Multilevel Latent Transition Mixture Modeling for Organizational Models: A Simulation

Grant B. Morgan
Kevin E. Wells
R. Noah Padgett
Baylor University

Abstract

Person-centered methodologies generally refer to those that take unobserved heterogeneity of populations into account. The use of person-centered methodologies has proliferated, which is likely due to a number of factors, such as methodological advances coupled with increased personal computing power and ease of software use. The purpose of the proposed paper is to (1) present and discuss multilevel latent transition analysis and considerations for its use and (2) conduct a simulation study to assist researchers with both how to use measures of model-data fit and when to use them, and (3) demonstrate variance decomposition. With respect to the model parameters, the estimates tended to be unbiased, on average, but were prone to considerable variability. The standard errors were consistently over- or under-estimated and should not likely be trusted/interpreted. The problems with the standard errors naturally affects the confidence one might have in her parameter estimates in a negative way. Additional simulation will be necessary to find the boundaries of where the standard error estimation via the built-in algorithm can be trusted.Ultimately, we cannot yet recommend use of these models within the conditions that were simulated in this study.

Multilevel Latent Transition Mixture Modeling for Organizational Models: A Simulation

Efforts to classify individual cases into homogeneous groups have long been used in order to better understand complex sets of information. Classification of cases into homogeneous groups has important implications in the social sciences, such as education, medicine, psychology, or economics, where identifying smaller subsets of like cases may be of particular interest. Person-centered methodologies generally refer to those that take unobserved heterogeneity of populations into account. That is, rather than treat all individuals as if they originated from a single underlying population, as is true with variable-centered methodologies person-centered methodologies allow for multiple subpopulations to underlie a set of data. The challenge with these methods is identifying the correct number (i.e., frequency) of subpopulations, or classes, and the parameters (i.e., form) associated with each, when the frequency and form are not known *a priori*. Although, multilevel mixture models are available to researchers, they have not been methodologically examined or discussed as extensively as other cross-sectional and longitudinal *mixture* models.

Introduction to Multilevel Latent Transition Analysis

Using latent class analysis and its extension for longitudinal data, (latent transition analysis [LTA]), multiple underlying, homogeneous subgroups can be inferred from a set of categorical and/or continuous observed variables within a large heterogeneous data set. Such analyses allow researchers to statistically treat members of different subgroups separately, which may provide researchers with more power to detect effects of interest and closer alignment between statistical modeling and one's guiding theory. Furthermore, the hierarchical structure of organizational data must also be taken into account; that is, students (i.e., level-1 units) are nested within teacher/schools (i.e., level-2 units). Finally, multilevel LTA can be used to estimate the number of classes in each structured unit and the potential movement, or transitions, participants make between classes across time. The transitions/stability between latent classes across time can be treated as the outcome in and of itself, or the transitions/stability can be used as a correlate or predictor of some other, distal outcome.

From Asparouhov and Muthén (2008), the multinomial logistic regression for the latent classification variable at time 1, C_1, can be expressed:

$$\mathrm{P}(C_{1ij} = c) = \frac{\exp(\alpha_{1cj} + \beta_{1cj}x_{1ij})}{\sum_c \exp(\alpha_{1cj} + \beta_{1cj}x_{1ij})} \tag{1}$$

The multinomial logistic regression for the latent classification variable at time 2 using C_1 as a covariate can be expressed:

$$\mathrm{P}(C_{2ij} = d | C_{1ij} = c) = \frac{\exp(\alpha_{2dj} + \gamma_{dcj} + \beta_{2dj}x_{2ij})}{\sum_d \exp(\alpha_{2dj} + \gamma_{dcj} + \beta_{2dj}x_{2ij})} \tag{2}$$

The simplest case in which there are two data collection points is shown in Figure 1. The decomposition of variance for this model (i.e., proportion of variance in C_2 explained by the model) can be conducted in 11 steps.

Statement of Purpose

Due to a number of programming constraints, systematic methodological investigations of multilevel LTA have not been conducted, which poses a problem for applied analysts. That is, little is known about the accuracy with which parameters are estimated in these models under virtually any real world conditions. The purpose of this paper is address this gap in the literature and knowledge base by employing Monte Carlo methods to examine the parameter accuracy under known conditions. In addition to providing the model output and interpretations, we also explain the variance decomposition, which is more complex than in traditional hierarchical linear models.

Methods

The population structure used in the proposed study is a mixture model with known number of latent classes, class prevalence, and class-specific indicator distributions. The population structure also includes local independence. Models will then be estimated to determine the extent to which the true structure was recovered. Mplus was used to generate all data and estimate each model.

Design Factors. We manipulated four factors in this study – the level-2 variance at time 2, the transition between latent classes at time 1 and time 2, the effect of the level-2 unit across time, and the sample sizes. The level-2 variances at time 2 were 0.0, 0.25, or 0.5. The transition effect of latent class membership at time 1 on latent class membership at time 2 was 1, 2, or 3.5 (in odds ratio units). The effect of the level-2 units across time was 0.1, 0.3, or 0.5. The overall sample size was 500 or 1,000 but were based on different combinations of level-2 and level-1 units. That is, the level-2/level-1 within level-2 sample sizes were 50/10, 25/20, 100/5, and 50/20. All factors were fully crossed, which produced 108 (3 level-2 variances $\times$ 3 transition parameters $\times$ 3 level-2 effects $\times$ 4 sample sizes= 108) cells in the design. Each cell was replicated 100 times. The design factors and their associated levels are shown in Figure 2.

Each simulated model 10 latent class indicators at times 1 and 2, and class enumeration was limited to two. The latent class prevalence used for the two class population models was .50-.50 at time 1. Classes will be separated based on class-specific response probabilities to the dichotomous class indicators.

Evaluating parameter estimation. Summary information was output from Mplus and summarized across the 108 different conditions. The estimation accuracy was evaluated with respect to the parameter estimates as well as the standard errors. For parameter estimates, relative bias (RB) values were computed by comparing the differences between the average and true parameter values divided by the true values. For the standard errors, RB values were computed by comparing the standard deviation of the parameter estimates within each cell against the mean of estimated standard errors with each cell and then dividing by the standard deviation of the parameter estimates within each cell. Positive bias in RB indicates overestimation of a parameter and negative bias indicates underestimation. The biases in the parameter estimates were used as the outcome variables in this study.

To guide the presentation of results, we conducted factorial analyses of variance (ANOVA) using relative bias of parameters estimates and relative bias of standard errors as dependent variables with main and two-way interactions of all design factors as independent variables (Bandalos & Leite, 2013). An ANOVA was conducted for each outcome variable separately, and the magnitudes of these effects on each outcome were assessed using effect size estimates from the ANOVA tables. Only those effect sizes that reached or exceeded Cohen's (1988) large effect size of .14 were considered to be practically significant.

Variance Decomposition

Ultimately, these models tend to focus on explaining latent class membership at time 2. Latent class membership at time 2 is affected by α_1, C_1, and $\alpha 2$ so the variability of latent class membership at time 2 has to be carefully decomposed using a series of steps. The parameters chosen for simulation produced models that accounted for between 9.8% and 62.9% of the variability in latent class membership at time 2. Below we demonstrate the variance decomposition for the model that explained the maximum variability in latent class membership at time 2. For ease of explanation, we refer to latent class membership at times 1 and 2 as C_1 and C_2, respectively, and the level-2 variance at times 1 and 2 as α_1 and α_2, respectively. We use the terms schools to refer to level-2 units and students to refer to level-1 units.

1. Proportion of variability in C_1 by α_1.

The proportion of variability in C_1 explained by α_1 is the same as the ICC for a multilevel logistic model.

$$ICC = \frac{\text{var}(\alpha_1)}{\text{var}(\alpha_1) + \frac{\pi^2}{3}} = \frac{2}{2 + 3.29} = .378$$

The parameter estimate itself (var$(\alpha_1) = 2$) suggests the the proportion of students in classes 1 and 2 at time 1 varies depending on the school. In other words, class 1 contains about 50% of the students at time 1 on average, but this percentage may differ across schools. The larger the parameter estimate, the greater the school effect.

2. Find the variance of C_1.

   The proportion of students in each latent class at time 1 in the population is .5. Thus, the variance of C_1 is:

$$\text{var(C}_1) = .50 \times (1 - .50) = .25$$

3. Find the total contribution of C_1 on C_2.

$$\gamma^2 \times \text{var(C}_1) = 3.5^2 \times .25 = 3.06$$

4. Split the contribution of C_1 on C_2 based on the proportion explained by α_1:

$$3.06 \times .378 = 1.16$$

   and the residual of C_1:

$$3.06 \times (1 - .378) = 1.90$$

5. Next, we need to determine the contribution of α_1 to C_2; α_1 affects C_2 through C_1

and α_2. The contribution of α_1 on C_2 via α_2 is:

$$\beta^2 \times \text{var}(\alpha_1) = .5^2 \times 2 = .5$$

The total contribution of α_1 is:

$$1.16 + .5 + 2\sqrt{.1.16 \times .5} = 3.18$$

6. Find the total variance of C_2 by adding the individual contributions.

$$\text{level-1 residual} + 1.16 + 3.18 + \text{var(C}_2) = 3.29 + 1.16 + 3.18 + 0.5 = 8.87$$

7. Compute the proportion of variance in C_2 explained by α_1.

$$\frac{3.18}{8.87} = .358$$

8. Compute the proportion of variance in C_2 explained by residual variance of α_2.

$$\frac{0.5}{8.87} = .06$$

9. Compute the proportion of variance in C_2 explained by residual of C_1.

$$\frac{1.90}{8.87} = .215$$

10. Find total proportion of variance in C_2 explained by the model.

$$.358 + .06 + .215 = .629$$

11. Find total proportion of variance in C_2 explained by C_1.

$$\frac{3.06}{8.87} = .345$$

12. Find the additional proportion of variance in C_2 by adding the school-level effect at time 2.

$$.629 - .345 = .284$$

Results

Convergence Rates

Ninety-six of the 108 conditions had convergence rates of 100%, and nine conditions had convergence rates of 99%. The three conditions that did not reach 99% convergence had convergence rates of 95%, 91%, and 83%. These three conditions had sample size in common – 50 level-2 units and 20 level-1 units. Overall, convergence was not a problem with parameter estimation.

Parameter Estimate Bias

No systematic relative bias was noted in the estimates of model parameters across design factors. The effect size estimates for all main and two-way interaction effects of the design factors was $\leq .01$ across all four outcome variables. On average, the parameter estimates were unbiased. That said, there was considerable variability in the relative bias across conditions. For example, the relative bias in estimate of the β parameter ranged from -6770% to 9614%. The five number summary for the β parameter is -6770.0%, -43.3%, -1.4%, 46.1%, and 9613.7%. The relative bias distribution for the γ parameter was very skewed. The five number summary for this parameter was -124.6%, -17.9%, 3.4%, 31.3%, and 9387790%. The relative bias distribution for the level-2 variance at time 1 parameter (α_1) was heavily influenced by outliers. The five number summary for this parameter was -100.0%, -24.8%, -3.7%, 21.9%, and 571.1% The relative bias distribution for the level-2 variance at time 2 parameter (α_2) was also heavily influenced by outliers. The five number summary for this parameter was -100.0%, -49.6%, 0.0%, 20.0%, and 1437.0%. Again, none of the design factors had a systematic influence on any of these outcomes, on average, despite producing enormous variability. The conditions when bias was most extreme tended to be observed when there were five or 10 level-1 units within level-2 unit.

Standard Error Bias

Differences in average relative bias were observed in standard errors of the γ parameter for the true γ main effect ($\eta_P^2 = .82$), the interaction between true γ parameter and the sample size ($\eta_P^2 = .30$), and the interaction between true γ and the level-2 variance at time 2 ($\eta_P^2 = .19$). Biases tended to be negative, which indicates that the standard errors were underestimated and were sometimes grossly so.

Gamma (γ) Values. The standard error estimation deteriorated as the true γ parameter increased; that is, as the odds of transitioning from one class to another across time increased. The mean relative bias in the standard errors was 2.2% when $\gamma = 1$, -31.7% when $\gamma = 2$, and –80.9% when $\gamma = 3.5$. As was the case with the parameter estimates, there was considerable variability in the relative biases of the standard errors. The distributions of the standard error relative bias are presented in Figure 3.

Interaction Between Gamma (γ) Values & Sample Size. When the true γ parameter was 1.0, the relative bias in standard errors was unbiased on average across sample sizes although there were differences in variability. When the true γ parameter was 2.0, the standard error relative bias was least extreme when the level-1 sample sizes was largest (i.e., 20) and tended to be worst when the level-1 sample size was smallest (i.e., 5). The greatest variability was observed for the sample size condition of 100 level-2 units and five level-1 units regardless of the true γ value. When the true γ parameter was 3.5, the standard error relative bias was next within 50 percentage points of 0. That is, the standard error relative bias was quite large for all models with γ values of 3.5.

The distributions of the standard error relative bias are presented in Figure 4.

Interaction Between Gamma (γ) Values & Level-2 Variance. When the true γ parameter was 1.0, the relative bias in standard errors was unbiased on average across level-2 variance at time 2 although there were differences in variability. When the true γ parameter was 2.0, the standard error relative bias was closest to zero when the level-2 variance at time 2 was 0.5. When the true γ parameter was 3.5, again, the relative bias in standard errors was extreme. There was The greatest variability was observed when the true γ value was 2.0 regardless of the level-2 variance at time 2. The distributions of the standard error relative bias are presented in Figure 5.

Educational or scientific importance

The use of person-centered methodologies has proliferated, which is likely due to a number of factors, such as methodological advances coupled with increased personal computing power and ease of software use. We believe that these methods offer significant benefit to a host of fields provided that these tools are used appropriately in accordance with one's guiding theoretical framework. With this paper, we sought to help readers better understand these methods by presenting a fairly generalized model - multilevel latent transition analysis. The results of our study suggest that, with respect to the model parameters, the estimates tend to be unbiased, on average, but were prone to considerable variability. The standard errors were consistently over- or under-estimated and should not likely be trusted/interpreted. The problems with the standard errors naturally affects the confidence one might have in her parameter estimates in a negative way. It should be noted that the standard errors tended to be least biased when the level-1 sample size was larger, but additional simulation will be necessary to find the boundaries of where the standard error estimation via the built-in algorithm can be trusted. It may be necessary to use some other approach, such as bootstrapping, to construct the sampling distributions of the parameter estimates. This is will require significant increases in computing times. Ultimately, we cannot yet recommend use of these models within the conditions that were simulated in this study.

References

Asparouhov, T., & Muthén, B. (2008). Multilevel mixture models. In *Advances in latent variable mixture models* (pp. 27–51). Information Age Charlotte, NC.

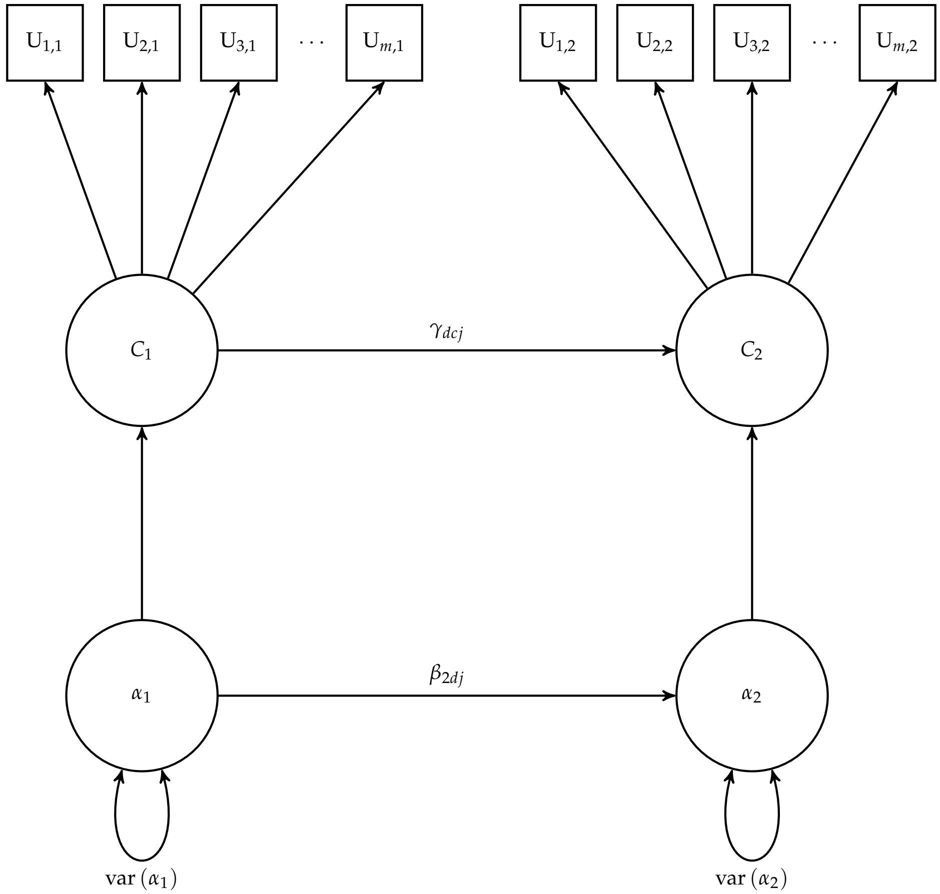

Figure 1. Conceptual diagram of a multilevel latent transition model with two time points and *m* latent class indicators at each time point.

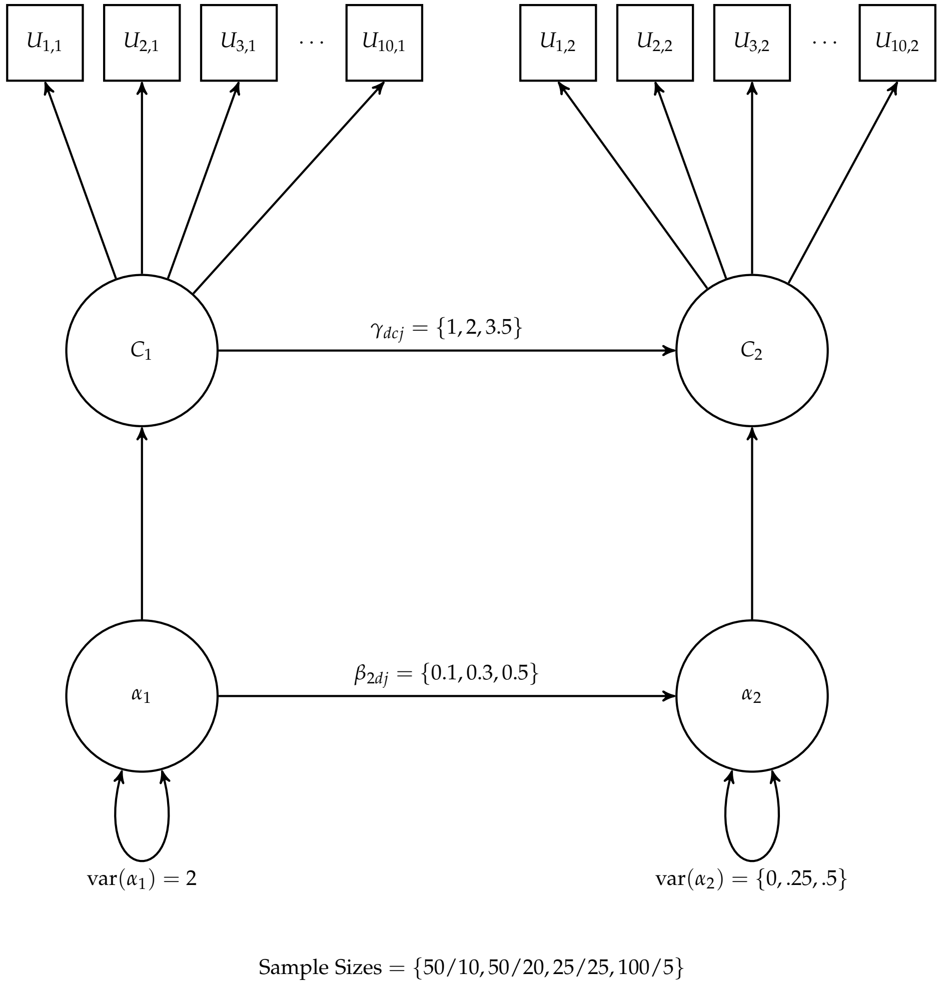

Figure 2. Path diagram of a multilevel latent transition model with manipulated levels of design factors.

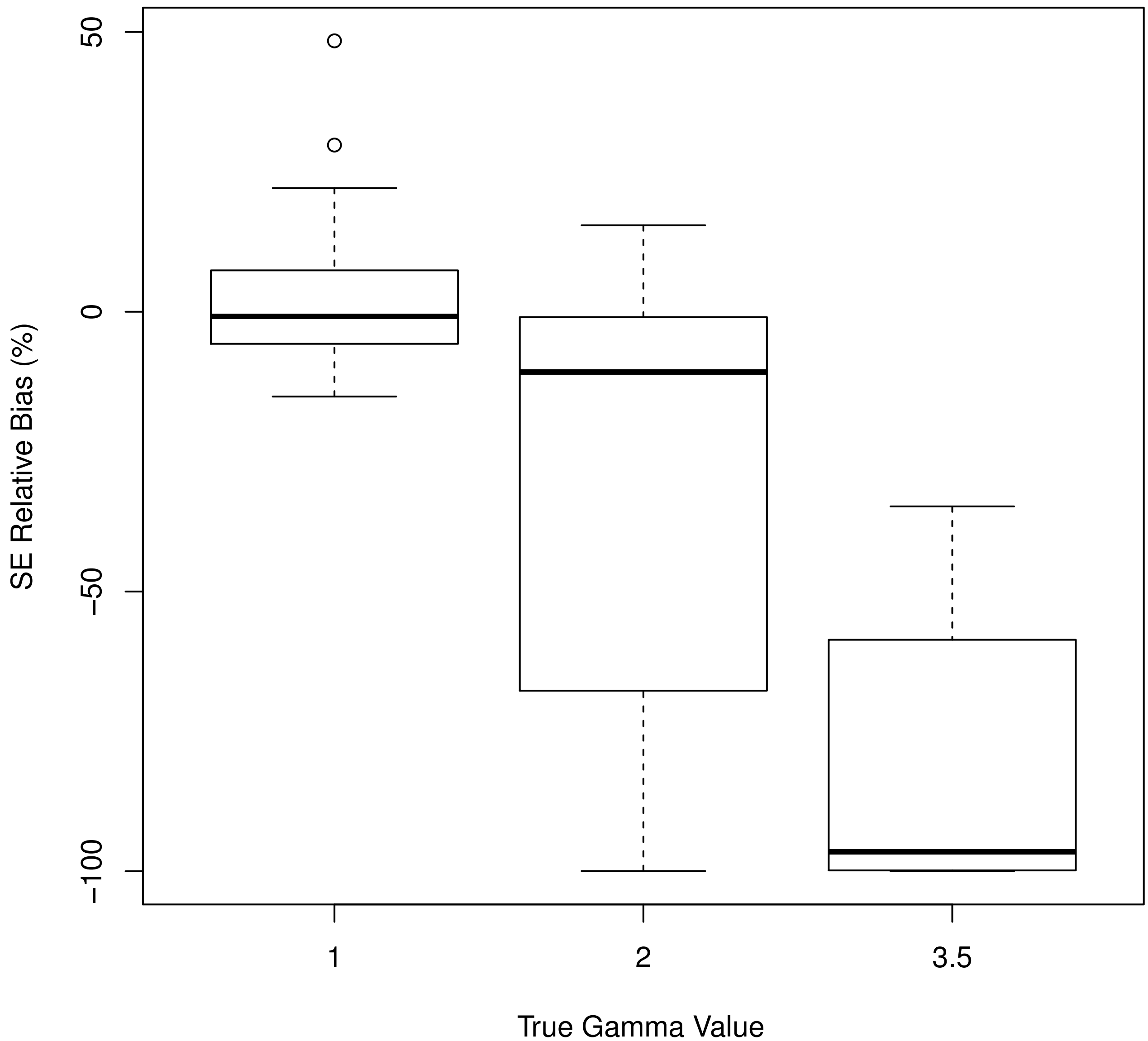

Figure 3. Boxplots of relative bias in standard errors for the γ parameter.

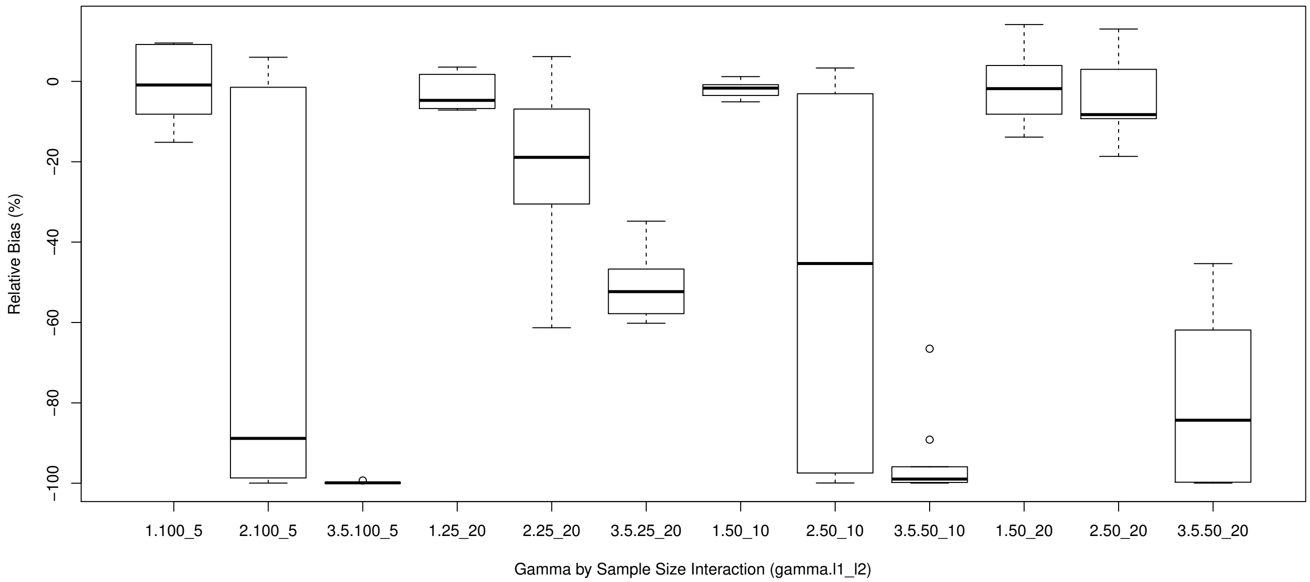

Figure 4. Boxplots of relative bias in standard errors for the γ parameter by sample size interaction.

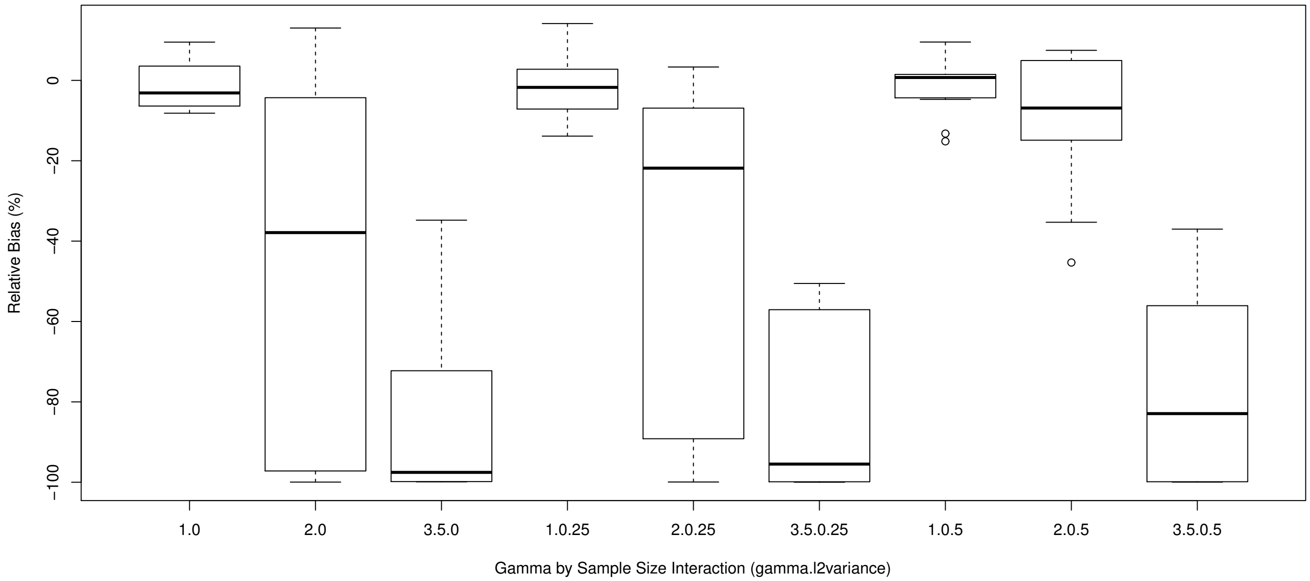

Figure 5. Boxplots of relative bias in standard errors for the γ parameter by level-2 variance interaction.