IMPLICATIONS**

In some circumstances such as the lack of information of group membership of the crossed factor, researchers might not be able to use cross-classified random-effects models. If the underlying assumptions are similar to those used in this study, ignoring a crossed factor and analyzing cross-classified data using hierarchical linear models will not cause spurious significant results for predictors that are not associated with the ignored crossed factor. Nevertheless, the price of ignoring a crossed factor is the reduction in the statistical power on detecting the target (fixed) effects because of the redistribution of the variance component of the ignored crossed factor, which in turn results in overestimation of the standard error and reduces the statistical power. It should be noted that the standard errors of the regression coefficients of predictors associated with the ignored crossed factor will be underestimated, resulting in inflated Type I error rates.

Both our empirical demonstration and previous studies (Fielding, 2002; Morebeek, 2004) show that when the data are unbalanced, ignoring a crossed factor may have an effect not only on the standard errors but also on the point estimates of regression coefficients themselves. The mechanism underlying this effect is still unclear and further research on this issue is needed. Researchers should try to prevent ignoring a crossed factor, especially when they have severely unbalanced data.

**ACKNOWLEDGMENTS**

This work was partially supported by Project Achieve, Grant R01HD039367, funded by the National Institute of Child Health and Human Development. The contents of this study are solely our responsibility and do not necessarily represent the official views of the funding agency. We are grateful to Victor...
Willson for his invaluable suggestions on earlier versions of this article as part of Wen Luo’s dissertation.

REFERENCES


APPENDIX

In the case of two random factors (F1 and F2) crossed at the top level, let $k$ index clusters in F1 ($k = 1, \ldots, n_1$), $l$ index clusters in F2 ($l = 1, \ldots, n_2$), $j$ index Level 2 clusters within each of the $n_1 \times n_2$ cells ($j = 1, \ldots, n_3$), and $i$ index Level 1 units within each of the Level 2 clusters ($i = 1, \ldots, n_4$). The formulas for the sum of squares of residuals at Level 1, Level 2, Level 3 crossed factor F1, and Level 3 crossed factor F2 are

\[
\begin{align*}
SS_R &= \sum \sum \sum [Y_{ij(kl)} - \bar{Y}_{j(kl)} - (\Lambda_{ij(kl)} - \bar{\Lambda}_{j(kl)}) \Gamma]^2 \\
SS_B &= \sum \sum \sum n_4 [\bar{Y}_{..(kl)} - \bar{Y}_{..l} - (\bar{\Lambda}_{..(kl)} - \bar{\Lambda}_{..l}) \Gamma]^2 \\
SS_{F1} &= \sum n_2 n_3 n_4 [\bar{Y}_{..l} - \bar{\Lambda}_{..l} \Gamma]^2 \\
SS_{F2} &= \sum n_1 n_3 n_4 [\bar{Y}_{..l} - \bar{\Lambda}_{..l} \Gamma]^2,
\end{align*}
\]

where $\Lambda_{ij(kl)}$ is a row vector containing the predictor variables, and $\Gamma$ is a column vector containing the regression coefficients (including the intercept).

In the case of two random factors (F1 and F2) crossed at the intermediate level, let $l$ index Level 3 clusters ($l = 1, \ldots, n_1$), $j$ index clusters in F1 within each Level 3 cluster ($j = 1, \ldots, n_2$), $k$ index clusters in F2 within each Level 3 cluster ($k = 1, \ldots, n_3$), and $i$ index Level 1 units within each unique combination of F1 and F2 ($i = 1, \ldots, n_4$). The formulas for the sum of squares of residuals at Level 1, Level 2 crossed factor F1, Level 2 crossed factor F2, and Level 3 are

\[
\begin{align*}
SS_R &= \sum \sum \sum \sum [Y_{i(jk)l} - \bar{Y}_{.j.l} - \bar{Y}_{..l} + \bar{Y}_{..l}] \\
&\quad - (\Lambda_{i(jk)l} - \bar{\Lambda}_{..jl} - \bar{\Lambda}_{..kl} + \bar{\Lambda}_{..l}) \Gamma]^2 \\
SS_{F1} &= \sum \sum n_3 n_4 \bar{Y}_{.j..l} - \bar{\Lambda}_{..jl} - \bar{\Lambda}_{..kl} + \bar{\Lambda}_{..l} \Gamma]^2 \\
SS_{F2} &= \sum \sum n_2 n_4 \bar{Y}_{..l..} - \bar{\Lambda}_{..jl} - \bar{\Lambda}_{..kl} + \bar{\Lambda}_{..l} \Gamma]^2 \\
SS_B &= \sum n_2 n_3 n_4 \bar{Y}_{..l..} - \bar{\Lambda}_{..jl} \Gamma]^2,
\end{align*}
\]

where $\Lambda_{i(jk)l}$ is a row vector containing the predictor variables, and $\Gamma$ is a column vector containing the regression coefficients (including the intercept).