Health and Place 52 (2018) 121–126

Contents lists available at ScienceDirect

Health & Place

journal homepage: www.elsevier.com/locate/healthplace

Does an uneven sample size distribution across settings matter in cross-classified multilevel modeling? Results of a simulation study

Carly E. Milliren[a,b,*], Clare R. Evans[c], Tracy K. Richmond[b,d], Erin C. Dunn[e,f,g]

[a] Center for Applied Pediatric Quality Analytics, Boston Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA
[b] Division of Adolescent/Young Adult Medicine, Boston Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA
[c] Department of Sociology, University of Oregon, 736 PLC 1291, Eugene, OR 97403, USA
[d] Department of Pediatrics, Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA
[e] Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA
[f] Department of Psychiatry, Harvard Medical School, 401 Park Drive, Boston, MA 02215, USA
[g] Stanley Center for Psychiatric Research, The Broad Institute of Harvard and MIT, 75 Ames Street, Cambridge, MA 02142, USA

ARTICLE INFO

Keywords:
Cross-classified multilevel modeling
Contextual effects
Simulation
Schools
Neighborhoods

ABSTRACT

Background: Recent advances in multilevel modeling allow for modeling non-hierarchical levels (e.g., youth in non-nested schools and neighborhoods) using cross-classified multilevel models (CCMM). Current practice is to cluster samples from one context (e.g., schools) and utilize the observations however they are distributed from the second context (e.g., neighborhoods). However, it is unknown whether an uneven distribution of sample size across these contexts leads to incorrect estimates of random effects in CCMMs.
Methods: Using the school and neighborhood data structure in Add Health, we examined the effect of neighborhood sample size imbalance on the estimation of variance parameters in models predicting BMI. We differentially assigned students from a given school to neighborhoods within that school's catchment area using three scenarios of (im)balance. 1000 random datasets were simulated for each of five combinations of school- and neighborhood-level variance and imbalance scenarios, for a total of 15,000 simulated data sets. For each simulation, we calculated 95% CIs for the variance parameters to determine whether the true simulated variance fell within the interval.
Results: Across all simulations, the “true” school and neighborhood variance parameters were estimated 93–96% of the time. Only 5% of models failed to capture neighborhood variance; 6% failed to capture school variance.
Conclusions: These results suggest that there is no systematic bias in the ability of CCMM to capture the true variance parameters regardless of the distribution of students across neighborhoods. Ongoing efforts to use CCMM are warranted and can proceed without concern for the sample imbalance across contexts.

1. Introduction

Multilevel modeling (MLM) has become a staple of social science and public health research, allowing researchers to examine macro-level contextual effects across multiple settings, including students within schools (Munoz and Chang, 2007; Kim and McCarthy, 2006; Sellstrom and Bremberg, 2006), residents within neighborhoods (Tendulkar et al., 2010; Pickett and Pearl, 2001; Leventhal and Brooks-Gunn, 2000), and patients within hospitals (Rice and Alastair 1996). In MLM, both fixed and random effects account for the clustering of individuals within context, while also generating effect estimates for the contexts themselves (Diez-Roux, 2000). For more than two decades, studies using MLM have demonstrated that contexts are important determinants of health and behavior, even after accounting for individual characteristics and composition.

Recently, MLM researchers have begun to recognize the importance of considering multiple contexts *simultaneously*. For instance, there is growing interest in cross-classified multilevel modeling (CCMM) (Goldstein, 1994; Rabash and Browne, 2001), which allows researchers to examine instances when individuals are nested in non-hierarchical contexts, such as when students attending the same school live in different neighborhoods and conversely when students from the same neighborhood attend different schools. To date, CCMM has been used to examine the impact of schools and neighborhoods on a variety of health and behavioral outcomes (Dunn et al., 2015a, 2016, 2017, 2015b; Townsend et al., 2012; De Clercq et al., 2014; Evans et al., 2016), as

* Corresponding author at: Center for Applied Pediatric Quality Analytics, Boston Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA.
E-mail addresses: carly.milliren@childrens.harvard.edu (C.E. Milliren), cevans@uoregon.edu (C.R. Evans), tracy.richmond@childrens.harvard.edu (T.K. Richmond), edunn2@mgh.harvard.edu (E.C. Dunn).

https://doi.org/10.1016/j.healthplace.2018.05.009
Received 22 December 2017; Received in revised form 4 May 2018; Accepted 23 May 2018
1353-8292/

well as contextual effects of classrooms and teachers on educational outcomes (Heck, 2009; Kim et al., 2010). A major advantage of CCMM relative to MLM is that it enables researchers to avoid the "omitted context bias", wherein variance in a random effects model is misattributed from the missing context to the included context, as the included context "soaks-up" the effect of the missing context (Dunn et al., 2015b; Evans et al., 2016).

Sample size requirements for MLM are well established (Dedrick et al., 2009; McNeish and Stapleton, 2014). To generate unbiased estimates of random effects variance parameters, methodologists recommend between 5 and 20 lower level units (e.g., students) as a minimum for each higher level unit (e.g., school) (McNeish and Stapleton, 2014). Including some schools with a smaller sample size in the data set is not problematic, however, because estimates for contexts with small sample size are automatically down-weighted in MLM estimation. Thus, *most* schools in the sample would need a minimum of 5–10 students to provide a reasonable estimate of the school-level variance. However, similar guidelines are not yet available for CCMM, raising questions about the minimum sample size required per unit of analysis in the CCMM setting.

Further, there is uncertainty about whether random effect estimates are sensitive to the sampling strategy and the potential imbalance of sample size across units of analysis. Many researchers conducting CCMM studies use data drawn from samples where only one context was originally intended to be studied. For instance, school-based researchers intentionally sample students by school, ignoring the distribution of students across neighborhoods. Because the dataset contained information about both school features and neighborhood of residence, researchers could fit a CCMM to estimate both school and neighborhood-level random effects – even though neighborhoods were not the primary sampling unit. As a result, the distribution of the sample across neighborhood catchment areas may be uneven due to schools being the primary sampling unit, potentially biasing estimates of neighborhood-level effects. Small and imbalanced neighborhood sizes could result in higher variability and imprecise estimates for random effects and possible bias leading to inaccurate conclusions regarding contextual effects. As CCMM becomes more popular with researchers encountering more non-nested data structures – particularly in the case of group randomized control trials – it is essential to determine whether estimates of contextual-level effects are biased when the sample sizes are unevenly distributed across the two contexts studied. If contextual effects are biased, it is also important to describe the direction of that bias, whether toward or away from the null.

The current study aimed to address these questions by performing a series of simulation analyses based on data from the National Longitudinal Study of Adolescent to Adult Health (Add Health) (Harris et al., 2009a, 2009b), one of the largest nationally representative surveys in the U.S (Harris et al., 2015). Our goal was to determine the extent to which a sample can be distributed unevenly across one higher-level context before random effects variance estimates become biased. Add Health was an ideal empirical dataset in which to ground these simulations because it is widely used in public health and has already linked contextual measures of schools and neighborhoods to health and behavior. Further, it intentionally sampled from one context (i.e., schools were the primary sampling unit) and the sample was distributed unevenly across a second context (i.e., neighborhoods). Additionally, because Add Health was drawn to be nationally representative, the distribution of students across schools and neighborhoods is a realistic sample of school catchment areas within the U.S. While schools in the sample each had a reasonable sample size, the neighborhoods those students came from were not always well represented, with many having only a single respondent. Furthermore, because of the rich information contextual information available, CCMMs are increasingly being used in Add Health papers despite unanswered questions of their validity prompted by the small neighborhood sample sizes. By anchoring these simulations to a realistic example and commonly used dataset, we ensure that our examination of CCMM validity is conducted within a relevant parameter space with practical implications for future Add Health studies. Body mass index (BMI) was chosen as the outcome for this simulation because of its clarity for analysis purposes (measured continuously and has an approximately Gaussian distribution) as well as its salience as a public health issue (Baskin et al., 2005; Lawrence, 2004).

2. Methods

Empirical data from the Wave 1 in-home sample of Add Health was used as a basis for the school and neighborhood data structure in our simulations. There were 20085 students who attended 132 unique schools and lived in 2410 unique neighborhoods. The school and neighborhood data structure in the Add Health is cross-classified because students attending the same school often resided in different neighborhoods and students living in the same neighborhood attended different schools. Specifically, there were 2979 unique combinations of school and neighborhood, with a median of 1 school per neighborhood (range 1–3) and a median of 14 neighborhoods per school (range 1–234). Thus, the data were not purely hierarchical, but rather schools in particular drew students from many neighborhoods.

Overall, school sizes in Add Health ranged from 20 to 1720 with median 126.5 (interquartile range 85–174.5). Neighborhood sizes ranged from 1 to 276 (median 2; interquartile range 1–5); 45% of neighborhoods had only a single student while only 8% had 25 or more. These values indicate a wide distribution in neighborhood sizes with most falling in the lower range. While Add Health schools would appear to have sufficient sample sizes, at least according to the rules for hierarchical MLM, it was unclear whether this highly imbalanced neighborhood design affects random effects variance estimates for neighborhoods in CCMM.

2.1. Assignment of students to neighborhoods for the simulation (determining balance)

To remain consistent with the existing cross-classified data structure, we maintained the number of students nested within each school (range 20–1720; mean 152; median 126), as well as the number of neighborhoods feeding into each school (range 1–234). With the structure defined, we sorted students into neighborhoods for three different levels of sample size balance across neighborhoods: *perfectly balanced*, *mildly imbalanced*, and *very imbalanced*.

For the perfectly balanced scenario, the number of students within each school was divided evenly across the neighborhoods sending students to that school. Due to rounding, some schools had too many or too few students; this was addressed by randomly subtracting or adding from neighborhoods so that the number of students in each school was consistent with the empirical data, each neighborhood still had at least one student, and as close to perfect balance as possible was achieved.

For both imbalanced scenarios, we utilized a geometric distribution to assign students to neighborhoods given the number of neighborhoods per school. The probability of assignment to a given neighborhood k given the initial proportion p, was calculated as:

$$P(X = k) = (1-p)^{k-1}p \tag{1}$$

where p = initial proportion (probability of assignment to first neighborhood) and k = given neighborhood sending students to a specific school. P was set at 0.25 for the mildly imbalanced and 0.7 for the imbalanced scenario, meaning that the first neighborhood for each school was assigned 25% of students and 70% of students, respectively. Fig. 1 illustrates the assignment of students to neighborhoods under the balance scenarios for a hypothetical school with 60 students from 12 neighborhoods.

In practice, under both the mildly imbalanced and very imbalanced scenarios this resulted in some neighborhoods with zero students

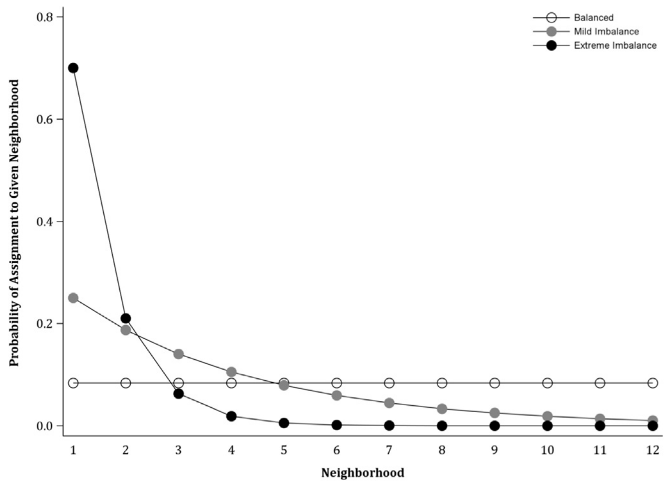

Fig. 1. Probability of assignment to a given neighborhood under different balance scenarios in a hypothetical example school with 60 students from 12 different neighborhoods.

assigned. To maintain the original data structure, each neighborhood was forced to have at least one student, resulting in most schools having either slightly too few or too many students relative to the original data structure. This discrepancy between school sizes for our assignments relative to the original data ranged from 97 fewer to 213 extra for the mildly imbalanced scenario (median − 1) and 22 fewer to 228 extra in the very imbalanced scenario (median + 9). Therefore, students were subtracted or added from neighborhoods using the geometric distribution so that the total number of students in each school was consistent with the empirical data while maintaining the minimum of one student per neighborhood. Following these adjustments, the average percent of students in the largest neighborhood for each school was 26% (range 5–100%; 56% included only one student) for the mildly imbalanced case and 61% (range 9–100%; 84% included only one student) for the very imbalanced case. Comparatively, the distribution of students across neighborhoods in the empirical data fell close to our mildly imbalanced scenario with the average percent of students in the largest neighborhood for each school of 36% (range 2–100%; 45% included only one student). Therefore, the distribution of students across neighborhoods in the real Add Health data is encompassed by our chosen range of simulated values allowing us to explore a relevant parameter space.

2.2. Simulation of individual data and partitioning variance

To evaluate the impact of balance (or imbalance) on our ability to estimate a "true" random effect (RE) variance estimate for both neighborhood and school contexts, BMI values were assigned to the simulated agents of the model, such that the "true" RE values for schools and neighborhoods were known. Individual BMI values were simulated to be consistent with the empirical data with mean 22.57 and standard deviation 4.47 and following a Gaussian distribution. The total variance in BMI was constant across all simulations at $\sigma^2_{total} = 20.0$ along with the variance attributable to individuals at $\sigma^2_{individual} = 19.0$ so that the total percent of variability accounted for by individual BMI differences was 95%, which is consistent with the empirical Add Health data. Thus the intra-class correlation (ICC) for the effect of neighborhoods and schools was the remaining 5% of the total variance which was differentially partitioned between schools and neighborhoods to examine the effect of variance partitioning on parameter re-capture. This degree of variance at the contextual level is comparable to other studies with continuous outcomes, which have generally found that 4–7% of the variance is attributable to higher-level contexts (Dunn et al., 2015a, 2015b, 2016). We chose five different scenarios to partition this 5% variance: 1) attributed primarily to neighborhood, very little to school (ICC = 4.5% and 0.5%, respectively); 2) attributed majority to neighborhood, small proportion to school (ICC = 3.75%; 1.25%); 3) split equally between neighborhood and school (ICC = 2.5%; 2.5%); 4) attributed majority to school, small proportion to neighborhood (ICC = 3.75%; 1.25%); 5) attributed primarily to school, very little to neighborhood (ICC = 4.5%; 0.5%). For each variance combination and level of balance, 1000 datasets were simulated resulting in 15000 total simulations.

2.3. Statistical analysis

CCMM were fit using the GLIMMIX procedure in SAS (v9.3; Cary, NC), which uses restricted maximum likelihood estimation. A null cross-classified model (no predictors) was fit for each simulation, predicting BMI with a Gaussian response distribution and identity link function. Random intercepts were included for school and neighborhood. The fitted CCMM model predicting BMI (denoted y) with individuals (denoted i) simultaneously belonging to non-nested contexts for school (denoted j) and neighborhood (denoted k) was the following for a person i in school j and neighborhood k:

$$y_{i(jk)} = \beta_0 + u_{0j} + u_{0k} + e_{0i(jk)} \tag{2}$$

In Eq. (2) above, β_0 refers to the overall mean BMI (y) across all schools and neighborhoods; u_{0j} refers to the random effect for schools, u_{0k} refers to the random effect for the neighborhood and $e_{0i(jk)}$ refers to the random effect for the individual with the combination of school j and neighborhood k.

Variance parameters from each model at the individual-, school-, and neighborhood-level and their 95% confidence intervals were used to calculate *the percentage of times the true simulated variance parameter was re-captured across 1000 simulations* for each scenario of balance and variance partitioning. Chi-squared tests were used to test for systematic bias in variance parameter capture by balance and variance partitioning scenario.

3. Results

Table 1 presents results from our simulations demonstrating the percent of simulations in which the true neighborhood and school variance parameters were successfully captured by the 95% CI. Across all simulations, variance re-capture ranged from 93% to 96% for both neighborhood and school variance. Overall, 5% of the simulations failed to capture the neighborhood parameter (N = 804/15,000); of this, 65% were underestimates and 35% were overestimates with median bias in the variance of − 0.12 (range − 0.38 to 0.53). In addition, 6% of simulations failed to capture the school variance (N = 832/15,000) and of these, 85% were underestimates and 15% were overestimates with median bias in the variance of − 0.15 (range − 0.43 to 0.48). Fewer than 1% of simulations (N = 69/15,000) failed to re-capture both the school and neighborhood variance parameter. Regardless of balance level, there was no difference in capture of the true neighborhood variance parameter across the variance partitioning scenarios (p = 0.59). Rates of school variance re-capture varied across variance partitioning (p = 0.009); this difference was driven by slightly lower rates of capture than expected when nearly all of the variance was attributable to the school. Similarly, for each of the neighborhood balance scenarios, there was no difference in the capture of the true neighborhood variance parameter (p = 0.07) or the true school variance parameter (p = 0.11) regardless of variance partitioning across levels.

3.1. Balanced neighborhood size

When neighborhood size was close to perfectly balanced, true

Table 1
Percentage of times the true variance parameter is captured by 95% confidence intervals across 1000 simulations.

School Variance σ^2_{school} (ICC)	**Neighborhood Variance** $\sigma^2_{neighborhood}$ (ICC)									
	0.90 (4.50%)		0.75 (3.75%)		0.50 (2.50%)		0.25 (1.25%)		0.10 (0.50%)	
	Neighbor-hood (%)	School (%)	Neighbor-hood (%)	School (%)	Neighbor-hood (%)	School (%)	Neighbor-hood (%)	School (%)	Neighbor-hood (%)	School (%)
Balanced										
0.90 (4.50%)									**96%**	**93%**
0.75 (3.75%)							**94%**	**95%**		
0.50 (2.50%)					**96%**	**93%**				
0.25 (1.25%)			**96%**	**95%**						
0.10 (0.50%)	**95%**	**94%**								
Mild imbalance										
0.90 (4.50%)									**94%**	**94%**
0.75 (3.75%)							**95%**	**96%**		
0.50 (2.50%)					**94%**	**95%**				
0.25 (1.25%)			**94%**	**95%**						
0.10 (0.50%)	**95%**	**94%**								
Extreme imbalance										
0.90 (4.50%)									**93%**	**94%**
0.75 (3.75%)							**94%**	**95%**		
0.50 (2.50%)					**95%**	**94%**				
0.25 (1.25%)			**94%**	**93%**						
0.10 (0.50%)	**95%**	**96%**								

neighborhood variance was captured 95% of the time when almost all of the variance was attributable to the neighborhood ($ICC_{neighborhood}$ = 4.5%); and 96% of the time when almost all of the variance was attributable to the school (ICC_{school} = 4.5%). Among those where the true neighborhood variance parameter was not captured (N = 240/5000), neighborhood variance was underestimated 50% of the time by an average of 0.18 (SD 0.04) and overestimated in 50% of simulations by an average of 0.20 (SD 0.04). Capture of school variance parameter was 94% when neighborhood variance was large and 93% when school variance was large. Across simulations where the true school variance parameter was not captured (N = 303/5000), school variance was underestimated in 84% of simulations by an average of 0.17 (SD 0.08) and overestimated 16% of the time by an average of 0.25 (SD 0.09). While these results are somewhat unsurprising given the advantages of a balanced data structure, a setup rarely encountered in real social science research, these results are consistent in the other imbalance-scenarios as well.

3.2. Mildly imbalanced neighborhood size

Under the mildly imbalanced neighborhood size scenario (average of 26% of students in the largest neighborhood), true neighborhood variance was captured 95% of the time when neighborhood variance was large and 94% when school variance was large. Among those where the true neighborhood parameter was not captured (N = 279/5000), neighborhood variance was underestimated 68% of the time by an average of 0.15 (SD 0.05) and overestimated in 32% of simulations by an average of 0.21 (SD 0.05). School variance capture was 94% when either neighborhood or school variance was large. Across all simulations where the true school variance was not captured (N = 254/5000), school variance was underestimated in 87% of simulations by an average of 0.17 (SD 0.08) and overestimated 13% of the time by an average of 0.26 (SD 0.10).

3.3. Extremely imbalanced neighborhood size

When neighborhood size was extremely imbalanced (average of 61% of students from a school in a single neighborhood), the true neighborhood variance was captured 95% of the time when neighborhood variance was large and 93% of the time when school variance was large. Among simulations where the neighborhood variance was not captured (N = 285/5000), neighborhood variance was underestimated 75% of the time by an average of 0.18 (SD 0.07) and overestimated in 25% of simulations by an average of 0.25 (SD 0.09). School variance was captured 96% of the time when school variance was small and 94% of the time when school variance was large. Across all simulations where the true school parameter was not captured (N = 275/5000), school variance was underestimated in 84% of simulations by an average of 0.19 (SD 0.07) and overestimated 16% of the time by an average of 0.28 (SD 0.09).

4. Discussion

In this study, we examined the effect of neighborhood sample size imbalance on the estimation of variance parameters in cross-classified multilevel models (CCMMs) predicting BMI in a series of simulations across a variety of neighborhood size/balance and variance partitioning scenarios. Our results provide compelling evidence that true variance parameters are captured by the 95% CI regardless of either the level of (im)balance or the partitioning of variance across levels in CCMMs. Furthermore, when this imbalance resulted in many neighborhoods having very small sample sizes, this did not adversely affect the capture of "true" random effect values. Across all levels of balance and variance partitioning, capture of both neighborhood-level and school-level variance were very high, ranging from 93% to 96%. Our findings indicate that variance parameters are relatively unbiased by context size and balance as the true variance parameters were captured in almost every simulated sample and capture probabilities were all close to nominal coverage of 95%. There were slight variations in parameter capture depending on variance partitioning, but no differences by level of neighborhood size balance.

Importantly, most of the simulations where either school or neighborhood variance parameters were *not* successfully captured resulted in *underestimates* of the parameter. This lends additional credence to the validity of studies utilizing CCMM to detect contextual effects because bias towards the null in those cases would be preferable to over-estimation.

In our simulations, neighborhood variance was captured 95% and 94% of the time for the mild imbalance and extreme imbalanced scenarios, respectively while school variance was captured 96% and 95% of the time. These results are reassuring and indicate that true contextual-level variance parameters have a high chance of being captured regardless of the relative imbalance of students across neighborhoods or partitioning of the variance. The fact that reasonable estimates are obtained for neighborhoods and schools even when the degree of imbalance is more extreme than the empirical data is also encouraging.

This study has several limitations. First, we examined a single, continuous and approximately normally-distributed outcome. Whether these results generalize to a binary, ordinal, or non-normally-distributed outcome is unknown and warrant further exploration. Second, we examined only three levels of neighborhood balance, which were chosen to provide two relatively extreme scenarios (perfect balance and extreme imbalance) and a middle-ground which is reflective of the true Add Health data. Data sets with more extreme imbalance may still pose problems for estimation. Third, we used the existing school and neighborhood data structure of Add Health and did not consider varying numbers of schools or neighborhoods per school, which may be of concern for school and neighborhood studies other than Add Health. However, grounding our example in Add Health provided a realistic example of school and neighborhood distributions in the United States and a wide range of neighborhoods per school (range 1–234), though this data structure is certainly not representative of all possible school and neighborhood allocations, especially those schools outside of the U.S. Future studies should examine the impact of different numbers of schools and neighborhoods identified in other cross-classified samples or manipulate the theoretical number of neighborhoods per school. Additionally, when simulating data, we assumed a total variance due to school and neighborhood clustering of only 5% and did not consider possibilities with larger or smaller contextual-level contributions; however, the total contextual variance contribution was chosen to be consistent with the literature for other continuous health outcomes (Dunn et al., 2015b; Townsend et al., 2012). Finally, it is unknown how sensitive variance capture is to choice of modeling algorithm; results may vary slightly between restricted maximum likelihood estimates (like those from SAS) and Bayesian Markov Chain Monte Carlo (MCMC) techniques such as that employed by MLwiN software (Goldstein, 1994; Browne, 2012; Rasbash et al., 2012) to fit cross-classified data structures. Comparison between variance parameters from restricted maximum likelihood using SAS and MCMC from MLwiN in the empirical data yielded no material difference between either variance parameter estimates or ICCs indicating modeling algorithm has little to no impact on variance estimates.

In conclusion, our results suggest no inherent bias in the ability of CCMM to capture true variance parameters at both neighborhood and school-levels across a variety of neighborhood balance scenarios and variance partitioning. Though our example focuses on schools and neighborhoods, the results are applicable more broadly to non-nested contexts. Thus, researchers using CCMMs may feel reassured that the true variance is captured regardless of relative balance of context size when modeling normally-distributed outcomes. Such insights are informative, as CCMM is becoming increasingly used, software continues to develop and algorithms become more efficient, resulting in CCMMs becoming commonplace in epidemiologic research.

Acknowledgements

This research uses the data structure from The National Longitudinal Study of Adolescent to Adult Health (Add Health), a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

Author contributions

Author 1, Author 2, Author 3 and Author 4 conceived of and designed the study and analysis. Author 1 analyzed the data and wrote the first draft of the manuscript. All authors contributed to the writing of the manuscript and agree with the manuscript results and conclusions. All authors made critical revisions and approved the final manuscript.

Competing interests

The authors report no conflicts of interest, competing interests or financial disclosures related to this study.

Funding

This study was supported by the National Institutes of Health (K01MH102403, Author 4).

References

Dunn, E.C., Milliren, C.E., Evans, C.R., Subramanian, S.V., Richmond, T.K., 2015a. Disentangling the relative influence of schools and neighborhoods on adolescents' risk for depressive symptoms. Am. J. Public Health 105 (4), 732–740.
Dunn, E.C., Richmond, T.K., Milliren, C.E., Subramanian, S.V., 2015b. Using cross-classified multilevel models to disentangle school and neighborhood effects: an example focusing on smoking behaviors among adolescents in the United States. Health Place 31, 224–232.
Richmond, T.K., Dunn, E.C., Milliren, C.E., Rosenfeld Evans, C., Subramanian, S.V., 2016. Disentangling overlapping influences of neighborhoods and schools on adolescent body mass index. Obesity (Silver Spring) 24 (12), 2570–2577.
Milliren, C.E., Richmond, T.K., Evans, C.R., Dunn, E.C., Johnson, R.M., 2017. Contextual Effects of Neighborhoods and Schools on Adolescent and Young Adult Marijuana Use in the United States. Substance Abuse: Res. Treat. 11, 1–10.
Baskin, M.L., Ard, J., Franklin, F., Allison, D.B., 2005. Prevalence of obesity in the United States. Obes. Rev. 6 (1), 5–7.
Browne, W.J., 2012. MCMC Estimation in MLwiN, version 2.26. Centre for Multilevel Modeling, University of Bristol, United Kingdom.
De Clercq, B., Pfoertner, T.K., Elgar, F.J., Hublet, A., Maes, L., 2014. Social capital and adolescent smoking in schools and communities: a cross-classified multilevel analysis. Soc. Sci. Med. 119, 81–87.
Dedrick, R.F., Ferron, J.M., Hess, M.R., et al., 2009. Multilevel Modeling: a review of methodological issues and applications. Rev. Educ. Res. 79 (1), 69–102.
Diez-Roux, A.V., 2000. Multilevel analysis in public health research. Annu. Rev. Public Health 21, 171–192.
Evans, C.R., Williams, D.R., Onnela, J.P., Subramanian, S.V., 2016. Multiple contexts and adolescent body mass index: schools, neighborhoods, and social networks. Soc. Sci. Med. 162, 21–31.
Goldstein, H., 1994. Multilevel cross-classified models. Sociol. Methods Res. 22, 364–375.
Harris K.M., 2009a. The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007–2009 machine-readable data file and documentation. In: Carolina Population Center University of North Carolina atChapel Hill, ed. Chapel Hill, NC.
Harris K.M., Halpern C.T., Whitsel E. et al. 2009b. The National Longitudinal Study of Adolescent to Adult Health: Research Design; ⟨http://www.cpc.unc.edu/projects/addhealth/design⟩ (Accessed 10 November 2014).
Harris K.M., Halpern E., Whitsel J. et al. 2015. Add Health Publications: Measuring the Impact. (Accessed 12 April 2016).
Heck, R., 2009. Teacher effectiveness and student achievement: investigating a multilevel cross-classified model. J. Educ. Adm. 47 (2), 227–249.
Kim, J., McCarthy, W.J., 2006. School-level contextual influences on smoking and drinking among Asian and Pacific Islander adolescents. Drug Alcohol Depend. 84 (1), 56–68.
Kim, Y., Petscher, Y., Foorman, B., Zhou, C., 2010. The contributions of phonological awareness and letter-name knowledge to letter-sound acquisition - a cross-classified multilevel model approach. J. Educ. Psychol. 102 (2), 313–326.
Lawrence, R., 2004. Framing obesity: the evolution of news discourse on a public health issue. Int. J. Press/Polit. 9 (3), 56–75.
Leventhal, T., Brooks-Gunn, J., 2000. The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. Psychol. Bull. 126 (2), 309–337.
McNeish, D., Stapleton, L., 2014. The effect of small sample size on two level model estimates: a review and illustration. Educ. Psychol. Rev. 1–20.
Munoz, M.A., Chang, F.C., 2007. The elusive relationship between teacher characteristics and student academic growth: a longitudinal multilevel model for change. J. Pers. Eval. Educ. 20 (3), 147–164.
Pickett, K.E., Pearl, M., 2001. Multilevel analyses of neighborhood socioeconomic context and health outocmes: a critical review. J. Epidemiol. Community Health 55,

111–122.
Rabash, J., Browne, W.J., 2001. Non-hierarchical multilevel models. In: Leyland, A.H., Goldstein, H. (Eds.), Modeling Non-hierarchical Structures. Wiley and Sons, New York, NY, pp. 93–105.
Rasbash, J., Steele, F., Browne, W.J., Goldstein, H., 2012. A User's Guide to MLwiN, Version 2.26. Centre for Multilevel Modeling, University of Bristol, United Kingdom.
Rice, N., Alastair, L., 1996. Multilevel models: applications to health data. J. Health Serv. Res. Policy 1 (3), 154–164.
Sellstrom, E., Bremberg, S., 2006. Is there a "school effect" on pupil outcomes? A review of multilevel studies. J. Epidemiol. Community Health 60 (2), 149–155.
Tendulkar, S., Buka, S., Dunn, E., Subramanian, S., Koenen, K., 2010. A multi-level investigation of neighborhood effects on parental warmth. J. Community Psychol. 38 (5), 557–573.
Townsend, N., Rutter, H., Foster, C., 2012. Age differences in the association of childhood obesity with area-level and school-level deprivation: cross-classified multilevel analysis of cross-sectional data. Int. J. Obes. (Lond.) 36, 45–52.