

# X Matters Too: How the Blended Slope Problem Manifests Differently in Unilevel vs. Multilevel Models

Elizabeth A. Sanders<sup>1</sup> , Timothy R. Konold<sup>2</sup> 

[1] *Measurement & Statistics, College of Education, University of Washington, Seattle, WA, USA.* [2] *Research, Statistics, & Evaluation, School of Education and Human Development, University of Virginia, Charlottesville, VA, USA.*

---

Methodology, 2023, Vol. 19(1), 1–23, <https://doi.org/10.5964/meth.9925>

Received: 2022-07-16 • Accepted: 2023-02-08 • Published (VoR): 2023-03-31

Handling Editor: Francisco Abad, Universidad Autónoma de Madrid, Madrid, Spain

Corresponding Author: Elizabeth A. Sanders, University of Washington, 312G Miller Hall, Box 353600, Seattle, WA 98195-3600, USA. E-mail: [lizz@uw.edu](mailto:lizz@uw.edu)

---

## Abstract

Aside from multilevel models (MLMs), several analytic approaches are available for handling cluster-induced dependencies. Nevertheless, the literature on MLM alternatives has called less explicit attention to the potential bias in level-1 (L1) slope coefficients resulting from the “blended slope” problem—a problem that arises when dependencies in predictors (Xs) exist and when L1 predictor-outcome (X-Y) relations differ from those at level-2 (L2). As such, applied researchers may be drawing incorrect inferences about their L1 predictor effects when they specify models without considering clustering in Xs. The present paper reviews this “blended slope” problem and uses Monte Carlo simulation to illustrate how the problem manifests more for unilevel models compared with MLMs. In short, analyses of clustered data should always: 1) report outcome *and predictor* ICCs, 2) cluster-mean center L1 predictors or incorporate L2 aggregate predictors, and 3) employ a model that takes clustered residuals into account.

## Keywords

blended slope, clustering, centering, between-within, predictor ICC, multilevel model, cluster-robust standard errors, fixed vs. random effects

Multilevel models (also known as “hierarchical,” “random effects,” and “mixed effects” models) for clustered data<sup>1</sup> have become a staple analytic approach in the psychological and educational sciences (Antonakis et al., 2021; Bauer & Sterba, 2011; Luo et al., 2021) as well as other social science disciplines (Bell & Jones, 2015; Bell et al., 2019). Not only can multilevel models (MLMs) account for non-independence in residual errors, they are also able to flexibly disentangle within- and between-cluster relationships, allowing



researchers to test context-driven research questions, such as “Does therapist training predict individual-level patient mental health outcomes?” or “Does school climate predict individual-level student achievement?” Moreover, MLMs can model more than two levels of data, model data that are cross-classified or partially nested, and be useful in estimating both slope and intercept variance (i.e., relaxing the assumption of homogeneity of slopes).

Despite their popularity, quantitative methodologists have pointed out that the use of MLM can be unnecessarily complex for clustered data when MLM model assumptions are untenable (especially in smaller samples) or when the substantive research questions are focused primarily on level 1 (L1) in a 2-level hierarchy, such as “Does a patient’s own treatment fidelity predict their mental health?” or “Does a student’s own perception of school belongingness predict their academic achievement?” While a few MLM alternatives exist<sup>2</sup>, unilevel regression (also known as single-level and fixed effects models) with cluster-robust standard errors (CR-SEs) is perhaps the most convenient and flexible alternative for handling residual error non-independence without MLM-specific assumptions (Bell et al., 2019; Hayes & Cai, 2007; Huang, 2016; McNeish et al., 2017). Nonetheless, applied researchers may ignore clustering completely in circumstances when they find the variance due to clustering in the outcome (Y) is not statistically significant (Raudenbush & Bryk, 2002, pp. 63–64; Snijders & Bosker, 2012, p. 22), or if the “design effect” computed for Y is  $< 1.5$  (Lai & Kwok, 2015). Yet, the assumption of independence for any model pertains to the residual errors, not Y itself, and dependencies in X will contribute to inducing correlated residuals, also known as the endogeneity problem (e.g., Grilli & Rampichini, 2011). Further, even when correlated residuals stemming from Y or X are appropriately handled using MLM or a unilevel alternative, researchers may be quite unaware that their L1 predictor slope estimates are still prone to bias as a result of what we call the “blended slope” problem (e.g., Hamaker & Muthén, 2020)—an issue more well known in the multilevel literature that occurs when there is clustering in X (a non-zero intraclass correlation) and different X–Y relations at L1 and L2 (which cannot be readily known a priori).

---

1) Clustered data can include scores from measurements within individuals (i.e., repeated measures data, such as panel data) as well as scores from individuals within structural units (i.e., cross-sectional data), or some blend of the two. For clarity but without loss of generality, we focus on the 2-level cross-sectional hierarchical data.

2) There are three main approaches for analyzing 2-level hierarchical data: 1) a random intercept multilevel model in which the residual error variance is decomposed into within- and between-cluster variance (clustering is treated as a random effect); 2) a unilevel fixed effects model in which clustering is treated as a fixed effect using a set of dummy- or effect-coded predictors that control for all variation due to clustering in the outcome and other L1 predictors, disallowing inclusion of cluster-level L2 predictors (but see Bell & Jones, 2015 for how to allow L2 predictor inclusion); and 3) a unilevel fixed effects model where clustering is not treated as a random or fixed effect, but instead is used in the residual standard error adjustment, after OLS estimation of parameter estimates (e.g., cluster-robust SEs, design-effect adjusted SEs, or algorithmically adjusted SEs in a generalized estimating equation).

The present paper examines the “blended slope” problem in terms of how it applies differentially to unilevel regression models (Mundlak, 1978) that do or do not take clustering into account, as compared with multilevel models (Raudenbush & Bryk, 2002), with specific attention on the independent contribution of cluster dependencies in X that might be ignored by researchers if there are little or no observed dependency in Y. First, we review the problem mathematically for the 2-level case; next, we review how to fix the problem. Thereafter, we use results from a Monte Carlo simulation to illustrate how the problem manifests in different analytic approaches to modeling clustered data<sup>3</sup>. Consistent with others (Antonakis et al., 2021; Bell et al., 2019; Enders & Tofighi, 2007; Hamaker & Muthén, 2020), our simulation results also demonstrate how cluster-mean centering L1 predictors can solve the problem in both MLMs and their unilevel alternatives<sup>4</sup>.

## The “Blended Slope” Problem

Fairly recent simulation evidence by McNeish (2014) appeared to suggest that X–Y regression slope coefficient estimates are equivalent for MLM and unilevel fixed effects regression models; however, the simulations were designed so that within-cluster (L1) and between-cluster (L2) X–Y relations were identical. In instances where this is not the case, other research has shown that regression coefficient estimates of L1 predictors can be greatly distorted, to varying degrees, whenever X–Y relationships differ<sup>5</sup> at L1 and L2 (Mundlak, 1978; Scott & Holt, 1982). In this circumstance, if an L1 predictor’s slope is estimated using a regression, the estimate will be a hodgepodge of the true within- and between-cluster slopes, which we call a “blended slope” ( $\beta_{Total}$ ), expressed as follows (Raudenbush & Bryk, 2002, p. 137):

$$\beta_{Total} = (1 - \psi)\beta_W + \psi\beta_B = \beta_W + \psi(\beta_B - \beta_W) \quad (1)$$

In Equation 1,  $\beta_{Total}$  is a weighted sum of the within-cluster L1 X–Y relationship ( $\beta_W$ ) and the between-cluster L2 X–Y relationship ( $\beta_B$ ). It is further assumed that  $\beta_W$  and  $\beta_B$  are estimated as follows:

3) We also provide an applied analysis example in the [Supplementary Materials](#).

4) Irrespective of whether a unilevel regression with a SE adjustment or a multilevel regression is employed, inclusion of L1 aggregates as L2 predictors, and/or cluster-mean centering L1 predictors, will equivalently ensure L1 predictor slopes are unbiased (Antonakis et al., 2021; Bell & Jones, 2015; Enders & Tofighi, 2007). However, estimating slopes of the L2 aggregates themselves can suffer from unreliability/sampling error, especially when cluster sizes are small (e.g., Grilli & Rampichini, 2011). Given that it is unnecessary to include aggregates when the focus is on L1 relations, we focus on using cluster-mean centering. This all said, researchers must be aware of dependencies in their data to know to use these methods, including looking at the ICC in Xs as well as Y.

5) Differences in X–Y relations at L1 and L2 can only occur when non-independence is present in X.

$$\hat{\beta}_W = \frac{\sum_{j=1}^J \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_{.j})(Y_{ij} - \bar{Y}_{..})}{\sum_{j=1}^J \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_{.j})^2} \quad (2)$$

and

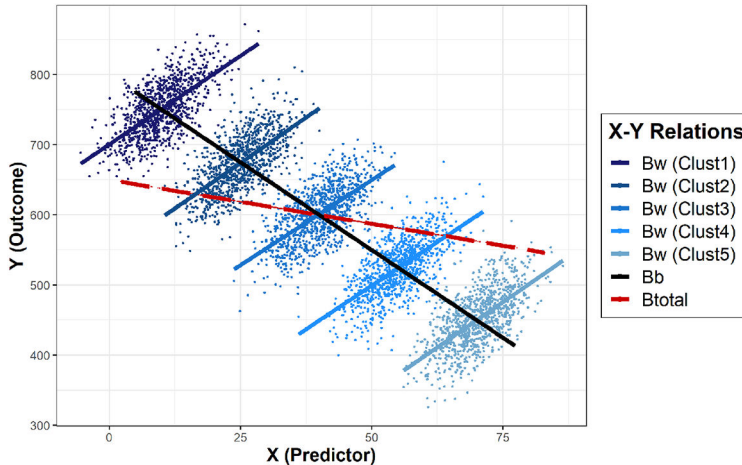
$$\hat{\beta}_B = \frac{\sum_{j=1}^J \sum_{i=1}^{n_j} (\bar{X}_{.j} - \bar{X}_{..})(Y_{ij} - \bar{Y}_{..})}{\sum_{j=1}^J \sum_{i=1}^{n_j} (\bar{X}_{.j} - \bar{X}_{..})^2} \quad (3)$$

where  $i = 1, 2, 3, \dots, n_j$  individuals within clusters, and clusters as  $j = 1, 2, 3, \dots, J$  clusters.

Inspecting Eq. 1 makes clear that greater between-cluster relations relative to within-cluster relations results in a blended slope that is larger than the true within-cluster slope, whereas the reverse is true when between-cluster relations are smaller than within-cluster relations. As such, the blended slope will be a biased estimator of the within-cluster slope whenever within- and between-cluster relationships differ (Figure 1A). By contrast, the blended slope is not problematic when the within- and the between-cluster relationships are similar (Figure 1B).

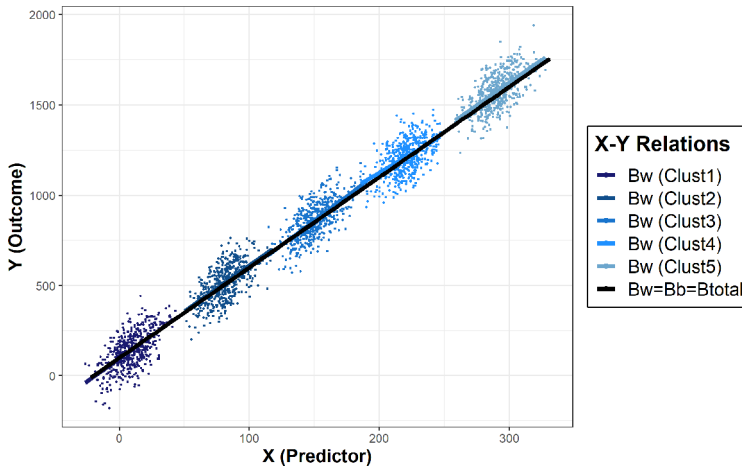
**Figure 1A**

*Hypothetical X-Y Relationships for Data with  $J = 5$  Clusters by Between-Within Relations: Within- and Between-Cluster Slopes Differ*



**Figure 1B**

*Hypothetical X-Y Relationships for Data with J = 5 Clusters by Between-Within Relations: Within- and Between-Cluster Slopes are Same*



*Note.* In Panels 1A and 1B, the blue data clouds and individual blue lines represent within-cluster (L1) X–Y slopes (Bw), and the black solid line represents the between-cluster (L2) X–Y slopes (Bb). The red dashed line in Panel A represents the “blended slope” (Btotal) that would be estimated when either the L1 predictor is not cluster-mean centered and/or when the L2 cluster means (aggregates) of the L1 predictor are not included in the model.

Importantly, the weight,  $\psi$ , in Equation 1 plays a role in the magnitude of slope bias. For unilevel regression models that ignore clustering altogether, this weight is defined by Raudenbush and Bryk (2002, p. 137; see also Snijders & Bosker, 2012, p. 31) as:

$$\psi = \eta_X^2 \tag{4}$$

where:

$$\eta_X^2 = \frac{\sum_{j=1}^J \sum_{i=1}^{n_j} (\bar{X}_{.j} - \bar{X}_{..})^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_{..})^2} \tag{5}$$

Here,  $\eta_X^2$  is the ratio of sum of squared deviations between clusters ( $SS_B$ ) to the total sum of squared deviations ( $SS_{Total}$ ) in the predictor. That is, the proportion of total variability in X explained by clustering.

For 2-level (random intercept) MLMs, Scott and Holt (1982) showed the population weight for estimating slope bias is:

$$\psi = \frac{(1 - \rho_{Residual})(\rho_X)}{(1 - \rho_{Residual}) + (m)(\rho_{Residual})(1 - \rho_X)} \quad (6)$$

Here,  $\rho_{residual}$  represents the population intraclass correlation (ICC) among residual errors due to clustering when unilevel regression is conducted,  $\rho_X$  represents the ICC among predictor scores due to clustering, and  $m$  denotes cluster size, assuming that the within-cluster sample sizes ( $n_j$ 's) are equal across clusters.

In practice, both  $\rho$ s in Equation 6 are  $\eta^2$  values (as in Equation 5), rather than ICCs. Further, the population weight in Equation 6 can be computed for samples as:

$$\psi = \frac{(1 - \hat{\eta}_{Residual}^2)(\hat{\eta}_X^2)}{(1 - \hat{\eta}_{Residual}^2) + (m)(\hat{\eta}_{Residual}^2)(1 - \hat{\eta}_X^2)} \quad (7)$$

The foregoing now leads to the connection between the blended slope weight given in Equation 4 and Equation 7. Specifically, the term involving the correlated residual error ( $\hat{\eta}_{Residual}^2$ ) can be reinterpreted as the proportion of observed correlated residual error that *would be accounted for by the modeling method chosen*. In the case of unilevel regression, correlated residual error would *not* be taken into account during model parameter estimation, whereas in MLM regression, residual clustering due to dependencies in Y are taken into account via the random intercept term. Thus, the formula for the blended slope weight in Equation 7 reduces to the unilevel regression weight shown in Equation 4 because the  $\hat{\eta}_{residual}^2$  term is zero in this context:

$$\psi_{Unilevel} = \frac{(1 - 0)(\hat{\eta}_X^2)}{(1 - 0) + (m)(0)(1 - \hat{\eta}_X^2)} \quad (8)$$

therefore:

$$\psi_{Unilevel} = \hat{\eta}_X^2 \quad (9)$$

To summarize, the two common factors that drive the slope bias in both MLMs and MLM alternatives include: 1) greater differences in the within- and between-cluster X-Y relations (i.e., L1 vs. L2 slopes), and 2) greater dependencies in the predictor (X) due to clustering ( $\hat{\eta}_X^2$ ). In addition, a careful comparison of the bias weight formulas for MLMs (Equation 7) and unilevel models (Equation 8) shows that *L1 slope bias in MLMs will be less pronounced than in unilevel models* when: (a) there are dependencies present in Y, and (b) when the data involve larger cluster sizes ( $m$ ). Specifically, in MLM regression, residual clustering due to ICCY is explicitly modeled via the random intercept term such that the  $(1 - \hat{\eta}_{Residual}^2)$  term yields a value less than 1 (depending on ICCY); in contrast, this term is always 1 in a unilevel regression. In short, as ICCY increases, for MLMs, the

numerator in Equation 7 will become smaller and the denominator will become larger, resulting in a smaller  $\psi$  weight of the difference in L1 and L2 slopes (Equation 1), all else being equal. Second, the impact of the decrease in the  $(1 - \hat{\eta}_{Residual}^2)$  term on  $\psi$  as ICCY increases (for MLMs) is amplified by the average cluster size ( $m$ ). As  $m$  increases, all other things equal, the denominator becomes larger, resulting in an even smaller  $\psi$  weight and a lower level of L1 slope bias. In short, when ICCY or  $m$  are relatively high,  $\psi$  for MLM models with uncentered L1 predictors will tend toward zero, and the blended slope reduces to just the within-cluster slope.

## Fixing the Problem

As shown in Equation 1, the “blended slope” ( $\beta_{Total}$ ) is simply a weighted linear combination of within- and between-cluster components of an L1 predictor. However, if an L1 predictor ( $X$ ) is cluster-mean (or “group-mean”) centered—as distinct from grand-mean centered—all clustering effects are removed and the predictor becomes the within-cluster component of  $X$ ,  $X_W$ . When the original predictor is aggregated into cluster means, it becomes the between-cluster component of  $X$ ,  $X_B$ , which now contains all cluster-specific information about  $X$ . Further, the separated components of  $X$  ( $X_W$  and  $X_B$ ) are uncorrelated with each other. In fact, including either just  $X_W$ , a combination of  $X$  and  $X_B$ , or both  $X_W$  and  $X_B$ , eliminates the slope bias problem, both for unilevel and MLM regression approaches. Although Enders and Tofighi (2007) and others have discussed this in the multilevel literature (Bell et al., 2019; Hamaker & Muthén, 2020; Raudenbush & Bryk, 2002, pp. 31–35; Snijders & Bosker, 2012, p. 58), in practice, researchers are still not using/reporting how they are dealing with the problem, whether it be cluster-mean centering and/or use of aggregates, even in MLMs (Antonakis et al., 2021; Luo et al., 2021). More importantly, we find very rare discussion regarding the blended slope problem (or how to solve it) within the methodological literature on CR-SEs as an alternative approach to MLMs for modeling clustered data (exceptions include Bell et al., 2019, and McNeish, 2019).

## Correlated Residual Error Due to X (Also Known as Endogeneity)

Putting aside the blended slope problem, it is important to note that dependencies in L1 (within-cluster) predictors, not just outcomes, can cause correlated residual errors (Cronbach, 1976; Mundlak, 1978), which is also known as the endogeneity problem (e.g., Bell & Jones, 2015; Grilli & Rampichini, 2011). *Even when there are no dependencies in Y*, residual error dependencies can be induced by dependencies in X via the shared relationship between X and Y. This may occur, for example, in outcomes that have more biologically based etiologies such as information processing, memory, or attention (Fiedorowicz et al., 2001), which are unlikely to vary across L2 units such as therapists or schools. Yet, individualized (L1) treatment characteristics or interventions used as

predictors would likely vary across L2 units (e.g., differences among provider techniques or school policies).

Specifically, the magnitude of cluster dependency in X ( $\hat{\eta}_X^2$ ) will be directly transferred to dependencies in the predicted values of the outcome, since  $\hat{Y} = X\hat{\beta}$ . Thus, the magnitude of the residual error dependencies,  $\hat{\eta}_{Residual}^2$ , will be a function of  $\hat{\eta}_X^2$  and the proportion of variance in Y explained by X,  $\hat{\eta}_{Y.X}^2$ . Of course, it is more complex in typical analyses where there are many predictors with differing dependency levels and relationships with Y and each other. However, as noted above, when L1 predictor cluster-mean centering is used, and/or L2 aggregates are added, dependencies in Xs will be removed such that dependencies in Y associated with dependencies in Xs will also be removed.

## Monte Carlo Simulation

To illustrate how the blended slope problem (and associated standard error distortion) manifests in both multilevel/random effects models (MLMs) and unilevel/fixed effects models (with and without cluster-robust standard errors), we conducted a Monte Carlo simulation (an applied analysis example is featured in the [Supplementary Materials](#)). Our results will show that ignoring cluster dependencies in X when there are different Level 1 (L1) and Level 2 (L2) X–Y relations yields biased slopes and distorted standard errors, especially for unilevel models. Further, because unilevel regression with CR-SEs only adjusts standard errors for residual error clustering, the bias in the slope coefficient will still be identical to that of a regression model that does not employ CR-SEs (although models using CR-SEs have more accurate standard errors). Last, we also demonstrate how simple cluster-mean centering L1 predictors resolves the slope bias problem.

## Data Generation

Data were generated using *Mplus 8* (Muthén & Muthén, 2017) within the *MplusAutomation* package in *R* (Hallquist & Wiley, 2018) as 2-level, 2-predictor models in which Y and both Xs are continuous. We varied seven crossed factors, with 500 replications per cell. Factor levels were chosen to capture null conditions in addition to realistic research scenarios to maximize generalizable results. Manipulated conditions were:

1.  $J$  number of clusters (30, 100).
2.  $m$  cluster sizes (5, 100).
3. Intraclass correlation (ICC) in X (.01, .20, .40).
4. ICC in Y (.01, .20, .40).
5. Within-cluster total  $R_{y.x}^2$  (.00, .20, .50), with each predictor contributing half the proportion each, and with positive and negative slopes for non-null conditions.



6. between-cluster total  $R^2_{y \cdot x}$  (.00, .20, .50), with each predictor contributing half the proportion each, and with positive and negative slopes for non-null conditions.
7. predictor-predictor correlations at each level  $r_{xx}$  (.00, .20, .40).

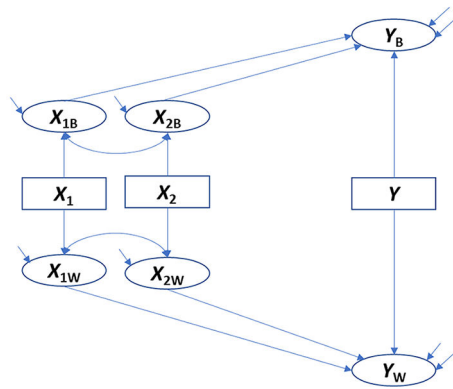
**Intraclass Correlations**

ICCs were induced by setting specific values at L2 for the variance of X and the residual variance of Y, each calculated a priori for every design cell. Specifically, we assumed predictors had separable within- and between-cluster effects on the outcome using latent disaggregation (Figure 2), as follows:

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{1ij} - \mu_{1,j}) + \gamma_{01}(\mu_{1,j}) + \gamma_{20}(X_{2ij} - \mu_{2,j}) + \gamma_{02}(\mu_{2,j}) + u_{0j} + e_{ij} \quad (10)$$

**Figure 2**

2-Level, 2-Predictor Model With Predictors (Xs) and Outcome (Y) Measured at L1



where the  $i^{th}$  score in the  $j^{th}$  cluster is a function of the sum of the conditional mean ( $\gamma_{00}$ ), the within-cluster (cluster-mean centered, L1) and between-cluster (aggregate, L2) fixed effects of the first predictor ( $\gamma_{10}$  and  $\gamma_{01}$ , respectively), the within- and between-cluster fixed effects of the second predictor ( $\gamma_{20}$  and  $\gamma_{02}$ , respectively), the deviation between the cluster mean and the grand mean intercept ( $U_{0j}$ ), and the deviation between the  $i^{th}$  score and its  $j^{th}$  cluster mean ( $e_{ij}$ ).

Each of the L1 variables were fixed to have unit variances ( $\sigma^2_W$ ), and L2 variances ( $\sigma^2_B$ ) were set to achieve a desired ICC using the Jak (2019) approach:

$$\sigma^2_B = \frac{ICC (\sigma^2_W)}{1 - ICC} \quad (11)$$

The residual outcome variances were calculated from a desired  $R^2_{y \cdot x}$  value:

$$\sigma_{Y \text{ residual}, W}^2 = \sigma_{Y, W}^2 - (R_W^2 * \sigma_{Y, W}^2) \quad (12)$$

$$\sigma_{Y \text{ residual}, B}^2 = \sigma_{Y, B}^2 - (R_B^2 * \sigma_{Y, B}^2) \quad (13)$$

Given the data generation process, we could not set X or Y ICCs to be exactly zero; instead, our null ICCs were set to .01.

### Regression Coefficients

Regression coefficients were derived for each combination of  $R_{y,x}^2$  and  $r_{xx}$  conditions using eigenvalue decomposition (e.g., Pedhazur, 1997). Both predictors were assumed to have the same population slope values; parameter values are provided in the [Supplementary Materials](#).

### Analyses

For each condition, 500 datasets were simulated and saved using *Mplus*' Monte Carlo utility (via *MplusAutomation* in *R*) assuming a 2-level, 2-predictor linear latent variable model shown in [Equation 10](#) ([Figure 2](#)). Each dataset was analyzed using six approaches<sup>6</sup>:

1. Unilevel linear regression model without standard error adjustment, *predictors un-centered* (L1uc).
2. Unilevel linear regression model with cluster-robust standard errors (CR-SEs), *predictors un-centered* (L1CRuc).
3. Multilevel linear regression (random intercept model), *predictors un-centered* (L2uc).
4. Unilevel linear regression model without standard error adjustment, *predictors cluster-mean centered* (aggregate predictors not included) (L1c).
5. Unilevel linear regression model with cluster-robust standard errors (CR-SEs), *predictors cluster-mean centered* (aggregate predictors not included) (L1CRc).
6. Multilevel linear regression (random intercept model), and *predictors cluster-mean centered* (aggregate latent predictors were included in this analysis because it was the population generation model with which to compare with the other analyses; nevertheless, aggregates are uncorrelated with cluster-mean centered predictors and thus L1 slopes and their standard errors are unaffected by their inclusion) (L2c).

In *Mplus*, TYPE = GENERAL was employed for analysis approaches (1) and (4); for approaches (2) and (5) we used TYPE = COMPLEX, and for approaches (3) and (6) we

6) We did not include unilevel models with L2 aggregate predictors as an alternative to L1 predictor cluster-mean centering because L2 aggregate slope estimates can suffer from unreliability (e.g., Grilli & Rampichini, 2011), and in unilevel models that do not adjust SEs for cluster dependencies, L2 aggregate slope standard errors become downwardly biased.

used TYPE = TWOLEVEL. Maximum likelihood estimation was used for all analyses. Example code files are provided in the [Supplementary Materials](#).

## Results

For each analysis approach and condition, slope coefficient and standard error estimates were averaged across the 500 replications. Evaluation of the correlation between the results for the two predictors across all analysis approaches and conditions showed that the two estimated slopes and their standard errors were correlated at .99 or higher; thus, without loss of generality we report slope parameter estimate and standard error bias<sup>7</sup> for the first predictor throughout.

### Slope Coefficient Relative Bias

Relative bias for the L1/within-cluster predictor slope estimates was tabulated for non-null conditions in which 10% or 25% variance in Y was explained by each L1/within-cluster predictor in isolation (20% and 50% total variance explained jointly by both within-cluster predictors). Specifically, relative bias was computed as the difference between the estimated slope and the true within-cluster slope, divided by the true within-cluster slope; a relative bias of zero indicates no bias, whereas positive values indicate overestimation and negative values indicate underestimation. (Hoogland & Boomsma, 1998 suggest that relative bias exceeding  $\pm 5\%$  for point estimates is cause for concern.)

Table 1 displays mean relative bias across conditions for each of the six analyses, paneled by X and Y ICCs. As can be observed, for un-centered predictor analyses (top three rows), the unilevel approaches exhibited identical bias, with increased bias solely a reflection of increased dependencies in X. For MLM with un-centered predictors, the bias due to dependencies in X was much lower than in the unilevel approaches. Not surprisingly, when we visualize the bias by the true differences in the within- and between-cluster slopes (Figure 3), we see that: 1) there is no bias when the ICC in X is close to zero, 2) there is no bias when cluster-mean centering is used, and, 3) for un-centered predictor analyses, the magnitude of bias (underestimation of the within-cluster slope) increases markedly as the ICC X increases and also as the differences between L1 and L2 slopes increases (identically for positive or negative differences). As can also be observed in both Figure 3 and Figure 4A/Figure 4B (which collapses across slope difference conditions), the bias in the un-centered MLM slope disappears for larger clusters whereas the

---

7) We focus here on bias in slope point estimates and their standard errors; Type I error and power, which depend jointly on point estimates and their standard errors, are given in the [Supplementary Materials](#). We do note that the Type I error rate was unacceptably inflated for unilevel models that did not use cluster-mean centering—even those employing cluster-robust standard errors.

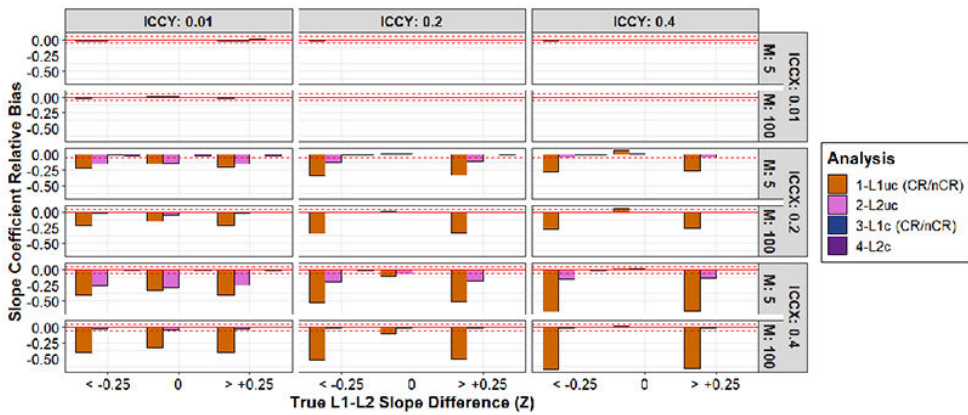
bias in the un-centered unilevel analyses is large and constant across cluster sizes (Figure 4A/Figure 4B).

Across all of these results, the relative bias we observed was negative (underestimation) because our simulation design assumed the within-cluster predictor had a greater relationship with the outcome than the between-cluster aggregate predictor. Had we designed the simulation to have greater between-cluster (L2) predictor-outcome relations than within-cluster (L1) relations, bias would be in a positive direction (overestimation).

Last but not least, slope bias was completely removed for MLM and unilevel regression analysis approaches by cluster-mean centering the predictors.

**Figure 3**

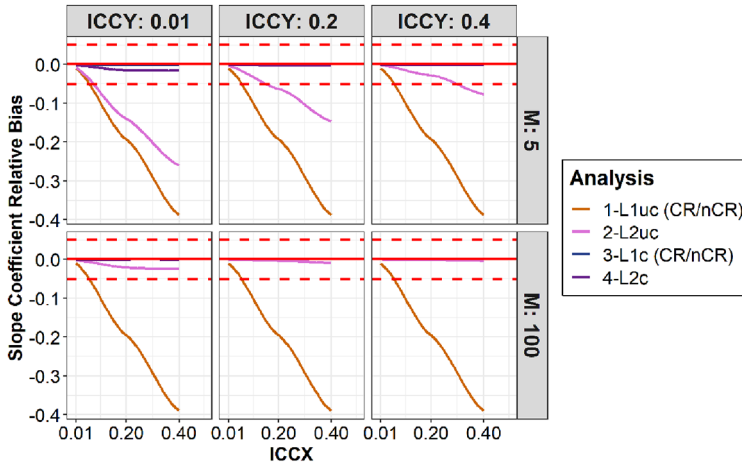
*Slope Coefficient Relative Bias for Analyses by Within- and Between-Cluster (L1-L2) Slope Differences when  $r_{xx} = 0$*



*Note.*  $N = 40$  non-null conditions/cell with X-X correlations = 0 (for ease of interpretation), per analysis, 500 replications/condition. Horizontal axis reflects standardized L1-L2 true slope differences grouped as: < -0.25, -0.25 to +0.25, and > +0.25 standard deviations in differences. L1uc (CR/nCR) = unilevel regression, un-centered predictors, with and without cluster-robust standard errors; L2uc = multilevel random intercept model, un-centered predictors; L1c (CR/nCR) = unilevel regression, cluster-mean centered predictors, with and without cluster-robust standard errors; L2c = multilevel random intercept model, cluster-mean centered predictors. Solid red line delineates 0 = no bias; values within  $0 \pm .05$  (red dashed lines) are considered acceptable.

**Figure 4A**

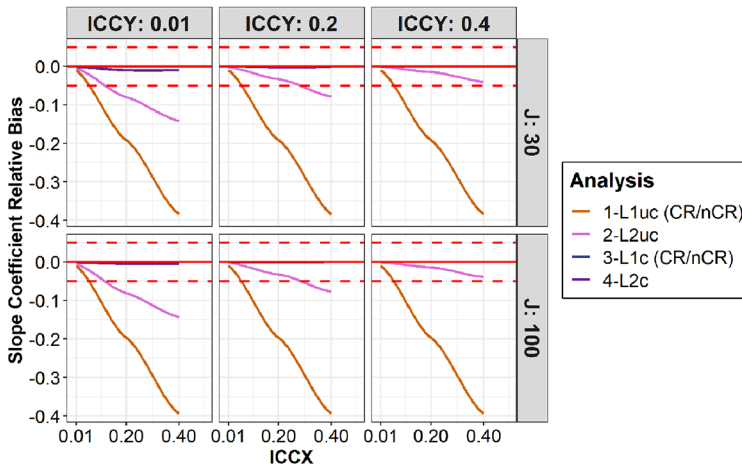
*Slope Coefficient Relative Bias for Analyses by Cluster Size: ICCs by Cluster Size*



Note. See note to Figure 4B.

**Figure 4B**

*Slope Coefficient Relative Bias for Analyses by Number of Clusters: ICCs by Number of Clusters*



Note.  $N = 360$  non-null conditions/cell, per analysis, 500 replications/condition. L1uc (CR/nCR) = unilevel regression, un-centered predictors, with and without cluster-robust standard errors; L2uc = multilevel random intercept model, un-centered predictors; L1c (CR/nCR) = unilevel regression, cluster-mean centered predictors, with and without cluster-robust standard errors; L2c = multilevel random intercept model, cluster-mean centered predictors. Solid red line delineates 0 = no bias; values within  $0 \pm .05$  (red dashed lines) are considered acceptable.

**Table 1**  
*Within-Cluster (L1) Slope Coefficient Relative Bias for Analysis Approach by Y and X ICC Combinations*

Analysis	ICC Y = .01			ICC Y = .20			ICC Y = .40											
	ICC X = .01		ICC X = .20	ICC X = .01		ICC X = .20	ICC X = .40		ICC X = .20	ICC X = .40								
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)								
<b>Un-centered Predictor</b>																		
1L	-.01	(.01)	-.19	(.04)	-.39	(.05)	-.01	(.05)	-.19	(.19)	-.39	(.24)	-.01	(.08)	-.19	(.32)	-.39	(.39)
1L-CRSE	-.01	(.01)	-.19	(.04)	-.39	(.05)	-.01	(.05)	-.19	(.19)	-.39	(.24)	-.01	(.08)	-.19	(.32)	-.39	(.39)
2L	-.01	(.01)	-.08	(.06)	-.14	(.13)	.00	(.02)	-.03	(.06)	-.08	(.09)	.00	(.01)	-.01	(.06)	-.04	(.08)
<b>Cluster-Mean Centered Predictor</b>																		
1L	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)
1L-CRSE	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)
2L	.00	(.00)	-.01	(.01)	-.01	(.01)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)	.00	(.00)

Note. N = 240 non-null conditions/cell. 500 replications/condition. 1L = unlevel regression; 1L-CRSE = unlevel regression with cluster-robust standard errors; 2L = multilevel random intercept model. Values closer to 0 are better; values within 0±.05 are considered acceptable.

## Slope Standard Error Empirical Bias

Because standard errors are unique to each condition and analysis approach, we computed empirical bias as the ratio of the mean of the model-implied standard errors to the expected standard error, with the expected standard error based on the empirical standard deviation of the slope parameters (e.g., Bell & Jones, 2015). Values of 1 indicate no bias, positive values indicate upward bias, and negative values indicate downward bias (Hoogland & Boomsma, 1998 suggest that relative bias exceeding  $\pm 10\%$  (below 0.9 or above 1.1) for standard errors is cause for concern).

Table 2 summarizes empirical bias for the mean slope standard error estimates. Across analyses, MLM and unilevel CR-SE approaches showed reasonable slope standard errors. Only the unilevel approach without adjustment exhibited poor standard errors: when no centering was used, the standard errors were biased downward as a function of both X and Y ICCs (Figure 5A/Figure 5B); when cluster-mean centering was used, the bias from X ICC vanished (but not bias from Y ICC).

## Discussion

The present study demonstrates how the “blended slope” problem manifests differentially in unilevel and multilevel models (MLMs) for varied levels of within- and between-cluster (L1 and L2) slope values, and varied levels of dependencies in predictors (Xs) and outcomes (Ys). Our results show clearly that bias can occur in either analysis when clustering in X is ignored. What both our mathematical review and simulation results also reveal is that ignoring dependencies in predictors in MLMs is less pronounced than in unilevel models as cluster sizes increase, and as dependencies in Y increase. We consider these findings in the context of three likely data analytic circumstances that researchers might currently be implementing: 1) when a lack of dependency in Y is the decision rule for ignoring clustering altogether; 2) when a unilevel model with adjusted standard errors is used, but the researcher unwittingly uses grand-mean centering or no centering for their L1 predictors; and 3) when a multilevel model (MLM) is used but again, the researcher is unaware that there may be differences in L1 and L2 X-Y relations and fails to cluster-mean center their L1 predictors (or add L2 aggregates).

**Table 2**  
*Within-Cluster (L1) Slope Standard Error Empirical Bias for Analysis Approach by Y and X ICC Combinations*

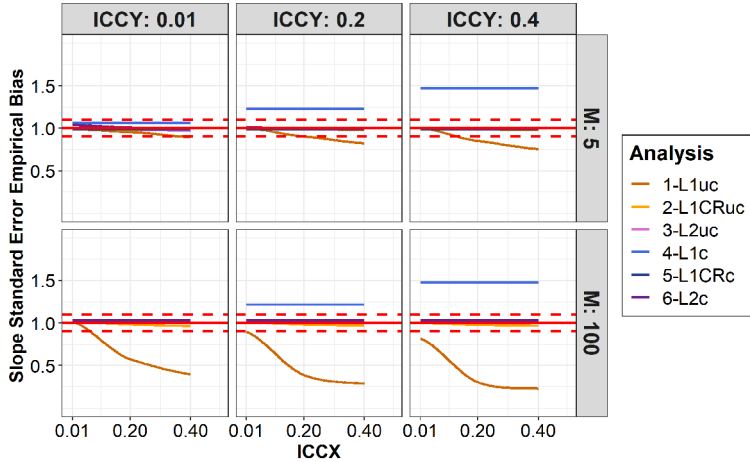
Analysis	ICC Y = .01						ICC Y = .20						ICC Y = .40					
	ICC X = .01		ICC X = .20		ICC X = .40		ICC X = .01		ICC X = .20		ICC X = .40		ICC X = .01		ICC X = .20		ICC X = .40	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
<b>Un-centered Predictor</b>																		
1L	1.01	(.03)	.77	(.21)	.65	(.26)	.95	(.06)	.65	(.26)	.56	(.27)	.91	(.10)	.58	(.28)	.49	(.27)
1L-CRSE	1.01	(.03)	.98	(.03)	.97	(.03)	1.01	(.03)	.99	(.03)	.98	(.03)	1.00	(.03)	.98	(.03)	.97	(.03)
2L	1.01	(.03)	1.00	(.03)	.99	(.04)	1.01	(.03)	1.01	(.03)	1.00	(.03)	1.01	(.02)	1.01	(.02)	1.01	(.03)
<b>Cluster-Mean Centered Predictor</b>																		
1L	1.05	(.03)	1.05	(.03)	1.05	(.03)	1.22	(.06)	1.22	(.06)	1.22	(.06)	1.47	(.10)	1.47	(.10)	1.47	(.10)
1L-CRSE	1.01	(.03)	1.01	(.03)	1.01	(.03)	1.01	(.03)	1.01	(.03)	1.01	(.03)	1.01	(.03)	1.01	(.03)	1.01	(.03)
2L	1.03	(.02)	1.02	(.02)	1.02	(.02)	1.02	(.02)	1.01	(.02)	1.01	(.02)	1.02	(.02)	1.01	(.02)	1.01	(.02)

*Note.* N = 240 non-null conditions/cell, per analysis, and 500 replications/condition. 1L = unilevel regression; 1L-CRSE = unilevel regression with cluster-robust standard errors; 2L = multilevel random intercept model. Values closer to 1 are better; values within 1 ± .10 are considered acceptable.



Figure 5A

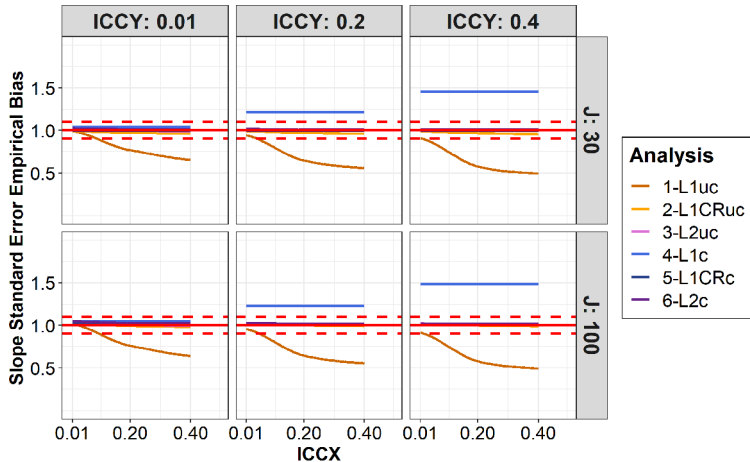
Standard Error Empirical Bias for Analyses by Cluster Size: ICCs by Cluster Size



Note. See note to Figure 5B.

Figure 5B

Standard Error Empirical Bias for Analyses by Number of Clusters: ICCs by Number of Clusters



Note.  $N = 360$  non-null conditions/cell, per analysis, 500 replications/analysis. L1uc (CR/nCR) = unilevel regression, un-centered predictors, with and without cluster-robust standard errors; L2uc = multilevel random intercept model, un-centered predictors; L1c (CR/nCR) = unilevel regression, cluster-mean centered predictors, with and without cluster-robust standard errors; L2c = multilevel random intercept model, cluster-mean centered predictors. Solid red line delineates 1 = no bias; values within  $1 \pm .10$  (red dashed lines) are considered acceptable.

In the first circumstance, dependencies in X will cause distorted slope coefficient estimates and distorted slope standard error estimates. With respect to coefficient distortion, the predictor slope will be underestimated when the true within-cluster (L1) X-Y relation is larger than the true between-cluster (L2) X-Y relation, which is more likely to be the case when the theoretical focus is on individual differences. In other words, researchers will erroneously infer that the X-Y relationship effect size is smaller than it is in truth. In our Monte Carlo simulation design, the within-cluster relation was designed to be stronger than the between-cluster relation and hence under-estimation was observed as the difference in within- and between-cluster slopes increased (Table 1). Irrespective of slope bias, when clustering is completely ignored using unilevel regression, the slope standard errors will be downwardly biased (assuming a positive ICC). In our simulations, the standard error was underestimated by 35% when X ICC = .40 and Y ICC = .01, all else constant (Table 2). When both X and Y ICC = .40, the standard errors were underestimated by 51%. As is well known, downward standard error bias leads to a false sense of confidence in hypothesis tests and confidence interval coverage.

In the second analytic circumstance, when a researcher chooses to use unilevel regression with CR-SEs as a means for handling correlated residual errors but is unaware of the need to cluster-mean center their L1 predictors (or add L2 aggregates; e.g., Bell et al., 2019), the L1 slope coefficient bias will still be identical to that for the unilevel regression without standard error adjustment (Table 1). This is because the adjustment occurs after coefficients are estimated. As observed in our simulations, the standard errors for the CR-SE approach were similar to those for MLM (Table 2).

In the third analytic circumstance, when a researcher uses MLM to analyze their data but fails to cluster-mean center L1 predictors (or add L2 aggregates; e.g., Antonakis et al., 2021), the L1 slope bias will exhibit the same directional pattern as the other two analysis methods, but the bias will likely be smaller for two reasons. First, the MLM regression coefficient estimation directly takes into account cluster differences in the residuals (see again Equation 7). Second, as the average cluster size becomes large, especially as dependencies in Y increase, the weight for L1-L2 slope differences tends toward zero and the blended slope becomes the within-cluster slope (see again Figure 4A). In our simulations, the bias in MLM slopes when failing to cluster-mean center predictors was much lower than that of bias in unilevel slopes that did not use cluster-mean centered predictors (-14% compared to -39% on average when ICC X = .40 and ICC Y = .01; Table 1).

## Why is There a Discrepancy With Earlier Work?

Applied researchers may be surprised to learn that L1 predictor slopes estimated with unilevel regression and MLM are not the same. This is because recent methodological work showed that both analysis approaches could produce the same coefficient estimates (McNeish, 2014). However, this is only the case when data are generated to have the

same within- and between-cluster X-Y relationships (Figure 1B). As we review and demonstrate in this paper, L1 slope bias arises only when there is a difference in the L1 and L2 X-Y relations (Figure 1A), which in practice may be far more realistic than assuming they are the same.

## Isn't Everyone Already Using Cluster-Mean Centering?

Some readers may be under the impression that everyone is appropriately cluster-mean centering lower-level predictors (or including their aggregates) when analyzing clustered data. We do not believe this is the case. Although the centering issue has been discussed in prominent MLM textbooks (e.g., Raudenbush & Bryk, 2002; Snijders & Bosker, 2012) as well as MLM methodological literature (e.g., Enders & Tofighi, 2007), two recent meta-analyses of multilevel modeling research by Luo et al. (2021) and Antonakis et al. (2021) showed that very few researchers reported either cluster-mean centering their L1 predictors and/or including L2 aggregate predictors to ensure slope results were unbiased. More importantly, we found little guidance in the MLM alternatives literature (unilevel models for analyzing clustered data structures) about the issues involved, particularly for unilevel regression with CR-SEs (exception is Bell et al., 2019). Thus, we remain skeptical that researchers have received the message yet.

## Two Caveats in Simulation Conditions

Besides the specificity of our design conditions, our simulation results were limited in two regards. First, we used *latent* disaggregation of lower-level regressors into within- and between-cluster (L1 and L2) predictors in our correctly specified multilevel models. Latent disaggregation mitigates against score unreliability that can result from sampling error when estimating X-Y relations at L1 and L2 (Asparouhov & Muthén, 2019). However, we did not build in extra measurement error, and so L1 estimates would not differ from L1 estimates in univariate multilevel software, such as *R* lme4 (Bates et al., 2015). However, L2 aggregates themselves are automatically subject to sampling error (Grilli & Rampichini, 2011), and so one would expect the difference between latent and univariate multilevel models to be located in the L2 aggregate slope estimates (which are not the focus of our paper). Indeed, when we subjected the simulated data to 2-level models estimated with *R* lme4 (with observed cluster means used for L1 predictor cluster-mean centering and creating L2 aggregates), we found the L1 slope estimates to be nearly perfectly correlated between *Mplus* and lme4 results,  $r > .99$  (within and across all sample sizes) but the L2 aggregate slopes were less correlated, at  $r \geq .86$  across sample sizes ( $r \geq .81$  for the smallest sample size).

The second point to note is that we used continuous (normally distributed) predictors and outcomes in our simulation, as well as in our applied analysis provided in the [Supplementary Materials](#). Nevertheless, there is no reason to believe the same biases

would not persist with categorical data. In fact, categorical predictors can also be either cluster-mean centered (Enders & Tofighi, 2007, Appendix B), and/or aggregated as L2 cluster mean (percentages) that are then added to the model to ensure that slope inferences are correct.

## Conclusion

Although MLM has become a staple analytic technique in the psychological and educational sciences (Antonakis et al., 2021; Bauer & Sterba, 2011; Luo et al., 2021), alternative unilevel regression approaches for analyzing clustered data when only L1 questions are of interest are also possible—and sometimes preferable—to MLM, such as unilevel regression with cluster-robust standard errors (CR-SEs) (Hayes & Cai, 2007; Huang, 2016; McNeish et al., 2017). Yet, in any clustered data, the predictor-outcome relation at L1 may easily differ from that of the relation at L2. If lower-level predictors are not cluster-mean centered, or if their aggregates are not added as predictors to the model, the L1 slope estimate will be an unintelligible mixture of the L1 and L2 components. The current study reviews and illustrates this “blended slope” problem, and its differential impact on unilevel models compared to MLMs, when left unchecked. Importantly, the problem only occurs when clustering in X is ignored, resulting in L1 slope under- or over-estimation that is especially pronounced in unilevel regression models, and to a lesser extent, in MLMs with smaller cluster sizes.

In our view, there is no utility in trying to draw inferences from a “blended slope,” as it serves neither the individual level nor the cluster level; further, it is prone to inflated Type I error rates<sup>8</sup> in both unilevel and multilevel analyses. Simply put, researchers should always check and report both X and Y ICCs when they are dealing with clustered data, regardless of whether the nature of their data is cross-sectional or repeated measures/time series. Second, given that most datasets employ numerous predictors that vary in their magnitude of cluster dependencies as well as their correlations with the outcome and each other, it should be a *routine procedure* for researchers to either cluster-mean center their lower-level predictors and/or add cluster mean aggregates as additional predictors (see [Supplementary Materials](#) for R code to obtain cluster-mean centered and aggregate predictors), irrespective of whether unilevel or multilevel modeling is used to handle residual error dependencies.

---

8) See the [Supplementary Materials](#) for Type I error rates.

---

**Funding:** The authors have no funding to report.

---

**Competing Interests:** We have no known conflict of interest to disclose.

---

**Author Contributions:** E.A.S. and T.R.K. conceptualized the study and wrote the paper; E.A.S. wrote code for the Monte Carlo simulations and secondary analysis of simulation results.

---

**Data Availability:** Data is freely available at [Supplementary Materials](#).

---

## Supplementary Materials

For this article, an applied data analysis example, parameter values, example code files, Monte Carlo simulation Type I error and power results, and R codes for predictor cluster-mean centering and aggregating are available via PsychArchives (for access see [Index of Supplementary Materials](#) below):

### Index of Supplementary Materials

Sanders, E. A., & Konold, T. R. (2023). *Supplementary materials to "X matters too: How the blended slope problem manifests differently in unilevel vs. multilevel models"* [Data analysis example, parameter values, Mplus code files, Monte Carlo error and power results, R code for predictor cluster-mean centering and aggregating]. *PsychOpen GOLD*.  
<https://doi.org/10.23668/psycharchives.12593>

## References

- Antonakis, J., Bastardo, N., & Rönkkö, M. (2021). On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations. *Organizational Research Methods, 24*(2), 443–483. <https://doi.org/10.1177/1094428119877457>
- Asparouhov, T., & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling, 26*(1), 119–142. <https://doi.org/10.1080/10705511.2018.1511375>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bauer, D. J., & Sterba, S. K. (2011). Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation. *Psychological Methods, 16*(4), 373–390. <https://doi.org/10.1037/a0025813>
- Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity, 53*(2), 1051–1074. <https://doi.org/10.1007/s11135-018-0802-x>

- Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effect modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1), 133–153.  
<https://doi.org/10.1017/psrm.2014.7>
- Cronbach, L. J. (1976). *Research on classrooms and schools: Formulations of questions, design and analysis*. Stanford Evaluation Consortium.
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121–138.  
<https://doi.org/10.1037/1082-989X.12.2.121>
- Fiedorowicz, C., Benezra, E., MacDonald, W., McElgunn, B., Wilson, A., & Kaplan, B. (2001). Neurobiological basis of learning disabilities: An update. *Learning Disabilities*, 11(2), 61–74.
- Grilli, L., & Rampichini, C. (2011). The role of sample cluster means in multilevel models: A view on endogeneity and measurement error issues. *Methodology*, 7(4), 121–133.  
<https://doi.org/10.1027/1614-2241/a000030>
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling*, 25(4), 621–638.  
<https://doi.org/10.1080/10705511.2017.1402334>
- Hamaker, E. L., & Muthén, B. (2020). The fixed versus random effects debate and how it relates to centering in multilevel modeling. *Psychological Methods*, 25(3), 365–379.  
<https://doi.org/10.1037/met0000239>
- Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, 39(4), 709–722. <https://doi.org/10.3758/BF03192961>
- Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, 26(3), 329–367.  
<https://doi.org/10.1177/0049124198026003003>
- Huang, F. L. (2016). Alternatives to multilevel modeling for the analysis of clustered data. *Journal of Experimental Education*, 84(1), 175–196. <https://doi.org/10.1080/00220973.2014.952397>
- Jak, S. (2019). Cross-level invariance in multilevel factor models. *Structural Equation Modeling*, 26(4), 607–622. <https://doi.org/10.1080/10705511.2018.1534205>
- Lai, M. H. C., & Kwok, O. (2015). Examining the rule of thumb of not using multilevel modeling: The “design effect smaller than two” rule. *Journal of Experimental Education*, 83(3), 423–438.  
<https://doi.org/10.1080/00220973.2014.907229>
- Luo, W., Li, H., Baek, E., Chen, S., Lam, K. H., & Semma, B. (2021). Reporting practice in multilevel modeling: A revisit after 10 years. *Review of Educational Research*, 91(3), 311–355.  
<https://doi.org/10.3102/0034654321991229>
- McNeish, D. (2019). Effect partitioning in cross-sectionally clustered data without multilevel models. *Multivariate Behavioral Research*, 54(6), 906–925.  
<https://doi.org/10.1080/00273171.2019.1602504>

- McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114–140.  
<https://doi.org/10.1037/met0000078>
- McNeish, D. M. (2014). Analyzing clustered data with OLS regression: The effect of a hierarchical design. *Multiple Linear Regression Viewpoints, 40*(1), 11–16.  
[http://www.glmj.org/archives/articles/McNeish\\_v40n1.pdf](http://www.glmj.org/archives/articles/McNeish_v40n1.pdf)
- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica, 46*(1), 69–85.  
<https://doi.org/10.2307/1913646>
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. Thompson Learning.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2nd ed.). SAGE.
- Scott, A. J., & Holt, D. (1982). The effect of two-stage sampling on ordinary least squares methods. *Journal of the American Statistical Association, 77*(380), 848–854.  
<https://doi.org/10.1080/01621459.1982.10477897>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis* (2nd ed.). SAGE.



*Methodology* is the official journal of the European Association of Methodology (EAM).



leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.