A NOVEL MIXED EFFECTS MODELING FRAMEWORK FOR LONGITUDINAL ORDINAL SUBSTANCE USE DATA

James S. McGinley

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology (Quantitative).

Chapel Hill
2014

Approved by:

Daniel J. Bauer

Patrick J. Curran

Amy H. Herring

Andrea M. Hussong

David Thissen

ABSTRACT

JAMES S. MCGINLEY: A Novel Mixed Effects Modeling Framework for Longitudinal Ordinal Substance Use Data
(Under the direction of Patrick J. Curran)

Substance use is a serious public health concern. Despite advances in the theoretical conceptualization of within and between person pathways to substance use, researchers are limited by the longitudinal models currently available. Researchers often fit linear mixed effects models (L-MM) and proportional odds mixed effects models (PO-MM) to longitudinal ordinal data with many response categories defined by collapsed count data (e.g., 0 drinking days, 1-2 days, 3-6 days, etc.). Consequently, existing models ignore the underlying count process, resulting in a disjoint between the construct of interest and the models being fitted. My proposed novel ordinal-count mixed effects modeling framework overcomes this limitation by explicitly linking ordinal responses to a suitable underlying count distribution. In doing so, researchers can fit ordinal negative binomial mixed effects models (ONB-MM) and ordinal zero-inflated negative binomial mixed effects models (OZINB-MM) to ordered data as if they had directly observed the underlying discrete counts. The utility the ONB-MM and OZINB-MM was verified by simulation studies. The simulation studies demonstrated that the proposed ordinal-count models recovered the underlying unobserved count process across a range of conditions that may arise in the study of substance use. Results also showed the advantages of the proposed ordinal-count mixed effects models compared to existing L-MM and PO-MM. In sum, my ordinal-count mixed effects modeling framework offers several quantitative and substantive advantages over currently available methods.

To Dad and Jessica.

ACKNOWLEDGEMENTS

This project was funded by NIDA (Grant F31 DA033688). I would like to thank my wife, Melissa, for supporting me through my graduate career despite not knowing what I do. I am also thankful for the exceptional graduate students and faculty of the L. L. Thurstone Psychometric Lab. I am especially grateful for the guidance provided by my mentor, Dr. Patrick Curran.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1: INTRODUCTION

Addiction research often aims to describe how between and within person differences in substance use behaviors unfold over time. For example, clinical psychologists have posited ways a variety of mental health constructs such as depression, stress, and anxiety are related to substance use (Fleming, Mason, Mazza, Abbott, & Catalano, 2008; Hussong, Jones, Stein, Baucom, & Boeding, 2011; King, Molina, & Chassin, 2009). Public health researchers have examined demographic differences in longitudinal trends of alcohol consumption (Moore et al., 2005). Other topics such as the influence of peers on substance use intersect multiple disciplines including public health, psychology, and human development (Curran, Stice, & Chassin, 1997; Norton, Lindrooth, & Ennett, 1998; Simons-Morton & Chen, 2006).

Despite diversity in the goals and focuses of substance use research, there exists a common need for flexible longitudinal data analytic techniques. Over the past two decades, accessible pedagogical references and advances in user-friendly statistical software have led to an increase in the utilization of longitudinal models for continuous outcomes in substance use research (e.g., Curran, 2000; Hedeker, Gibbons, & Flay, 1994; Singer & Willett, 2003). However, in most situations, these linear models are inappropriate because indicators of substance use constructs are largely discrete counts such as quantity or frequency of use.

Advances in Modeling Count Substance Use Outcomes

Statisticians have outlined several longitudinal mixed effects models for count data such as the Poisson, negative binomial, and zero-inflated mixed effects models (Hall, 2000; Hur, Hedeker, Henderson, Khuri, & Daley, 2002; Min & Agresti, 2005; Siddiqui, 1996, Yau, Wang,

& Lee, 2003). These models are slowly being integrated into substance use research with the help of recent dissemination efforts (e.g., Atkins, Baldwin, Zheng, Gallop, & Neighbors, 2013; Desantis et al., 2013; Xie, Tao, McHugo, & Drake, 2013). Most of these count modeling approaches have yet to be rigorously evaluated for the study of substance use, but they show significant advantages over widely used linear models. For example, substance use count data usually show common distributional characteristics such as being positively skewed with a large number of responses at the lower bound of zero (e.g., no use). Standard linear mixed effects models are not well-suited to handle these types of distributions and, consequently, can lead to violations of model assumptions (e.g., normality of errors), poor model fit, and invalid predictions (e.g., negative predictions). In turn, these shortcomings restrict substance use researchers' ability to effectively test theoretically derived research hypotheses.

Count mixed effects models are ideal for the distributional characteristics of substance use data and are, consequently, less prone to the limitations of linear mixed effects models. Further, there have been several useful methodological developments in the area of longitudinal models for zero-inflated count data, which are counts that exhibit an abundance of zeroes beyond what would be expected by standard count distributions (Hall, 2000; Liu & Powers, 2007; Min & Agresti, 2005). These zero-inflated models usually address inflation from a mixture framework with a binary process (modeled through a logistic or probit regression) and a count process (modeled through standard or truncated Poisson or negative binomial regression). As a result, zero-inflated methods are different from typical count models from both quantitative and substantive standpoints. These methods are especially promising for adolescent substance use research because subjects frequently report no drug use over the assessed time-periods. A small number of studies have applied zero-inflated mixed effects models to adolescent substance use count data (e.g., Buu, Li, Tan, & Zucker, 2012; Otten, van Lier, & Engels, 2011), but these models are still relatively new. It is clear that count mixed effects models provide advantages over standard linear mixed effects models for modeling substance use data. However, an ignored and complicating characteristic of real substance data is that count constructs are often collected as ordinal items, not open-ended responses.

Ordinal-Count Substance Use Outcomes

In practice, substance use researchers collect what I call ordinal-count data instead of open-ended discrete counts. Ordinal-count data are ordinal data in which the response categories represent unique ranges of counts (e.g., 1-2 days, 3-5 days, etc.). This approach is recommended by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and has been used in large scale studies such as Monitoring the Future (MTF) and the Health Behaviour of School Aged Children (HBSC) (Currie et al., 2012; Johnston et al., 2012; NIAAA, 2003). Table 1 provides a sample alcohol use frequency item used by the MTF and HBSC studies.

Table 1. *An example past 30 day alcohol frequency item used by the MTF and HBSC studies*

On how many occasions did you drink alcohol in the last 30 days?
0. 0 Occasions
1. 1 to 2 Occasions
2. 3 to 5 Occasions
3. 6 to 9 Occasions
4. 10 to 19 Occasions
5. 20 to 39 Occasions
6. 40+ Occasions

While these ordinal-count items may offer measurement benefits (e.g., minimizing participant burden and recall error), they pose a significant challenge for statistical modeling. For instance, ordinal-count items usually have a large number of categories (e.g., 5-12 response categories) and the distributions of those ordinal responses mirror that of the underlying unobserved counts (e.g., positive skew and large number of responses in the first "zero" category). These characteristics make modeling ordinal-count substance use data challenging. Substance use researchers typically use standard linear mixed effects models or ordinal (e.g., proportional odds) mixed effects models to accommodate the nesting of time within individuals, but both of these models have significant limitations.

Existing Models for Ordinal-Count Substance Use Data

Linear Mixed Effects Model (L-MM)

The linear mixed effects model (L-MM) is likely the most popular method for handling ordinal-count data. Let Y_{it} denote the observed ordinal-count response for a sample of $i = 1,2,\ldots,N$ individuals at $t = 1,2,\ldots,T_i$ time-points. The basic L-MM can be expressed as

$$Y_{it} = \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{z}'_{it}\boldsymbol{\nu}_i + \varepsilon_{it} \tag{1}$$

where $\boldsymbol{x}_{it}$ is a p x 1 vector of covariates (typically including "1" as the first element for the intercept) with the corresponding p x 1 vector of regression coefficients, $\boldsymbol{\beta}$. Additionally, $\boldsymbol{z}_{it}$ represents a r x 1 vector of random effect variables (again, usually including "1" as the first element for the intercept) with a corresponding r x 1 vector of random subject effects, $\boldsymbol{\nu}_i$. The errors, ε_{it}, are assumed to be normally distributed with a mean of 0 and variance of σ^2 and the random effects are assumed to follow a multivariate normal distribution with a r x 1 mean vector $\mathbf{0}$ and a r x r variance-covariance matrix, $\boldsymbol{\Sigma}$.

The L-MM has two major shortcomings for ordinal-count substance use data. First, the substantive interpretation of model parameters is usually ambiguous, assuming that the true construct of interest is the underlying substance use counts. Researchers must apply scores to the observed ordinal data in order to fit linear mixed effects models. The most commonly applied scoring approach is to use the integers representing the response categories as ordinal scores. For example, using the item from Table 1, the ordinal scores used in linear models would be integers ranging from 0-6. Upon close inspection, it is clear that these scores provide little information about the underlying count construct. Further, applying the same scoring approach to different ordinal-count items can imply measurement differences. For instance, suppose that a second study also assessed alcohol frequency, but the response categories differed (e.g., 0=0 occasions, 1=1 occasion, 2=2 occasions, 3=3 occasions, 4=4 occasions, 5= 5-6 occasions, 6=7 or more occasions). This is an extreme example but it clearly illustrates that the same integer scores imply different levels of substance use depending on the measure, which impedes the accumulation of research findings. Most important, inferences are made relative to the arbitrarily defined ordinal scoring metric, not the count construct of true substantive interest.

The second limitation of linear mixed effects models applied to ordinal-count substance use data is that model assumptions can easily be violated. As previously stated, the distribution of ordinal-count items often looks similar to that of the true underlying count process. As a result, fitting linear mixed effects models to ordinal-count data is similar to fitting linear models to count data. This, in turn, can lead to violations of key model assumptions such as normally distributed errors. Additional negative consequences of fitting linear models to ordinal-count data may include erroneous inferences and invalid predictions (e.g., negative predicted values). Previous work has identified several reasons why ordinal models are preferable to linear models for ordered categorical data (e.g., Hedeker & Gibbons, 2006; McKelvey & Zavoina, 1975; Winship & Mare, 1984). However, the limitations outlined here demonstrate that these concerns are further compounded with ordinal-count data because standard linear mixed effects models are not optimal from quantitative or substantive standpoints.

Proportional Odds Mixed Effects Model (PO-MM)

The ordinal proportional odds mixed effects model (PO-MM) is an alternative modeling strategy used by substance use researchers for ordinal-count data (Hedeker & Gibbons, 1994; Hedeker, & Mermelstein, 2000). The logistic PO-MM for ordinal data with $c = 1,2, \ldots, M$ response categories can be expressed in terms of cumulative log-odds (or logits) as

$$log\left[\frac{P_{itc}}{1 - P_{itc}}\right] = \tau_c + [\boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{z}'_{it}\boldsymbol{\nu}_i] \quad (c = 1, \ldots, M - 1) \tag{2}$$

where $log[.]$ is the natural log function, $P_{ijc} = \Pr(Y_{it} \geq c)$, τ_c are the M - 1 intercepts, $\boldsymbol{x}_{it}$ is a p x 1 vector of covariates with the corresponding p x 1 vector of regression coefficients, $\boldsymbol{\beta}$. Additionally, $\boldsymbol{z}_{it}$ represents a r x 1 vector of random effect variables (again, usually including "1" as the first element for the intercept) with a corresponding r x 1 vector of random subject effects, $\boldsymbol{\nu}_i$. The random effects, $\boldsymbol{\nu}_i$, are assumed to be distributed $MVN(\mathbf{0}, \boldsymbol{\Sigma})$.

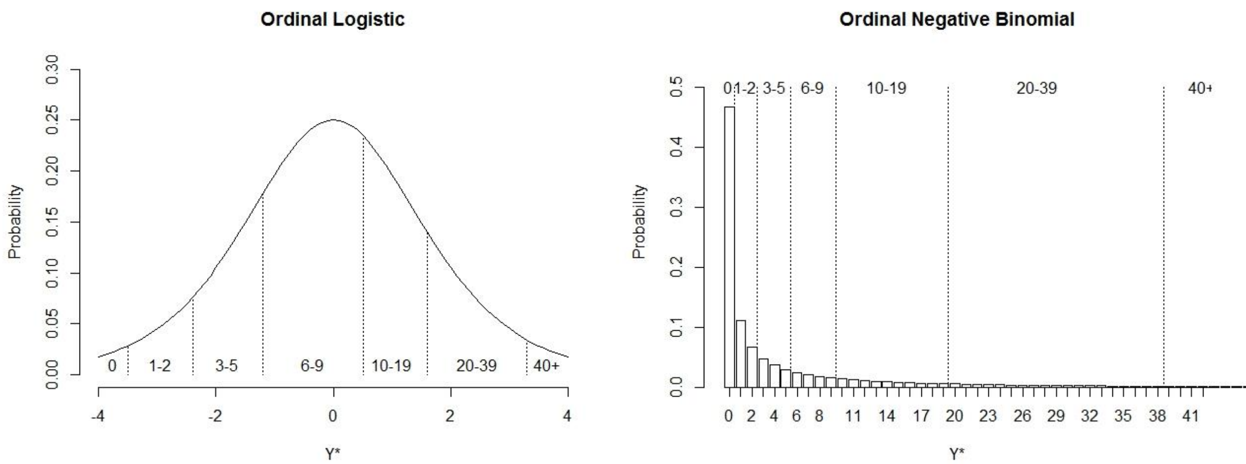
There are several reasons why these mixed effects models are not ideal for substance use research. First, like the linear mixed effects model, the proportional odds mixed effects model does not provide ideal substantive interpretations. The model parameters are interpreted at the level of the observed response categories and reflect covariate effects across the cumulative log-odds. As a result, these models provide no substantive insight on the underlying count process.

Second, the standard PO-MM is not parsimonious. Ordinal-count substance use data often have many response categories (e.g., 5-12 categories). This requires researchers to fit proportional odds mixed effects models with many parameters because M -1 intercepts (or thresholds) must be estimated. Third, the key assumption of proportional odds may not hold for substance use ordinal-count data. The proportional odds assumption requires that the same covariate effect holds across all of the cumulative logits. This assumption may be particularly untenable in substance use research given the large number of response categories. For instance, with a ten category ordinal-count scale, covariates would be assumed to have exactly the same effect across nine different cumulative logits.

From a substantive standpoint, violations of the proportional odds assumption could be expected. For example, it is conceivable that the effect of covariates on the log-odds of use versus non-use (e.g., first category vs. all other categories) could be different from other cumulative log-odds comparisons. In this case, the proportional odds assumption would not hold. Alternative models such as non-proportional odds and partial proportional odds mixed effects models exist, but they add significant complexity in estimation and produce results that are difficult to interpret. In order for a model to be consistent with the count construct of actual substantive interest, the parameter estimates should reflect how changes in covariates impact the mean of the underlying unobserved count construct.

Fourth, from a conceptual standpoint, the latent response variable motivation for the proportional odds model is inconsistent with the true nature of ordinal-count data. Although this latent response variable motivation is not a strict requirement for the PO-MM (McCullagh & Nelder, 1989), it is widely used for pedagogical purposes in substance use research so it should be addressed (Feldman, Masyn, & D. Conger, 2009; Hedeker & Mermelstein, 2000). The standard latent variable motivation states that underlying the observed ordinal responses is an unobserved latent variable, Y^*, that follows a logistic distribution (ordered logistic). This underlying continuous latent variable is then divided by estimated thresholds, which are the negative of the intercepts, corresponding to the response categories. The left panel of Figure 1 displays a visualization of the standard latent response variable motivation assuming a seven category substance use measure like in Table 1.

Figure 1. *Comparison of underlying logistic and negative binomial latent response variables.*



Note: Cut-points are typically estimated in the ordinal logistic whereas they are treated as fixed and known in the ordinal negative binomial.

Conceptually, there is a clear disjoint between this underlying logistic/normal latent variable and the ordinal-count scale that is known to represent collapsed counts. Thus, ordinal-count data should be conceptualized from an underlying count latent response variable motivation as shown in the right panel of Figure 1. The count latent response variable is characterized by fixed and known thresholds, or intercepts, (denoted by the vertical bars in Figure 1) and assumes that underlying the observed ordinal responses is an unobserved count variable, Y^*. In Figure 1, the underlying latent count variable follows a negative binomial distribution, but other distributions are plausible. In sum, these shortcomings illustrate that the mixed effects proportional odds model lacks the flexibility necessary to handle ordinal-count substance use data. For this reason, we recently developed the ordinal-count modeling framework.

Proposed Ordinal-Count Mixed Effects Modeling Framework

Statisticians and economists have discussed methods for handling grouped, or collapsed, count data for over four decades (Carter, Bowen, & Myers, 1971; Carter & Myers, 1973; Moffatt, 1995; Moffatt & Peters, 2000). McGinley, Curran, and Hedeker (under review) recently analytically and empirically demonstrated how this work can be expanded and cast more broadly within an ordinal modeling perspective motivated by an underlying latent variable methodology. Our ordinal-count modeling framework explicitly links the observed ordinal responses to a suitable underlying count distribution through the known response scale cut-points. In doing so, researchers can fit count models to ordinal-count data as if the underlying discrete counts were actually observed. This methodology appropriately matches the statistical model with the true substance use constructs of interest and permits thorough evaluations of intricate hypotheses.

My dissertation significantly expands on this earlier work in two ways. First, I outline two innovative longitudinal ordinal-count models that include random effects so that the models explicitly account for the clustering of time-points within individuals. I focus on an ordinal negative binomial mixed effects model (ONB-MM) and an ordinal zero-inflated negative binomial mixed effects model (OZINB-MM) but, the logic that I describe holds for applying other mixed effects count models to ordinal-count data. To my knowledge, no previous studies have proposed longitudinal ordinal models with random effects that assume an underlying count distribution instead of the typical underlying continuous normal or logistic distributions. Second, I evaluate and discuss results from simulation studies that considered several critical factors directly relevant to longitudinal substance use research. These factors include different numbers of time-points, ordinal-count response scales, and data generating mechanisms. In summary, my dissertation provides a comprehensive evaluation of two novel longitudinal models that are highly applicable to substance use research.

Ordinal Negative Binomial Mixed Effects Model (ONB-MM)

In order to model the ordinal-count responses as a function of a negative binomial distribution, we first assume that underlying the observed ordinal responses, Y_{it}, is an unobserved count latent variable, Y_{it}^*. The probability mass function (PMF) for the negative binomial distribution can then be expressed in terms of ${Y_{it}}^*$

$$f(y_{it}^*) = P(\mathrm{Y}_{it}^* = y_{it}^*) = \frac{\Gamma(y_{it}^* + \alpha^{*-1})}{\Gamma(\alpha^{*-1})\Gamma(y_{it}^* + 1)} \left(\alpha_{y^*}^{-1}\mu_{it}^*\right)^{y_{it}^*} (1 + \alpha^{*-1}\mu_{it}^*)^{-(y_{it}^* + \alpha^{*-1})}, \quad y_{it}^* = 0, 1, 2, \ldots \tag{3}$$

where $E(Y_{it}^*) = \mu_{it}^*$, $Var(Y_{it}^*) = \mu_{it}^* + \alpha^*\mu_{it}^{*2}$ and α^* is the dispersion parameter for the negative binomial distribution underlying the observed ordinal responses. The negative binomial distribution reduces to the better known Poisson distribution when the dispersion parameter is zero (e.g., α^*= 0). The cumulative distribution function (CDF) for the negative binomial distribution is the sum of the PMFs such that

$$F(y_{it}^*) = \sum_{v=0}^{y_{it}^*} f(v) \tag{4}$$

where the cumulative probability is evaluated at y_{it}^*. We then can use the natural log function to model the mean of Y_{it}^* as a function of covariates.

$$\log(\mu_{it}^*) = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{z}_{it}'\boldsymbol{v}_i \tag{5}$$

where $\boldsymbol{x}_{it}$ is a p x 1 vector of covariates (typically including “1” as the first element for the intercept) with the corresponding p x 1 vector of regression coefficients, $\boldsymbol{\beta}$. Additionally, $\boldsymbol{z}_{it}$ represents a r x 1 vector of random effect variables (again, usually including “1” as the first element for the intercept) with a corresponding r x 1 vector of random subject effects, $\boldsymbol{v}_i$. The random effects are assumed to follow a multivariate normal distribution with a r x 1 mean vector $\mathbf{0}$ and a r x r variance-covariance matrix $\mathbf{\Sigma}$. We can also solve for μ_{it}^{*} using the inverse link

$$\mu_{it}^{*} = e^{\boldsymbol{x}'_{it}\boldsymbol{\beta}+\boldsymbol{z}'_{it}\boldsymbol{v}_i} \tag{6}$$

Next, we link the observed ordinal outcome, Y_{it}, to the unobserved underlying count latent variable, Y_{it}^{*}. This linking procedure occurs through the fixed and known cut-points, which are defined by the ordinal response scale. Specifically,

$$Y_{it} = c \quad if\ \kappa_{c-1} < Y_{it}{}^{*} \leq \ \kappa_c \tag{7}$$

where κ_c is the count number that defines the upper bound of ordinal response category *c*. We can thus express the probability of observing a response in category *c* for subject *i* at time *t* as a function of the cumulative probabilities from the underlying Y_{it}^{*} distribution.

$$P(Y_{it} = c) = P(\kappa_{c-1} < Y_{it}{}^{*} \leq \kappa_c) = F(\kappa_c) - F(\kappa_{c-1}) \tag{8}$$

with $F(\kappa_c)$ and $F(\kappa_{c-1})$ designating the CDFs evaluated at the known upper count numbers for categories *c* and *c-1* for a distribution with a mean of μ_{it}^{*} and dispersion of α^{*}.

The ordinal negative binomial mixed effects model can be estimated using marginal maximum likelihood estimation (MML). The marginal likelihood for the ONB-MM is

$$L_{ONB} = \prod_{i=1}^{N} \int l_{ONB}(\boldsymbol{Y}_i|\boldsymbol{v})g(\boldsymbol{v})d\boldsymbol{v} \tag{9}$$

where $l_{ONB}(\boldsymbol{Y}_i|\boldsymbol{v})$ is the conditional likelihood for the ordinal-count data using and underlying negative binomial distribution (hence the *ONB* subscript) weighted by the multivariate normal density for random effects, $g(\boldsymbol{v})$. For the conditional likelihood, $l_{ONB}(\boldsymbol{Y}_i|\boldsymbol{v})$, we expand on Equation 8, which expresses the probability of observing a response in category *c* for subject *i* at time *t* as function of the cumulative probabilities from the underlying Y_{it}^{*} distribution. Letting $(y_{it1}, \ldots, y_{itc})$ represent binary reference codes indicative of the response for subject *i* at time *t* (e.g., $y_{itc} = 1$ if $Y_{it} = c$ and 0 otherwise), the conditional likelihood for the ordinal data following an underlying negative binomial distribution is

$$l_{ONB}(\boldsymbol{Y}_i|\boldsymbol{v}) = \prod_{t=1}^{T_i}\left[\prod_{c=1}^{M}[F(\kappa_c) - F(\kappa_{c-1})]^{y_{itc}}\right] \tag{10}$$

The log-likelihood is often used for optimization, which is the natural log of Equation 9.

$$LogL_{ONB} = \sum_{i=1}^{N}\log(L_{ONB}) \tag{11}$$

Computation of the integral over the random effects can be done using adaptive Gaussian quadrature. Next, I describe how the ordinal zero-inflated negative binomial model can be extended to accommodate random effects.

Ordinal Zero-Inflated Negative Binomial Mixed Effects Model (OZINB-MM)

In substance use research, especially involving adolescents, researchers encounter ordinal-count data with a large proportion of participants reporting "no use" over the assessed time-period. For example, consider an outcome that represents the number of drinking occasions over the past 30 days. Many adolescent will report "zero days" for this outcome. In the count literature, these data are often referred to as being zero-inflated. Over the past several decades, statisticians have proposed zero-inflated Poisson (ZIP) and negative binomial models (ZINB) for cross-sectional data that exhibit this characteristic (Greene, 1994; Lambert, 1992). The ZIP and ZINB models function similarly except the ZINB extends the ZIP by incorporating the dispersion parameter, which allows the variance of the count response to differ from its mean. These models assume a mixture distribution with a binary process (e.g., logistic or probit regression) and a count process (e.g., Poisson or negative binomial regression). As a result, zeroes arise from two distinct sources; individuals that are not at-risk for a given outcome (e.g., lifetime non-smokers report no smoking) or individuals who are at-risk that do not experience the outcome over the assessment period (e.g., smokers who did not smoke over the assessed timeframe). Various sources have demonstrated how the ZIP and ZINB can be extended to accommodate random effects (e.g., Hall, 2000; Hur, Hedeker, Henderson, Khuri, & Daley, 2002; Min & Agresti, 2005, Yau, Wang, & Lee, 2003), but no prior studies have shown how these models can be used with ordinal-count data.

We can logically extend the outlined ONB-MM to account for zero-inflation. Similar to Lambert (1992), we can express the ordinal zero-inflated negative binomial mixed effects model (OZINB-MM) with regard to the underlying unobserved ${Y_{it}}^*$

$$\begin{array}{ll} {Y_{it}}^* \sim 0 & \textit{with probability} \;\; \pi_{it} \\ {Y_{it}}^* \sim f(y_{it}^*) & \textit{with probability} \;\; (1-\pi_{it}) \end{array} \tag{12}$$

where $f(y_{it}^*)$, is the negative binomial PMF from Equation 3 and π_{it} is the probability of being an inflated, or "excess", zero for subject *i* at time *t*. The PMF can then be expressed as

$$\begin{aligned} P(Y_{it}^* = 0) &= \pi_{it} + (1-\pi_{it})f(0) \\ P(Y_{it}^* = y_{it}^*) &= (1-\pi_{it})f(y_{it}^*), \quad y_{it} = 1, 2, \dots \end{aligned} \tag{13}$$

where $f(0)$ is the probability that $Y_{it}^* = 0$ based on the negative binomial PMF shown in Equation 3.

I will model the binary zero process through a logistic model.

$$\text{logit}(\pi_{it}) = \boldsymbol{w}_{it}'\boldsymbol{\gamma} + \boldsymbol{m}_{it}'\boldsymbol{u}_i \tag{14}$$

where $\boldsymbol{w}_{it}$ is a q x 1 vector of covariates (typically including "1" as the first element for the intercept) with the corresponding q x 1 vector of regression coefficients, $\boldsymbol{\gamma}$. Additionally, $\boldsymbol{m}_{it}$ represents a s x 1 vector of random effect variables (again, usually including "1" as the first element for the intercept) with a corresponding s x 1 vector of random subject effects, $\boldsymbol{u}_i$. We can also solve for π_{it} using the inverse link

$$\pi_{it} = \frac{1}{1 + e^{-(\boldsymbol{w}'_{it}\boldsymbol{\gamma} + \boldsymbol{m}'_{it}\boldsymbol{u}_i)}} \tag{15}$$

The count process for the OZINB-MM is modeled as explained in Equations 5 and 6. It is important to highlight that the count and binary processes can have the same, or different, predictors. The random effects are assumed to follow a multivariate normal distribution with a $(r + s)$ x 1 mean vector $\mathbf{0}$ and a $(r + s)$ x $(r + s)$ variance-covariance matrix, $\boldsymbol{\Sigma}$. This is because there are r count-part random effects and s logistic-part random effects, all of which potentially covary. Clearly, this leads to a substantial increase in model complexity. For instance, consider a plausible theoretical model in which there are random intercepts and slopes on the count negative binomial process and the logistic process. In this case, $\boldsymbol{\Sigma}$ is a symmetric 4 x 4 matrix.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{v_0} & & & \\ \sigma_{v_1 v_0} & \sigma^2_{v_1} & & \\ \sigma_{u_0 v_0} & \sigma_{u_0 v_1} & \sigma^2_{u_0} & \\ \sigma_{u_1 v_0} & \sigma_{u_1 v_1} & \sigma_{u_1 u_0} & \sigma^2_{u_1} \end{bmatrix} \tag{16}$$

Thus, we would need to estimate four variances and six covariances. However, a reasonable way to simplify the random effects structure is to exclude the random slopes from both the count and logistic processes (e.g., Hedeker & Gibbons, 2006). This results in correlated random intercepts from the count and logistic processes and $\boldsymbol{\Sigma}$ is a symmetric 2 x 2 matrix.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{v_0} & \\ \sigma_{u_0 v_0} & \sigma^2_{u_0} \end{bmatrix} \tag{17}$$

I focus on this random effects structure for my dissertation because it is used in applied research, but other structures are possible. Although outside the scope of my dissertation, a much needed direction for future research is investigating how to specify these random effects structures for longitudinal substance use research.

The OZINB-MM can be estimated using marginal maximum likelihood estimation (MML). The marginal likelihood is

$$L_{OZI} = \prod_{i=1}^{N} \int l_{ozi}(\boldsymbol{Y}_i|\boldsymbol{\theta})g(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{18}$$

where $l_{OZI}(\boldsymbol{Y}_i|\boldsymbol{\theta})$ is the conditional likelihood for the ordinal-count data assuming an underlying zero-inflated negative binomial distribution weighted by the multivariate normal density for random effects, $g(\boldsymbol{\theta})$. I used $\boldsymbol{\theta}$ to denote random effects from the count, $\boldsymbol{v}$, and binary processes, $\boldsymbol{u}$. The conditional likelihood, $l_{OZI}(\boldsymbol{Y}_i|\boldsymbol{\theta})$, is

$$l_{OZI}(\boldsymbol{Y}_i|\boldsymbol{\theta}) = \prod_{t=1}^{T_i}\left[\prod_{c=1}^{M}\left[(1-\pi_{it})\big(F(\kappa_c) - F(\kappa_{c-1})\big) + I(Y_{it} = 1)\pi_{it}\right]^{y_{itc}}\right] \tag{19}$$

where $I(Y_{it} = 1)$ is an indicator function that equals one when $Y_{it} = 1$ (e.g. when the first response category representing zero is selected) and zero otherwise.

The marginal log-likelihood is then simply the natural log of Equation 18.

$$LogL_{OZI} = \sum_{i}^{N} \log(L_{OZI}) \tag{20}$$

Like the ONB-MM, computation of the integral over the random effects can be done using adaptive Gaussian quadrature. In sum, I have outlined how the novel ordinal-count modeling framework can be extended to accommodate random effects with specific emphasis on the ONB-MM and OZINB-MM. Next, I discuss the specific research goals of my dissertation.

Current Research

I had three general goals for the current research. The first goal was to establish that the proposed ordinal-count mixed effects modeling framework effectively recovers the true underlying count process. Second, I evaluated which situations improve or worsen the performance of the proposed ordinal-count mixed effects models. Third, I compared the proposed ordinal-count mixed effects models to existing mixed effects models used for substance use research. Taken together, these three components demonstrated the usefulness of these novel models for substance use research above and beyond statistical methods currently available.

My simulation studies evaluated the performance of ordinal-count models with random effects for longitudinal substance use research across a number of conditions. These ordinal-count models are novel and, to my knowledge, have not been discussed in the quantitative or substantive literatures. In my simulations, I considered important simulation factors such as different underlying data generating mechanisms, sample sizes, number of time-points, and response scales. In addition to the ONB-MM and OZINB-MM, I evaluated the standard linear mixed effects model (L-MM) and proportional odds mixed effects models (PO-MM) for longitudinal ordinal-count data. Despite the previously outlined shortcomings of L-MM and PO-MM, these models were important to consider since they are widely used by substance use researchers. My simulation studies focused on testing four specific hypotheses. The first two hypotheses pertained to the ONB-MM and the second two focused on the OZINB-MM.

Hypotheses Regarding the ONB-MM

First, I hypothesized that the ONB-MM fitted to longitudinal ordinal-count data would accurately recover the population generating values from the underlying count process across all conditions. Both sample size and number of time-points were expected to influence the ONB-MM model performance in predictable ways. More precisely, the ONB-MM should have higher convergence rates, more accurate estimates, and smaller standard errors as the number of time-points and subjects increase (e.g., n=1000 and t=7). With a smaller number of time-points (e.g., t=3), the ONB-MM should have more non-converged models because of difficulty in estimating the random slope variance component. Similar convergence issues were also expected to arise with the PO-MM and L-MM. I did not anticipate that the number of response categories (5pt scale versus 10pt scale) would impact the accuracy of parameter estimates. However, I did expect slightly larger standard errors for the 5pt scale compared to the 10pt scale due to the reduced amount of available information.

Second, I hypothesized that the ONB-MM would outperform the existing PO-MM and L-MM. I evaluated this hypothesis by examining model fit, empirical power, Type I error, and relative efficiency of predictions. In terms of model fit, I expected the ONB-MM to have better model fit criteria (e.g., lower AIC and BIC values) compared PO-MM and L-MM fitted to the same ordinal-count data. The proposed ONB-MM was hypothesized to have superior model fit because it was the only model that properly specified the underlying count data generating mechanism. I also expected the PO-MM to provide better model fit than the L-MM because it explicitly accounts for the ordinal nature of the response variable. I did not expect large differences in empirical power among the ONB-MM, PO-MM, and L-MM. However, it was plausible for the ONB-MM and L-MM to have slightly higher empirical power than the PO-MM because less total parameters were estimated (the PO-MM required the estimation of several intercepts). The predictions produced by the ONB-MM should be superior to both the PO-MM and the L-MM because the ONB-MM accounted for the true underlying count structure. I expected the PO-MM to again outperform the L-MM because the PO-MM accounted for the ordinal scale.

Hypotheses Regarding the OZINB-MM

Third, like the ONB-MM, I hypothesized that the OZINB-MM fitted to longitudinal zero-inflated ordinal-count data would accurately recover the underlying zero-inflated count process across all conditions. The performance of the OZINB-MM was expected to be best with more subjects and time-points. Given the structure of the random effects (random intercepts in the zero and count processes with no random time effects), the smaller number of time-points (e.g., t=3) should not impact model convergence with a large sample size (e.g., n=1000). However, I anticipated convergence issues when both the sample size and number of time-points are small (e.g., n=300, t=3) because of the high degree of model complexity. Similar to the ONB-MM, I expected smaller standard errors with increased sample size and time-points. I did not expect the number of response categories to impact model performance in any substantively meaningful way beyond slightly smaller standard errors.

Fourth, I hypothesized that the OZINB-MM would outperform the existing PO-MM and L-MM. Like the ON-MM, I evaluated this hypothesis by investigating model fit, empirical power, Type I error, and relative efficiency of predictions. Because the PO-MM and L-MM were fitted with correlated random intercept and time effects and the zero-inflated data were generated without a random time effect, I expected low rates of model convergence for these models compared to the OZINB-MM. I anticipated that the OZINB-MM would have substantially better model fit compared PO-MM and L-MM because the OZINB-MM explicitly accounted for the underlying zero-inflated process.

Evaluating empirical power and Type I errors across the OZINB-MM, PO-MM, and L-MM was not straightforward because the models have substantially different interpretations. However, I hypothesized that the PO-MM and L-MM could produce unexpected patterns of significant effects (e.g., spurious effects). For example, a predictor with no effect in the generating zero-inflated population model may have elevated Type I error rates when fitting the PO-MM or L-MM because of model misspecification and violations of key model assumptions.

Relatedly, I expected empirical power for the PO-MM and L-MM to be heavily influenced by model misspecification (e.g., ignoring the zero process). For instance, suppose that increases in predictor A lead to increases in the log-odds of being non-user (e.g. zero process). However, among users, increases in predictor A lead to decreases in use (e.g., count process). This pattern of effects should reduce the level of empirical power for predictor A in the L-MM and PO-MM compared to the count process in the OZINB-MM. This is because the effect in the binary process effectively washes out the effect in count process when fitting the L-MM and PO-MM. The opposite pattern could also occur (e.g., decrease in log-odds/increase in use) and result in higher levels of empirical power for the L-MM and PO-MM compared to the OZINB-MM count process. Lastly, I posited that predictions from the OZINB-MM would be superior to those produced by the PO-MM and L-MM because the OZINB-MM explicitly models the true underlying zero-inflated count structure. The PO-MM was hypothesized to outperform the L-MM because it accounts for the ordinal nature of the response variable.

Summary of Current Research

In order to test my research hypotheses, I conducted two simulation studies. The first simulation study evaluated hypotheses one and two which pertained to the ONB-MM. The second simulation study evaluated hypotheses three and four which related to the OZINB-MM. I next provide a brief summary of the studies and how they address my research hypotheses.

I generated longitudinal ordinal-count data from an underlying negative binomial distribution for the first simulation study and longitudinal ordinal-count data from an underlying zero-inflated negative binomial distribution for the second simulation study. In both studies, I incorporated important design factors such as response scale, sample size, and number of time-points. Additionally, I considered commonly used PO-MM and L-MM to compare my proposed modeling framework to current best practices in substance use research.

I evaluated hypotheses one and three, which focused on the general performance of the ON-MM and OZINB-MM, using measures of bias and accuracy for the parameter estimates. These measures provided insight on how well the proposed ordinal-count mixed effects models recovered the true underlying count process of substantive interest. I evaluated hypotheses two and four, which compared the proposed ordinal-count mixed effects models to existing models, by investigating model fit (e.g., -2 log-likelihood, AIC, BIC), empirical power, Type I error, and relative efficiency.

Collectively, my simulation studies investigated ordinal outcomes with underlying counts across a variety of conditions consistent with substance use research. They provided information about the general utility of the proposed models and helped to identify factors that improve and worsen model performance. The simulations studies also served as a means for assessing the advantages of the proposed ordinal-count methods over models widely used in practice.

CHAPTER 2: STUDY 1: EVALUATING THE ORDINAL NEGATIVE BINOMIAL MIXED EFFECTS MODEL (ONB-MM)

I evaluated the proposed ONB-MM by simulating longitudinal ordinal-count data that were consistent with those observed in adult substance use studies (e.g., ages from mid-twenties into early thirties). I then fitted ONB-MM, PO-MM, and L-MM to the ordinal-count data. My first hypothesis focused on evaluating whether or not the proposed ONB-MM could recover the true underlying count population generating values even though the response variable was ordinal. I hypothesized that the ONB-MM would adequately recover the population generating values from the underlying negative binomial count process. I expected the ONB-MM to perform best (e.g., highest convergence rates, smallest standard errors, most accurate estimates) with more subjects and time-points (e.g., n=1000, t=7). To quantify the general performance of the ONB-MM, I examined several measures of bias and accuracy. I expected that, with less data, model performance would deteriorate slightly but still be acceptable. I also hypothesized that the ONB-MM would provide superior fit to the ordinal-count data and produce more efficient predictions relative to the existing proportional odds and linear mixed effects models.

The longitudinal ordinal-count data were generated from a negative binomial distribution with correlated random intercepts and time effects. The selection of my population generating parameters was guided by count mixed effects models fitted to real empirical substance use data. I included time-varying predictors, time-invariant predictors, and time varying-by-time invariant predictor interactions in the population generating model. Given my proposed hypotheses, I employed eight unique condition combinations. There were two sample sizes (n=1000, 300), two numbers of time-points (t=3, 7), and two response scales (5pt, 10pt). I next describe the specifics of my simulation design with regards the population generating model, data generation, fitted models, and simulation outcome measures.

Simulation Study Design

The population generating parameters were partially based on results obtained by fitting count negative binomial mixed effects models to adult (e.g., ages 24 to 30) frequency of past 30 day alcohol use data from the 1997 National Longitudinal Survey of Youth (NLSY97). These parameters value were intended to represent those likely obtained in practice by adult substance use researchers. In order to establish a range of effects, the fixed effects values were based on empirical power levels derived by fitting count mixed effects models in my preliminary pilot work. The population model included correlated random intercept and linear time effects. The covariates included a time-varying linear time predictor, t_t, two standard normal continuous time-invariant predictors x_{1i} and x_{2i} correlated .3, and interactions between the time-invariant predictors and time. A summary of population generating parameters is listed in the first column of Table 4.

For each condition, there were r=250 replications. I used SAS 9.3 to generate open-ended longitudinal count data from a negative binomial distribution and these counts were collapsed into ordinal-count responses according the two unique response scales in Table 2. I fitted ordinal negative binomial (ONB-MM), proportional odds (PO-MM), and linear (L-MM) mixed effects models to each of the datasets. All of the models had properly specified fixed and random effects and were fitted in PROC NLMIXED in SAS using adaptive quadrature with five quadrature points and dual quasi-Newton optimization. This resulted in a total of 6,000 fitted models. The variance components for the random effects were modeled in terms of standard deviations to help with convergence in PROC NLMIXED. This is common practice with these types of models (Kiernan, Tao, & Gibbs, 2012).

Table 2. *Ordinal response scales used in the simulation studies.*

Five Point Scale	Ten Point Scale
On how many occasions did you drink alcohol in the last 30 days?	**On how many occasions did you drink alcohol in the last 30 days?**
0. Never	0. Never
1. 1 to 2 Occasions	1. 1 Occasion
2. 3 to 6 Occasions	2. 2 Occasions
3. 7 to 14 Occasions	3. 3 Occasions
4. 15+ Occasions	4. 4 to 5 Occasions
	5. 6 to 7 Occasions
	6. 8 to 10 Occasions
	7. 11 to 15 Occasions
	8. 16 to 24 Occasions
	9. 25+ Occasions

I evaluated my first hypothesis that focused on the general performance of the ONB-MM, using several criteria including raw bias (Bias), standardized bias (SB), root mean square error (RMSE), and 95% CI coverage probabilities. These measures provided valuable insights on how well the proposed ordinal-count mixed effects model recovered the true underlying count process of substantive interest. The bias measures were critical for evaluating the deviation in estimates from the true population value. Raw bias represented the difference between the mean estimate and the population value (e.g., $\bar{\hat{\beta}} - \beta$). Researchers have suggested that raw bias may be problematic at levels anywhere from one-half to two times the empirical standard error (Rubinstein, 1981; Schafer & Graham, 2002). The standardized bias measure calculates bias as a function of the empirical standard error. This provided a measure of bias that depends directly on the uncertainty in the parameter estimate (e.g., $\left(\frac{\bar{\hat{\beta}}-\beta}{SE(\hat{\beta})}\right)*100$). Researchers have suggested that a standardized bias of +/- 40 percent can be problematic (Collins, Schafer, & Kam, 2001). The RMSE provided a measure of overall accuracy that incorporates both bias and variability on the same scale as the parameter (e.g., $\sqrt{\left(\bar{\hat{\beta}}-\beta\right)^2+\left(SE(\hat{\beta})\right)^2}$). The 95% CI coverage provided an indicator of the proportion of times the confidence interval contains the true population value. This helped determine if the proposed ordinal-count mixed effects model produces results that are too liberal or conservative with regard to Type I error rate and power.

I evaluated my second hypothesis that compared the proposed ordinal negative binomial mixed effects models to existing models by examining model fit, empirical power, Type I error, and relative efficiency of predictions. For model fit, I examined the average -2 log-likelihoods, Akaike information criterion (AIC), and Bayesian information criterion (BIC) across the replications with properly converged solutions (Akaike, 1973; Schwarz, 1978). Empirical power was computed by recording the proportion of significant effects across the converged replications using a standard alpha level of .05. The relative efficiency of predictions was computed by taking the respective ratios of the efficiency of the existing PO-MM and L-MM to the proposed ONB-MM. Efficiency was computed by taking the average of $\sum_i \sum_t \left(\hat{Y}_{it}-\mu_{it}\right)^2$ across the converged replications. Here, μ_{it} denotes the $E[Y_{it}|\boldsymbol{x}_{it}, \boldsymbol{\nu}_i]$ based on the true population generating model. In the L-MM, $\hat{Y}_{it}$ was the predicted value of the ordinal-count outcome scored as category numbers (e.g., 0, 1, 2, etc.) for person *i* at time *t*. For the ONB-MM and PO-MM, $\hat{Y}_{it}$ was $\sum_c Y_{itc}\hat{p}_{itc}$ where $\hat{p}_{itc}$ was the subject specific predicted probability of person *i* at time *t* being in category c. Predictions were computed using the parameter estimates and empirical Bayes estimates of the random effects. Since I defined relative efficiency as a ratio of efficiencies (e.g., $RE_{\frac{PO}{ONB}}$, $RE_{\frac{L}{ONB}}$), a relative efficiency greater than 1 suggests that the proposed ordinal-count models are preferred.

I used meta-models to help examine the performance of the proposed ONB-MM. The mixed effects meta-models included random intercepts because different models were fitted to the same datasets and the 5pt and 10pt response scales had the same underlying count data. I created a pseudo Cohen's f^2 to provide a measure of local effect size for linear meta-models fitted to raw bias. Specifically, Cohen's f^2 is defined as $f^2 = \frac{R_{AB}^2 - R_A^2}{1 - R_{AB}^2}$ where B is the variable of interest above and beyond the other set of variables designated as A. I defined R^2 as $R^2 = \frac{V_{null} - V_{full}}{V_{null}}$ where V_{null} is the total variance (between variance + within variance) for the intercept-only model and V_{full} is the total variance for the full model that had several variants depending on if R_A^2 or R_{AB}^2 was calculated for a given predictor. I compared this measure of effect size to the standard Cohen's f^2 from standard linear regression models ignoring clustering and the general results did not meaningfully change. Cohen (1988) suggested that f^2 values 0.02, 0.15, and 0.35 constitute small, moderate, and large effects. The meta-models evaluating empirical power were logistic mixed effects models that modeled the probability of obtaining a significant effect. The meta-models considered simulation factors such as sample size, number of time-points, number of response categories, model fitted, and their respective interactions. In building the meta-models, I included all condition main effects and significant interactions using an alpha .01 to coarsely adjust for multiple testing and high statistical power.

Taken together, my simulation study provided a thorough evaluation of the proposed ordinal negative binomial mixed effects model for longitudinal substance use research. In addition, my simulation offered insights on how the proposed model performs relative to widely utilized existing techniques. Finally, this research considers critical factors such as response scales, sample size, and number of time-points to help determine the necessary conditions for effectively implementing this novel longitudinal modeling approach.

Results

Overall Performance of ONB-MM

Model non-convergence was only notable when the number of time-points was small (e.g., t=3) and the convergence rate improved with added subjects. The rates of convergence differed slightly by number of response categories. When n=300, the 5pt scale had a smaller number of converged models compared to the 10pt scale (5pt: 69.6%, 10pt: 70.8%). When n=1000, the 5pt scale had a larger number of converged models compared to the 10 points scale (5pt: 86.8%, 10pt: 83.3%).

Table 3 through Table 5 show that, across all conditions, the ONB-MM recovered the assigned parameter values adequately. Results showed small biases, RMSE, and close to 95% coverage rates. The meta-model results displayed in Table 3 show little impact of the simulation conditions on raw bias. The most notable effect was an interaction between sample size and number of time-points on bias in the random time standard deviation ($p<.0001$, f^2=.04). Specifically, when n=300, the number of time-points had a small-moderate effect such that more time-points predicted less bias ($p<.0001$, f^2=.09). However, when n=1000, the number of time-points did not predict raw bias (p=.58, f^2=0). Examination of Table 4 and Table 5 shows that bias in estimating the random time effect standard deviation was large when n=300 and t=3 and decreases with more observations and subjects. A similar pattern also occurred with the dispersion parameter and random intercept-time covariance, but these trends did not appear to be overly problematic. As expected, the standard errors were smaller as the number of time-points and subjects increased. Results suggested that the number of response categories did not impact model performance with regard to bias, efficiency, or coverage.

Table 3. *Results from linear mixed effects meta-models fitted to raw bias for the ONB-MM.*

	β_0			$\beta_1: (t_t)$			$\beta_2: (x_{1i})$			$\beta_3: (x_{2i})$			$\beta_4: (t_t x_{1i})$			$\beta_5: (t_t x_{2i})$		
	F(1,851)	**p**	f^2	**F(1,852)**	**p**	f^2	**F(1,852)**	**p**	f^2	**F(1,852)**	**p**	f^2	**F(1,852)**	**p**	f^2	**F(1,852)**	**p**	f^2
N	.20	.67	0	1.75	.19	.002	.48	.49	.001	.22	.64	0	.81	.37	0	.16	.69	0
TP	.04	.84	0	2.29	.13	.002	1.42	.23	.002	2.19	.14	.002	.03	.85	0	.14	.70	0
Scale	4.99	.03	0	1.36	.24	0	.02	.88	0	.06	.81	0	.11	.73	0	.44	.51	0
N*Scale	7.37	.007	0	-		-	-	-	-	-	-	-	-		-	-	-	-

	$\sigma_{v_{00}}$			$\sigma_{v_{11}}$			$\sigma_{v_{01}}$			α		
	F(1,851)	**p**	f^2	**F(1,852)**	**p**	f^2	**F(1,851)**	**p**	f^2	**F(1,852)**	**p**	f^2
N	.05	.83	0	36.21	<.0001	.07	9.99	.002	.02	3.23	.07	.004
TP	.57	.45	0	50.82	<.0001	.09	12.80	.0004	.02	19.39	<.0001	.02
Scale	3.78	.05	0	4.38	.04	.001	4.07	.04	0	2.01	.16	0
N*Scale	7.21	.007	0	-	-	-	6.84	.003	.0003	-	-	-
N*TP	-	-	-	40.36	<.0001	.04	9.01	.009	.01	-	-	-

Note: N denotes sample size and TP denotes time-points. f^2 denotes the pseudo Cohen's f^2.

Table 4. *Recovery of population generating values for the ONB-MM when n=300.*

	***n*=300**												
	3 Time-points												
	5 Response Categories							**10 Response Categories**					
	True	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**
$\boldsymbol{\beta_0}$	**1.00**	1.00	0.12	0.00	2.07	0.12	0.95	1.01	0.12	0.01	4.98	0.12	0.94
$\boldsymbol{\beta_1: (t_t)}$	**-0.05**	-0.06	0.06	-0.01	-10.44	0.06	0.95	-0.06	0.06	-0.01	-13.01	0.06	0.95
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	-0.01	0.11	-0.01	-10.32	0.11	0.98	-0.01	0.11	-0.01	-12.19	0.11	0.94
$\boldsymbol{\beta_3: (x_{2i})}$	**0.22**	0.23	0.12	0.01	12.86	0.12	0.94	0.23	0.11	0.01	9.06	0.11	0.95
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.03**	0.03	0.06	0.00	-0.58	0.06	0.94	0.03	0.06	0.00	5.62	0.06	0.94
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.05	0.00	-2.18	0.05	0.96	0.00	0.06	0.00	2.31	0.06	0.95
$\boldsymbol{\alpha}$	**0.75**	0.72	0.09	**-0.03**	**-32.30**	0.09	0.96	0.73	0.08	**-0.02**	**-28.88**	0.09	0.97
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.42	0.12	0.01	9.29	0.12	0.95	1.41	0.11	0.00	2.96	0.11	0.97
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.29	0.10	**0.06**	**60.56**	0.12	0.94	0.28	0.09	**0.05**	**59.02**	0.10	0.95
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.19	0.14	-0.04	-28.72	0.15	0.97	-0.18	0.12	-0.03	-27.05	0.12	0.98
	7 Time-points												
$\boldsymbol{\beta_0}$	**1.00**	1.00	0.10	0.00	-3.61	0.10	0.94	1.00	0.10	0.00	-2.13	0.10	0.96
$\boldsymbol{\beta_1: (t_t)}$	**-0.05**	-0.05	0.02	0.00	1.75	0.02	0.93	-0.05	0.02	0.00	0.03	0.02	0.94
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.00	0.10	0.00	5.00	0.10	0.97	0.00	0.10	0.00	4.10	0.10	0.97
$\boldsymbol{\beta_3: (x_{2i})}$	**0.22**	0.22	0.10	0.00	-4.20	0.10	0.96	0.22	0.10	0.00	-4.75	0.10	0.95
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.03**	0.03	0.02	0.00	-2.30	0.02	0.94	0.03	0.02	0.00	-1.17	0.02	0.94
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.02	0.00	-8.78	0.02	0.95	0.00	0.02	0.00	-8.49	0.02	0.96
$\boldsymbol{\alpha}$	**0.75**	0.75	0.05	0.00	3.28	0.05	0.95	0.75	0.05	0.00	2.44	0.05	0.94
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.40	0.09	-0.01	-6.33	0.09	0.94	1.40	0.09	-0.01	-10.57	0.09	0.94
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.23	0.02	0.00	-8.28	0.02	0.95	0.23	0.02	0.00	-10.39	0.02	0.96
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.15	0.04	0.00	-4.12	0.04	0.96	-0.15	0.04	0.00	0.40	0.04	0.96

Note. Est is the average estimate, SE is the empirical standard error, Bias is raw bias, SB is standardized bias, RMSE is root mean squared error, and 95% CI is the coverage for the 95% CI. Values are bolded to highlight particularly interesting results.

Table 5. *Recovery of population generating values for the ONB-MM when n=1000.*

	***n*=1000**												
	3 Time-points												
	5 Response Categories							**10 Response Categories**					
	True	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**
$\boldsymbol{\beta_0}$	**1.00**	1.00	0.07	0.00	-7.54	0.07	0.95	0.99	0.06	-0.01	-8.64	0.06	0.96
$\boldsymbol{\beta_1: (t_t)}$	**-0.05**	-0.05	0.03	0.00	-1.03	0.03	0.96	-0.05	0.03	0.00	1.65	0.03	0.96
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.01	0.06	0.01	8.56	0.06	0.95	0.00	0.06	0.00	3.78	0.06	0.94
$\boldsymbol{\beta_3: (x_{2i})}$	**0.22**	0.22	0.06	0.00	-4.20	0.06	0.97	0.22	0.06	0.00	2.92	0.06	0.97
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.03**	0.03	0.03	0.00	-7.71	0.03	0.92	0.03	0.03	0.00	-6.51	0.03	0.91
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.03	0.00	-5.37	0.03	0.95	0.00	0.03	0.00	-7.92	0.03	0.96
$\boldsymbol{\alpha}$	**0.75**	0.74	0.05	-0.01	-18.43	0.05	0.96	0.74	0.05	-0.01	-16.56	0.05	0.94
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.41	0.07	0.00	-4.49	0.07	0.93	1.41	0.07	0.00	-6.72	0.07	0.94
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.24	0.07	0.01	11.36	0.07	0.96	0.23	0.07	0.00	3.01	0.07	0.95
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.16	0.08	-0.01	-7.15	0.08	0.96	-0.15	0.07	0.00	-5.64	0.07	0.94
	7 Time-points												
$\boldsymbol{\beta_0}$	**1.00**	1.00	0.05	0.00	-1.39	0.05	0.96	1.00	0.05	0.00	-0.07	0.05	0.96
$\boldsymbol{\beta_1: (t_t)}$	**-0.05**	-0.05	0.01	0.00	-0.41	0.01	0.97	-0.05	0.01	0.00	-1.68	0.01	0.97
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.00	0.05	0.00	-1.01	0.05	0.97	0.00	0.05	0.00	-2.13	0.05	0.97
$\boldsymbol{\beta_3: (x_{2i})}$	**0.22**	0.22	0.06	0.00	3.76	0.06	0.94	0.22	0.06	0.00	3.87	0.06	0.94
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.03**	0.03	0.01	0.00	-2.17	0.01	0.92	0.03	0.01	0.00	-1.82	0.01	0.94
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.01	0.00	-2.88	0.01	0.95	0.00	0.01	0.00	-4.37	0.01	0.95
$\boldsymbol{\alpha}$	**0.75**	0.75	0.03	0.00	7.70	0.03	0.94	0.75	0.03	0.00	7.75	0.03	0.92
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.41	0.05	0.00	2.83	0.05	0.98	1.41	0.04	0.00	2.54	0.04	0.97
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.23	0.01	0.00	-4.66	0.01	0.96	0.23	0.01	0.00	-8.65	0.01	0.96
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.15	0.02	0.00	-0.60	0.02	0.98	-0.15	0.02	0.00	0.90	0.02	0.97

Note. Est is the average estimate, SE is the empirical standard error, Bias is raw bias, SB is standardized bias, RMSE is root mean squared error, and 95% CI is the coverage for the 95% CI. Values are bolded to highlight particularly interesting results.

Comparing ONB-MM with Existing Models

Model Convergence and Model Fit

Next, I will discuss the performance of the ONB-MM compared to commonly-used longitudinal methods. In general, across all models fitted, model non-convergence was only notable with a small number of time-points, see Table 6. In these conditions, the PO-MM had a slightly higher rate of model convergence compared the proposed ONB-MM. For instance, when n=300 and t=3, the 5pt and 10pt response scales had 72.4% and 73.6% of PO-MM converge compared to 70.8% and 69.4% ONB-MM. When n=1000 and t=3, the respective convergence rates for the 5pt and 10pt scales were 86% and 90.4% for the PO-MM compared to 83.2% and 86.8% for the ONB-MM. Interestingly, the model convergence rates for the L-MM were much lower than the ONB-MM and PO-MM, especially with 10 response categories. For n=300 and t=3, the 5pt scale and 10pt scale convergence rates were 68.8% and 42.4%. For n=1000 and t=3, the 5pt and 10pt scale convergences rates were 77.6% and 37.6%. Table 6 also provides the average -2ll, AIC, and BIC values across the converged replications for the ONB-MM, PO-MM, and L-MM. Results indicated that across all conditions the proposed ordinal negative binomial mixed effects model fitted the ordinal-count data better than both the existing models (e.g., smaller -2ll, AIC, and BIC). Further, the proportional odds mixed effects model provided better fit to the ordinal-count data compared to the commonly used linear mixed effects model.

Table 6. *Number of converged models and model fit for the ONB-MM, PO-MM, and L-MM across conditions.*

	***n*=300**											
	3 Time-points						**7 Time-points**					
	5 Response Categories			**10 Response Categories**			**5 Response Categories**			**10 Response Categories**		
	ONB-MM	**PO-MM**	**L-MM**	**ONB-MM**	**PO-MM**	**L-MM**	**ONB-MM**	**PO-MM**	**L-MM**	**ONB-MM**	**PO-MM**	**L-MM**
# Converged (*max*=250)	177	181	172	174	184	106	250	250	250	250	249	250
-2LL	2522.28	2532.07	2799.40	3488.09	3494.29	4168.87	5573.85	5622.10	6232.83	7738.55	7792.32	9376.35
AIC	2542.28	2556.07	2819.40	3508.09	3528.29	4188.87	5593.85	5646.10	6252.83	7758.55	7826.32	9396.35
BIC	2579.31	2600.52	2856.44	3545.13	3591.25	4225.91	5630.89	5690.54	6289.87	7795.59	7889.29	9433.39
	***n*=1000**											
	ONB-MM	**PO-MM**	**L-MM**	**ONB-MM**	**PO-MM**	**L-MM**	**ONB-MM**	**PO-MM**	**L-MM**	**ONB-MM**	**PO-MM**	**L-MM**
# Converged (*max*=250)	208	215	194	217	226	95	250	250	250	249	249	250
-2LL	8415.56	8461.12	9331.13	11622.88	11672.27	13856.06	18583.70	18754.88	20800.72	25787.83	25988.65	31279.11
AIC	8435.56	8485.12	9351.13	11642.88	11706.27	13876.06	18603.70	18778.88	20820.72	25807.83	26022.65	31299.11
BIC	8484.64	8544.01	9400.21	11691.95	11789.71	13925.13	18652.78	18837.78	20869.80	25856.91	26106.08	31348.19

Empirical Power and Type I Error

I used meta-models to investigate differences in empirical power across conditions. Specifically, I used logistic mixed effects models to assess the probability of finding significant effects for the non-zero effects (e.g., β_1, β_3, β_4). Table 7 provides a summary of the omnibus effects for the non-zero parameters. Results showed that sample size, number of time-points, and models fitted impacted empirical power. Further, Tables 8 and 9 provide the average parameter estimates, average standard errors, empirical power, and Type I error rates for the converged ordinal-negative binomial, proportional odds, and linear mixed effects models. I next discuss the effects of the simulation conditions on empirical power from the meta-models.

Table 7. *Results from logistic mixed effects meta-models for empirical power across the ONB-MM, PO-MM, and L-MM.*

		$\boldsymbol{\beta_1: (t_t)}$		$\boldsymbol{\beta_3: (x_{2i})}$		$\boldsymbol{\beta_4: (t_t x_{1i})}$	
Effect	**Num. DF**	**F(4184)**	**p**	**F(4186)**	**p**	**F(4186)**	**p**
N	1	85.90	<.0001	233.09	<.0001	70.09	<.0001
TP	1	178.03	<.0001	0.01	.94	164.62	<.0001
Scale	1	5.75	.02	0.31	.57	0.42	.52
Models	2	18.84	<.0001	1.41	.24	9.06	<.0001
N*TP	1	17.54	<.0001	-	-	31.43	<.0001
TP *Models	2	55.81	<.0001	-	-	-	-
		$\boldsymbol{\sigma_{v_{00}}}$		$\boldsymbol{\sigma_{v_{11}}}$		$\boldsymbol{\sigma_{v_{01}}}$	
	Num. DF	**F(Den.df)**	**p**	**F(1689)**	**p**	**F(1689)**	**p**
N	1	-	-	19.89	<.0001	89.19	<.0001
TP	-	-	-	-	-	-	-
Scale	1	-	-	16.39	<.0001	33.94	<.0001
Models	2	-	-	4.36	.013	8.79	.0002

Note. N denote sample size and TP denote time-points. F denotes the F-statistic with the Denominator DF in parentheses. Empirical power for the random intercept standard deviation $\sigma_{v_{00}}$ was 1.0 so a meta-model was not used. Empirical power for the random slope standard deviation and intercept-slope covariance was also almost always 1.0 across conditions when *t*=7. For this reason, meta-models for these random effect components examined the effects of sample size, response scale, and models when *t*=3.

Table 8. *Average estimate, average standard error, and proportion significant across the ONB-MM, PO-MM, and L-MM for n=300.*

		***n*=300**											
		5 Response Categories						**10 Response Categories**					
		3 Time-points											
		ONB-MM		**PO-MM**		**L-MM**		**ONB-MM**		**PO-MM**		**L-MM**	
	True	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**
$\boldsymbol{\beta_0}$	**1.00**	1.00(.12)	1.00	-	-	1.50(.07)	1.00	1.01(.11)	1.00	-	-	2.88(.16)	1.00
$\boldsymbol{\beta_1}: (\boldsymbol{t_t})$	**-0.05**	-0.06(.06)	0.15	-0.09(.09)	0.18	-0.05(.04)	0.28	-0.06(.06)	0.17	-0.09(.09)	0.19	-0.11(.08)	0.26
$\boldsymbol{\beta_2}: (\boldsymbol{x_{1i}})$	**0.00**	-0.01(.11)	0.02	-0.03(.17)	0.04	0.00(.08)	0.04	-0.01(.11)	0.06	-0.03(.17)	0.05	-0.03(.17)	0.08
$\boldsymbol{\beta_3}: (\boldsymbol{x_{2i}})$	**0.22**	0.23(.11)	0.59	0.36(.17)	0.54	0.15(.08)	0.52	0.23(.11)	0.56	0.36(.17)	0.57	0.35(.17)	0.60
$\boldsymbol{\beta_4}: (\boldsymbol{t_t x_{1i}})$	**0.03**	0.03(.06)	0.11	0.05(.09)	0.12	0.02(.04)	0.12	0.03(.06)	0.10	0.05(.09)	0.10	0.05(.09)	0.11
$\boldsymbol{\beta_5}: (\boldsymbol{t_t x_{2i}})$	**0.00**	0.00(.06)	0.04	-0.01(.09)	0.03	0.00(.04)	0.03	0.00(.06)	0.05	0.00(.09)	0.03	0.00(.09)	0.04
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.42(.13)	1.00	2.16(.25)	1.00	1.00(.07)	1.00	1.41(.12)	1.00	2.19(.23)	1.00	2.21(.15)	1.00
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.29(.18)	0.51	0.49(.31)	0.43	0.21(.14)	0.46	0.28(.16)	0.51	0.48(.29)	0.52	0.46(.28)	0.54
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.19(.15)	0.18	-0.50(.39)	0.14	-0.09(.06)	0.26	-0.18(.13)	0.22	-0.51(.36)	0.21	-0.47(.30)	0.30
		7 Time-points											
$\boldsymbol{\beta_0}$	**1.00**	1.00(.10)	1.00	-	-	1.47(.07)	1.00	1.00(.10)	1.00	-	-	2.80(.15)	1.00
$\boldsymbol{\beta_1}: (\boldsymbol{t_t})$	**-0.05**	-0.05(.02)	0.68	-0.07(.03)	0.67	-0.03(.01)	0.65	-0.05(.02)	0.73	-0.07(.03)	0.69	-0.07(.03)	0.65
$\boldsymbol{\beta_2}: (\boldsymbol{x_{1i}})$	**0.00**	0.00(.10)	0.03	0.01(.15)	0.03	0.01(.07)	0.03	0.00(.10)	0.03	0.01(.15)	0.04	0.01(.15)	0.03
$\boldsymbol{\beta_3}: (\boldsymbol{x_{2i}})$	**0.22**	0.22(.10)	0.52	0.31(.15)	0.50	0.15(.07)	0.53	0.22(.10)	0.52	0.32(.15)	0.51	0.33(.15)	0.52
$\boldsymbol{\beta_4}: (\boldsymbol{t_t x_{1i}})$	**0.03**	0.03(.02)	0.32	0.04(.03)	0.25	0.02(.01)	0.26	0.03(.02)	0.30	0.04(.03)	0.27	0.04(.03)	0.29
$\boldsymbol{\beta_5}: (\boldsymbol{t_t x_{2i}})$	**0.00**	0.00(.02)	0.05	0.00(.03)	0.06	0.00(.01)	0.05	0.00(.02)	0.04	0.00(.03)	0.05	-0.01(.03)	0.05
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.40(.09)	1.00	2.09(.15)	1.00	1.00(.05)	1.00	1.40(.09)	1.00	2.12(.14)	1.00	2.22(.12)	1.00
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.23(.02)	1.00	0.34(.04)	1.00	0.16(.01)	1.00	0.23(.02)	1.00	0.35(.03)	1.00	0.35(.03)	1.00
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.15(.04)	0.98	-0.36(.10)	1.00	-0.08(.02)	1.00	-0.15(.04)	0.99	-0.37(.10)	1.00	-0.43(.09)	1.00

Note. The dispersion parameter, α, had an empirical power of 1.0 across all conditions. Est is the average estimate, SE is the average standard error, and Sig is the proportion of significant effects at alpha=.05.

Table 9. *Average estimate, average standard error, and proportion significant across the ONB-MM, PO-MM, and L-MM for n=1000.*

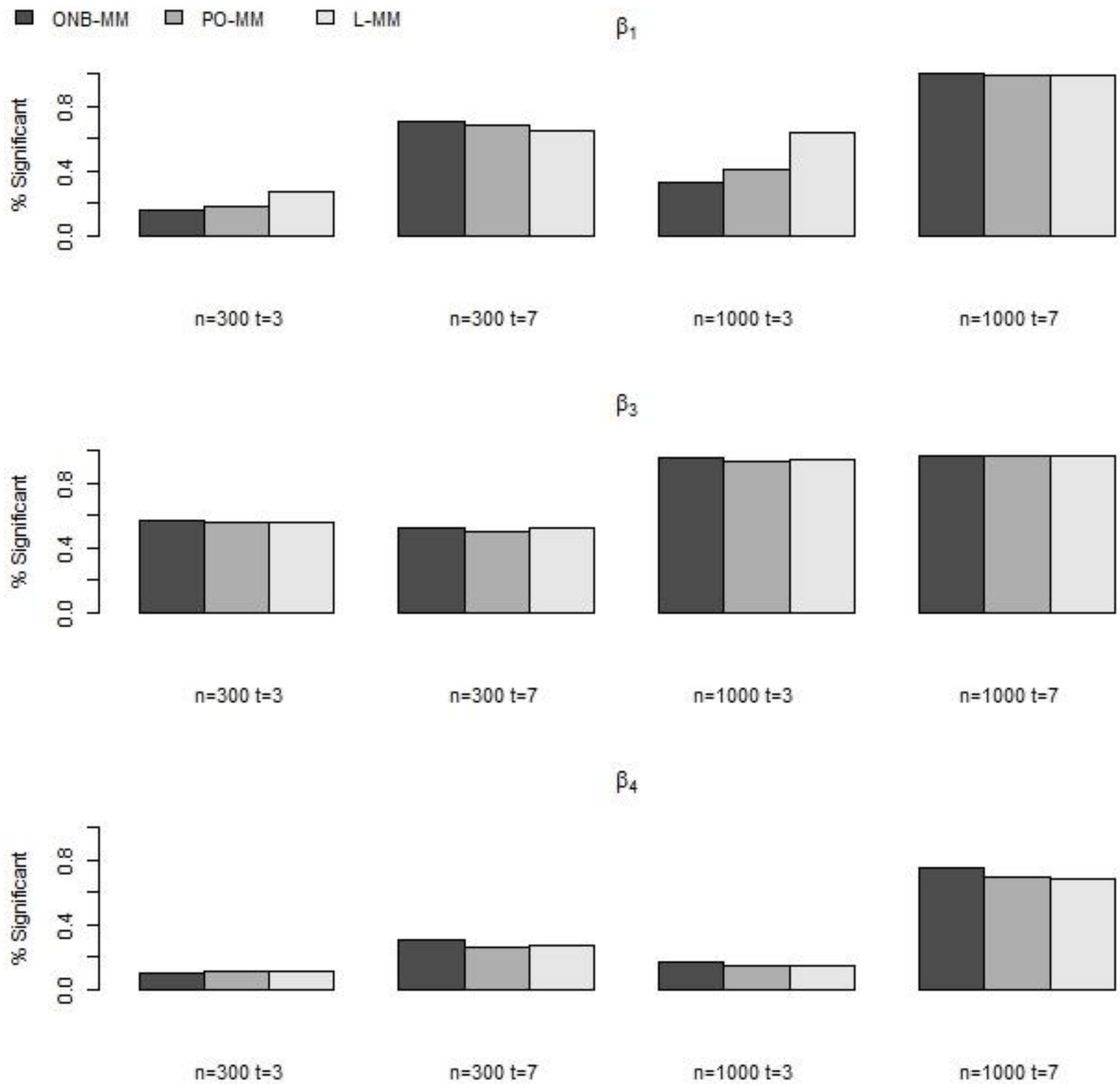
	***n*=1000**													
		5 Response Categories						**10 Response Categories**						
		3 Time-points												
		ONB-MM		**PO-MM**		**L-MM**		**ONB-MM**		**PO-MM**		**L-MM**		
	True	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**	**Est(se)**	**Sig**	
$\boldsymbol{\beta_0}$	**1.00**	1.00(.06)	1.00	-	-	1.49(.04)	1.00	0.99(.06)	1.00	-	-	2.83(.09)	1.00	
$\boldsymbol{\beta_1}: (\boldsymbol{t_t})$	**-0.05**	-0.05(.03)	0.33	-0.08(.05)	0.40	-0.05(.02)	0.59	-0.05(.03)	0.34	-0.08(.05)	0.42	-0.11(.04)	0.67	
$\boldsymbol{\beta_2}: (\boldsymbol{x_{1i}})$	**0.00**	0.01(.06)	0.05	0.01(.09)	0.05	0.00(.04)	0.04	0.00(.06)	0.06	0.00(.09)	0.06	0.01(.09)	0.01	
$\boldsymbol{\beta_3}: (\boldsymbol{x_{2i}})$	**0.22**	0.22(.06)	0.96	0.32(.09)	0.94	0.15(.04)	0.95	0.22(.06)	0.96	0.33(.09)	0.94	0.33(.09)	0.95	
$\boldsymbol{\beta_4}: (\boldsymbol{t_t x_{1i}})$	**0.03**	0.03(.03)	0.14	0.04(.05)	0.15	0.02(.02)	0.16	0.03(.03)	0.19	0.04(.05)	0.15	0.04(.05)	0.14	
$\boldsymbol{\beta_5}: (\boldsymbol{t_t x_{2i}})$	**0.00**	0.00(.03)	0.05	0.00(.05)	0.05	0.00(.02)	0.06	0.00(.03)	0.04	-0.01(.05)	0.05	0.00(.05)	0.07	
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.41(.07)	1.00	2.11(.13)	1.00	0.98(.04)	1.00	1.41(.06)	1.00	2.16(.13)	1.00	2.20(.08)	1.00	
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.24(.11)	0.60	0.40(.19)	0.63	0.17(.09)	0.55	0.23(.10)	0.67	0.41(.17)	0.69	0.18(.16)	0.75	
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.16(.08)	0.49	-0.41(.20)	0.54	-0.07(.04)	0.58	-0.15(.07)	0.59	-0.44(.19)	0.62	-0.41(.16)	0.74	
		7 Time-points												
$\boldsymbol{\beta_0}$	**1.00**	1.00(.05)	1.00	-	-	1.47(.04)	1.00	1.00(.05)	1.00	-	-	2.80(.08)	1.00	
$\boldsymbol{\beta_1}: (\boldsymbol{t_t})$	**-0.05**	-0.05(.01)	1.00	-0.07(.02)	0.99	-0.03(.01)	0.99	-0.05(.01)	0.99	-0.07(.02)	0.99	-0.07(.02)	0.99	
$\boldsymbol{\beta_2}: (\boldsymbol{x_{1i}})$	**0.00**	0.00(.06)	0.03	0.00(.08)	0.03	0.00(.04)	0.03	0.00(.06)	0.03	0.00(.08)	0.03	0.00(.09)	0.04	
$\boldsymbol{\beta_3}: (\boldsymbol{x_{2i}})$	**0.22**	0.22(.06)	0.97	0.33(.08)	0.97	0.15(.04)	0.98	0.22(.06)	0.98	0.33(.08)	0.97	0.34(.09)	0.97	
$\boldsymbol{\beta_4}: (\boldsymbol{t_t x_{1i}})$	**0.03**	0.03(.01)	0.75	0.04(.02)	0.68	0.02(.01)	0.69	0.03(.01)	0.75	0.04(.02)	0.71	0.04(.02)	0.67	
$\boldsymbol{\beta_5}: (\boldsymbol{t_t x_{2i}})$	**0.00**	0.00(.01)	0.05	0.00(.02)	0.05	0.00(.01)	0.07	0.00(.01)	0.05	0.00(.02)	0.05	-0.01(.02)	0.07	
$\boldsymbol{\sigma_{v_{00}}}$	**1.41**	1.41(.05)	1.00	2.09(.08)	1.00	1.00(.03)	1.00	1.41(.05)	1.00	2.13(.08)	1.00	2.23(.07)	1.00	
$\boldsymbol{\sigma_{v_{11}}}$	**0.23**	0.23(.01)	1.00	0.34(.02)	1.00	0.16(.01)	1.00	0.23(.01)	1.00	0.35(.02)	1.00	0.36(.02)	1.00	
$\boldsymbol{\sigma_{v_{01}}}$	**-0.15**	-0.15(.02)	1.00	-0.36(.05)	1.00	-0.08(.01)	1.00	-0.15(.02)	1.00	-0.37(.05)	1.00	-0.42(.05)	1.00	

Note. The dispersion parameter, α, had an empirical power of 1.0 across all conditions. Est is the average estimate, SE is the average standard error, and Sig is the proportion of significant effects at alpha=.05

Empirical Power for Non-Zero Effects: β_1, β_3, β_4

For the time-varying time predictor effect, β_1, the impact of sample size depended on the number of time-points. Specifically, increased sample size had a greater impact with more time-points (*t*=3: OR=7.27, p<.0001, *t*=7: OR=191.22, p<.0001). There was no effect of response category using an alpha=.01 (p=.02). Results also indicated that the ONB-MM, PO-MM, and L-MM differed in empirical power and these relationships depended on number of time-points. The top panel of Figure 2 displays these differences. For instance, when *t*=3, the odds of a significant β_1 effect were larger for the PO-MM versus the ONB-MM and the L-MM versus the ONB-MM (PO-MM vs. ONB-MM: OR=2.00, p=.0002; L-MM vs. ONB-MM: OR=12.19, p<.0001). The odds were also larger for the L-MM compared to the PO-MM (OR=6.08, p<.0001) when *t*=3. When *t*=7, the odds of a significant β_1 effect did not differ between the PO-MM and ONB-MM (p=.11) but, the odds were lower for the L-MM compared to the ONB-MM (OR=.49, p=.002). The odds for the L-MM compared to the PO-MM did not significantly differ (p=.12).

Figure 2. *Visualizing differences in empirical power for the non-zero effects across the ONB-MM, PO-MM, and L-MM*



Note. The plotted values are collapsed over the 5 and 10 response category conditions.

Results from the meta-model examining the β_3 main effect for x_{2i} were straightforward. The number of time-points, number of response categories, and models fitted did not impact empirical power for β_3 effect (time-points: p=.94; scale: p=.57; models: p=.24). However, there was an effect of sample size such that the odds of significant β_3 effect were larger for *n*=1000 than *n*=300 (OR=88.36, p<.0001). The empirical power levels for *n*=300 were .50-.60 compared to .94-.98 for *n*=1000. The middle panel of Figure 2 shows these differences in empirical power. Next, I address empirical power for the $\beta_4 t_t x_{1i}$ effect.

The final meta-model examined the t_t-by-x_{1i} interaction effect, β_4. Results from this meta-model indicated that the number of response categories did not impact the odds of a significant β_4 interaction effect (p=.52). The effect of sample size depended on the number of time-points (p<.0001). Specifically, when *t*=3, there was no effect of sample size (p=.08) but, when *t*=7, the sample size of *n*=1000 had larger odds compared to *n*=300 (OR=24.47, p<.0001). Results indicated that the odds of a significant β_4 effect differed across the ONB-MM, PO-MM, and L-MM. While the L-MM and PO-MM did not differ (p=.95), the odds of significant β_4 were lower for both the PO-MM and the L-MM compared to the ONB-MM (PO-MM vs. ONB-MM: OR=.61, p<.0001; L-MM vs. ONB-MM: OR=.60, p<.0001).

Empirical Power for the Random Effects Components: $\sigma_{v_{00}}$, $\sigma_{v_{11}}$, $\sigma_{v_{01}}$

Across all conditions and models the random intercept standard deviation had an empirical power of 1.00. Similarly, when *t*=7, the random slope standard deviation and intercept-slope covariance almost always had empirical power of 1.00. When *t*=3, the odds of significant variability in the random time effect were larger for *n*=1000 than *n*=300 (OR=2.60, p<.0001) and for the 10pt scale than the 5pt scale (OR=1.60, p<.0001). The odds did not vary across the ONB-MM, L-MM, and PO-MM (p=.013). When *t*=3, the odds of a significant intercept-slope covariance were higher for *n*=1000 than *n*=300 (OR=10.75, p<.0001) and 10pt scale than 5pt scale (OR=2.18, p<.0001). The L-MM had a higher odds than both the ONB-MM (OR=2.00, p<.0001) and PO-MM (1.70, p=.002), but the ONB-MM and PO-MM did not differ (p=.29).

Type 1 Error for Null Effects: β_2, β_5

Tables 8 and Table 9 also shows that for the null effects, the Type I error rates were acceptable across models and conditions. The majority of the false positive rates were equal to or less than .05. The L-MM had a few cases of slightly higher Type I error rates (e.g., .07, .08), but these did not appear problematic.

Summarizing Empirical Power and Type I Errors

The results did not show unexpected patterns of effects as a function of the conditions or models fitted. Results indicated reasonable Type I error rates for the null effects (e.g., β_2 and β_5) across conditions and models. Unsurprisingly, there was a clear relationship between empirical power and the amount of data present. Further, the meta-models reveled that empirical power varied in slightly complicated ways for the β_1 and β_4 effects. I next discuss the relative efficiency of predictions across the evaluated models.

Comparing the Relative Efficiency of Predictions

Table 10 displays the relative efficiency of predictions produced from the existing and proposed mixed effects models. Results showed that the predictions from the proposed ONB-MM were more efficient than both the PO-MM and L-MM (e.g., the relative efficiencies were greater than 1). The L-MM and PO-MM displayed more inefficiency relative to the ONB-MM as the number of time-points and number of response categories increased, but sample size had little impact. For example, with *t*=3 and the 5pt scale, the efficiency of predictions for the PO-MM was 8% larger than the ONB-MM compared to 13% when *t*=7 with a 10pt scale. A similar trend held for the L-MM compared to the ONB-MM, but with ratios closer to 1. The efficiency ratios for the PO-MM relative to the L-MM were between 1.03 and 1.06, suggesting that the L-MM was more efficient.

Table 10. *Relative efficiency of predictions from the ONB-MM, PO-MM, and L-MM across conditions.*

	3 Time-Points		7 Time-Points	
	5 Categories	10 Categories	5 Categories	10 Categories
***n*=300**				
$RE_{\frac{PO}{ONB}}$	1.08	1.10	1.10	1.13
$RE_{\frac{L}{ONB}}$	1.03	1.04	1.06	1.07
$RE_{\frac{PO}{L}}$	1.05	1.06	1.03	1.05
***n*=1000**				
$RE_{\frac{PO}{ONB}}$	1.09	1.10	1.10	1.13
$RE_{\frac{L}{ONB}}$	1.04	1.04	1.06	1.07
$RE_{\frac{PO}{L}}$	1.04	1.06	1.04	1.05

Note. Relative efficiencies greater than 1 indicate that the model in the denominator produced more efficient predictions.

Summary of Simulation Study Results

Taken together, the result showed that the proposed ONB-MM performed well across the full range of conditions. As expected, the proposed model performed best with more subjects and time-points. However, the number of response categories did not meaningfully impact model performance. The ordinal negative binomial mixed effects model provided better fit to the ordinal-count data than existing models and produced more efficient predictions. Meta-models also suggested that empirical power differed across the evaluated models and simulation conditions.

CHAPTER 3: STUDY 2: EVALUATING THE ORDINAL ZERO-INFLATED NEGATIVE BINOMIAL MIXED EFFECTS MODEL (OZINB-MM)

I assessed the proposed ordinal zero-inflated negative binomial mixed effects model (OZINB-MM) by simulating longitudinal zero-inflated ordinal-count data consistent with those arising in adolescent substance use research. I then fitted the OZINB-MM along with commonly used proportional odds and linear mixed effects models to the simulated data to evaluate my hypotheses. First, I tested whether the proposed OZINB-MM recovered the true underlying count population generating values despite the response variable being defined by ordinal categories. I expected the OZINB-MM to adequately recover the population generating values from the underlying zero-inflated count process. I anticipated optimal performance from the OZINB-MM (e.g., highest convergence rates, smallest standard errors, most accurate estimates) with more subject and time-points (e.g., n=1000, t=7). I posited that model performance in conditions with less data would slightly worsen. In comparison to the existing proportional odds and linear mixed effects models, I expected my proposed OZINB-MM to provide superior fit and produce more efficient predictions.

The ordinal-count data were generated from a zero-inflated negative binomial distribution with correlated random intercepts in the binary (logistic) and count (negative binomial) processes. The selection of my population generating parameters was guided by zero-inflated count mixed effects models fitted to real adolescent substance use data. I included time-varying predictors, time-invariant predictors, and time-varying by time-invariant predictor interactions in the population generating model for both the binary and count processes. I created eight unique condition combinations. There were two sample sizes (*n*=1000, 300), two numbers of time-points (*t*=3, 7), and two response scales (5pt, 10pt). Three mixed effects models (OZINB-MM, PO-MM, and L-MM) were fitted to each of the unique datasets.

Simulation Study Design

The overall design of my simulation study mirrored that described in Chapter 2 for the ONB-MM. The population generating parameters were partially based on results obtained by fitting count zero-inflated negative binomial mixed effects models to adolescent (e.g., ages 13 to 17) frequency of past 30 day alcohol use data from the 1997 National Longitudinal Survey of Youth (NLSY97). I tailored the generating values to represent realistic values that adolescent substance use researchers may encounter in practice. In order to establish a range of effects, the population generating fixed effects were partly based on empirical power levels derived by fitting count mixed effects models in pilot work. The population model included correlated random intercepts from the binary and count processes and no random time effects. The covariates included a time-varying linear time predictor, t_t, two standard normal continuous time-invariant predictors x_{1i} and x_{2i} correlated .3, and interactions between the time-invariant predictors and time. A summary of population generating parameters is listed in the first column of Table 12.

For each condition, there were *r*=250 replications. I used SAS 9.3 to generate open-ended longitudinal count data from a zero-inflated negative binomial distribution and the counts were collapsed into ordinal-count responses according the two unique response scales in Table 2. I fitted the ordinal zero-inflated negative binomial (OZINB-MM), proportional odds (PO-MM), and linear (L-MM) mixed effects models to each of the datasets. The OZINB-MM had properly specified fixed and random effects. The PO-MM and L-MM had the same set of fixed effects predictors without the "zero" class. The L-MM and PO-MM included correlated random intercept and time effects to be consistent with what is commonly done in adolescent substance use research. I fitted all models using PROC NLMIXED in SAS using adaptive quadrature with nine points and dual quasi-Newton optimization. The variance components for the random effects were modeled in terms of standard deviations to help with convergence in PROC NLMIXED. This is recommended with these types of models (Kiernan, Tao, & Gibbs, 2012).

First, I assessed the general performance of the OZINB-MM using raw bias (Bias), standardized bias (SB), root mean square error (RMSE), and 95% CI coverage probabilities. Each of these outcome measures are described in detail in Chapter 2. I used meta-models to examine raw bias in the OZINB-MM. Second, I compared the proposed ordinal zero-inflated negative binomial mixed effects models to existing proportional odds and linear mixed effects models. I accomplished this by examining model fit, empirical power, Type I error, and relative efficiency of predictions. I defined each of these aspects in Chapter 2. I used logistic mixed effects meta-models to examine empirical power across the models and conditions. However, it is important to note that the parameters from the OZINB-MM are substantively different from the PO-MM and L-MM. For this reason, I also examined empirical power across the models from a qualitative perspective. For example, I evaluated whether or not the average pattern and direction of effects were similar across models.

In sum, my simulation study provided a realistic evaluation of the proposed ordinal zero-inflated negative binomial mixed effects model for longitudinal adolescent substance use research. My simulation offered insights into how the proposed ordinal-count zero-inflated model performs relative to widely utilized existing techniques. Importantly, my simulation study considered critical factors such as response scales, sample size, and number of time-points to help determine the necessary conditions for effectively implementing this innovative longitudinal modeling approach.

Results

Overall Performance of OZINB-MM

Model convergence rates across conditions were generally high across all conditions. Within sample size by time-point conditions, the 10pt scale consistently had higher convergence rates than the 5pt scale. When *n*=300 and *t*=3, convergence rates for the 5pt and 10pt scales were 86.8% and 92%. Across the other conditions, the model convergence rate was no less than 96.4%. It was interesting that there were more non-converged models when *n*=1000 and *t*=7 compared when *n*=300 and *t*=7 or *n*=1000 and *t*=3.

Tables 11 through Table 15 show that, across all conditions, the OZINB-MM generally recovered the assigned parameter values. There were small biases, RMSE, and close to 95% coverage. Results from the meta-models fitted to raw bias also confirmed the strong performance of the OZINB-MM across simulation conditions. The meta-model results showed that raw bias did not differ across simulation conditions for the fixed effects. Raw bias was only notable for the logistic random intercept standard deviation σ_{u_0}, the covariance between the random intercepts $\sigma_{v_0 u_0}$, and the dispersion parameter. These raw biases were larger when *n*=300 and *t*=3 and generally became less severe, or non-existent, as the number of time-points and sample size increased. It was interesting that standardized bias was slightly higher for these parameters when *n*=1000 and *t*=7 because of smaller empirical standard errors. However, these standard bias values were still not overly concerning at less than +/- 40%. The pseudo Cohen's f^2 values from the meta-models fitted to the raw bias of the random effect standard deviations and covariance suggested that the significant condition effects were small in magnitude. Further, the 95% coverage for dispersion parameter was low, especially when n=300 and t=3. The standard errors for the parameters became smaller as the number of time-points and subjects increased. Results also suggested that having more response categories resulted in slightly smaller standard errors, especially when t=3

Table 11. *Results from linear mixed effects meta-models fitted to raw bias for the OZINB-MM.*

Count Process

	β_0			$\beta_1: (t_t)$			$\beta_2: (x_{1i})$			$\beta_3: (x_{2i})$			$\beta_4: (t_t x_{1i})$			$\beta_5: (t_t x_{2i})$			α		
	F	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2
N	.43	.51	0	.01	.93	0	3.42	.06	.003	2.26	.13	.002	.75	.39	.001	.01	.91	0	1.96	.16	.002
TP	4.83	.03	.004	4.92	.03	.004	1.85	.17	.002	.01	.94	0	1.15	.28	.001	0	.96	0	.01	.94	0
Scale	.40	.40	0	0	.96	0	.45	.50	0	2.29	.13	0	.11	.74	0	1,54	.21	0	.55	.46	0

Zero Process

	γ_0			$\gamma_1: (t_t)$			$\gamma_2: (x_{1i})$			$\gamma_3: (x_{2i})$			$\gamma_4: (t_t x_{1i})$			$\gamma_5: (t_t x_{2i})$		
	F	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2
N	.94	.33	.001	.28	.60	0	.06	.80	0	.54	.46	.001	.08	.78	0	.17	.68	0
TP	.41	.52	0	1.59	.21	.002	3.68	.06	.004	1.55	.21	.002	.72	.40	.001	.53	.47	0
Scale	.39	.53	0	1.04	.31	0	1.22	.27	0	5.30	.02	0	.22	.64	0	3.43	.06	0

Random Effects

	σ_{v_0}			σ_{u_0}			$\sigma_{v_0 u_0}$		
	F	**p**	f^2	**F**	**p**	f^2	**F**	**p**	f^2
N	.33	.57	.001	9.77	.002	.009	.16	.69	0
TP	3.53	.29	.002	13.73	.0002	.01	5.08	.02	.004
Scale	11.53	.0007	.001	2.12	.15	0	0	.97	0
TP*Scale	9.36	.002	.001	-	-	-	-	-	-

Note. N denote sample size, TP denote time-points, F denote the F-statistic, p denote the p-value, and f^2 denote the pseudo Cohen's f^2. All F statistics had a Numerator DF of 1 and Denominator DF of 948 except for the meta-model fitted to the σ_{v_0}, which had a Numerator DF of 1 and Denominator DF of 947.

Table 12. *Recovery of population generating values for the OZINB-MM when n=300 and t=3.*

	***n*=300 and *t*=3**												
	5 Response Categories (*r*=217)							**10 Response Categories (*r* =230)**					
	Count Process												
	True	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**
$\boldsymbol{\beta_0}$	**0.20**	0.19	0.40	-0.01	-1.74	0.40	0.97	0.18	0.39	-0.02	-4.82	0.39	0.95
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.36	0.14	0.01	3.91	0.14	0.98	0.36	0.14	0.01	7.50	0.14	0.97
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.01	0.28	0.01	1.85	0.28	0.92	0.00	0.26	0.00	1.80	0.26	0.93
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.33	0.27	-0.02	-6.52	0.27	0.94	0.34	0.24	-0.01	-3.20	0.24	0.96
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.06	0.18	0.00	1.55	0.18	0.92	0.06	0.16	0.00	0.57	0.16	0.93
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.16	0.00	-1.52	0.16	0.93	0.00	0.15	0.00	-1.23	0.15	0.93
$\boldsymbol{\alpha}$	**1.00**	0.96	0.40	-0.04	-9.02	0.41	**0.85**	0.96	0.35	-0.04	-11.27	0.35	**0.85**
$\boldsymbol{\sigma_{v_0}}$	**1.00**	1.01	0.21	0.01	2.56	0.21	0.97	0.99	0.18	-0.01	-4.13	0.18	0.97
	Zero Process												
$\boldsymbol{\gamma_0}$	**2.00**	2.04	0.52	0.04	8.40	0.52	0.96	2.05	0.50	0.05	9.70	0.51	0.97
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.67	0.32	-0.02	-4.86	0.33	0.95	-0.67	0.32	-0.02	-5.29	0.32	0.95
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	-0.01	0.44	-0.01	-2.23	0.44	0.97	-0.02	0.42	-0.02	-4.72	0.42	0.97
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.54	0.48	0.01	2.11	0.48	0.95	-0.52	0.44	0.03	6.91	0.44	0.95
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.12	0.29	0.00	0.97	0.29	0.95	0.12	0.29	0.00	1.66	0.29	0.95
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	-0.01	0.32	-0.01	-2.08	0.32	0.96	-0.02	0.31	-0.02	-6.41	0.31	0.97
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.62	0.84	**0.17**	**20.19**	0.86	0.92	2.60	0.80	**0.15**	18.34	0.81	0.93
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.08	1.49	**-0.08**	-5.30	1.49	1.00	-1.13	1.17	**-0.13**	-11.40	1.17	0.98

Note. Est is the average estimate, SE is the empirical standard error, Bias is raw bias, SB is standardized bias, RMSE is root mean squared error, and 95% CI is the coverage for the 95% CI. Values are bolded to highlight particularly interesting results.

Table 13. *Recovery of population generating values for the OZINB-MM when n=300 and t=7.*

		***n*=300 and *t*=7**											
		5 Response Categories (*r*=248)						**10 Response Categories (*r*=249)**					
		Count Process											
	True	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**
$\boldsymbol{\beta_0}$	**0.20**	0.21	0.17	0.01	3.81	0.17	0.97	0.20	0.16	0.00	2.57	0.16	0.96
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.35	0.03	0.00	-2.56	0.03	0.95	0.35	0.03	0.00	-3.78	0.03	0.95
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.03	0.16	0.03	16.94	0.16	0.94	0.02	0.15	0.02	15.21	0.15	0.95
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.34	0.16	-0.01	-5.84	0.16	0.96	0.34	0.15	-0.01	-4.97	0.15	0.95
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.05	0.03	-0.01	**-20.09**	0.03	0.92	0.05	0.03	-0.01	-17.94	0.03	0.95
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.03	0.00	0.03	0.03	0.97	0.00	0.03	0.00	-0.33	0.03	0.96
$\boldsymbol{\alpha}$	**1.00**	0.98	0.13	-0.02	-15.36	0.13	**0.91**	0.98	0.11	-0.02	-18.71	0.11	0.92
$\boldsymbol{\sigma_{v_0}}$	**1.00**	0.99	0.09	-0.01	-8.42	0.09	0.98	1.00	0.09	0.00	-4.35	0.09	0.96
		Zero Process											
$\boldsymbol{\gamma_0}$	**2.00**	2.00	0.28	0.00	-1.57	0.28	0.94	2.00	0.28	0.00	-1.24	0.28	0.94
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.65	0.08	0.00	0.96	0.08	0.92	-0.65	0.08	0.00	0.72	0.08	0.91
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	0.02	0.25	0.02	8.08	0.25	0.97	0.02	0.25	0.02	7.42	0.25	0.96
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.57	0.28	-0.02	-8.10	0.28	0.95	-0.57	0.28	-0.02	-7.82	0.29	0.94
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.11	0.06	-0.01	-8.02	0.06	0.96	0.12	0.06	0.00	-7.27	0.06	0.96
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	0.00	0.06	0.00	7.17	0.06	0.97	0.00	0.06	0.00	6.47	0.06	0.96
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.43	0.27	-0.02	-6.81	0.27	0.93	2.43	0.27	-0.02	-6.88	0.27	0.93
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.00	0.35	0.00	1.02	0.35	0.94	-1.00	0.34	0.00	-0.25	0.34	0.94

Note. Est is the average estimate, SE is the empirical standard error, Bias is raw bias, SB is standardized bias, RMSE is root mean squared error, and 95% CI is the coverage for the 95% CI. Values are bolded to highlight particularly interesting results.

Table 14. *Recovery of population generating values for the OZINB-MM when n=1000 and t=3.*

	***n*=1000 and *t*=3**												
		5 Response Categories (*r*=249)						**10 Response Categories (*r*=250)**					
		Count Process											
	True	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**
$\boldsymbol{\beta_0}$	**0.20**	0.16	0.25	-0.04	-15.49	0.25	0.93	0.16	0.22	-0.04	-16.62	0.23	0.95
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.36	0.08	0.01	10.31	0.08	0.96	0.36	0.08	0.01	13.47	0.08	0.94
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	-0.01	0.13	-0.01	-6.96	0.13	0.96	-0.01	0.12	-0.01	-8.83	0.12	0.94
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.36	0.14	0.01	4.00	0.14	0.93	0.36	0.13	0.01	7.98	0.13	0.92
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.06	0.08	0.00	5.72	0.08	0.96	0.07	0.07	0.01	8.28	0.07	0.97
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.08	0.00	2.00	0.08	0.95	0.00	0.08	0.00	-4.14	0.08	0.96
$\boldsymbol{\alpha}$	**1.00**	1.00	0.25	0.00	0.67	0.25	**0.91**	1.00	0.21	0.00	0.16	0.21	0.93
$\boldsymbol{\sigma_{v_0}}$	**1.00**	1.01	0.11	0.01	8.26	0.11	0.96	1.00	0.09	0.00	3.48	0.09	0.94
		Zero Process											
$\boldsymbol{\gamma_0}$	**2.00**	1.99	0.24	-0.01	-3.69	0.24	0.96	1.99	0.23	-0.01	-4.05	0.23	0.97
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.66	0.14	-0.01	-7.30	0.15	0.96	-0.66	0.14	-0.01	-4.46	0.14	0.95
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	-0.02	0.20	-0.02	-9.36	0.21	0.94	-0.02	0.20	-0.02	-10.17	0.20	0.94
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.54	0.22	0.01	5.31	0.22	0.96	-0.53	0.21	0.02	7.92	0.21	0.95
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.13	0.14	0.01	4.92	0.14	0.94	0.13	0.14	0.01	5.66	0.14	0.94
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	0.00	0.14	0.00	1.70	0.14	0.95	0.00	0.13	0.00	-1.39	0.13	0.96
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.46	0.33	0.01	1.59	0.33	0.97	2.45	0.31	0.00	-1.23	0.31	0.95
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.08	0.57	**-0.08**	-14.57	0.57	0.96	-1.07	0.50	**-0.07**	-14.97	0.51	0.96

Note. Est is the average estimate, SE is the empirical standard error, Bias is raw bias, SB is standardized bias, RMSE is root mean squared error, and 95% CI is the coverage for the 95% CI. Values are bolded to highlight particularly interesting results.

Table 15. *Recovery of population generating values for the OZINB-MM when n=1000 and t=7.*

	***n*=1000 and *t*=7**												
		5 Response Categories (*r*=241)						**10 Response Categories (*r*=246)**					
		Count Process											
	True	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**	**Est**	**SE**	**Bias**	**SB**	**RMSE**	**95%CI**
$\boldsymbol{\beta_0}$	**0.20**	0.20	0.10	0.00	-1.86	0.10	0.97	0.19	0.10	-0.01	-6.74	0.10	0.95
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.35	0.02	0.00	-6.21	0.02	0.95	0.35	0.02	0.00	0.24	0.02	0.96
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.00	0.08	0.00	-0.45	0.08	0.95	0.00	0.08	0.00	-2.35	0.08	0.95
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.35	0.08	0.00	4.53	0.08	0.95	0.35	0.08	0.00	5.23	0.08	0.96
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.06	0.02	0.00	0.77	0.02	0.94	0.06	0.02	0.00	4.36	0.02	0.93
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00	0.02	0.00	-10.16	0.02	0.95	0.00	0.02	0.00	-8.37	0.02	0.95
$\boldsymbol{\alpha}$	**1.00**	0.99	0.07	-0.01	-16.72	0.07	0.93	0.99	0.06	-0.01	-11.47	0.06	0.93
$\boldsymbol{\sigma_{v_0}}$	**1.00**	0.99	0.05	-0.01	-11.98	0.05	0.97	1.00	0.05	0.00	-5.51	0.05	0.95
		Zero Process											
$\boldsymbol{\gamma_0}$	**2.00**	2.02	0.16	0.02	9.97	0.16	0.95	2.01	0.16	0.01	5.67	0.16	0.96
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.65	0.04	0.00	3.58	0.04	0.94	-0.65	0.04	0.00	2.42	0.04	0.95
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	0.01	0.14	0.01	9.95	0.14	0.95	0.01	0.14	0.01	10.05	0.14	0.95
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.55	0.15	0.00	2.52	0.15	0.96	-0.55	0.15	0.00	3.25	0.15	0.96
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.12	0.03	0.00	-8.82	0.03	0.95	0.12	0.03	0.00	-5.74	0.03	0.95
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	0.00	0.03	0.00	-5.66	0.03	0.95	0.00	0.03	0.00	-4.07	0.03	0.94
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.41	0.14	**-0.04**	**-25.97**	0.15	0.95	2.42	0.14	**-0.03**	**-24.10**	0.14	0.96
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.02	0.18	-0.02	-13.55	0.18	0.96	-1.03	0.17	-0.03	-15.74	0.17	0.96

Note. Est is the average estimate, SE is the empirical standard error, Bias is raw bias, SB is standardized bias, RMSE is root mean squared error, and 95% CI is the coverage for the 95% CI. Values are bolded to highlight particularly interesting results.

In sum, the proposed ordinal zero-inflated negative binomial mixed effects model generally recovered the underlying count process. Model convergence rates for the OZINB-MM were high across conditions. At a smaller sample size and number time-points, there was bias for the random effect components and the dispersion parameter. The dispersion parameter also had lower 95% coverage, particularly when n=300 and t=3. However, the raw biases and coverage rates improved with added subjects and time-points.

<u>Comparing OZINB-MM with Existing Models</u>

Model Convergence and Model Fit

Next, I will discuss the performance of the OZINB-MM compared to commonly-used longitudinal methods. Table 16 displays the number of converged models and average model fit across conditions and fitted models. The OZINB-MM had higher model convergence rates compared to the PO-MM and L-MM and the PO-MM had higher convergence rates than the L-MM across all simulation conditions. Although the convergence rates were satisfactory for the OZINB-MM across conditions, the proportional odds mixed effects models required more time-points to achieve acceptable convergence rates. For instance, when n=300 and t=3, the model respective convergence rates for the 5pt and 10pt scales were 31.2% and 25.6% compared to 97.6% and 99.6% when n=1000 and t=7. The rates of model convergence for the linear mixed effects models were low across all conditions (e.g., max: 28.8%; min: 6.8%). The L-MM convergence rates were lowest when n=1000 and the response scale had 10 levels (e.g., t=3: 8.8%; t=7: 6.8%).

Table 16 provides the average -2ll, AIC, and BIC values across the converged replications for the OZINB-MM, PO-MM, and L-MM. Results indicated that across all conditions the proposed ordinal negative binomial mixed effects model fitted the ordinal-count data better than both the existing models (e.g., smaller -2ll, AIC, and BIC). Further, the proportional odds mixed effects model provided better fit to the ordinal-count data compared to the linear mixed effects model.

Table 16. *Number of converged models and model fit for the OZINB-MM, PO-MM, and L-MM across conditions.*

	***n*=300**											
	***t*=3**						***t*=7**					
	5 Response Categories			**10 Response Categories**			**5 Response Categories**			**10 Response Categories**		
Fit Measure	**OZINB-MM**	**PO-MM**	**L-MM**	**OZINB-MM**	**PO-MM**	**L-MM**	**OZINB-MM**	**PO-MM**	**L-MM**	**OZINB-MM**	**PO-MM**	**L-MM**
# Converged (r=250)	217	78	48	230	64	34	248	209	71	249	216	27
-2LL	1354.83	1386.86	2291.24	1679.57	1711.54	3535.70	4197.71	4271.23	6376.72	5467.84	5549.68	9479.31
AIC	1386.83	1410.86	2311.24	1711.57	1745.54	3555.70	4229.71	4295.23	6396.72	5499.84	5583.68	9499.31
BIC	1446.09	1455.31	2348.27	1770.83	1808.51	3592.74	4288.97	4339.68	6433.76	5559.10	5646.64	9536.35
	***n*=1000**											
Fit Measure	**OZINB-MM**	**PO-MM**	**L-MM**	**OZINB-MM**	**PO-MM**	**L-MM**	**OZINB-MM**	**PO-MM**	**L-MM**	**OZINB-MM**	**PO-MM**	**L-MM**
# Converged (r=250)	249	123	34	250	91	22	241	244	54	246	249	17
-2LL	4551.89	4585.62	7713.09	5633.23	5676.24	11914.65	13991.59	14239.60	21316.77	18224.31	18503.84	31830.57
AIC	4583.89	4609.62	7733.09	5665.23	5710.24	11934.65	14023.59	14263.60	21336.77	18256.31	18537.84	31850.57
BIC	4662.42	4668.52	7782.16	5743.75	5793.67	11983.73	14102.12	14322.49	21385.84	18334.83	18621.27	31899.65

Empirical Power and Type I Error

Table 17 and Table 18 provide results from logistic mixed effects meta-models predicting the log-odds of a significant effect for the count and zero process parameters. Table 19 through Table 22 provide the average parameter estimates, average standard errors, and empirical power for the converged ordinal zero-inflated negative binomial, proportional odds, and linear mixed effects models. I compared results from the count process of the OZINB-MM to those from the linear and proportional odds mixed effects models. In practice, these results would be similarly interpreted from a substantive perspective net the zero class in OZINB-MM. It is important to note that few of the L-MM models converged and care must be taken in interpreting the results.[1]

Table 17. *Results from logistic mixed effects meta-models for empirical power and Type I errors in count process across the OZINB-MM, PO-MM, and L-MM.*

Non-Zero Effects (Count Process)							
		$\boldsymbol{\beta_1}: (\boldsymbol{t_t})$		$\boldsymbol{\beta_3}: (\boldsymbol{x_{2i}})$		$\boldsymbol{\beta_4}: (\boldsymbol{t_t x_{1i}})$	
Effect	**Num. DF**	**F(Den DF)**	**p**	**F(2518)**	**p**	**F(2514)**	**p**
N	1	-	-	278.65	<.0001	6.63	.01
TP	1	-	-	131.37	<.0001	57.16	<.0001
Scale	1	-	-	3.80	.05	10.50	.001
Models	2	-	-	110.04	<.0001	89.09	<.0001
N*TP	1	-	-	-	-	64.90	<.0001
N*Models	2	-	-	-	-	33.60	<.0001
TP*Models	2	-	-	-	-	30.08	<.0001
Null Effects (Count Process)							
		$\boldsymbol{\beta_2}: (\boldsymbol{x_{1i}})$		$\boldsymbol{\beta_5}: (\boldsymbol{t_t x_{2i}})$			
Effect	**Num. DF**	**F(2514)**	**p**	**F(2514)**	**p**		
N	1	0.41	.52	10.16	0.002		
TP	1	9.27	.002	14.12	0.0002		
Scale	1	1.35	.24	0.85	0.36		
Models	2	2.26	.10	183.37	<.0001		
N*Models	2	27.23	<.0001	16.77	<.0001		
TP*Models	2	34.47	<.0001	24.82	<.0001		

Note. N denotes sample size and TP denotes time-points. I could not test a meta-model for β_1 because the empirical power level was often 1 across the conditions. Effect size measures (e.g., odds ratios) corresponding to the probed interactions are presented in the text.

[1] I did not generate more datasets and refit L-MM to adjust for the low convergence rates because I wanted to examine the relative performance of models when fitted to exactly the same data.

Table 18. *Results from logistic mixed effects meta-models for empirical power in the binary zero process for the OZINB-MM.*

Non-Zero Effects (Binary Process)							
		$\gamma_1: (t_t)$		$\gamma_3: (x_{2i})$		$\gamma_4: (t_t x_{1i})$	
Effect	**Num. DF**	**F(Den. DF)**	**p**	**F(948)**	**p**	**F(948)**	**p**
N	1	-	-	228.48	<.0001	117.06	<.0001
TP	1	-	-	96.38	<.0001	287.44	<.0001
Scale	1	-	-	0.29	.59	0.47	.49
N*TP	1	-	-	-	-	38.99	<.0001

Note: N denotes sample size and TP denotes time-points. I could not test a meta-model for γ_1 because the empirical power level was often 1 across conditions. Meta-models for the null effects ($\gamma_2$and γ_5) were not conducted because the Type I error rates were acceptable across all conditions.

Table 19. *Average estimate, average standard error, and proportion significant across the OZINB-MM, PO-MM, and L-MM for n=300 and t=3.*

	***n*=300 and *t*=3**												
	5 Response Categories							**10 Response Categories**					
	Count Process												
		OZINB		**PO-MM**		**L-MM**		**OZINB**		**PO-MM**		**L-MM**	
	True	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**
$\boldsymbol{\beta_0}$	**0.20**	0.19(.44)	0.07	-	-	0.30(.04)	1.00	0.18(.40)	0.07	-	-	0.53(.09)	1.00
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.36(.15)	0.72	0.57(.22)	0.76	0.16(.03)	1.00	0.36(.14)	0.78	0.52(.22)	0.70	0.31(.07)	1.00
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.01(.24)	0.08	-0.04(.22)	0.03	-0.01(.05)	0.02	0.00(.22)	0.07	-0.03(.22)	0.03	0.01(.09)	0.06
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.33(.24)	0.27	0.55(.23)	0.65	0.11(.05)	0.63	0.34(.23)	0.33	0.53(.23)	0.66	0.21(.09)	0.50
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.06(.15)	0.10	0.00(.13)	0.09	0.00(.03)	0.13	0.06(.14)	0.10	-0.01(.13)	0.08	0.00(.07)	0.09
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00(.15)	0.07	0.03(.14)	0.12	0.04(.03)	0.27	0.00(.14)	0.07	0.06(.14)	0.14	0.09(.07)	0.24
$\boldsymbol{\alpha}$	**1.00**	0.96(.38)	0.93	-	-	-	-	0.96(.32)	0.98	-	-	-	-
$\boldsymbol{\sigma_{v_0}}$	**1.00**	1.01(.22)	1.00	1.95(.41)	1	0.41(.07)	1.00	0.99(.19)	1.00	1.91(.40)	1.00	0.81(.15)	1.00
$\boldsymbol{\sigma_{v_1}}$	-	-	-	.60(.47)	.45	.26(.07)	0.94	-	-	0.66(.37)	0.52	0.55(.13)	0.94
$\boldsymbol{\sigma_{v_{01}}}$	-	-	-	-0.15(.62)	0	0.04(.03)	0.17	-	-	-0.10(.59)	0.00	0.17(.14)	0.18
	Zero Process												
$\boldsymbol{\gamma_0}$	**2.00**	2.04(.52)	0.96					2.05(.50)	0.97				
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.67(.29)	0.68					-0.67(.28)	0.77				
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	-0.01(.42)	0.03					-0.02(.40)	0.03				
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.54(.44)	0.25					-0.52(.42)	0.26				
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.12(.27)	0.05					0.12(.26)	0.06				
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	-0.01(.27)	0.04					-0.02(.26)	0.03				
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.62(.76)	0.96					2.60(.71)	.98				
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.08(1.18)	0.23					-1.13(1.01)	.28				

Note. Est is the average estimate, SE is the average standard error, and Sig is the proportion of significant effects at alpha=.05.

Table 20. *Average estimate, average standard error, and proportion significant across the OZINB-MM, PO-MM, and L-MM for n=300 and t=7.*

	***n*=300 and *t*=7**												
		5 Response Categories						**10 Response Categories**					
		Count Process											
		OZINB		**PO-MM**		**L-MM**		**OZINB**		**PO-MM**		**L-MM**	
	True	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**
$\boldsymbol{\beta_0}$	**0.20**	0.21(.18)	0.22	-	-	0.22(.05)	1.00	0.20(.16)	0.24	-	-	0.34(.09)	1.00
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.35(.03)	1.00	0.60(.04)	1.00	0.25(.01)	1.00	0.35(.03)	1.00	0.61(.04)	1.00	0.51(.03)	1.00
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.03(.15)	0.06	-0.15(.17)	0.11	-0.01(.05)	0.07	0.02(.14)	0.05	-0.15(.17)	0.14	-0.01(.10)	0.04
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.34(.15)	0.60	0.57(.17)	0.89	0.14(.05)	0.86	0.34(.14)	0.64	0.56(.17)	0.88	0.28(.10)	0.89
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.05(.03)	0.40	0.02(.03)	0.12	0.00(.01)	0.07	0.05(.03)	0.49	0.03(.03)	0.14	0.02(.03)	0.07
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00(.03)	0.03	0.00(.03)	0.05	0.03(.01)	0.70	0.00(.03)	0.04	0.00(.03)	0.06	0.07(.03)	0.70
$\boldsymbol{\alpha}$	**1.00**	0.98(.12)	1.00	-	-	-	-	0.98(.11)	1.00	-	-	-	-
$\boldsymbol{\sigma_{v_0}}$	**1.00**	0.99(.10)	1.00	1.90(.21)	1.00	.42(.06)	1.00	1.00(.09)	1.00	1.90(.20)	1.00	0.82(.14)	1.00
$\boldsymbol{\sigma_{v_1}}$	-	-	-	0.15(.07)	0.63	.14(.02)	1.00	-	-	0.16(.07)	0.65	0.08(.03)	1.00
$\boldsymbol{\sigma_{v_{01}}}$	-	-	-	-0.09(.10)	0.05	.04(.01)	0.96	-	-	-0.08(.10)	0.05	0.17(.05)	0.96
		Zero Process											
$\boldsymbol{\gamma_0}$	**2.00**	2.00(.28)	1.00					2.00(.28)	1.00				
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.65(.07)	1.00					-0.65(.07)	1.00				
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	0.02(.27)	0.03					0.02(.27)	0.04				
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.57(.27)	0.56					-0.57(.27)	0.58				
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.11(.06)	0.44					0.12(.06)	0.45				
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	0.00(.06)	0.03					0.00(.06)	0.04				
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.43(.28)	1.00					2.43(.27)	1.00				
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.00(.34)	.85					-1.00(.32)	0.85				

Note. Est is the average estimate, SE is the average standard error, and Sig is the proportion of significant effects at alpha=.05.

Table 21. *Average estimate, average standard error, and proportion significant across the OZINB-MM, PO-MM, and L-MM for n=1000 and t=3.*

		***n*=1000 and *t*=3**											
		5 Response Categories						**10 Response Categories**					
		Count Process											
		OZINB		**PO-MM**		**L-MM**		**OZINB**		**PO-MM**		**L-MM**	
	True	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**
$\boldsymbol{\beta_0}$	**0.20**	0.16(.24)	0.24	-	-	0.29(.02)	1.00	0.16(.22)	0.12	-	-	0.53(.05)	1.00
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.36(.08)	1.00	0.61(.12)	1.00	0.17(.02)	1.00	0.36(.08)	1.00	0.60(.12)	1.00	0.34(.04)	1.00
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	-0.01(.13)	0.04	-0.01(.12)	0.02	0.00(.03)	0.06	-0.01(.12)	0.06	-0.02(.12)	0.03	0.00(.05)	0.05
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.36(.13)	0.76	0.57(.13)	0.99	0.11(.03)	1.00	0.36(.13)	0.80	0.57(.13)	1.00	0.22(.05)	1.00
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.06(.08)	0.13	-0.02(.07)	0.06	0.00(.02)	0.09	0.07(.07)	0.18	-0.01(.07)	0.02	0.01(.04)	0.05
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00(.08)	0.05	0.02(.07)	0.02	0.04(.02)	0.59	0.00(.08)	0.04	0.03(.07)	0.02	0.08(.04)	0.59
$\boldsymbol{\alpha}$	**1.00**	1.00(.22)	1.00	-	-	-	-	1.00(.19)	1.00	-	-	-	-
$\boldsymbol{\sigma_{v_0}}$	**1.00**	1.01(.11)	1.00	1.99(.23)	1.00	0.39(.04)	1.00	1.00(.10)	1.00	1.97(.22)	1.00	0.76(.09)	1.00
$\boldsymbol{\sigma_{v_1}}$	-	-	-	0.46(.29)	0.50	0.25(.04)	1.00	-	-	0.48(.25)	0.65	0.52(.07)	1.00
$\boldsymbol{\sigma_{v_{01}}}$	-	-	-	-0.07(.33)	0.01	0.07(.02)	0.91	-	-	-0.09(.32)	0.02	0.28(.08)	1.00
		Zero Process											
$\boldsymbol{\gamma_0}$	**2.00**	1.99(.26)	1.00					1.99(.25)	1.00				
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.66(.14)	1.00					-0.66(.14)	1.00				
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	-0.02(.20)	0.06					-0.02(.20)	0.06				
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.54(.21)	0.76					-0.53(.21)	0.77				
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.13(.13)	0.16					0.13(.13)	0.18				
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	0.00(.13)	0.05					0.00(.13)	0.04				
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.46(.36)	1.00					2.45(.34)	1.00				
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.08(.51)	.64					-1.07(.47)	0.68				

Note. Est is the average estimate, SE is the average standard error, and Sig is the proportion of significant effects at alpha=.0

Table 22. *Average estimate, average standard error, and proportion significant across the OZINB-MM, PO-MM, and L-MM for n=1000 and t=7.*

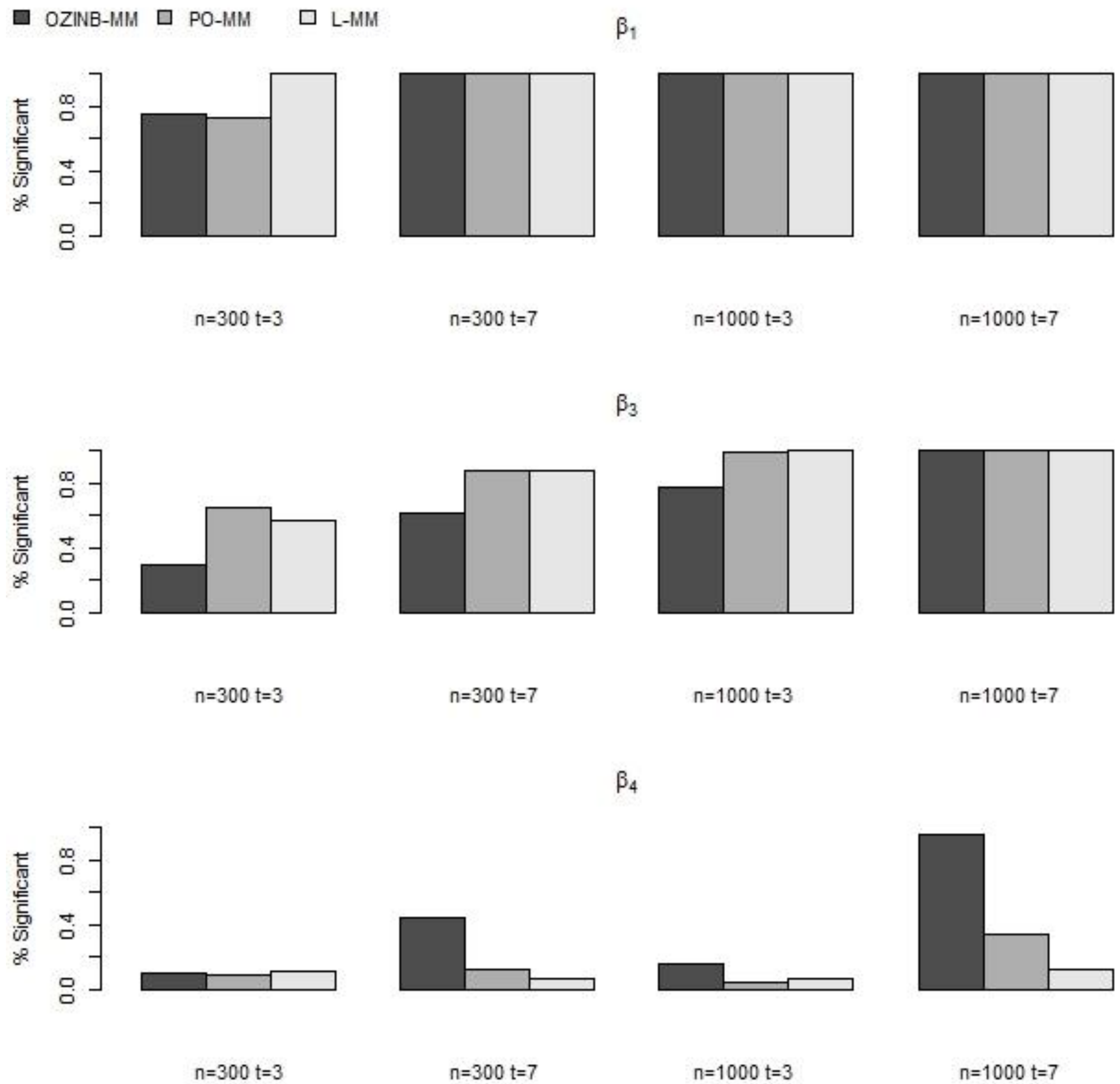
	***n*=1000 and *t*=7**												
	5 Response Categories							**10 Response Categories**					
	Count Process												
		OZINB		**PO-MM**		**L-MM**		**OZINB**		**PO-MM**		**L-MM**	
	True	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**	**Est(SE)**	**Sig**
$\boldsymbol{\beta_0}$	**0.20**	0.20(.10)	0.53	-	-	0.21(.02)	1.00	0.19(.09)	0.57	-	-	0.35(.05)	1.00
$\boldsymbol{\beta_1: (t_t)}$	**0.35**	0.35(.02)	1.00	0.61(.02)	1.00	0.25(.01)	1.00	0.35(.02)	1.00	0.61(.02)	1.00	0.52(.02)	1.00
$\boldsymbol{\beta_2: (x_{1i})}$	**0.00**	0.00(.08)	0.05	-0.16(.09)	0.39	-0.01(.03)	0.02	0.00(.08)	0.05	-0.17(.09)	0.45	-0.02(.05)	0.00
$\boldsymbol{\beta_3: (x_{2i})}$	**0.35**	0.35(.08)	1.00	0.55(.09)	1.00	0.13(.03)	1.00	0.35(.08)	1.00	0.55(.09)	1.00	0.27(.05)	1.00
$\boldsymbol{\beta_4: (t_t x_{1i})}$	**0.06**	0.06(.02)	0.95	0.02(.02)	0.29	0.00(.01)	0.06	0.06(.02)	0.98	0.03(.02)	0.40	0.02(.02)	0.18
$\boldsymbol{\beta_5: (t_t x_{2i})}$	**0.00**	0.00(.02)	0.05	0.00(.02)	0.05	0.03(.01)	0.98	0.00(.02)	0.05	0.01(.02)	0.07	0.08(.02)	1.00
$\boldsymbol{\alpha}$	**1.00**	1.00(.07)	1.00	-	-	-	-	0.99(.06)	1.00	-	-	-	-
$\boldsymbol{\sigma_{v_0}}$	**1.00**	0.99(.05)	1.00	1.92(.11)	1.00	0.40(.04)	1.00	1.00(.05)	1.00	1.91 (.11)	1.00	0.76(.08)	1.00
$\boldsymbol{\sigma_{v_1}}$	-	-	-	0.16(.04)	0.95	0.13(.01)	1.00	-	-	0.16(.04)	0.94	0.31(.02)	1.00
$\boldsymbol{\sigma_{v_{01}}}$	-	-	-	-0.09(.06)	0.32	0.04(.01)	1.00	-	-	-0.08(.06)	0.31	0.20(.03)	1.00
	Zero Process												
$\boldsymbol{\gamma_0}$	**2.00**	2.02(.15)	1.00					2.01(.15)	1.00				
$\boldsymbol{\gamma_1: (t_t)}$	**-0.65**	-0.65(.04)	1.00					-0.65(.04)	1.00				
$\boldsymbol{\gamma_2: (x_{1i})}$	**0.00**	0.01(.14)	0.05					0.01(.14)	0.05				
$\boldsymbol{\gamma_3: (x_{2i})}$	**-0.55**	-0.55(.15)	0.98					-0.55(.14)	0.97				
$\boldsymbol{\gamma_4: (t_t x_{1i})}$	**0.12**	0.12(.03)	0.98					0.12(.03)	0.98				
$\boldsymbol{\gamma_5: (t_t x_{2i})}$	**0.00**	0.00(.03)	0.05					0.00(.03)	0.06				
$\boldsymbol{\sigma_{u_0}}$	**2.45**	2.41(.15)	1.00					2.42(.14)	1.00				
$\boldsymbol{\sigma_{v_0 u_0}}$	**-1.00**	-1.02(.18)	1.00					-1.03(.17)	1.00				

Note. Est is the average estimate, SE is the average standard error, and Sig is the proportion of significant effects at alpha=.05.

Empirical Power for Count Non-Zero Parameters: β_1, β_3, β_4

The models showed high empirical power in identifying the positive linear age effect, β_1, see top panel of Figure 3 below for a graphical depiction. In fact, all of the models achieved an empirical power of 1 across all conditions except the OZINB-MM and PO-MM when *n*=300 and *t*=3. In this case, the OZINB-MM had the smallest empirical power for the 5pt scale whereas the PO-MM had the smallest empirical power for the 10pt scale. The increase in empirical power for the OZINB-MM in 10pt scale versus the 5pt scale appeared to be due to a decrease in estimated standard errors.

Figure 3. *Visualizing differences in empirical power for the non-zero effects in the count process across the OZINB-MM, PO-MM, and L-MM.*



Note. The plotted values are collapsed over the 5 and 10 response category conditions.

For the non-zero β_3 effect, the meta-model indicated that odds of obtaining a significant effect depended on the sample size, number of time-points, and models fitted. The odds of significant β_3 were larger when *n*=1000 than *n*=300 (OR=24.85, p<.0001) and *t*=7 than *t*=3 (OR=7.91, p<.0001). Further the odds of a significant effect were higher for the PO-MM than

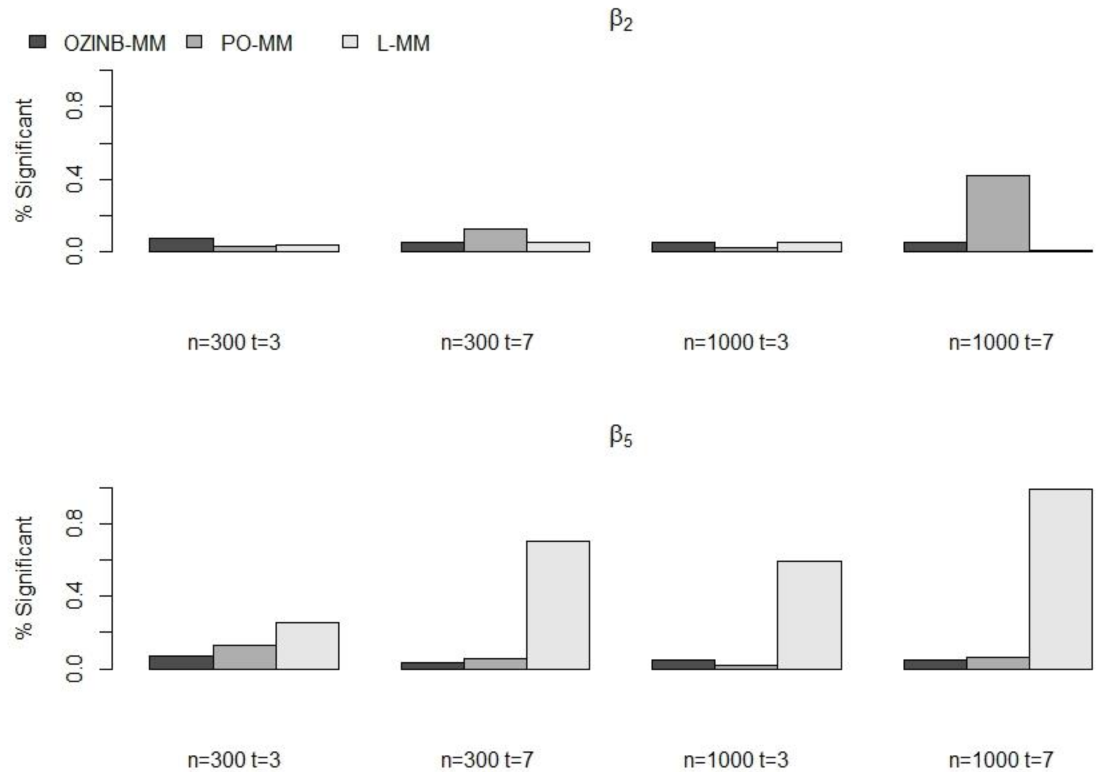
OZINB-MM (OR=9.18, p<.0001) and L-MM than OZINB-MM (OR=6.82, p<.0001). The odds of significant effect did not differ between the L-MM and PO-MM (p=.25). The middle panel of Figure 3 above shows how these differences in empirical power unfolded across conditions.

The meta-model indicated that empirical power for the β_4 interaction effect depended on a series of interactions between conditions. More precisely, sample size interacted with both number of time-points and model fitted and time-points interacted with model fitted in predicting the log-odds of a significant β_4 effect. When *n*=300 and *t*=3, the proposed OZINB-MM did not differ from the PO-MM (p=.99) or L-MM (p=.21) and the PO-MM did not differ from the L-MM (p=.31). However, when *n*=1000 and *t*=7, the odds of significant β_4 were larger for OZINB-MM than the PO-MM (OR=89.74, p<.0001) and L-MM (OR=614.99, p<.0001) and the odds were greater for the PO-MM than the L-MM (OR=6.85, p<.0001). Figure 3 above clearly shows the differences in empirical power across conditions.

Type I for Count Null Parameters: β_2 and β_5

Table 17 shows the results from the logistic mixed effects meta-models fitted to the null population parameters. The meta-model for the β_2 effect showed that the differences in Type I errors among the models depended on both sample size and number of time-points. When *n*=300 and *t*=3, the odds of significant effect were higher for the OZINB-MM than the PO-MM (OR=13.74, p<.0001), but not compared to the L-MM (p=.46). The L-MM odds were higher than those of the PO-MM (OR=8.98, p=.002). The Type I error rate for the OZINB-MM was only .08 so this did not seem to be overly problematic. When *n*=1000 and *t*=7, the odds of a significant β_2 were higher for the PO-MM than the OZINB-MM (OR=14.25, p<.0001) and L-MM (OR=20.80, p=.005), but there was no difference between the OZINB-MM and L-MM (p=.54). The Type I error rate was high for the PO-MM (.39-.45) given that there was no main effect of x_{1i} in the generating model. The Type I error rates for the OZINB-MM and L-MM were acceptable. The top panel of Figure 4 also illustrates these findings.

Figure 4. *Visualizing differences in Type I errors for the null effects in the count process across the OZINB-MM, PO-MM, and L-MM*



Note. The plotted values are collapsed over the 5 and 10 response category conditions.

The meta-model for the β_5 effect also showed that differences in Type I errors among the models depended on both sample size and number of time-points. However, the L-MM had a greater proportion of spurious interaction effects. When *n*=300 and *t*=3, the odds of significant effect were higher for the L-MM than the OZINB-MM (OR=4.28, p<.0001) and PO-MM (OR=3.65, p=.0009) whereas the PO-MM and OZINB-MM did not differ (p=.36). The Type I error rate was acceptable for the OZINB-MM and PO-MM, but the rate for the L-MM was high.

This same general pattern occurred for the other sample size by time-point combinations except the Type I error rate for the L-MM increased with more data while the Type I error rates for the OZINB-MM and PO-MM remained in an acceptable range. For instance, when *n*=300 and *t*=7, the odds of significant effect were substantially higher for the L-MM than the OZINB-MM (OR=81.17, p<.0001) and PO-MM (OR=52.03, p<.0001) whereas the PO-MM and OZINB-MM did not differ (p=.11). The bottom panel of Figure 4 shows this trend in Type I error rates across conditions.

With regards to the random effects components, the random intercept standard deviation σ_{v_0} was statistically significant across all models and conditions. Interestingly, for the converged PO-MM and L-MM, there was quite often significant variability in the time effect σ_{v_1}. The odds of significant random slope variability, σ_{v_1}, increased with more subjects (OR=3.78, p<.0001) and time-points (OR=4.11, p<.0001), but the number of response categories had no impact (p=.44). The linear mixed effects model had a higher proportion of significant random slope variance components but, again, it is important to recognize that far fewer L-MM models converged. These finding are especially notable because the population generating model did not include a random time effect. Empirical power for the random intercept and time covariance, $\sigma_{v_{01}}$, was generally low for the PO-MM (except when *n*=1000 and *t*=7), but quite high for the L-MM across conditions.

Summary of Count Process Results

In sum, the results for the count process indicated that empirical power and Type I error rates varied substantially across the ordinal zero-inflated negative binomial, proportional odds, and linear mixed effects models. These differences often depended on conditions such as sample size and number of time-points, but number of response categories had no meaningful impact.

Results showed that null generating effects were often statistically significant using the existing modeling techniques, which ignore the zero process. I next discuss empirical power and Type I errors for the logistic process of the OZINB-MM.

Empirical Power and Type I Errors in Logistic Process

The empirical power and Type I error results for the binary zero-process were as expected. Empirical power for the γ_1 time effect was 1.0 for all conditions except *n*=300 and *t*=3, where it was .96-.97. For the γ_3 effect, the odds of significant effect were higher for *n*=1000 than *n*=300 (OR=21.74, p<.0001) and *t*=7 than *t*=3 (OR=7.08, p<.0001), but the number of response categories had no effect (p=.59). Results from the meta-model for the γ_4 effect indicated that sample size interacted with number of time-points such that, when *n*=1000, having time-points had a stronger impact (OR=572.09, p<.0001) compared to when *n*=300 (OR=18.76, p<.0001). The empirical power for the logistic random intercept component was 1.00 for all conditions except when *n*=300 and *t*=3, where it was .97-.98. The odds of a significant covariance between the random count and logistic intercepts increased with more subjects (OR=10.55, p<.0001) and time-points (OR=37.19, p<.0001), but the number of response categories had no effect (p=.10). Further, the Type I error rates for the null $\gamma_2$and $\gamma_5$effects were acceptable across all conditions.

Comparing the Relative Efficiency of Predictions

Table 23 displays the relative efficiency of predictions produced by the existing and proposed mixed effects models. Results showed that the predictions from the proposed OZINB-MM were more efficient than both the PO-MM and L-MM (e.g., the relative efficiencies were greater than 1). The L-MM and PO-MM displayed more inefficiency relative to the OZINB-MM as the number of time-points and number of response categories increased. The efficiency ratio for the PO-MM relative to the L-MM were similar when *n*=300 and *t*=3. However, the PO-MM had better efficiency than the L-MM for all other conditions, particularly with seven time-points. These results were difficult to interpret because both the PO-MM and L-MM had low rates of model convergence when *n*=300 and *t*=3 and the L-MM had very low rates of model conversion when *t*=7.

Table 23. *Relative efficiency of predictions from the OZINB-MM, PO-MM, and L-MM across conditions.*

	n=300			
	3 Time-Points		7 Time-Points	
	5 Categories	10 Categories	5 Categories	10 Categories
$RE_{\frac{PO}{OZINB}}$	1.11	1.14	1.10	1.15
$RE_{\frac{L}{OZINB}}$	1.11	1.12	1.33	1.39
$RE_{\frac{PO}{L}}$	1.00	1.02	.83	.83
	n =1000			
$RE_{\frac{PO}{OZINB}}$	1.08	1.12	1.10	1.15
$RE_{\frac{L}{OZINB}}$	1.11	1.19	1.35	1.45
$RE_{\frac{PO}{L}}$	.97	.94	.81	.79

Note. Relative efficiencies greater than 1 indicate that the model in the denominator produced more efficient predictions.

Summary of Simulation Study Results

Taken together, the results showed that the proposed OZINB-MM performed well across all conditions. The proposed model performed best with more subjects and time-points. The number of response categories did not have a significant impact on model performance. Unsurprisingly, the ordinal zero-inflated negative binomial mixed effects model fitted the ordinal-count data better than the existing proportional odds and linear mixed effects models and produced more efficient predictions. The convergence rates for the L-MM were systematically low and the PO-MM had lower convergence rates with less subjects and time-points. Results also showed that the existing models can lead to spurious effects, partially because they disregard the zero process.

CHAPTER 4: DISCUSSION

I have introduced a novel modeling framework for longitudinal ordinal data that represent ranges of underlying counts. The framework explicitly links ordinal responses to an underlying count construct of substantive interest through cumulative probabilities. My simulation studies evaluated the ordinal negative binomial mixed effects model (ONB-MM) and ordinal zero-inflated negative binomial mixed effects model (OZINB-MM) for substance use data. Each of the studies assessed the general performance of the ordinal-count mixed effects models with regards to parameter recovery and the relative performance of proposed models compared to existing linear (L-MM) and proportional odds (PO-MM) mixed effects models. I will next discuss how the simulation studies supported my research hypotheses pertaining to the general performance of the ONB-MM and OZINB-MM. Then, I will discuss how the simulation results supported my hypotheses about the advantages of the proposed models and summarize the unique contributions of the present studies. Finally, I will provide recommendations for applied researchers, highlight limitations of the current studies, and offer directions for future research.

General Performance of ONB-MM and OZINB-MM

Results from the simulation studies supported my hypotheses that the ordinal negative binomial and ordinal zero-inflated negative binomial mixed effects models would adequately recover the population generating values across the full range of conditions. Slight biases in the random effects variance-covariance parameters and dispersion parameters occurred when the number of time-points and subjects were low. However, the performance of both models improved with more time-points and subjects (e.g., less bias, better accuracy, smaller standard errors, and higher rate of model convergence). The OZINB-MM showed low coverage for the 95% confidence interval of the dispersion parameter, but these coverage rates improved with added subjects and time-points. The number of response categories had a minimal impact on performance beyond slightly smaller standard errors for some parameters, which in turn influenced empirical power. The general pattern of results were consistent with analytic theory and prior research on ordinal and ordinal-count models with larger numbers of response categories (Bauer & Sterba, 2011; McGinley, Curran, & Hedeker, under review). In sum, my simulation studies showed that, assuming the underlying construct of interest truly follows the specified count distribution (e.g., negative binomial, zero-inflated negative binomial), the proposed ordinal-count models performed well.

Comparing ONB-MM and OZINB-MM with Existing Models

Results from the simulation studies also supported my hypotheses that the proposed ordinal-count mixed effects models would outperform commonly used linear and proportional odds mixed effects models. The model convergence rates for the ONB-MM were similar to the PO-MM and higher than the L-MM. When t=3, the convergence rates of the L-MM were low. Across all conditions, the ONB-MM had better model fit than the PO-MM and the PO-MM had better model fit than the L-MM. The empirical power rates differed across the models and conditions, but there were not meaningful substantive differences between the models (e.g., different patterns of effects). The number of subjects had a strong effect on empirical power for the β_3 effect corresponding to the time-invariant predictor, x_{2i}, whereas both the number of subjects and the number of time-points influenced empirical power for the $t_t x_{1i}$ interaction effect, β_4. Type I error rates were not concerning across the conditions. Results also demonstrated that the proposed ordinal negative binomial mixed effects model produced more efficient predictions relative to the existing models. Interestingly, the linear mixed effects model produced more efficient predictions than the proportional odds mixed effects model. This finding was unexpected and is worth future investigation for researchers that are interested in the efficiency of predictions.

As hypothesized, the OZINB-MM outperformed the existing models across all of the outcomes. The proposed ordinal zero-inflated negative binomial mixed effects model had higher convergence rates than the PO-MM and the PO-MM had higher convergence rates than the L-MM. The convergence rates for the L-MM were very low. This was not surprising because the zero-inflated ordinal-count data were not generated with a random time effect. As a result, fitting L-MM with correlated random intercept and slope effects caused difficulties in estimation. Given these low convergence rates, the L-MM results with the zero-inflated ordinal-count data may lack generalizability. I did not expect the PO-MM to have such high convergence rates or variability in the random time effect. The OZINB-MM provided better model fit and more efficient predictions than the PO-MM and L-MM. The PO-MM provided better model fit and more efficient predictions than the L-MM. However, I am uncertain whether the PO-MM was truly more efficient than the L-MM because such a small number of L-MM converged. This finding also conflicted with results from the underlying negative binomial distribution where the L-MM produced more efficient predictions than the PO-MM.

Results from my investigation of empirical power and Type I errors for the proposed OZINB-MM and existing PO-MM and L-MM fitted to the zero-inflated ordinal-count data were notable. The β_3 effect had substantially higher empirical power in the existing L-MM and PO-MM compared to the proposed OZINB-MM because the existing models ignored the zero process that also had a non-zero negative effect for x_{2i} (e.g., γ_3). For example, the positive β_3

effect indicated that increases in x_{2i} led to increases in the log expected counts for drinkers and the negative γ_3 indicated that increases in x_{2i} led to decreases in the log-odds of being a non-drinker. The opposite trend happened with the β_4 interaction effect. Specifically, the positive β_4 suggested that the log expected counts increased more over time for individuals with higher levels of x_{1i}, but the positive γ_4 from the zero process indicated that log-odds of being a non-drinker decreased less over time for individuals with higher levels of x_{1i}. This pattern of effects led to decreased empirical power for the β_4 interaction in the existing linear and proportional odds mixed effects models compared to the proposed OZINB-MM. In other words, the γ_4 effect from the zero process often washed out the count process β_4 effect causing the empirical power to decrease. These findings are especially important because prior substance use researchers have found that the direction of effects may take the same or different directional signs (e.g., positive or negative) for the count and binary zero processes (Sheu, Hu, Keeler, Ong, Sung, 2004; Mrug & McCay, 2013). Thus, substantive conclusions drawn from existing models may be biased if a true underlying zero-inflated distribution is ignored.

Type I error rates for the OZINB-MM were adequate across all conditions. However, the L-MM and PO-MM had high Type I error rates when fitted to the zero-inflated data. Specifically, the proportional odds model had high Type I error rates for the null β_2 effect and the L-MM had high Type I error rates for the β_5 null interaction effect. These spurious effects were unexpected given that x_{1i} and $t_t x_{2i}$ had null generating effects in both the count and zero processes. It seems plausible that assumption violations (e.g., proportional odds) and model misspecification (e.g., ignoring the zero process) impacted these Type I error rates. These findings are significant because applied researchers could draw incorrect inferences from standard models fitted to zero-inflated ordinal-count data. However, it should be noted that the

OZINB-MM has a unique interpretation compared to the existing models because it is a mixture of negative binomial (count) and logistic distributions (binary). Model non-convergence was also a concern with zero-inflated data, especially for the L-MM and the PO-MM when t=3. In sum, my two simulation studies demonstrated that the proposed ordinal-count mixed effects models effectively recovered the underlying count processes and showed advantages over commonly used mixed effects models.

Unique Contribution of the Present Studies

Methodologists have previously discussed statistical models for longitudinal substance use data (Rose, Chassin, Presson, Sherman, 2000). This work often aims to disseminate proportional odds models for ordinal outcomes, linear models fitted to ordinal outcomes, and various count mixed effects models for count outcomes using empirical demonstrations (Atkins, Baldwin, Zheng, Gallop, & Neighbors, 2013; Curran, 2000; Hedeker, Gibbons, & Flay, 1994). However, few studies focus on rigorously evaluating the appropriateness of these methods for substance use research using simulation studies. Further, to my knowledge, no previous studies have outlined how to fit count mixed effects models to longitudinal ordinal data. In fact, methodologists have suggested that this is not even possible (Koran & Hancock, 2010).

A small number of researchers have developed methods for handling grouped counts (Carter, Bowen, & Myers, 1971; Carter & Myers, 1973; Moffatt, 1995; Moffatt & Peters, 2000). However, we were the first to embed this methodology within the broader ordinal modeling perspective motivated by an underlying count latent response variable (McGinley, Curran, & Hedeker, under review).My work here is unique from prior research in three ways. First, I described how the ordinal-count framework can be extended to longitudinal data. I focused on outlining two novel ordinal-count mixed effects models (the ONB-MM and OZINB-MM) that are applicable for substance use research. Second, I evaluated the performance of these models across a variety of conditions using simulation studies. Third, I demonstrated the substantive and quantitative benefits of the proposed ordinal-count mixed effects models compared to existing methods currently used by substance use researchers.

Recommendations for Applied Researchers

Substance use outcomes are regularly collected as ordinal items with underlying counts (Currie et al., 2012; Johnston et al., 2012; NIAAA, 2003). However, researchers are rarely interested in these ordinal variables at the level of the response categories themselves (e.g., category 1 compared to all other categories; categories 1 and 2 compared to all other categories, etc.). This disjoint between the fitted models and the substantive construct of interest is a significant limitation of the proportional odds mixed effects model. Similarly, substance use researchers regularly utilize standard linear mixed effects models with ordinal category number scores to make inferences even though this strategy lacks a clearly defined metric. In addition to this interpretational disconnect, my simulation studies suggested that there may be situations when fitting these existing models to ordinal-count data leads to spurious effects (e.g., high Type I error rates) and erroneous inferences. Together, these limitations impact the researchers' ability to accurately and reliably test theory.

Fortunately, my proposed ordinal-count mixed effects models permit researchers to make inferences about the count constructs that underlie the ordinal categories. My simulation studies showed that the ordinal negative binomial and ordinal zero-inflated negative binomial mixed effects models performed well across a variety of conditions. Results suggested that researchers should be able to reliably fit these models with a reasonable amount of data. More complicated random effects structures (e.g., polynomial time trends for the ONB-MM or random time effects for the OZINB-MM) would likely require more data for satisfactory performance. As always, increased subjects and repeated measures will improve model performance but increasing the number of response categories will likely have little impact if the categories are reasonably defined. Researchers should be conscientious about how the underlying count distribution is specified in these models (e.g., Poisson vs. negative binomial vs. zero-inflated). The selection of an appropriate underlying count distribution can be guided by both empirical theory and model comparisons.

Applied researchers should also be aware of potential difficulties associated with these ordinal-count mixed effects models. First, the proposed mixed effects models were fitted using PROC NLMIXED in SAS, which is not especially user friendly. These models cannot currently be fitted in other more accessible statistical software packages. Second, these models may require good starting values to reach convergence. Prior research and experience can be useful to help guide the selection of these values. Third, fitting these models is computationally burdensome. In the simulation studies, convergence for a single replication of the ONB-MM usually took between 30 minutes and 90 minutes and the OZINB-MM took upwards to four hours. Applied researchers can use some strategies to reduce this burden (e.g., decreasing the number of quadrature points, using different optimizations methods, loosening convergence criteria), however these changes can also negatively impact model performance. This computational burden may restrict researchers' ability to test a large number of competing models. In sum, although there are currently a few practical concerns with ordinal-count mixed effects models, these novel methods are generally quite robust and well suited for substance use data.

Limitations and Future Directions

My simulation studies considered two potential ordinal-count mixed effects models. I focused on the ordinal negative binomial and ordinal zero-inflated negative binomial mixed effects models because they aligned with substance use research. However, there many other possible models that I did not consider including Poisson models, negative binomial models with heterogeneous dispersion, Poisson and negative binomial Hurdle models, and more (Hilbe, 2011). Further, empirical theory may posit alternative random effects structures (e.g., random time effects for the ordinal zero-inflated model). It is unknown whether or not the findings here generalize to these other models. I also examined commonly employed linear and proportional odds mixed effects models, but other lesser used models could have been considered (e.g., partial proportional odds, non-proportional odds, nominal models). It should be restated that, in certain conditions, the existing mixed effects models frequently failed to converge and were not re-estimated. Thus, it is unknown to what extent using only the converged models influenced the model comparison findings.

A limitation of the ordinal-count modeling framework is that it assumes that the cut-points are fixed and known. If these cut-points are not known, the modeling strategy cannot be used. Further, ordinal-count models face limitations similar to their standard count model counterparts. For instance, if the underlying distribution is misspecified the accuracy of results will be impacted. Although it does not undermine the findings here, future research should investigate the performance of a wider array of ordinal-count mixed effects models under various conditions (e.g., missing data, varying effect sizes, alternative generating distributions, different random effects structures, unreliability in participant responses). Despite these potential limitations, my studies provide several unique contributions above and beyond the existing literature. I outlined two new ordinal models that link observed longitudinal ordered data to underlying count constructs and used simulation studies to evaluate the performance of these models across a variety of conditions. Finally, I demonstrated that ordinal-count mixed effects models offer both substantive and quantitative advantages compared to existing techniques that can help advance substance use research in meaningful ways.

REFERENCES

Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle. In Second international symposium on information theory* (pp. 267-281). Akademinai Kiado.

Atkins, D. C., Baldwin, S. A., Zheng, C., Gallop, R. J., & Neighbors, C. (2013). A tutorial on count regression and zero-altered count models for longitudinal substance use data. *Psychology of Addictive Behaviors*, *27*(1), 166-177.

Bauer, D.J. & Sterba, S.K. (2011). Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation. *Psychological Methods*, *16*, 373-390.

Buu, A., Li, R., Tan, X., & Zucker, R. A. (2012). Statistical models for longitudinal zero-inflated count data with applications to the substance abuse field. *Statistics in Medicine*, *31*(29), 4074-4086.

Carter Jr, W. H., Bowen Jr, J. V., & Myers, R. H. (1971). Maximum likelihood estimation from grouped Poisson data. *Journal of the American Statistical Association*, *66*(334), 351-353.

Carter Jr, W. H., & Myers, R. H. (1973). Maximum likelihood estimation from linear combinations of discrete probability functions. *Journal of the American Statistical Association*, *68*(341), 203-206.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

Collins, L. M., Schafer, J. L., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*(4), 330-351

Curran, P. J. (2000). A latent curve framework for the study of developmental trajectories in adolescent substance use. In J. S. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions* (pp. 1-42). Mahwah, NJ: Erlbaum.

Curran, P. J., Stice, E., & Chassin, L. (1997). The relation between adolescent alcohol use and peer alcohol use: A longitudinal random coefficients model. *Journal of Consulting and Clinical Psychology*, *65*(1), 130-140.

Currie C et al. (2012). Social determinants of health and well-being among young people. *Health Behaviour in School-aged Children (HBSC) study: international report from the 2009/2010 survey*. WHO Regional Office for Europe, (Health Policy for Children and Adolescents, No. 6): Copenhagen.

Desantis, S. M., Bandyopadhyay, D., Baker, N. L., Randall, P. K., Anton, R. F., & Prisciandaro, J. J. (2013). Modeling longitudinal drinking data in clinical trials: An application to the COMBINE study. *Drug and Alcohol Dependence, 132*(1-2):244-250.

Feldman, B. J., Masyn, K. E., & Conger, R. D. (2009). New approaches to studying problem behaviors: a comparison of methods for modeling longitudinal, categorical adolescent drinking data. *Developmental psychology*, *45*(3), 652.

Fleming, C. B., Mason, W. A., Mazza, J. J., Abbott, R. D., & Catalano, R. F. (2008). Latent growth modeling of the relationship between depressive symptoms and substance use during adolescence. *Psychology of Addictive Behaviors*, *22*(2), 186-197.

Greene, W.H. (1994). *Accounting for excess zeros and sample selection in Poisson and negative binomial regression models.* Working Paper, Department of Economics, Stern School of Business, New York University.

Hall, D. B. (2000). Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics*, *56*(4), 1030-1039.

Hedeker, D., & Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics*, 933-944.

Hedeker, D. R., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Hoboken, N.J: Wiley.

Hedeker, D., Gibbons, R. D., & Flay, B. R. (1994). Random-effects regression models for clustered data with an example from smoking prevention research. *Journal of Consulting and Clinical Psychology*, *62*(4), 757-765.

Hedeker, D., & Mermelstein, R. J. (2000). Analysis of longitudinal substance use outcomes using ordinal random-effects regression models. *Addiction*, *95*(11s3), 381-394.

Hur, K., Hedeker, D., Henderson, W., Khuri, S., & Daley, J. (2002). Modeling clustered count data with excess zeros in health care outcomes research.*Health Services and Outcomes Research Methodology*, *3*(1), 5-20.

Hussong, A. M., Jones, D. J., Stein, G. L., Baucom, D. H., & Boeding, S. (2011). An internalizing pathway to alcohol use and disorder. *Psychology of Addictive Behaviors*, *25*(3), 390-404.

Johnston, L. D., O'Malley, P. M, Bachman, J. G., & Schulenberg J E (2012). *Monitoring the Future national results on adolescent drug use: Overview of key findings, 2011*. Institute for Social Research, The University of Michigan: Ann Arbor.

Kiernan, K., Tao, J., & Gibbs, P. (2012). Tips and strategies for mixed modeling with SAS/STAT® Procedures. *In SAS Global Forum* (pp. 332-2012).

King, K. M., Molina, B. S., & Chassin, L. (2009). Prospective relations between growth in drinking and familial stressors across adolescence. *Journal of abnormal psychology*, *118*(3), 610-622.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, *34*(1), 1-14.

Lesaffre, E., & Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random effects model: an example. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *50*(3), 325-335.

Liu, H., & Powers, D. A. (2007). Growth curve models for zero-inflated count data: An application to smoking behavior. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(2), 247-279.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. Psychological methods, 7(1), 19.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models (2nd ed.)*. London: Chapman & Hall.

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, *4*(1), 103-120.

Min, Y., & Agresti, A. (2005). Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling*, *5*(1), 1-19.

Moffatt, P. G. (1995). Grouped Poisson regression models: Theory and an application to public house visit frequency. *Communications in Statistics-Theory and Methods*, *24*(11), 2779-2796.

Moffatt, P. G., & Peters, S. A. (2000). Grouped zero-inflated count data models of coital frequency. *Journal of Population Economics*, *13*(2), 205-220.

Moore, A. A., Gould, R., Reuben, D. B., Greendale, G. A., Carter, M. K., Zhou, K., & Karlamangla, A. (2005). Longitudinal patterns and predictors of alcohol consumption in the United States. *American journal of public health*, *95*(3), 458-465.

Mrug, S., & McCay, R. (2013). Parental and peer disapproval of alcohol use and its relationship to adolescent drinking: Age, gender, and racial differences. *Psychology of Addictive Behaviors*, *27*(3), 604.

National Institute on Alcohol Abuse and Alcoholism [NIAAA] (2003). Recommended alcohol questions. Retrieved 6/3/2013 from http://www.niaaa.nih.gov/research/ guidelines-and-resources/recommended-alcohol-questions.

Norton, E. C., Lindrooth, R. C., & Ennett, S. T. (1998). Controlling for the endogeneity of peer substance use on adolescent alcohol and tobacco use. *Health Economics*, *7*(5), 439-453.

Otten, R., van Lier, P. A., & Engels, R. C. (2011). Disentangling two underlying processes in the initial phase of substance use: Onset and frequency of use in adolescent smoking. *Addictive Behaviors*, *36*(3), 237-240.

Rose, J. S., Chassin, L., Presson, C., & Sherman, S. J. (Eds.). (2000). *Multivariate applications in substance use research: New methods for new question*s. Hillsdale, NJ: Erlbaum.

Rubinstein, R. Y. (2008). *Simulation and the Monte Carlo method (2nd ed.)*. Hoboken, N.J.: John Wiley & Sons.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, *7*(2), 147-177.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464.

Sheu, M. L., Hu, T. W., Keeler, T. E., Ong, M., & Sung, H. Y. (2004). The effect of a major cigarette price change on smoking behavior in California: a zero-inflated negative binomial model. *Health Economics, 13*(8), 781-791.

Siddiqui, O. (1996). Modeling clustered count and survival data with an application to a school-based smoking prevention study (thesis/dissertation). *Chicago, University of Illinois at Chicago Center for Health Statistics*.

Siddiqui, O., Mott, J., Anderson, T., & Flay, B. (1999). The application of Poisson random-effects regression models to the analyses of adolescents' current level of smoking. *Preventive Medicine*, *29*(2), 92-101.

Simons-Morton, B., & Chen, R. S. (2006). Over time relationships between early adolescent and peer substance use. *Addictive Behaviors*, *31*(7), 1211-1223.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. New York: Oxford University Press.

Taylor, A. B., West, S. G., & Aiken, L. S. (2006). Loss of power in logistic, ordinal logistic, and probit regression when an outcome variable is coarsely categorized. *Educational and Psychological Measurement, 66*(2), 228-239.

Winship, C., & Mare, R. D. (1984). Regression models with ordinal variables. *American Sociological Review*, *49*(1), 512-525.

Xie, H., Tao, J., McHugo, G. J., & Drake, R. E. (2013). Comparing statistical methods for analyzing skewed longitudinal count data with many zeros: An example of smoking cessation. *Journal of Substance Abuse Treatment, 45*(1), 99-108.

Yau, K. K., Wang, K., & Lee, A. H. (2003). Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal*, *45*(4), 437-452.