

# Impact of an equality constraint on the class-specific residual variances in regression mixtures: A Monte Carlo simulation study

Minjung Kim<sup>1</sup> · Andrea E. Lamont<sup>2</sup> · Thomas Jaki<sup>3</sup> · Daniel Feaster<sup>4</sup> · George Howe<sup>5</sup> · M. Lee Van Horn<sup>6</sup>

Published online: 3 July 2015  
© Psychonomic Society, Inc. 2015

**Abstract** Regression mixture models are a novel approach to modeling the heterogeneous effects of predictors on an outcome. In the model-building process, often residual variances are disregarded and simplifying assumptions are made without thorough examination of the consequences. In this simulation study, we investigated the impact of an equality constraint on the residual variances across latent classes. We examined the consequences of constraining the residual variances on class enumeration (finding the true number of latent classes) and on the parameter estimates, under a number of different simulation conditions meant to reflect the types of heterogeneity likely to exist in applied analyses. The results showed that bias in class enumeration increased as the difference in residual variances between the classes increased. Also, an inappropriate equality constraint on the residual variances greatly impacted on the estimated class sizes and showed the potential to greatly affect the parameter estimates in each class. These results suggest that it is important to make assumptions about residual variances with care and to carefully report what assumptions are made.

**Keywords** Regression mixture · Differential effects · Effect heterogeneity · Residual variances

An important problem in behavioral research is understanding heterogeneity in the effects of a predictor on an outcome. Traditionally, the primary method for assessing this type of differential effect has been the use of interactions. A new approach for assessing effect heterogeneity is regression mixture models, which allow different patterns in the effects of a predictor on an outcome to be identified empirically without regard to a particular explaining variable.

Regression mixture models have been increasingly applied to different research areas, including marketing (Wedel & Desarbo, 1994, 1995), health (Lanza, Cooper, & Bray, 2013; Yau, Lee, & Ng, 2003), psychology (Van Horn et al., 2009; Wong, Owen, & Shea, 2012), and education (Ding, 2006; Silinskas et al., 2013). Where traditional regression analyses model the single average effect of a predictor on an outcome for all subjects, regression mixtures model heterogeneous

✉ Minjung Kim  
mjkim.epsy@gmail.com

✉ M. Lee Van Horn  
MLVH@unm.edu

Andrea E. Lamont  
alamont082@gmail.com

Thomas Jaki  
jaki.thomas@gmail.com

Daniel Feaster  
DFeaster@biostat.med.miami.edu

George Howe  
ghowe@email.gwu.edu

<sup>1</sup> Department of Psychology, University of Alabama, Tuscaloosa, AL, USA

<sup>2</sup> Department of Psychology, University of South Carolina, Columbia, South Carolina, USA

<sup>3</sup> Department of Mathematics and Statistics, Lancaster University, Lancaster, UK

<sup>4</sup> Department of Epidemiology and Public Health, University of Miami, Miami, FL, USA

<sup>5</sup> Department of Psychology, George Washington University, Washington, DC, USA

<sup>6</sup> Department of Individual, Family, & Community Education, University of New Mexico, Albuquerque, NM 87131, USA

effects by empirically identifying two or more subpopulations present in the data for which each subpopulation differs in the effects of a predictor or predictors on the outcome(s). Three types of parameters are estimated in regression mixture models: latent-class proportions (probability of class membership), class-specific regression coefficients (intercepts and slopes), and residual structures. Although all of these elements are crucial to establish a regression mixture model, most methodological research in the area of mixture modeling has focused on detection of the true number of latent classes (Nylund, Asparouhov, & Muthén, 2007; Tofiqhi & Enders, 2008) or on recovering the true effects of covariates on the latent classes and outcome variables (Bolck, Croon, & Hagenaars, 2004; Vermunt, 2010). The residual variance components are often simplified without thorough examination of the consequences.

In this study, we aimed to evaluate the effects of ignoring differences in residual variances across latent classes in regression mixture models. Before examining our research question, we briefly review regression mixture models in the framework of finite mixture modeling.

## Regression mixture models

Regression mixture models (Desarbo, Jedidi, & Sinha, 2001; Wedel & Desarbo, 1995) allow researchers to investigate unobserved heterogeneity in the effects of predictors on outcomes. Regression mixture models are part of the broader family of finite mixture models, which also include latent-class analyses, latent-profile analyses, and growth mixture models (see McLachlan & Peel, 2000, for a review of finite mixture modeling). In regression mixture models, subpopulations are identified by class-specific differences in the regression weights, which characterize class means on the outcomes (intercepts) and on the relationship between predictors and outcomes (slopes). Thus, subjects identified to be in the same latent class share a common regression line, whereas those in another latent class have a different regression line. The overall distribution of the outcome variable(s) is conceived of as a weighted sum of the distribution of outcomes within each class.

Take a sample of  $N$  subjects drawn from a population with  $K$  classes. The general regression mixture model within class  $k$  for a single continuous outcome can be written as:

$$y_{i|x,k} = \beta_{0k} + \sum_{p=1}^P \beta_{pk} x_{ip} + \varepsilon_{ik}, \quad (1)$$

$$\varepsilon_{ik} \sim N(0, \sigma_k^2),$$

where  $y_i$  represents the observed value of  $y$  for subject  $i$ ,  $k$  denotes the group or class index,  $\beta_{0k}$  is the class-specific intercept coefficient,  $p$  is the number of predictors,  $\beta_{pk}$  is the

class-specific slope coefficient for the corresponding predictor,  $x_{ip}$  is the observed value of predictor  $x$  for subject  $i$ , and  $\varepsilon_{ik}$  denotes the class-specific residual error, which may be allowed to follow a class-specific variance,  $\sigma_k^2$ . The value of  $K$  is specified in advance, but the class-specific regression coefficients and the proportion of class membership are estimated. Hence, regression mixture models formulated in this way allow each subgroup in the population to have a set of unique regression coefficients and potentially unique residual variances, which represent the differential effects of the predictors on outcomes.

The between-class portion of the model specifies the probability that an individual is a member of a given class. This model may include covariates to predict class membership, and can be specified as

$$Pr(c_i = k|z_i) = \frac{\exp\left(\alpha_k + \sum_{q=1}^Q \gamma_{qk} z_{iq}\right)}{\sum_{s=1}^K \exp\left(\alpha_s + \sum_{q=1}^Q \gamma_{qs} z_{iq}\right)}, \quad (2)$$

where  $c_i$  is the class membership for subject  $i$ ,  $k$  is the given class,  $z_i$  is the observed value of predictor  $z$  of latent class  $k$  for subject  $i$ ,  $\alpha_k$  denotes the class-specific intercept, and  $\gamma_k$  is the class-specific effect of  $z$ , which explains the heterogeneity captured by latent classes. In this study, we are not considering the effect of  $z$ , so the equation can be simplified to

$$Pr(c_i = k) = \frac{\exp(\alpha_k)}{\sum_{s=1}^K \exp(\alpha_s)}.$$

Since there are many parameters to be estimated, model convergence can sometimes be a problem, and even when models do converge, they may not converge to a stable solution. To simplify model estimation, the class-specific residual variances are often constrained to be equal across classes. Referring to multilevel regression mixtures, Muthén and Asparouhov (2009) stated that, “For parsimony, the residual variance  $\theta_c$  is often held class invariant” (p. 640). This can be suitable in some cases, such as when the residual variances are very similar for all latent classes. However, the effects of this constraint on the latent-class enumeration and model results when there are differences in residual variances between classes has not been thoroughly examined.

We know of no existing research examining the effects of misspecifying the residual variances in regression mixture models, and of little research examining these effects with mixture models in general. McLachlan and Peel (2000) demonstrated the impact of specifying the common covariance matrix between two clusters in multivariate normal mixtures.

They found that the class proportions (and consequently, the assignment of individuals to a class) were poorly estimated under this condition, and they cautioned against the use of homoscedastic variance components. On the basis of this initial attempt, Enders and Tofghi (2008) investigated the impact of misspecifying the within-individual (Level-1) residual variances in the context of growth mixture modeling of longitudinal data. In their study, the class-varying within-individual residual variances were constrained to be equal across classes, and the impacts on latent-class enumeration and parameter estimates were assessed. They found some bias in the within-class growth trajectories and variance components when the residual variances were misspecified. In growth mixtures, intercepts and slopes are directly estimable for every individual, and the value of the model is to classify individuals who are similar in their patterns of these growth parameters. In regression mixture models, however, individual slopes are not directly estimable, and the mixture is used to allow us to estimate variability in the regression slopes that cannot otherwise be estimated. Thus, we expect that the impact of misspecifying the error variance structure on the parameter estimates would be more severe in regression mixture models than in growth mixture models.

Unlike growth mixture models, in which the means (intercepts), slopes over time, and variances are the main focus, regression mixture models focus on the regression weights characterizing the associations between the predictors and outcomes. In this case, we see a clear reason to expect differences in the residual variances between classes: If the regression weights are larger, there should be less residual variance, given the larger explained variance. Thus, even though residual variances are not the substantive focus when estimating these models, because differences between classes in residual variances are expected, it is especially important to understand the effects of misspecifying this portion of the model.

A review of the literature in which regression mixtures have been used showed a lack of consensus in the specification of residual variances; some authors freely estimated the class-specific residual variance (Daeppen et al., 2013; Ding, 2006; Lee, 2013), whereas others constrained them to be equal across classes (Muthén & Asparouhov, 2009). Interestingly, the majority of the studies employing regression mixtures have given no information about their residual variance specifications, whether the equality constraint had been imposed or not (Lanza et al., 2013; Lanza, Kugler, & Mathur, 2011; Liu & Lu, 2011; Schmeige, Levin, & Bryan, 2009; Wong & Maffini, 2011; Wong et al., 2012). The contribution of this article is to examine the degree to which this is a consequential decision that should be thoughtfully made and clearly reported, so that readers can understand regression mixture results and so that results may be replicated in the future. In the present study, we are focusing on the specification of  $\sigma_k^2$ , which is the variance of the residual error,  $\varepsilon_{ik}$ , and represents

the unexplained variance after taking into account the effect of all predictors in the model. We assume that the residual variances are normally distributed in this study, to avoid the complex issue of nonnormal errors in the regression mixture models (George et al., 2013; Van Horn et al., 2012).

## Study aims

The objective of this study was to examine the consequences of constraining the class-specific residual variances to be equal in regression mixture models under conditions that approximate those likely to be found in applied research. It is reasonable to expect that the unexplained/residual variances of the outcomes would differ across the classes between which the effect size of the predictor on the outcome varies. However, in practice, residual variances are of little interest substantively, and it has been recommended that they can be constrained to be equal across latent classes for the sake of model parsimony. In this study, we used Monte Carlo simulations to investigate the impact of this equality constraint.

Our first aim was to investigate whether imposing equality constraints on the residual variances across classes affects the results of class enumeration. We generated data from a population with two classes. The sizes of the residual variances differed across simulation conditions; however, we kept the differences in effect sizes between the classes the same across conditions. We examined how often the true two classes were detected when the residual variances were constrained to be equal for ten scenarios that differed in their numbers of predictors, the correlations between the predictors, and whether there was a difference in intercepts between the classes. Because class enumeration is mainly determined by the degree of separation between classes, we hypothesized that the equality constraints for small differences in variance would have minimal impact on selecting the correct number of latent classes. We also hypothesized that when models were misspecified by constraining variances to be equal, additional classes would increasingly be found as class separation and power increased.

The second aim was to examine parameter bias in the regression coefficients, variance estimates, and class proportions that would result from constraining the residual variances to be equal across classes. We hypothesized that the scenario with a large difference in residual variances between the classes would result in regression coefficients with substantial bias, because we had forced two very different values to be the same. With an additional class-varying predictor in the model, although the total residual variances were reduced for each latent class, the differences in the residual variances across latent classes would increase, because the difference in total effect sizes would be greater. Thus, we expected that there would be greater bias when there was an additional predictor in the regression mixture models. We expected no bias

when the true model contained two classes with equal variances.

The outcome of the first aim was the proportion of simulations that would select the true population model over a comparison model using the Bayesian information criterion (BIC) and sample-size-adjusted BIC (ABIC) as criteria. Although the Akaike information criterion (AIC) is also frequently used for model selection in finite mixture models, previous research has shown that the AIC has no advantage for latent class enumeration and tends to overestimate the number of latent classes in regression mixtures (Nylund et al., 2007; Van Horn et al., 2009). Therefore, we do not further discuss the AIC in this study. For the second aim, the accuracy of the parameter estimates of intercepts, regression coefficients of the predictors, residual variances, and percentages of subjects in each class were examined.

## Method

### Data generation

Data were generated using R (R Development Core Team, 2010) with 1,000 replications for each condition and a sample size of 3,000 in each dataset. (See Appendix 2 for the code used to generate the data.) Because regression mixtures rely on the shape of the residual distributions for identification, this is seen as a large-sample method (Fagan, Van Horn, Hawkins, & Jaki, 2013; Liu & Lin, 2014; Van Horn et al., 2009). We chose a sample size of 3,000 to be consistent with other research in the field (Smith, Van Horn, & Zhang, 2012) and because samples of this size are available in many publicly available datasets in behavioral research. Our starting point for finding differential effects was a population composed of two populations (which should be identified as classes), with a small effect size for the relationship between a single predictor and an outcome ( $r = .20$ ) in one population and a large effect size for this relationship in the other ( $r = .70$ ). The rationale for this condition was that we believe that a difference in correlations between subpopulations of .20 and .70 is the minimum needed for regression mixture models to be useful in capturing heterogeneity in the effects of  $X$  on  $Y$ . This corresponds to a small effect in one group and a large effect in the other, which can be found in some applied research employing regression mixture models (Lee, 2013; Silinskas et al., 2013). If the method cannot find a difference between a small and a large effect with a sample size of 3,000, then we argue that it has limited practical value for detecting differential

effects. If it meets this minimum criterion, then it has application at least in some situations. This condition was therefore chosen because it represents a threshold for the practical use of the method and is a good starting point for evaluating other features of the regression mixtures.

In this study, because we focused on the effect of misspecified residual variances, we held constant the class membership probabilities (.50) and held differences in the effect sizes between classes to be equal. The challenge in this situation was to create conditions in which the difference in effect sizes between the two classes was the same, but in which the residual variances differed. To achieve equal variances and have distinct regression weights, we chose the regression weights that had the same absolute value but differed in directionality, in which case the residual variances would be equal in each class. We also had a condition with a moderate difference in variances, in which regression weights were scaled to be closer to 0 than in the .20/.70 condition. In order to maintain the same effect size in each condition, we computed a Fisher's (1915)  $z$ -transformation<sup>1</sup> for each correlation, when  $r$  was .20 and .70 and  $z'$  was .203 and .867, which led to a difference between the two classes of  $z' = .664$ . Thus, the effect size difference was fixed to be a difference of  $z' = .664$  between the two classes for all conditions. The models can be written as

Large-difference condition:

$$Y_{i|c=1} = 0 + .200X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, .96),$$

$$Y_{i|c=2} = 0 + .700X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, .51);$$

Moderate-difference condition:

$$Y_{i|c=1} = 0 + (-.126)X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, .984),$$

$$Y_{i|c=2} = 0 + .491X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, .759);$$

No-difference condition:

$$Y_{i|c=1} = 0 + (-.321)X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, .897),$$

$$Y_{i|c=2} = 0 + .321X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, .897),$$

where  $X$  was generated from a standard normal distribution with a mean of 0 and standard deviation of 1. The differences in variances were set to be .45 for the large-difference, .225 for the moderate-difference, and 0 for the no-difference condition, whereas the total variance of  $Y$  was set to be 1 within each class across all conditions.

In order to assess the effects of variance constraints in situations more likely to mirror those observed in applied applications of regression mixtures, the simulations were expanded to include two predictors, the effects of which both differed between classes. This resulted in nine additional simulation conditions that differed in the correlations between these

<sup>1</sup> Fisher's  $z'$  was calculated with the formula  $z' = \frac{1}{2} \ln\left(\frac{1+z}{1-z}\right)$ .



predictors as well as in the means of the outcomes (intercepts) within each class. The general model for the multivariate conditions can be written as

$$Y_{i|c=1} = \beta_{01} + \beta_{11}X1_i + \beta_{21}X2_i + \varepsilon_{i1}, \quad \varepsilon_{i1} \sim N(0, \sigma^2_1),$$

$$Y_{i|c=2} = \beta_{02} + \beta_{12}X1_i + \beta_{22}X2_i + \varepsilon_{i2}, \quad \varepsilon_{i2} \sim N(0, \sigma^2_2).$$

Because the predictors in a multivariate model (especially where the predictors are operating in the same way) are typically correlated, we varied the relationship between  $X1$  and  $X2$  to range from having no relationship (Pearson correlation coefficient  $r = 0$ ), a moderate relationship ( $r = .5$ ), and a strong relationship ( $r = .7$ ). The population regression weights for the multivariate conditions were calculated to maintain the univariate relationships in light of the correlation between the predictors.<sup>2</sup> Additionally, the intercept values for the larger-effect class ( $\beta_{02}$ ) were varied to be 0, 0.5, and 1 for the condition with two predictors in the model, whereas the intercepts for the smaller-effect class ( $\beta_{01}$ ) were always zero. Therefore, we generated a total of 30 sets of simulations, including three variance difference (large, moderate, and no) conditions for the univariate model, and 27 conditions (3 variance difference  $\times$  3 correlations of predictors  $\times$  3 intercept difference) for the multivariate model. Larger intercept differences result in greater class separation and should increase the power to find two classes when the model is correctly specified, and to find more than two classes when the model is misspecified. The point of these analyses is to examine the effects of constraining the class variances to be equal as class separation increases. We note that with two predictors, the effect size when the predictors are both included in the model is not the same across conditions; specifically, there is less residual variance when the predictors are less correlated.

## Data analysis

Mplus 7.1 (Muthén & Muthén, 1998–2012), employing the maximum likelihood estimator with robust standard errors (MLR), was used for estimating the regression mixture models. We first fit the relaxed model (defined as the true model for the cases with a large and a moderate difference in variances between classes) by allowing the class-varying residual variances. This served two purposes: First, it validated the data-generating process by showing that the parameter estimates from the true model were as expected; second, it demonstrated that when there was no difference in variances between the classes (i.e., for the no-difference condition), it was still possible to estimate class-specific variances.

<sup>2</sup> The population parameters for the regression coefficients of the two predictors and the residual variances for both classes are presented in Appendix 1.

Then we examined the impact on class enumeration of constraining the residual variances to be the same between the classes. One-class, two-class, and three-class models were run for each of the 30 simulation conditions. The outcome was the percentage of simulations in which the true number of classes (two) was selected using the BIC (Schwarz, 1978) and the ABIC (Sclove, 1987). Both the BIC and ABIC have been shown to be effective for latent-class enumeration in regression mixture models (Van Horn et al., 2009).

Next, we compared the two-class constrained model with class-invariant residual variances to the two-class relaxed model with residual variances freely estimated in each class. A null hypothesis test was conducted to examine whether the restricted model fit worse than the relaxed model by using the Satorra–Bentler log-likelihood ratio test (SB LRT; Satorra & Bentler, 2001). The adequacy of parameter estimates was formally assessed using the root mean squared error (RMSE) and the coverage rate for the true population value for each parameter. RMSE is a function of both the bias and variability in the estimated parameter, and it is computed as the rooted squared value of the difference between the true population value and the estimated parameter (i.e.,  $RMSE = \sqrt{(True - Estimated)^2}$ ). We also report the average parameter estimates, standard errors, standard deviations, maximum and minimum values, and coverage rates. The coverage rate is the percentage of the 1,000 simulations for which the true parameter values fall in the 95 % confidence interval for each model parameter. If the parameter estimates and standard errors are unbiased, coverage should be 95 %. This shows the accuracy of statistical inference for each parameter in each condition.

## Results

### Class enumeration

All models converged properly across all simulation conditions. Before examining the impact of the equality constraints on the residual variances, we analyzed the regression mixtures that freely estimated the residual variances for both classes. The percentages selecting the true two-class model are presented in Table 1. For the single-predictor model, the two-class model was selected in 86.2 % and 87.2 % of the simulations using the BIC and the ABIC, respectively, when the difference in variances was large. Under the moderate-variance-difference condition (.225 difference), the true two-class model was selected in 52.3 % and 86.7 % of the simulations, respectively. Under the no-variance-difference condition when freely estimating the variances within classes, the two-class model was selected only in 46.7 % of simulations

**Table 1** Class enumeration results when the residual variances are freely estimated

Correlation	Difference in Intercepts	Difference in Variances	%BIC			%ABIC		
			2 vs. 1	3 vs. 1&2	2 vs. 1&3	2 vs. 1	3 vs. 1&2	2 vs. 1&3
<i>Single Predictor</i>								
		Large	86.40 %	0.20 %	86.20 %	98.40 %	11.20 %	87.20 %
		Moderate	52.30 %	0.00 %	52.30 %	91.40 %	4.70 %	86.70 %
		No	46.70 %	0.00 %	46.70 %	88.30 %	4.20 %	84.10 %
<i>Two Predictors</i>								
Zero	No	Large	100.00 %	0.10 %	99.89 %	100.00 %	5.81 %	94.19 %
		Moderate	83.58 %	0.00 %	83.58 %	86.99 %	6.71 %	80.28 %
		No	100.00 %	0.10 %	99.90 %	100.00 %	8.20 %	91.80 %
	0.5	Large	100.00 %	0.30 %	99.70 %	100.00 %	7.70 %	92.30 %
		Moderate	92.28 %	0.20 %	92.08 %	93.99 %	6.41 %	87.58 %
		No	100.00 %	0.00 %	100.00 %	100.00 %	7.71 %	92.29 %
	1	Large	100.00 %	0.40 %	99.60 %	100.00 %	7.60 %	92.40 %
		Moderate	99.50 %	0.30 %	99.20 %	99.70 %	9.60 %	90.10 %
		No	100.00 %	0.10 %	99.90 %	100.00 %	7.90 %	92.10 %
.5	No	Large	99.40 %	0.20 %	99.20 %	100.00 %	7.61 %	92.39 %
		Moderate	49.15 %	0.10 %	49.05 %	55.35 %	3.90 %	51.45 %
		No	46.55 %	0.00 %	46.55 %	91.69 %	7.81 %	83.88 %
	0.5	Large	100.00 %	0.00 %	100.00 %	100.00 %	6.90 %	93.10 %
		Moderate	63.30 %	0.00 %	63.30 %	69.00 %	4.60 %	64.40 %
		No	97.00 %	0.10 %	96.90 %	100.00 %	6.90 %	93.10 %
	1	Large	100.00 %	0.20 %	99.80 %	100.00 %	7.90 %	92.10 %
		Moderate	84.10 %	0.30 %	83.80 %	87.40 %	4.70 %	82.70 %
		No	100.00 %	0.30 %	99.70 %	100.00 %	6.90 %	93.10 %
.7	No	Large	89.60 %	0.20 %	89.40 %	99.50 %	7.00 %	92.50 %
		Moderate	51.50 %	0.00 %	51.50 %	57.10 %	4.50 %	52.60 %
		No	51.95 %	0.10 %	51.85 %	93.49 %	6.11 %	87.39 %
	0.5	Large	99.90 %	0.10 %	99.80 %	100.00 %	7.11 %	92.89 %
		Moderate	60.60 %	0.10 %	60.50 %	65.90 %	3.10 %	62.80 %
		No	97.30 %	0.00 %	97.30 %	100.00 %	5.10 %	94.90 %
	1	Large	100.00 %	0.20 %	99.80 %	100.00 %	6.60 %	93.40 %
		Moderate	84.18 %	0.00 %	84.18 %	87.79 %	6.11 %	81.68 %
		No	100.00 %	0.00 %	100.00 %	100.00 %	4.90 %	95.10 %

%BIC = percent selecting the true 2-class model using the Bayesian Information Criteria; %ABIC = percent selecting the true 2-class model using the sample-size-adjusted BIC

using the BIC, whereas they were correctly selected using the ABIC in 84.1 % of the simulations. The simulations that included two predictors suggest that the failure to select the two-class model is a function of power that is related to relatively low class separation. When there is greater class separation—larger differences between classes in variance and intercepts, and additional predictors with weak correlations—the two-class model is selected nearly all of the time using penalized information criteria.<sup>3</sup>

<sup>3</sup> The results are summarized in Table 1, a complete set of results is available from the first author on request.

The primary research questions for this article were assessed using one- through three-class regression mixture models with the variances constrained to be equal between classes under all conditions. We first examined whether the true two-class model would be selected over the one-class and three-class models in each simulation using the BIC and the ABIC. The results in Table 2 show that, as expected, the equality constraint does not impact class enumeration if the residual variances are actually the same. Under the equal-variance condition for the single-predictor model, the true two-class model is usually selected by the BIC (70.4 %) and the ABIC (94.4 %). These detection rates for finding two

**Table 2** Class enumeration results when the residual variances are constrained to be equal across classes

Correlation	Difference in Intercepts	Difference in Variances	%BIC			%ABIC			%Relaxed Model Favored		
			2c vs. 1	3c vs. 1&2c	2c vs. 1&3c	2c vs. 1	3c vs. 1&2c	2c vs. 1&3c	%BIC	%ABIC	SB LRT
<i>Single Predictor</i>											
		Large	71.50 %	3.50 %	68.00 %	94.30 %	38.70 %	55.60 %	80.70 %	92.10 %	69.40 %
		Moderate	68.00 %	0.00 %	68.00 %	94.20 %	1.70 %	92.50 %	11.60 %	26.40 %	33.90 %
		No	70.40 %	0.00 %	70.40 %	94.90 %	0.50 %	94.40 %	0.90 %	6.80 %	9.20 %
<i>Two Predictors</i>											
Zero	No	Large	100.00 %	100.00 %	0.00 %	100.00 %	100.00 %	0.00 %	100.00 %	100.00 %	100.00 %
		Moderate	78.60 %	9.90 %	68.70 %	81.50 %	52.80 %	28.70 %	97.70 %	99.40 %	99.60 %
		No	100.00 %	0.00 %	100.00 %	100.00 %	0.30 %	99.70 %	0.60 %	4.00 %	6.00 %
	0.5	Large	100.00 %	100.00 %	0.00 %	100.00 %	100.00 %	0.00 %	100.00 %	100.00 %	100.00 %
		Moderate	89.40 %	24.60 %	64.80 %	91.10 %	72.90 %	18.20 %	99.40 %	99.80 %	99.80 %
		No	100.00 %	0.00 %	100.00 %	100.00 %	0.00 %	100.00 %	0.40 %	3.20 %	5.20 %
	1	Large	100.00 %	100.00 %	0.00 %	100.00 %	100.00 %	0.00 %	100.00 %	100.00 %	100.00 %
		Moderate	99.10 %	56.80 %	42.30 %	99.30 %	91.00 %	8.30 %	99.90 %	100.00 %	100.00 %
		No	100.00 %	0.00 %	100.00 %	100.00 %	0.30 %	99.70 %	0.70 %	3.20 %	5.30 %
.5	No	Large	90.50 %	7.30 %	83.20 %	99.20 %	50.50 %	48.70 %	97.00 %	98.80 %	99.30 %
		Moderate	50.30 %	0.00 %	50.30 %	55.30 %	1.00 %	54.30 %	16.80 %	36.20 %	45.30 %
		No	70.10 %	0.00 %	70.10 %	95.30 %	0.20 %	95.10 %	1.10 %	4.60 %	7.40 %
	0.5	Large	100.00 %	19.80 %	80.20 %	100.00 %	71.00 %	29.00 %	98.80 %	99.80 %	99.80 %
		Moderate	63.90 %	0.10 %	63.80 %	68.50 %	2.40 %	66.10 %	21.20 %	43.20 %	53.40 %
		No	99.30 %	0.00 %	99.30 %	100.00 %	0.00 %	100.00 %	0.00 %	2.60 %	5.50 %
	1	Large	100.00 %	40.30 %	59.70 %	100.00 %	86.00 %	14.00 %	99.50 %	99.80 %	99.90 %
		Moderate	84.30 %	0.10 %	84.20 %	86.80 %	5.30 %	81.50 %	33.30 %	55.70 %	63.50 %
		No	100.00 %	0.00 %	100.00 %	100.00 %	0.30 %	99.70 %	0.30 %	2.80 %	5.30 %
.7	No	Large	69.90 %	1.00 %	68.90 %	95.90 %	27.20 %	68.70 %	89.60 %	95.60 %	97.30 %
		Moderate	52.70 %	0.00 %	52.70 %	57.40 %	0.70 %	56.70 %	14.10 %	33.30 %	42.10 %
		No	72.20 %	0.00 %	72.20 %	97.10 %	0.00 %	97.10 %	1.30 %	5.70 %	8.80 %
	0.5	Large	99.90 %	6.90 %	93.00 %	100.00 %	43.10 %	56.90 %	92.80 %	97.00 %	98.10 %
		Moderate	61.70 %	0.00 %	61.70 %	66.70 %	1.30 %	65.40 %	17.80 %	37.60 %	45.40 %
		No	99.30 %	0.00 %	99.30 %	100.00 %	0.10 %	99.90 %	1.00 %	3.30 %	5.30 %
	1	Large	100.00 %	17.70 %	82.30 %	100.00 %	66.40 %	33.60 %	95.70 %	98.60 %	99.10 %
		Moderate	84.80 %	0.10 %	84.70 %	87.90 %	3.80 %	84.10 %	25.80 %	48.70 %	58.00 %
		No	100.00 %	0.00 %	100.00 %	100.00 %	0.20 %	99.80 %	0.50 %	3.50 %	5.90 %

2c = two-class constrained, 3c = three-class constrained

classes are generally lower than those seen in previous research, possibly because differences in variances help to increase class separation. That this is due to decreased power due to less class separation is supported because the detection rate is increased in the multivariate model, in which the residual variances are smaller; when the two predictors are not correlated, the two-class model is selected in almost all simulations by both the BIC (100.0 %) and the ABIC (99.7 %); when the two predictors are related, which decreases class separation, the detection rate for two classes goes down to 70.1 % with the BIC, whereas it is quite high with the ABIC (>95.1 %).

For simulation conditions in which the constraint on the variance was inappropriate (i.e., a difference in variances existed, but variances were constrained to be equal; see Table 2), we hypothesized that the three-class result would be found. The results were more nuanced than this: When class separation was high, the BIC and ABIC both selected the three-class model over the one-class and two-class results in every simulation. However, when class separation decreased (i.e., with no difference between classes in the intercepts and a higher correlation between predictors, or generally when the predictors are more correlated) the models tended to select the one- or two-class solutions. In fact, the misspecified

model sometimes performed better than the correctly specified model, because the misspecification increased the probability of selecting the two- over the one-class result. The detection rate for selecting the correct number of classes was slightly decreased (BIC = 68.0 %, ABIC = 92.5 %) when the actual residual variances were moderately different (.225 difference) in the univariate model. Under the large-variance-difference condition (.45 difference), the detection rate was noticeably down to 55.6 % using the ABIC, whereas the BIC was relatively stable (68.0 %). When there were two uncorrelated predictors in the model, which had the biggest differences in residual variances between the two latent classes, a three-class model was selected in all simulations (100.0 %) by the BIC as well as the ABIC, showing that the equality constraint leads to selecting an additional latent class. On the other hand, when there is some relationship ( $r = .5$  and  $.7$ ) between the two predictors and no intercept differences between the two classes, the power to detect the additional latent class capturing the effect heterogeneity decreased.

### Model comparisons between the relaxed and restricted models

Next, we compared the restricted two-class models with the equality constraint to the relaxed two-class model with class-specific residual variances. The last three columns in Table 2 show the results of the model comparisons based on the BIC, ABIC, and SB LRT. Overall, the relaxed models were favored over the restricted models when there were large variance differences between the classes, whereas the restricted models were favored when equal variances were present. When there was a moderate difference in residual variances between the classes for the univariate model, all criteria tended to favor the restricted model. On the other hand, when the two predictors were not correlated, relaxed models were favored in most cases by all three criteria. Small differences were observed among the three model fit indices, with the BIC always selecting the restricted model more often and the ABIC selecting the relaxed model more. The SB LRT was best at choosing the relaxed model when the difference in variances was small.

### Parameter estimates

We next examined the accuracy of parameter estimates from the two-class regression mixture models when the residual variances were constrained to be equal across classes. Given that we aimed to know the consequence of constraining the variance for the parameter estimates and we knew that the population had two classes, we included all 1,000 replications in the assessment of estimation quality, rather than just the simulations in which the two-class model was selected using fit indices. Table 3 presents the results of the parameter estimates from the restricted two-class model with a single

predictor. The true population values for generating the simulated data are given in the table. Next to the true value, the mean of each parameter estimate across 1,000 replications, the standard deviation of the estimated coefficients across replications (an empirical estimate of the standard error), the mean of the estimated standard error across all simulations, and the minimum, maximum, RMSE, and coverage rates are presented.

Because they were constrained to be equal across classes, bias in the residual variances ( $\sigma^2$ ) between the classes was assured, and the observed estimates are between the two true population values. The primary purpose of this aim was to assess the consequences of residual variance constraints for the other model parameter estimates. First, the class mean (i.e., the log-odds of being in Class 1 vs. Class 2) was severely biased when the large variance difference was constrained to be equal across classes. The true value of the class mean was 0.00, reflecting the equal proportions (.50) for the two classes. Under the large-variance-difference condition, the average across simulations of the log-odds of being in Class 1 was  $-1.755$ , which corresponds to a probability of .147. In other words, when variances were constrained to be equal on average, 14.7 % of the 3,000 subjects were estimated as being in the small-effect-size class. The mean of the log-odds of class membership increased to  $-0.777$  (i.e., 31.5 % of subjects were estimated as being in Class 1) under the moderate-variance-difference condition, which is still considerably under the true value of 0. When the actual variances were equal between the classes (no-variance-difference condition), the estimated class mean was unbiased ( $-0.011$ ).

The regression coefficients of the predictor (i.e., the slopes of the regression lines) for both classes were severely biased when there was a large difference in residual variances that were constrained to be equal in estimating the model. In this case, the true regression coefficient of .20 for Class 1 was on average estimated to be  $-.22$ , which now is in the wrong direction (negative) from the true population model (positive). The average estimate of the slopes for Class 2 is also downward biased, from .70 to .57. On the other hand, the mean of the outcome variable (i.e., the intercepts of the regression lines) is correctly estimated to be zero for both classes. Because the average parameter estimates show substantial bias, the estimated standard errors are of little importance. However, we note that the standard deviation and minimum and maximum values of the parameter estimates across simulations provide the evidence of large variation and in some cases of extreme solutions, especially for the parameters estimates for Class 1.

The RMSEs increased as the magnitude of the variance difference increased (see Table 3), and were especially large for the slope coefficients in the large-variance-difference condition (RMSE for  $\beta_{11} = .42$ , RMSE for  $\beta_{12} = .13$ ), indicating that these parameters were severely biased when misspecifying the residual variances to be equal across



**Table 3** Parameter estimates from the single predictor model with equality constraints on the residual variances

Difference in Variance	Parameters	True	Estimated					
			Mean ( $SD^1$ )	$MSE^2$	Min	Max	RMSE <sup>3</sup>	Coverage <sup>4</sup> (%)
Large	Class mean <sup>5</sup>	0.000	-1.755 (0.57)	0.481	-4.994	0.315	1.76	6.4
	Intercept for Class 1 ( $\beta_{01}$ )	.000	.00 (.22)	.157	-1.940	1.495	.13	94.5
	Slope for Class 1 ( $\beta_{11}$ )	.200	-.219 (.20)	.147	-2.226	.311	.42	18.3
	Residual variance for Class 1 ( $\sigma^2_1$ )	.960	.717 (.02)	.025	.642	.799	.24	0.0
	Intercept for Class 2 ( $\beta_{02}$ )	.000	.00 (.03)	.027	-.112	.141	.02	94.5
	Slope for Class 2 ( $\beta_{12}$ )	.700	.569 (.04)	.038	.428	.693	.13	10.0
Moderate	Class mean	0.000	-0.777 (0.52)	0.519	-3.933	1.036	0.80	63.2
	Intercept for Class 1 ( $\beta_{01}$ )	.000	.00 (.08)	.083	-.623	.387	.06	97.3
	Slope for Class 1 ( $\beta_{11}$ )	-.126	-.290 (.12)	.120	-1.248	.019	.17	70.4
	Residual variance for Class 1 ( $\sigma^2_1$ )	.984	.866 (.03)	.029	.768	.943	.12	1.7
	Intercept for Class 2 ( $\beta_{02}$ )	.000	.00 (.04)	.038	-.146	.157	.03	95.9
	Slope for Class 2 ( $\beta_{12}$ )	.491	.401 (.06)	.065	.223	.682	.10	64.7
No	Class mean	0.000	-0.011 (0.51)	0.505	-1.776	2.836	0.39	91.2
	Intercept for Class 1 ( $\beta_{01}$ )	.000	.00 (.05)	.053	-.202	.195	.04	96.7
	Slope for Class 1 ( $\beta_{11}$ )	-.320	-.334 (.09)	.088	-.758	-.081	.07	91.8
	Residual variance for Class 1 ( $\sigma^2_1$ )	.897	.892 (.03)	.030	.802	1.000	.02	94.0
	Intercept for Class 2 ( $\beta_{02}$ )	.000	.00 (.05)	.053	-.195	.150	.04	96.0
	Slope for Class 2 ( $\beta_{12}$ )	.320	.329 (.09)	.089	.082	.930	.07	92.1
	Residual variance for Class 2 ( $\sigma^2_2$ )	.897	.892 (.03)	.030	.802	1.000	.02	94.0

<sup>1</sup> Standard deviation of all replications; <sup>2</sup> Mean of the estimated standard error; <sup>3</sup> RMSE = root mean squared error; <sup>4</sup> Coverage = coverage of true population value across 1,000 replications; <sup>5</sup> Class mean: small-effect-size class is the reference group

classes. The RMSE for class means indicates the extreme bias in this parameter when variances are incorrectly constrained to be equal. RMSEs for the model parameters under the equal-variance condition were small (range of .02 to .07), as they should be given that the data were generated such that the classes had equal variances in this condition.

The coverage rates for the equal-variance condition were above 90 % for all of the parameters, which indicates that the 95 % confidence interval for each parameter in the restricted model contained the true population value more than 90 % of the time. The coverage rates became worse as the difference between the two residual variances increased. When the variance difference was large, the coverage rates for the class mean, slope coefficients, and residual variances were very low; in this situation, it is unlikely that the correct inference would be made.

Table 4 presents the parameter estimates from the restricted two-class model with two strongly correlated ( $r = .7$ ) predictors and no intercept differences between the two classes. This is one scenario of nine total simulation scenarios of the multivariate model. All other result tables are available from the first author upon request, given the limited space in the article. Although the

results are not directly comparable to those from the univariate model because of differences in total effect sizes and the sizes of the residual variances, the overall results are similar. When the residual variances with large differences were constrained to be the same across classes, the regression weights for both classes are downward biased, causing the slope of the smaller effect to switch direction. The class mean is again downward biased, indicating that greater numbers of observations are incorrectly assigned to be in Class 2 (larger effect class). As expected, there is no bias in the parameter estimates when the equality constraint is held for the equal-variance conditions.

### Post hoc analyses for class identification

Previous analyses had found that an inappropriate equality constraint on the residual variances greatly impacted estimated class sizes and caused regression weights to switch directions. To better understand how this constraint impacts the model results, we examined individuals who were misclassified as a result of the constraint. This analysis used a single simulated dataset of 100,000

**Table 4** Parameter estimates from the two predictor model of .7 correlation and no intercept differences with equality constraint

Difference in Variance	Parameters	True	Estimated					
			Mean ( $SD^1$ )	$MSE^2$	Min	Max	RMSE <sup>3</sup>	Coverage <sup>4</sup> (%)
Large	Class mean <sup>5</sup>	0.000	-1.562 (0.59)	0.452	-4.158	2.651	1.562	7.4
	Intercept for Class 1 ( $\beta_{01}$ )	.000	-.002 (.14)	.131	-.779	.728	.002	96.0
	Slope 1 for Class 1 ( $\beta_{11}$ )	.118	-.093 (.18)	.178	-1.482	.316	.211	74.1
	Slope 2 for Class 1 ( $\beta_{21}$ )	.118	-.087 (.18)	.177	-.853	.885	.205	72.0
	Residual variance for Class 1 ( $\sigma^2_1$ )	.959	.708 (.02)	.025	.625	.780	.251	0.0
	Intercept for Class 2 ( $\beta_{02}$ )	.000	.002 (.04)	.029	-.135	.769	.002	95.4
	Slope 1 for Class 2 ( $\beta_{12}$ )	.412	.339 (.04)	.043	.225	.487	.073	52.5
	Slope 2 for Class 2 ( $\beta_{22}$ )	.412	.327 (.10)	.043	-.833	.467	.085	51.3
	Residual variance for Class 1 ( $\sigma^2_2$ )	.495	.708 (.02)	.025	.625	.780	.213	0.0
Moderate	Class mean	0.000	-0.704 (0.42)	0.421	-2.888	0.601	.704	57.9
	Intercept for Class 1 ( $\beta_{01}$ )	.000	-.004 (.07)	.073	-.266	.300	.004	96.6
	Slope 1 for Class 1 ( $\beta_{11}$ )	-.074	-.162 (.11)	.109	-.814	.128	.088	88.9
	Slope 2 for Class 1 ( $\beta_{21}$ )	-.074	-.155 (.11)	.110	-.802	.224	.081	89.2
	Residual variance for Class 1 ( $\sigma^2_1$ )	.984	.860 (.03)	.029	.760	.950	.124	1.2
	Intercept for Class 2 ( $\beta_{02}$ )	.000	.001 (.04)	.037	-.126	.143	.001	97.0
	Slope 1 for Class 2 ( $\beta_{12}$ )	.289	.239 (.06)	.058	.091	.475	.050	82.0
	Slope 2 for Class 2 ( $\beta_{22}$ )	.289	.242 (.06)	.057	-.004	.483	.047	83.0
	Residual variance for Class 1 ( $\sigma^2_2$ )	.751	.860 (.03)	.029	.760	.950	.109	2.1
No	Class mean	0.000	0.020 (0.43)	0.426	-2.178	1.373	.020	93.2
	Intercept for Class 1 ( $\beta_{01}$ )	.000	.00 (.05)	.049	-.204	.153	.000	95.2
	Slope 1 for Class 1 ( $\beta_{11}$ )	-.188	-.189 (.07)	.077	-.436	.031	.001	95.1
	Slope 2 for Class 1 ( $\beta_{21}$ )	-.188	-.191 (.08)	.077	-.621	.444	.003	93.8
	Residual variance for Class 1 ( $\sigma^2_1$ )	.894	.887 (.03)	.029	.790	1.003	.007	93.7
	Intercept for Class 2 ( $\beta_{02}$ )	.000	.00 (.05)	.051	-.368	.178	.001	96.3
	Slope 1 for Class 2 ( $\beta_{12}$ )	.188	.188 (.08)	.079	-.004	.573	.005	95.4
	Slope 2 for Class 2 ( $\beta_{22}$ )	.188	.188 (.08)	.079	-.509	.499	.006	94.9
	Residual variance for Class 1 ( $\sigma^2_2$ )	.894	.894 (.03)	.029	.790	1.003	.007	93.7

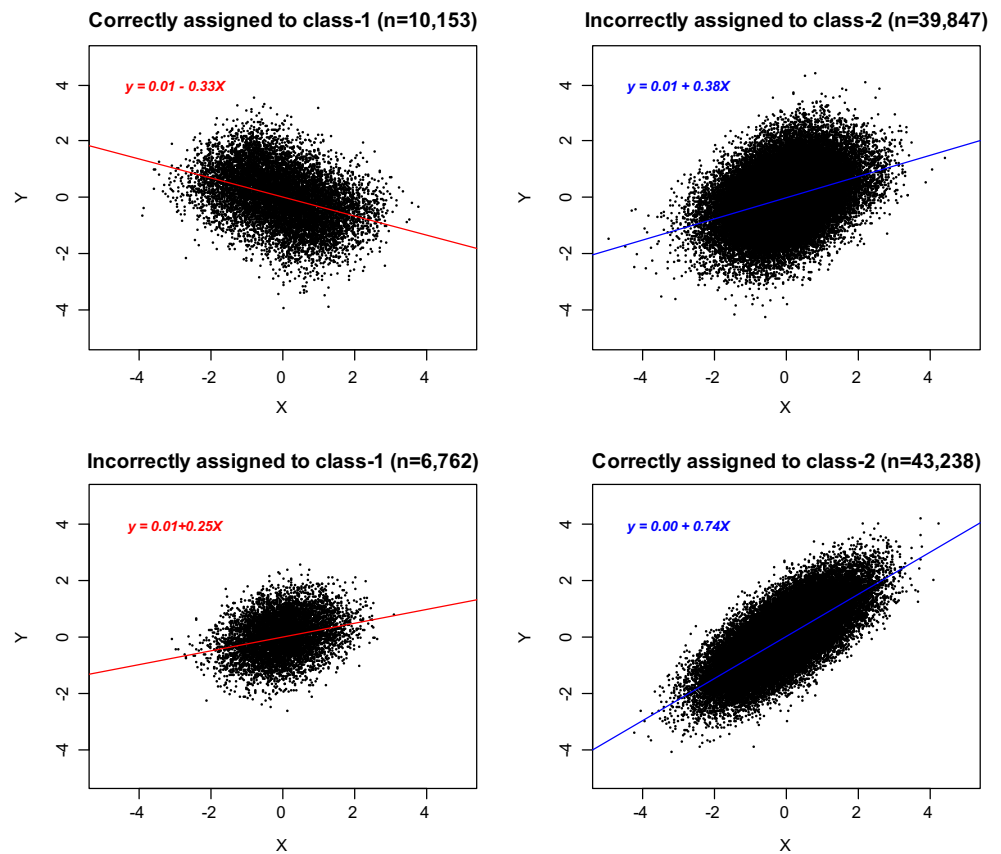
Note. <sup>1</sup> Standard deviation of all replications; <sup>2</sup> Mean of the estimated standard error; <sup>3</sup> RMSE = root mean squared error; <sup>4</sup> Coverage = coverage of true population value across 1,000 replications; <sup>5</sup> Class mean: small-effect-size class is the reference group.

subjects to fit the two-class restricted mixture model in which data were generated under the large-variance-difference condition. We then assigned individuals to latent classes using a pseudoclass draw (Bandein-Roche, Miglioretti, Zeger, & Rathoutz, 1997), in which individuals were assigned to each class with a probability equal to the model-estimated posterior probability of being in that class. Because the data were simulated, we also knew the true class assignment for each individual. We then examined which individuals were correctly versus incorrectly assigned as a result of constraining the residual variances. Figure 1 presents a scatterplot with a regression line for each of the four groups defined by the true and estimated class memberships. As can be seen in the figure, a considerable number of Class-1 subjects (about 80 %) were incorrectly assigned to Class 2, whereas most of the subjects in Class 2 (above

86 %) remained in the same class. The relationship between the predictor and outcome in Class 1 was now changed from a positive ( $\beta_{11} = .20$ ) to a negative ( $\beta_{11} = -.18$ ) direction.

This reassignment of individuals helps to explain the mechanism through which constraining residual variances leads to bias. Because the effect size is stronger in Class 2, Class 2 dominates the estimation. Extreme values from Class 1, which show the strong positive relationship between  $X$  and  $Y$ , are moved to Class 2 because the variance in Class 2 is forced to be increased. At the same time, the residual variance of Class 1 is reduced by allocating those extreme cases to Class 2. Because those who followed an upward slope in Class 1 have now been moved to Class 2, the remaining individuals follow a downward slope (seen in the first scatterplot of Fig. 1), and the effect

**Fig. 1** Assignments of individual observations when holding an equality constraint on the residual variances under the large-variance-difference condition



of  $X$  on  $Y$  in Class 1 has now effectively changed direction. The individuals who are incorrectly assigned to Class 1 have low variance because the variability in Class 1 must be decreased and the variability in Class 2 must increase. This demonstrates how a simple misspecification of residual variances can cause estimates of the regression weights to switch signs.

## Discussion

Regression mixture models allow the investigation of differential effects of predictors on outcomes. Although they have recently been applied to a range of research, the effect of misspecifying the class-specific residual variances has remained unknown. In this study, we examined the impacts of constraining the residual variances on the latent-class enumeration and on the accuracy of the parameter estimates, and found effects on both class enumeration and the class-specific regression estimates.

Class enumeration was not affected by the equality constraint when the residual variances were truly the same. As differences in the residual variances across classes increased, the detection rate for selecting the correct number of classes decreased. The ABIC seemed to be more sensitive to the misspecification of the residual variances, which is similar to

the findings of Enders and Tofghi (2008) when looking at growth mixture models. However, the differences in information criteria between the competing models were very small in many cases. In practice, an investigator who finds a very small difference in a penalized information criterion will need to use other methods to determine the correct number of classes; in this case, if the two-class model shows two large classes with meaningfully different regression weights between the classes, then the researcher would be correct to choose the two-class solution even if the BIC and ABIC slightly favored the one-class result.

These results for latent-class enumeration help to put previous research comparing the indices for class enumeration into perspective. Previous research with regression mixtures has found that in situations in which there was a large difference in variances between classes (as we used in this article) and a large sample size (6,000), the BIC performed very well and the ABIC showed no advantages (George et al., 2013). This study showed that none of these indices perform as well when the sample size is somewhat lower: With large differences in residual variances and the correct model specification, the two-class model was supported less than 90 % of the time; when the differences in residual variances were moderate or zero, the two-class model was supported using the ABIC, and the support for two classes was strongest when residuals were constrained to be equal. Thus, we do not

recommend using either the BIC or the ABIC as the sole criterion to decide the number of classes. Along with the information criteria, class proportions, the regression weights for each latent class, and previous research should be taken into account when deciding the number of latent classes.

The results for the parameter estimates were clearer than those for latent-class enumeration: Parameter estimates showed substantial bias in both class proportions and regression weights when the class-specific variances were inappropriately constrained. This is consistent with previous research evaluating the effects of the misspecification of variance parameters in other types of mixture models (Enders & Tofighi, 2008; McLachlan & Peel, 2000). Moreover, we hypothesized that the impact of misspecifying the error variance structure on the parameter estimates would be much more severe in regression mixture models than in growth mixture models. As expected, although there was relatively minor bias in parameter estimates in growth mixture models (Enders & Tofighi, 2008), we found substantial bias in the regression weights for both latent classes in regression mixture models. In light of these results, a reasonable recommendation is that in regression mixture models, residual variances should be freely estimated in each class by default unless models with constrained variances fit equally well and there are no substantive differences in the parameter estimates.

Although these simulations showed no problems with estimating class-specific variances, in practice there will be situations in which estimation problems emerge when class-specific variances are specified. One option is to compare models in which variances are constrained to be equal to those in which they are constrained to be unequal (such as the variance of Class 1 equaling half the variance of Class 2). If no model clearly fits the data better and when other model parameters change substantially, then any results should be treated with great caution.

As with most simulation studies, this study was limited to examining only a small number of conditions. Specifically, we limited the study to a two-class model with a 50/50 split in the proportion of subjects in each class, a sample size of 3,000, and constant effect size differences between the two classes.

The main design factor for this study was the amount of difference in the residual variances and intercepts between the two classes, as well as the degree of relationship between the two predictor variables. When class separation is stronger than in our simulation conditions (as indicated by larger class differences in the regression weights or intercepts, or more outcome variables), the models should perform better. The purpose of this study was to demonstrate the potential effects of inappropriate constraints on residual variances, so the actual effects in any one condition may differ substantially from those found here; however, this illustrates the potentially strong impact of misspecification of residual variances in regression mixtures. Users of regression mixture models should be aware of the potential for finding effects that are opposite from the true effects when the residuals, which are of little importance to most users, are misspecified.

**Author note** This research was supported by Grant Number R01HD054736, M.L.V.H. (PI), funded by the National Institute of Child Health and Human Development.

## Appendix 1

**Table 5** Parameter values for generating the multivariate population model

$r_{x1 \times x2}$	Difference in Variance	$\beta_{11}$	$\beta_{21}$	$\sigma^2_1$	$\beta_{11}$	$\beta_{21}$	$\sigma^2_2$
0	Large	.2	.2	.92	.7	.7	.02
	Moderate	-.126	-.126	.968	.491	.491	.518
	No	-.32	-.32	.795	.32	.32	.795
0.5	Large	.133	.133	.956	.467	.467	.456
	Moderate	-.084	-.084	.982	.327	.327	.732
	No	-.213	-.213	.886	.213	.213	.886
0.7	Large	.118	.118	.959	.412	.412	.495
	Moderate	-.074	-.074	.984	.289	.289	.751
	No	-.188	-.188	.894	.188	.188	.894

## Appendix 2: R code for generating the data

```
# Single predictor with large-variance-difference condition #
dat <- matrix(NA, ncol = 3, nrow = 3000)
dat[1:1500,3] <- -1
dat[1501:3000,3] <- -2
for(i in 1:1000){
  dat[,1] <- rnorm(3000)
  dat[1:1500,2] <- -dat[1:1500,1]*(-0.32) + rnorm(1500,sd = sqrt(0.898))
  dat[1501:3000,2] <- -dat[1501:3000,1]*0.32 + rnorm(1500,sd = sqrt(0.898))
}
```

```

write.table(dat,paste(C:/Temp/data',i',.dat',sep = ""), col.names = FALSE, row.names = FALSE)
}
Mplus code for analyzing regression mixture model with
equality constraint
#constraining the residual variances (by default)#
Title: 2-class model with an equality constraint;
  data: file = C:/Temp/data1.dat;
  variable:
  NAMES = X Y Group;
  USEVARIABLES = X Y;
  CLASSES = c(2);
  analysis:
    type = mixture;
    starts = 100 20;
  model:
  %overall%
  Y on X;
  Y;
  %c#2 %
  Y on X;
  ! Y;          !constraining the variance by not writing out this statement (by default)
  Output:
  TECH14;

```

## References

- Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., & Rathoutz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, *92*, 1375–1386.
- Bolck, A., Croon, M., & Hagenars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, *12*, 3–27.
- Daepfen, J., Faouzi, M., Sanglier, T., Sanchez, N., Coste, F., & Bertholet, N. (2013). Drinking patterns and their predictive factors in control: A 12-month prospective study in a sample of alcohol-dependent patients initiating treatment. *Alcohol and Alcoholism*, *48*, 189–195.
- Desarbo, W. S., Jedidi, K., & Sinha, I. (2001). Customer value analysis in a heterogeneous market. *Strategic Management Journal*, *22*, 845–857.
- Development Core Team, R. (2010). *R: A language and environment for statistical computing (Version 2.10)*. Vienna: R Foundation for Statistical Computing.
- Ding, C. S. (2006). Using regression mixture analysis in educational research. *Practical Assessment Research & Evaluation*, *11*(11), 1–11. Retrieved from <http://pareonline.net/getvn.asp?v=11&n=11>
- Enders, C. K., & Tofghi, D. (2008). The impact of misspecifying class-specific residual variances in growth mixture models. *Structural Equation Modeling*, *15*, 75–95. doi:10.1080/10705510701758281
- Fagan, A. A., Van Horn, M. L., Hawkins, J. D., & Jaki, T. (2013). Differential effects of parental controls on adolescent substance use: For whom is the family most important? *Journal of Quantitative Criminology*, *29*, 347–368. doi:10.1007/s10940-012-9183-9
- Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika*, *10*, 507–521.
- George, M. R. W., Yang, N., Van Horn, M. L., Smith, J., Jaki, T., Feaster, D. J., & Maysn, K. (2013). Using regression mixture models with non-normal data: Examining an ordered polytomous approach. *Journal of Statistical Computation and Simulation*, *83*, 757–770.
- Lanza, S. T., Kugler, K. C., & Mathur, C. (2011). Differential effects for sexual risk behavior: An application of finite mixture regression. *Open Family Studies Journal*, *4*, 81–88.
- Lanza, S. T., Cooper, B. R., & Bray, B. C. (2013). Population heterogeneity in the salience of multiple risk factors for adolescent delinquency. *Journal of Adolescent Health*, *54*, 319–325. doi:10.1016/j.jadohealth.2013.09.007
- Lee, E. J. (2013). Differential susceptibility to the effects of child temperament on maternal warmth and responsiveness. *Journal of Genetic Psychology*, *174*, 429–449.
- Liu, M., & Lin, T. (2014). A skew-normal mixture regression model. *Educational Psychological Measurement*, *74*, 139–162. doi:10.1177/0013164413498603
- Liu, Y., & Lu, Z. (2011). The Chinese high school student's stress in the school and academic achievement. *Educational Psychology*, *31*, 27–35. doi:10.1080/01443410.2010.513959
- McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.
- Muthén, B. O., & Asparouhov, T. (2009). Multilevel regression mixture analysis. *Journal of the Royal Statistical Society, Series A*, *172*, 639–657.
- Muthén, L. K., & Muthén, B. O. (1998–2012). Mplus (Version 7.1). Los Angeles, CA: Muthén & Muthén.



- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*, 535–569. doi:10.1080/10705510701575396
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*, 507–514.
- Schmeige, S. J., Levin, M. E., & Bryan, A. D. (2009). Regression mixture models of alcohol use and risky sexual behavior among criminally-involved adolescents. *Prevention Science, 10*, 335–344.
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464. doi:10.1214/aos/1176344136
- Sclove, L. S. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika, 52*, 333–343.
- Silinskas, G., Kiuru, N., Tolvanen, A., Niemi, P., Lerkkanen, M.-K., & Nurmi, J.-E. (2013). Maternal teaching of reading and children's reading skills in Grade 1: Patterns and predictors of positive and negative associations. *Learning and Individual Differences, 27*, 54–66.
- Smith, J., Van Horn, M. L., & Zhang, H. (2012). *The effects of sample size on the estimation of regression mixture models*. Vancouver: Paper presented at the American Educational Research Association.
- Tofighi, D., & Enders, C. K. (2008). *Identifying the correct number of classes in growth mixture models*. Greenwich: Information Age.
- Van Horn, M. L., Jaki, T., Masyn, K., Ramey, S. L., Antaramian, S., & Lemanski, A. (2009). Assessing differential effects: Applying regression mixture models to identify variations in the influence of family resources on academic achievement. *Developmental Psychology, 45*, 1298–1313.
- Van Horn, M. L., Smith, J., Fagan, A. A., Jaki, T., Feaster, D. J., Masyn, K., ... Howe, G. (2012). Not quite normal: Consequences of violating the assumption of normality in regression mixture models. *Structural Equation Modeling, 19*, 227–249. doi:10.1080/10705511.2012.659622
- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis, 18*, 450–469.
- Wedel, M., & Desarbo, W. S. (1994). A review of recent developments in latent class regression models. In R. P. Bagozzi (Ed.), *Advanced methods of marketing research* (pp. 352–388). Malden: Blackwell Publishers.
- Wedel, M., & Desarbo, W. S. (1995). A mixture likelihood approach for generalized linear models. *Journal of Classification, 12*, 21–55.
- Wong, Y. J., & Maffini, C. S. (2011). Predictors of Asian American adolescents' suicide attempts: A latent class regression analysis. *Journal of Youth and Adolescence, 40*, 1453–1464.
- Wong, Y. J., Owen, J., & Shea, M. (2012). A latent class regression analysis of men's conformity to masculine norms and psychological distress. *Journal of Counseling Psychology, 59*, 176–183. doi:10.1037/a0026206
- Yau, K. K., Lee, A. H., & Ng, A. S. (2003). Finite mixture regression model with random effects: Application to neonatal hospital length of stay. *Computational Statistics and Data Analysis, 41*, 359–366.