



Published in final edited form as:

Methodology (Gott). 2016 April ; 12(2): 33–43. doi:10.1027/1614-2241/a000106.

Modeling Likert scale outcomes with trend-proportional odds with and without cluster data

Ana W. Capuano^{1,*}, Jeffrey D. Dawson², Marizen R. Ramirez³, Robert S. Wilson¹, Lisa L. Barnes¹, R. William Fields³

¹Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA

²Department of Biostatistics, Iowa City, IA, USA

³Department of Occupational and Environmental Health, Iowa City, IA, USA

Abstract

Likert scales are commonly used in epidemiological studies employing surveys. In this tutorial we demonstrate how the proportional odds model and the trend odds model can be applied simultaneously to data measured in Likert scales, allowing for random cluster effects. We use two datasets as examples: an epidemiological study on aging and cognition among community-dwelling Black persons, and a clustered large survey data from 28,882 students in 81 middle schools. The first example models the Likert outcome from the question: “People act as if they think you are dishonest”. The trend-proportional odds model indicates that Black men have higher odds than Black women of reporting being perceived dishonest. The second example models the Likert outcome from the question: “How often have you been beaten up at school?”. The trend-proportional odds model indicates that children with disability have a higher odds of severe violence than other children. For both examples, the cumulative odds ratio increases by more than 60% at the higher Likert levels.

Keywords

trend odds; proportional odds; likert; survey; ordinal models; logistic regression

Introduction

Likert scales are commonly used in epidemiological studies. Individual Likert items are discrete and ordinal by definition, commonly with three to seven possible levels depending on the design. Likert questions are useful when relying on the participant’s report of experiences or symptoms (Huang et al., 2010). These questions provide a closed ordinal array of response options (Likert, 1932). It is relatively common practice to have multiple Likert questions of interest. Often, the main outcome is obtained by summing individual Likert scores. These scores are adopted values, for example they can go from 0 to 3 in a four-level Likert. This paper does not focus on such composite Likert scores but on individual Likert outcomes based on a single Likert item. Studies that employ a single Likert

*corresponding author: ana_capuano@rush.edu.

item may also be motivated by secondary hypothesis or by the need to generate preliminary data for future research.

Analysis of Likert outcomes should involve methods designed for ordinal data. However, these scales are often dichotomized for analyses, resulting in loss of information and reduction of power (Capuano, Dawson, & Gray, 2007). Another practice is to use models based on normal distribution. This approach is commonly criticized (Agresti, 2002; Clason & Dormody, 1994) given the discrete nature of the Likert (e.g. flooring, ceiling and no halves) and the fact that data are often skewed. A better alternative is, therefore, to use methods designed for this type of data, such as ordinal logistic regressions. Constrained cumulative ordinal models are a popular parsimonious method of analyzing ordinal outcomes. These include the proportional odds model (Aitchison & Silvey, 1957; McCullagh, 1980; Snell, 1964) and the trend odds model (Capuano & Dawson, 2013). These models are an extension of the binary logistic regression for ordinal outcomes with more than two levels.

In this paper, we first discuss the detailed formulas for the proportional odds model, the trend odds model, and a hybrid model known as the proportional-trend odds model. We then give a tutorial on how to fit these models. These models are also extended to clustered data, which, to our knowledge, has not been demonstrated in the literature for the case of trend odds models. Simple and complex examples with different 4-level Likert outcomes are used to illustrate the basic multivariate model and the extension. To illustrate the basic application, we model the Likert outcome from the question: “People act as if they think you are dishonest”. To illustrate the application to clustered data we model the Likert outcome from the question: “How often have you been beaten up at school?”. Sample SAS and R code are provided in the appendices.

The Proportional Odds Model

Consider a Likert outcome Y with $k+1$ categories. The “cumulative odds” are the odds of Y being at least equal to k . The cumulative odds can also be related to a predictor variable X_1 through a linear function based on the inverse log. We can express this function as

$$\log\left(\frac{\Pr(Y \geq k | X_1)}{\Pr(Y < k | X_1)}\right) = \alpha_k + \beta_{k,1}X_1, \quad (1)$$

where $\beta_{k,1}$ is the parameter associated with the log odds ratio at each level of the outcome Y for predictor variable X_1 .

In the unconstrained cumulative ordinal model, all $\beta_{k,x}$ are different and unrelated. That is $\beta_{1,x} \neq \beta_{2,x} \neq \beta_{3,x} \neq \beta_{4,x}$. In other words, there are no constraints imposed on the cumulative log odds ratio, $\beta_{k,x}$ (Agresti, 2002; McCullagh, 1980; Peterson, Harrell, & Brant, 1992; Peterson & Harrell, 1990). That there are multiple unrelated cumulative odds ratios makes interpretation of this model more challenging in practice. The number of parameters $\beta_{k,x}$ increases with the number of levels in the outcome. For example, a complex 7-level-Likert will have six cumulative odds ratios associated with each predictor, resulting in less parsimony and possible model fitting problems. The constrained cumulative ordinal models

are a family of models that try to identify a relationship between the cumulative odds ratios. These models are theoretically justified by the underlying logistic distribution. The proportional odds model and the trend odds models belong to the constrained cumulative ordinal family of models.

The proportional odds model was initially discussed by Aitchison and Silvey in 1957 and Snell in 1964, but popularized and named by McCullagh in 1982. The development of the method is based on an underlying (unobserved or latent) continuous variable that follows a logistic distribution. It is possible to calculate log odds ratios by assuming cut-points in the distribution (i.e. probability above or below any value in the distribution). When there are shifts in the location of the distribution with changes in a predictor (i.e. males may have lower mean than females), it has been proved that proportional log odds ratios are observed. Note that this latent variable is not directly modeled.

The proportional odds model works on the assumption that all $\beta_{k,x}$ are represented by a common $\beta_{\bullet,x}$, that is $\beta_{\bullet,x} = \beta_{1,x} = \beta_{2,x} = \beta_{3,x}$ (McCullagh, 1980). That leads to the proportional odds assumption that the odds of being above any given level are the same. In other words, instead of having several different odds ratio for a given independent variable, a single odds ratio is calculated. In this model, the cumulative odds is related to an independent variable X_1 through a linear function based on the inverse log. We can express this function as

$$\log\left(\frac{\Pr(Y \geq k | X_1)}{\Pr(Y < k | X_1)}\right) = \alpha_k + \beta_{\bullet,1}X_1. \quad (2)$$

Because the proportional odds model summarizes data more succinctly, results are easier to interpret and it is the preferred model if the proportional odds assumption is met. The proportional odds assumption can be tested using a score test for proportional odds when the sample is not too large (Scott, Goldberg, & Mayo, 1997; Stokes, Davis, & Koch, 2012). When the sample is large, graphical displays of observed odds and estimates from binary logistic regression can be used (Scott et al., 1997).

The Trend Odds Model

The trend odds model was developed by Capuano and Dawson in 2012 as an attempt to provide a parsimonious alternative to the proportional odds model. The development of the method is also based on an underlying continuous variable that follows a logistic distribution. When there are shifts in the scale of the distribution with changes in a predictor (i.e. males have more variability than females), it has been proved that trend log odds ratios are observed. The theoretical shifts in location and scale are of practical interest because they yield an increasing risk or protective effect that is represented respectively by all log odds below zero or all log odds above zero. As in the case for the proportional odds model, the latent variable is not directly modeled.

The trend odds models works on the assumption that all $\beta_{k,x}$ have a monotonic relationship. That is $\beta_{1,x} < \beta_{2,x} < \beta_{3,x}$ or $\beta_{1,x} > \beta_{2,x} > \beta_{3,x}$ (Capuano & Dawson, 2013). That means that the cumulative odds have a trend, increasing or decreasing monotonically with the increase of outcome level k . In this model, the cumulative odds is also related to an independent variable

through a linear function based on the inverse log. For an independent variable X_2 , the trend odds model can be expressed as

$$\log\left(\frac{\Pr(Y \geq k | X_2)}{\Pr(Y < k | X_2)}\right) = \alpha_k + \beta_{1,2}X_2 + \gamma_2 t_k X_2, t_k < t_{k+1}. \quad (3)$$

The scalar values t_k are assumed to be equal to $k-1$ when no *a priori* information is available, for example, from previous studies.

The Proportional-Trend Odds Model

The above models have being previously used with different types of ordinal data such as classification groups, levels of antibody titers and temperament scales (Capuano et al., 2007; Singh, Bard, & Jackson, 2014; Zerwas et al., 2012). Models often include multiple covariates (multivariate models). The proportional odds assumption may not hold for all the covariates in the model. Although some researchers are careful in examining the assumption, violations are sometimes ignored (Singh et al., 2014). Peterson and Harrell discussed the relief of the proportional odds assumption for some but not all covariates. This strategy has being used by some researchers so that some covariates have one beta (proportional odds) and others have multiple betas (unconstrained odds) (Zerwas et al., 2012). Alternatively, we can use a combination of the proportional odds and the trend odds, here called the proportional-trend odds model. This strategy implies that some predictors assume proportional odds and some assume trend odds.

Consider again a Likert outcome Y with $k+1$ categories. There may be two covariates of cumulative log odds of Y given X_2 be an increasing or decreasing trend. The proportional-trend odds model can be expressed as

$$\log\left(\frac{\Pr(Y \geq k | X_1, X_2)}{\Pr(Y < k | X_1, X_2)}\right) = \alpha_k + \beta_{\cdot,1}X_1 + \beta_{1,2}X_2 + \gamma_2 t_k X_2, t_k < t_{k+1} \quad (4)$$

In certain designs, randomization of data occurs at a cluster level. For example several participants may be recruited from a specific location, leading to within-location correlation. Random effects models are commonly used because they can handle unbalanced clusters as discussed by Hedeker and others (Hedeker, Gibbons, & Flay, 1994). Consider that k is a nested observation within cluster j . Equation (4) can be expanded to a random cluster effect with the trend-proportional odds expressed as

$$\log\left(\frac{\Pr(Y \geq k | X_{1,j}, X_{2,j})}{\Pr(Y < k | X_{1,j}, X_{2,j})}\right) = \alpha_k + \beta_{\cdot,1}X_{1,j} + \beta_{1,2}X_{2,j} + \gamma_2 t_k X_{2,j} + \varphi_j, t_k < t_{k+1} \quad (5)$$

The random cluster effect φ_j is assumed to follow a normal distribution with mean equal to zero and variance equal to σ .

The proportional-trend odds model assumes that the log odds are either proportional or have a specific patterns of increasing or decreasing trends. The proportional odds and the trend odds are theoretically justified by an underlying logistic distribution that has shifts

in location and scale, respectively. Less parsimonious parameterizations can be used for some covariates (i.e. unconstrained) for exploration analyses of covariates that violate both proportional and trend odds assumptions.

Tutorial

Analytical Strategy

To fit the proportional-trend odds model the researcher can follow some steps detailed in this section⁹. Preliminary analyses of the data should include an examination of the frequency distribution. If data are too sparse, ordinal modeling may not be possible (i.e. only two of the Likert levels have frequencies greater than 5), or it may be necessary to combine the last or first Likert level with the closest level. Assuming that there are enough observations per ordinal level, analyses can follow four steps.

Step 1 - Bivariable Analyses - Binary Logistic Regression—The objective of this step is to obtain cumulative odds for each possible cut-point in the initial inspection of the data. For example a 4-level Likert has three cut-points and three cumulative odds ratios: θ_1 , θ_2 , and θ_3 . Cumulative odds ratios for each possible cut-point can be obtained by fitting separate logistic regressions per cut-point (i.e. a Likert 0123 have three indicator variables with values: 0111, 0011, and 0001). Modeling can be performed using the Glm procedure in R (with family specified as binomial), or the Proc Logistic procedure in SAS. Alternatively one can fit the unconstrained model that will generate several parameters, one per cut-point. That can be accomplished by using the Proc Nlmixed procedure in SAS or the Vglm procedure in R (with family specified as cumulative and parallel specified as false). In contrast to SAS Proc Nlmixed, R Vglm does not require the user to select initial values and it is easier to use, especially for the less sophisticated user. Graphical displays of the log odds such as in Figures 1 and 2 can be used to investigate the proportional odds assumption and the trend odds assumption.

Step 2 - Bivariable Analyses - Proportional Odds Model—The objective of this step is to inspect the significance of the independent variables when entered individually in the proportional odds model and to assess the proportional odds assumption. The proportional odds model can be fit with the Proc Logistic procedure in SAS or the Glm procedure in R (with family specified as cumulative and parallel specified as true). The score test for proportional odds assumption is obtained from the SAS procedure logistic. In R the proportional odds assumption can be tested using likelihood ratio tests calculated from the difference in the deviance of proportional odds model and the deviance of the unconstrained model.

The proportional odds model provides a common odds ratio. That is the odds ratio of being above any level compared to the odds ratio of being below the level are the same ($\theta = \theta_1 = \theta_2 = \theta_3$). A significant or borderline score test provides partial support for rejecting the proportional odds assumption because large samples have sufficient power to detect very small deviations of proportionality of odds. A non-significant score test for proportional odd or likelihood ratio test results in failure to reject the proportional odds assumption. For

example, a p value above .2 provides strong evidence that the proportional odds model is adequate. In this case, step 3 is particularly important to verify if the proportional odds assumption is being rejected due to a trend in odds.

In SAS, parameters estimates from the proportional odds model with independent variables that present a trend can be stored to be used as initial values for modeling in step 3.

Step 3 - Bivariable Analyses - Trend Odds Model—The objective of this step is to inspect the significance of the independent variables when entered individually in the trend odds model and respective trend parameter γ . The trend odds model can be fit in SAS with the Proc Nlmixed procedure, and approximate calculation can be performed in R using parameters obtained from the Vglm procedure. The initials required for modeling with Proc Nlmixed are: (a) obtained from the proportional odds model, or (b) entered as zero. For example, estimated values for the three intercepts and beta from the proportional odds model of 4-level Likert can serve as initials for the three intercepts and beta of the trend odds model. Gamma can have zero as the initial value.

A significant or borderline score test for proportional odds ascertained in step 2 and a significant trend parameter γ obtained in this step together provide sufficient evidence to reject the proportional odds assumption and accept the trend odds assumption. A trend odds model provides the baseline odds and a multiplicative trend effect. In a 4-level Likert the scalars $t_1 = 0, t_2 = 1, t_3 = 2$ would yield the cumulative odd ratios: $\theta_1 = e^\beta, \theta_2 = e^\beta e^\gamma, \theta_3 = e^\beta e^{2\gamma}$. Note that there is a monotonic multiplicative effect on cumulative odds ratios ($\theta_1 = e^\beta, \theta_2 = \theta_1 e^\gamma, \theta_3 = \theta_2 e^\gamma$). The multiplicative trend effect is used to calculate the consecutive odds ratio when multiplying it to previous odds ratio. A multiplicative trend effect above one indicates that the odds ratios are increasing with each level of the Likert outcome (a positive trend, $\theta_3 > \theta_2 > \theta_1$). Similarly, a multiplicative trend effect below one indicates that the odds ratios are decreasing with each level of the Likert outcome (a negative trend, $\theta_3 < \theta_2 < \theta_1$). Again we can use graphical displays such as figures 1 and 2 to compare the observed odds to the odds calculated with the proportional odds model and odds calculated with the trend odds model.

Step 4 - Multivariable Proportional-Trend Analyses—The objective of this step is to identify the trend parameter and to build the final multivariate model. First, a proportional odds model including all the independent variables can be fit in SAS using Proc Logistic or in R using the Vglm procedure. In SAS, parameters estimates can be stored to be used as initials values of the proportional-trend model.

In SAS, the proportional-trend odds model can be fit with the Proc Nlmixed procedure. Similar to step 3 the seeds required for modeling with Proc Nlmixed are: (a) obtained from the multivariable proportional odds model using Proc Logistic, or (b) entered as zero. Multivariable modeling can start with a saturate trend odds model including all the independent variables and respective baseline and trend parameters. Using a backward elimination, the non-significant parameters can be gradually removed from the model. In R, an approximate calculation can be performed using parameters obtained from the

Vglm procedure. The Vglm procedure can be used with family specified as cumulative and parallel specified as false only for the independent variable that has a trend. The gamma related to the several parameters for the independent variable that was unconstrained can be approximated by assuming a scalar set such as 0, 1 and 2. Appendix A and B include the SAS and R code described in this section.

Example I - Data on perceived discrimination reported in the Minority Aging Research Study

Our first example uses data on perceived discrimination from the Minority Aging Research Study (MARS), an epidemiologic cohort study of aging and Alzheimer's disease in older African Americans. Participants were recruited from various community-based organizations, churches, and senior subsidized housing facilities in and around the Chicago metropolitan area. The Everyday Discrimination Scale (Williams, Yan, Jackson, & Anderson, 1997) is a 9-item questionnaire that measures chronic but minor forms of unfair treatment in everyday life. The scale, along with other psychosocial measures was given at baseline. A total of 603 participants without dementia at baseline are included in this analysis. The sample population is 24% men, and participants had a mean age of 73.6 years at baseline (range from 60.2 to 97.6). All participants self-reported their race as African American using the race classification question from the U.S. Census. For this illustrative example we focus on one question from the Everyday Discrimination Scale: "People act as if they think you are dishonest".

Dishonesty perception is based on a 4-level Likert: never, rarely, sometimes, and often. The first cumulative odds (baseline odds, θ_1) are the probability of experiencing a perception of dishonesty at least rarely over the probability of experiencing it less than rarely. The second cumulative odds (θ_2) are the probability of experiencing a dishonesty perception at least sometimes over the probability of experiencing it less than sometimes. The third cumulative odds (θ_3) are the probability of experiencing a dishonesty perception often over the probability of experiencing it less than often.

For the first step of the analysis, dishonesty perception is examined as the independent variable in the binary logistic regression, fitting three separate models:

$$\log\left(\frac{\Pr(Y \geq \text{rarely} \mid \text{gender})}{\Pr(Y < \text{rarely} \mid \text{gender})}\right) = \alpha_{\text{rarely}} + \beta_{\text{rarely,sex}} \text{sex},$$

$$\log\left(\frac{\Pr(Y \geq \text{sometimes} \mid \text{gender})}{\Pr(Y < \text{sometimes} \mid \text{gender})}\right) = \alpha_{\text{sometimes}} + \beta_{\text{sometimes,sex}} \text{sex},$$

and

$$\log\left(\frac{\Pr(Y \geq \text{often} \mid \text{gender})}{\Pr(Y < \text{often} \mid \text{gender})}\right) = \alpha_{\text{often}} + \beta_{\text{often,sex}} \text{sex}.$$

Sex is a dichotomous indicator for Black men. In the example above, the term $\beta_{rarely,sex}$ is interpreted as the log odds ratio of dishonesty perception at least rarely versus less than rarely for Black men compared to Black women. The term $\beta_{sometimes,sex}$ is interpreted as the log odds ratio of experiencing dishonesty perception at least sometimes versus less than sometimes for Black men compared to Black women. The term $\beta_{often,sex}$ is interpreted as the log odds ratio of experiencing dishonesty perception often versus less than often for Black men compared to Black women.

For the second step of the analysis, we fit a separate proportional odds model for each covariate:

$$\log\left(\frac{\Pr(Y \geq k \mid gender)}{\Pr(Y < k \mid gender)}\right) = \alpha_k + \beta_{\bullet, sex} \text{sex}.$$

For the third step of the analysis, we fit a separate trend odds model for each covariate:

$$\log\left(\frac{\Pr(Y \geq k \mid gender)}{\Pr(Y < k \mid gender)}\right) = \alpha_k + (\beta_{rarely,sex} + \gamma_{sex} t_k) \text{sex}.$$

Models under step 1 to 3 are repeated replacing age at baseline. Sex is a dichotomous indicator of men. The term γ_{sex} is interpreted as the trend in log odds ratios with frequencies of dishonesty perception for Black men compared to Black women. We use graphical displays to investigate the proportional odds assumption and the trend odds assumption. For example, figure 1 shows observed and estimated log odds by cut-points of the outcome. We can see that the log odds cannot be summarized by a flat line and that estimates from the trend odds model are less biased.

For the fourth step of the analysis, we use the effects obtained from previous steps as starting values. The initial trend odds model includes all the independent variables with parameters for baseline and trend:

$$\log\left(\frac{\Pr(Y \geq k \mid sex, age)}{\Pr(Y < k \mid sex, age)}\right) = \alpha_k + (\beta_{rarely,sex} + \gamma_{sex} t_k) \text{sex} + (\beta_{rarely,age} + \gamma_{age} t_k) \text{age}$$

Using a backward elimination, parameters are gradually removed from the model. The final model is reduced to:

$$\log\left(\frac{\Pr(Y \geq k \mid gender, age)}{\Pr(Y < k \mid gender, age)}\right) = \alpha_k + (\beta_{rarely,sex} + \gamma_{sex} t_k) \text{sex} + \beta_{\bullet, age} \text{age}$$

Results from these analysis show that sex and age are associated with dishonesty perception. There was a trend in the observed unadjusted cumulative odds ratios odds of reporting being perceived as dishonest at least rarely, at least sometimes and often that increased from 1.59 to 2.52 to 5.34.

The unadjusted proportional odds model estimated that the cumulative odds of increased frequency of dishonesty perception was 69% higher for Black men compared to Black women.

The unadjusted trend odds model estimated a trend multiplicative effect of about 1.66 ($p = .047$), although a score test for proportional odds assumptions provided weak evidence of deviation from proportionality (chi-square = 4.21, $p = .12$). That represents a series of odds ratios 1.59 (θ_1 , odds of dishonesty perception at least rarely versus less than rarely), 2.63 (θ_2 , odds of dishonesty perception at least sometimes versus less than sometimes), and 4.37 (θ_3 , odds of dishonesty perception often versus less than often). This series of odds ratios was obtained by multiplying the multiplicative trend effect by the previous odds ratio (i.e. $\theta_2 = \theta_1 \cdot 1.66 = 2.63$ and $\theta_3 = \theta_2 \cdot 1.66 = 4.37$).

The proportional-trend odds model included sex and age as covariates. Trend odds was assumed for sex and proportional odds was assumed for age. In the final model in Table 1, men had 61% higher odds of dishonesty perception at least rarely versus less than rarely compared with women. The final multivariate model shows a significant trend effect of about 1.67. That is the cumulative odds increased by 67% with each increase in Likert frequency. That represents a series of adjusted cumulative odds ratios of 1.61 (θ_1 , odds of dishonesty perception at least rarely versus less than rarely), 2.67 (θ_2 , odds of dishonesty perception at least sometimes versus less than sometimes), and 4.46 (θ_3 , odds of dishonesty perception often versus less than often).

Example II - Clustered data on severe violence victimization at school

In certain designs, randomization of data occurs at a cluster level. To illustrate, we use clustered data on school violence from 81 middle schools from a large multicultural urban school district. The schools were surveyed in the spring of 2004 and represent a total of 103 middle school programs (i.e. magnet, special education, and regular education) (Ramirez et al., 2012). The survey uses a questionnaire composed of items on exposure to violence. Respondents were asked if violence was experienced at specific locations such as school, neighborhood, and others. A total of 28,882 students completed self-administered paper-pencil scantron surveys at school, which were administered by school staff at each school site. The sample included 49% males. The racial/ethnic distribution was 68% Latinos, 11% Blacks/African Americans, 10% Whites, 4% Asians and 7% others. A total of 196 children (0.7%) were in special education. Variables on past violence exposure were combined into two distinct subscales: witnessed (Mdn = 3.0, $M = 3.10$, $SD = 2.63$) and victimized (Mdn = 1.0, $M = 1.32$, $SD = 1.65$). Past violence subscales, as well as disability, ethnicity and sex were considered potential risk factors in analyses (independent variables). In this example, we investigate risk factors for self-reported victimization to severe physical violence at school (dependent variable), based on the question “How often have you been beaten up at school?”.

Severe violence victimization was based on a 4-level Likert: never, sometimes, lots of times, and almost every day. The first cumulative odds (baseline odds, θ_1) are the probability of experiencing severe violence at least sometimes over the probability of experience it less

than sometimes. The second cumulative odds (θ_2) are the probability of experiencing severe violence at least lots of times over the probability of experience it less than lots of times. The third cumulative odds (θ_3) are the probability of experiencing severe violence almost every day over the probability of experiencing it less than almost every day.

For the first step of the analysis, disability was examined as the independent variable in the binary logistic regression, fitting three separate models:

$$\log\left(\frac{\Pr(Y \geq \textit{sometimes} \mid \textit{disability})}{\Pr(Y < \textit{sometimes} \mid \textit{disability})}\right) = \alpha_{\textit{sometimes}} + \beta_{\textit{sometimes,disability}} \textit{disability} + \varphi_j$$

$$\log\left(\frac{\Pr(Y \geq \textit{lots of times} \mid \textit{disability})}{\Pr(Y < \textit{lots of times} \mid \textit{disability})}\right) = \alpha_{\textit{lots of times}} + \beta_{\textit{lots of times,disability}} \textit{disability} + \varphi_j$$

and

$$\log\left(\frac{\Pr(Y \geq \textit{every day} \mid \textit{disability})}{\Pr(Y < \textit{every day} \mid \textit{disability})}\right) = \alpha_{\textit{every day}} + \beta_{\textit{every day,disability}} \textit{disability} + \varphi_j,$$

Disability is a dichotomous indicator for children in special education. Hence, in the example above, the term $\beta_{\textit{sometimes,disability}}$ is interpreted as the log odds ratio of experiencing severe violence at least sometimes versus less than sometimes for children in special education compared to others. The term $\beta_{\textit{lots of times,disability}}$ is interpreted as the log odds ratio of experiencing severe violence at least lots of times versus less than lots of times for children in special education compared to others. The term $\beta_{\textit{almost every day,disability}}$ is interpreted as the log odds ratio of experiencing severe violence almost every day versus less than almost every day for children in special education compared to others. The term $\gamma_{\textit{disability}}$ is interpreted as the trend in log odds ratios with frequencies of severe violence victimization for children in special education compared to others. Figure 2 shows that there is a trend in observed log odds ratios and that the trend odds will provide an estimate that is less biased.

For the second step of the analysis, we fit a separate proportional odds model for each covariate:

$$\log\left(\frac{\Pr(Y \geq k \mid \textit{disability})}{\Pr(Y < k \mid \textit{disability})}\right) = \alpha_k + \beta_{\bullet, \textit{disability}} \textit{disability} + \varphi_j,$$

For the third step of the analysis, we fit a separate trend odds model for each covariate:

$$\log\left(\frac{\Pr(Y \geq k \mid \textit{disability})}{\Pr(Y < k \mid \textit{disability})}\right) = \alpha_k + (\beta_{\textit{sometimes,disability}} + \gamma_{\textit{disability}} t_k) \textit{disability} + \varphi_j.$$

Models under step 1 to 3 were repeated replacing disability by sex, ethnicity, past violence victimization and past violence witnessed. Sex is a dichotomous indicator of males. Ethnicity is multinomial and was parameterized as 0/1 for each type of ethnicity, with

Black representing a reference group. Past violence subscales were considered as continuous variables.

For the fourth step of the analysis, we use the effects obtained from previous steps as starting values. The initial trend odds model included all the independent variables with parameters for respective baseline and trend:

$$\begin{aligned} \log\left(\frac{\Pr(Y \geq k \mid \text{disability})}{\Pr(Y < k \mid \text{disability})}\right) &= \alpha_k + (\beta_{\text{sometimes, disability}} + \gamma_{\text{disability } t_k}) \text{disability} \\ &+ (\beta_{\text{sometimes, sex}} + \gamma_{\text{sex } t_k}) \text{sex} \\ &+ (\beta_{\text{sometimes, white}} + \gamma_{\text{white } t_k}) \text{white} \\ &+ (\beta_{\text{sometimes, asian}} + \gamma_{\text{asian } t_k}) \text{asian} \\ &+ (\beta_{\text{sometimes, latino}} + \gamma_{\text{latino } t_k}) \text{latino} \\ &+ (\beta_{\text{sometimes, others}} + \gamma_{\text{others } t_k}) \text{others} \\ &+ (\beta_{\text{sometimes, witnessed}} + \gamma_{\text{witnessed } t_k}) \text{witnessed} \\ &+ (\beta_{\text{sometimes, victimized}} + \gamma_{\text{victimized } t_k}) \text{victimized} + \varphi_j. \end{aligned}$$

Using a backward elimination, parameters were gradually removed from the model. The final model was reduced to:

$$\begin{aligned} \log\left(\frac{\Pr(Y \geq k \mid \text{disability})}{\Pr(Y < k \mid \text{disability})}\right) &= \alpha_k + (\beta_{\text{sometimes, disability}} + \gamma_{\text{disability } t_k}) \text{disability} \\ &+ \beta_{\text{sometimes, sex}} \text{Sex} + \beta_{\text{sometimes, white}} \text{white} \\ &+ \beta_{\text{sometimes, asian}} \text{asian} + \beta_{\text{sometimes, latino}} \text{latino} \\ &+ \beta_{\text{sometimes, others}} \text{others} + \beta_{\text{sometimes, witnessed}} \text{witnessed} \\ &+ \beta_{\text{sometimes, victimized}} \text{victimized} + \varphi_j. \end{aligned}$$

Results from these analysis show that disability is the strongest risk factor for victimization to severe physical violence in these data. There is a significant random effect for schools. In the final model in Table 2, children in special education had 3.18 times higher odds of being the victim of severe violence at least sometimes versus less than sometimes compared to children in regular education. The final multivariate model showed a significant trend effect of about 1.68. That indicates a cumulative odd increase of 68% with each increase in Likert frequency, representing a series of adjusted cumulative odds ratios of 3.18 (θ_1 , odds of experiencing severe violence at least sometimes versus less than sometimes), 5.34 (θ_2 , odds of experiencing severe violence at least lots of times versus less than lots of times), and 8.98 (θ_3 , odds of experiencing severe violence almost every day versus less than almost every day).

For this tutorial, we repeated the models removing the random school effect to compare results. In the final model in Table 3, children in special education had 2.56 times higher odds of experiencing severe violence at least sometimes versus less than sometimes compared to children in regular education. The final multivariate model showed a significant trend effect of about 1.66, indicating a cumulative odd increase of 66% with each increase in Likert frequency. That represented a series of adjusted cumulative odds ratios of 2.56 (θ_1 , odds of experiencing severe violence at least sometimes versus less than sometimes), 4.26 (θ_2 , odds of experiencing severe violence at least lots of times versus less than lots of times), and 7.07 (θ_3 , odds of experiencing severe violence almost every day versus less than almost every day).

Final Remarks

In this paper, we detailed use of the proportional—trend odds model and its particular application to modeling of individual Likert outcomes. This tutorial is based on previous research by McCullagh (McCullagh, 1980), Capuano and Dawson (Capuano & Dawson, 2013), Peterson and Harrel (Peterson & Harrell, 1990), and others (Aitchison & Silvey, 1957; Hedeker et al., 1994; Snell, 1964). We provide instructions on how to fit the model and provide two didactic examples from real datasets. The scalar set 0,1,2 and 3 was used for both examples. As discussed previously, it is recommended to use this type of scalar when no a priori information is available. It is possible, however, to optimize the scalar set as a *priori* information becomes available.

Motivated by the situations commonly found in surveys, we extended the method to clustered data. This is an important addition to the literature as clustered data are common. This was the case of our second example where data were collected at multiple schools and resulted in within school correlation. Here we performed the analysis with and without a random effect for school. We detected a significant trend in disability with both approaches but odds were different. In addition, there was a differential effect of the previous violence subscales. It is possible that these differences are generated by the way the data were collected or, in other words, an artifact of the variability between schools.

This article outlines how to build a proportional-trend model. Such models provide interpretable measures of association. The proportional-trend model can provide additional insights into the extent of the risk factor. Both the proportional odds model and the trend odds model have a theoretical justification in underlying latent variables, making of them more than just a parameterization. The ability to flexibly use them in a single model can streamline analyses.

This paper demonstrates how to fit models for ordinal data in SAS and R; however, we recognize that there are a number of limitations to our work. First of all, the models make some specific assumptions, which can be challenging to prove or disprove. Second, sometimes the software may fail to converge to a specific answer if there are no reasonably accurate initial estimates for the parameters.

Appendix A-: A sample if analyses in SAS

```

*****SAMPLE DATA*****;
data f;
input f1 f2 f3 f4;
datalines;
65 85 97 120
23 42 23 42
23 22 19 18
50 5 30 5
;
run;
%macro convert_freq_into_data;
data _null_; set f; suffix=put(_n_,5); array xxx{*} _numeric_;
do i=1 to dim(XXX); call symput(cats(vname(XXX[i]),suffix),XXX[i]);end;
data prep;
%do j=1 %to 4 %by 1;%do i=1 %to 4 %by 1;
do i=1 to &&F&j.&i. by 1; x1=(&j<3); x2=(&j in (1,3)); likert=&i; retain x1 x2 likert; output; end;
%end;%end;
data sample_data; set prep (drop=i); id=_N_;
%do k=1 %to 4; likert&k.=(likert-&k.);%end;
run;
%mend convert_freq_into_data; %convert_freq_into_data;
*****STEP 1*****;
Data sample_data_desc; set sample_data; likert_desc=abs(likert-4);run;
proc catmod data=sample_data_desc;
response clogits;
model likert_desc= x1 x2;
run;

proc logistic data=sample_data_desc; model likert4= x1 x2;ods output ParameterEstimates=p4 ;run;
proc logistic data=sample_data_desc; model likert3= x1 x2;ods output ParameterEstimates=p3 ;run;
proc logistic data=sample_data_desc; model likert2= x1 x2;ods output ParameterEstimates=p2;run;
data p;
set p2 (in=in2) p3 (in=in3) p4 (in=in4);
if in2 then cutpoint=2;
else if in3 then cutpoint=3;
else if in4 then cutpoint=4;
if variable in ("x1","x2");
run;
symbol1 interpol=join value=dot color=black width=1 height=1.5;
symbol2 interpol=join value=dot color=gray width=1 height=1.5;
proc gplot data=p; plot estimate*cutpoint=variable;run;quit;
*****STEP 2*****;

proc logistic data=sample_data_desc; model likert= x1 ; run;
proc logistic data=sample_data_desc; model likert= x2 ;ods output ParameterEstimates=initialsA;run;
*****STEP 4*****;
proc print data=initialsA; run;
*example of a trend-proportional odds code for 4 level likert;
proc nlmixed data=sample_data ;
parms Int1 = -2.2725 Int2 = -1.4386 Int3 = -0.5527 beta_1 0.7780 trend_t = 0 ;
p1= 1/(1 + exp(-Int1 + trend_t*2*x2 + beta_1 * x2 ));
p2= 1/(1 + exp(-Int2 + trend_t*1*x2 + beta_1 * x2 ));
p3= 1/(1 + exp(-Int3 + trend_t*0*x2 + beta_1 * x2 ));
if likert =4 then p = p1;
else if likert =3 then p = p2-p1;
else if likert =2 then p = p3-p2;
else p = 1-p3;
p = (p>0 and p<=1)*p + (p<=0)*1e-8 + (p>1);
loglik = log(p);
model LIKERT ~ general(loglik);
*add this line for clustered data: random u ~ normal (0,varu) subject=ID;
run;
*****STEP 4*****;
proc logistic data=sample_data_desc; model likert= x1 x2;ods output ParameterEstimates=initialsB;run;
proc print data=initialsB; run;
*example of a trend-proportional odds code for 4 level likert;
proc nlmixed data=sample_data ;
parms Int1 = -2.5793 Int2 = -1.7325 Int3 = -0.8301 beta_p = 0.5798 beta_t = 0.7806 trend_t = 0 ;
p1= 1/(1 + exp(-Int1 + beta_p*x1 + trend_t*2*x2 + beta_t * x2 ));
p2= 1/(1 + exp(-Int2 + beta_p*x1 + trend_t*1*x2 + beta_t * x2 ));
p3= 1/(1 + exp(-Int3 + beta_p*x1 + trend_t*0*x2 + beta_t * x2 ));
if likert =4 then p = p1;
else if likert =3 then p = p2-p1;
else if likert =2 then p = p3-p2;
else p = 1-p3;
p = (p>0 and p<=1)*p + (p<=0)*1e-8 + (p>1);
loglik = log(p);
model LIKERT ~ general(loglik);
*add this line for clustered data: random u ~ normal (0,varu) subject=ID;
run;

```

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Appendix B.: A sample analyses in R

```

####DATA####
require(VGAM)
likert<-c(1, 2, 3, 4)
frequencies <-c(65, 23, 23, 50,85, 42, 22, 5,97, 23, 19, 30,120, 42, 18, 5)
values_likert<-c(1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4)
values_x1<-c(1,1,1,1,1,1,1,1,0,0,0,0,0,0,0)
values_x2<-c(1,1,1,1,0,0,0,1,1,1,1,0,0,0)
f <- data.frame(likert=rep(values_likert,frequencies),
x1=rep(values_x1,frequencies),x2=rep(values_x2,frequencies))
####STEP 1####
unconstrained<-vglm(fSlikert~fSx1+fSx2, family=cumulative(reverse = TRUE,parallel=FALSE))
summary(unconstrained)
#plot of the effects of x2
plot(Coeff(unconstrained)[7:9],type="l",xlab="cutpoint",ylab="estimate")
#plot of the effects of x1
par(new=T)
lines(Coeff(unconstrained)[4:6],type="l")
####STEP 2####
ux1<-vglm(fSlikert~fSx1, family=cumulative(parallel=FALSE))
ux2<-vglm(fSlikert~fSx2, family=cumulative(parallel=FALSE))
pomx1 <-vglm(fSlikert~fSx1, family=cumulative(parallel=TRUE))
pomx2 <-vglm(fSlikert~fSx2, family=cumulative(parallel=TRUE))
#likelihood ratio test
testing1<-1 - pchisq(deviance(pomx1)- deviance(ux1) , df=2)
testing2<-1 - pchisq(deviance(pomx2)- deviance(ux2) , df=2)
####STEP 3####
summary(ux2)
Coeff_x2<-Coeff(ux2)[5:7]
#idea of trend parameter for x2
gamma_biv_x2<-((Coeff_x2[3]-Coeff_x2[1])/2)+(Coeff_x2[2]-Coeff_x2[1])/2
####STEP 4####
pom<-vglm(fSlikert~fSx1+fSx2, family=cumulative(parallel=TRUE))
unconstrained2<-vglm(fSlikert~fSx1+fSx2, family=cumulative(parallel=FALSE~fSx2))
#likelihood ratio test
testing_mult_x2<-1 - pchisq(deviance(unconstrained2) - deviance(unconstrained), df=2)
#likelihood ratio test
summary(unconstrained2)
Coeff_x2<-Coeff(unconstrained2)[5:7]
#idea of trend parameter for x2 for the model with x1 and x2
gamma_x2<-((Coeff_x2[3]-Coeff_x2[1])/2)+(Coeff_x2[2]-Coeff_x2[1])/2

```

References

- Agresti A (2002). *Categorical data analysis*. New York: Wiley-Interscience.
- Aitchison J, & Silvey SD (1957). The generalization of probit analysis to the case of multiple responses. *Biometrika*, 44(1–2), 131–140.
- Capuano AW, & Dawson JD (2013). The trend odds model for ordinal data. *Statistics in Medicine*, 32(13), 2250–2261. doi:10.1002/sim.5689 [doi] [PubMed: 23225520]
- Capuano AW, Dawson JD, & Gray GC (2007). Maximizing power in seroepidemiological studies through the use of the proportional odds model. *Influenza and Other Respiratory Viruses*, 1(3), 87–93. doi:10.1111/j.1750-2659.2007.00014.x [doi] [PubMed: 18176626]
- Clason DL, & Dormody TJ (1994). Analyzing data measured by individual likert-type items. *Journal of Agricultural Education*, 35(4), 31–35.
- Hedeker D, Gibbons RD, & Flay BR (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. [PubMed: 7962879]
- Huang AJ, Subak LL, Wing R, West DS, Hernandez AL, Macer J, & Grady D(2010). An intensive behavioral weight loss intervention and hot flushes in women. *Archives of Internal Medicine*, 170(13), 1161–1167. doi:10.1001/archinternmed.2010.162 [doi] [PubMed: 20625026]
- Likert R (1932). A technique for the measurement of attitudes. *Archives of psychology*, 140, 4460.
- McCullagh P (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society.Series B (Methodological)*, 42(2), 109–142.
- Peterson B, Harrell FE, & Brant R (1992). Proportional odds model. *Biometrics Biometrics*, 48(1), 325–326.
- Peterson B, & Harrell FE (1990). Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society.Series C (Applied Statistics)*, 39(2), 205–217.

- Ramirez M, Wu Y, Kataoka S, Wong M, Yang J, Peek-Asa C, & Stein B (2012). Youth violence across multiple dimensions: A study of violence, absenteeism, and suspensions among middle school children. *The Journal of Pediatrics*, 161(3), 542–546. e2. [PubMed: 22521110]
- Scott SC, Goldberg MS, & Mayo NE (1997). Statistical assessment of ordinal outcomes in comparative studies. *Journal of Clinical Epidemiology*, 50(1), 45–55. doi:S0895435696003125 [pii] [PubMed: 9048689]
- Singh I, Bard I, & Jackson J (2014). Robust resilience and substantial interest: A survey of pharmacological cognitive enhancement among university students in the UK and Ireland. *PloS One*, 9(10), e105969. doi:10.1371/journal.pone.0105969 [doi] [PubMed: 25356917]
- Snell E (1964). A scaling procedure for ordered categorical data. *Biometrics*, , 592–607.
- Stokes ME, Davis CS, & Koch GG (2012). *Categorical data analysis using SAS* SAS institute.
- Williams DR, Yan Y, Jackson JS, & Anderson NB (1997). Racial differences in physical and mental health: Socio-economic status, stress and discrimination. *Journal of Health Psychology*, 2(3), 335–351. doi:10.1177/135910539700200305 [doi] [PubMed: 22013026]
- Zerwas S, Von Holle A, Torgersen L, Reichborn-Kjennerud T, Stoltenberg C, & Bulik CM (2012). Maternal eating disorders and infant temperament: Findings from the Norwegian mother and child cohort study. *The International Journal of Eating Disorders*, 45(4), 546–555. doi:10.1002/eat.20983 [doi] [PubMed: 22287333]

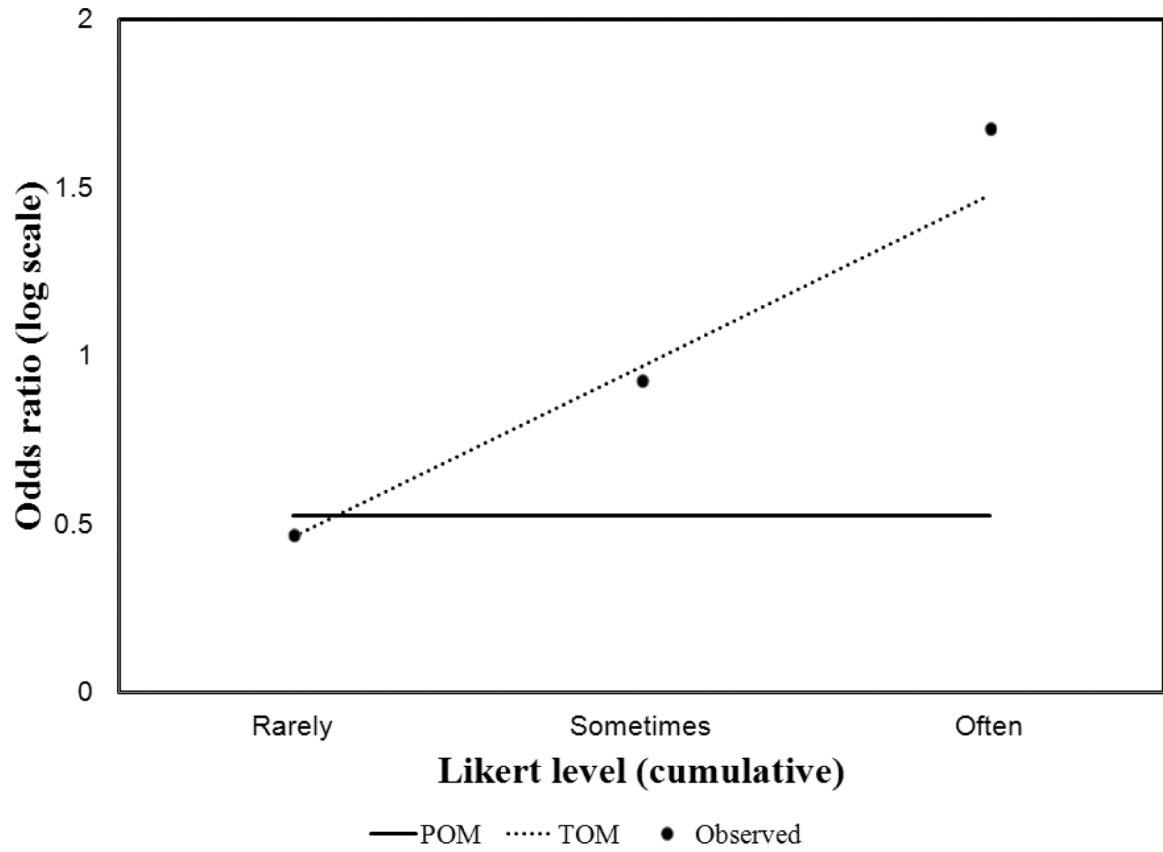


Figure 1. Unadjusted cumulative log odds ratio of dishonesty perception for Black men compared to Black women.

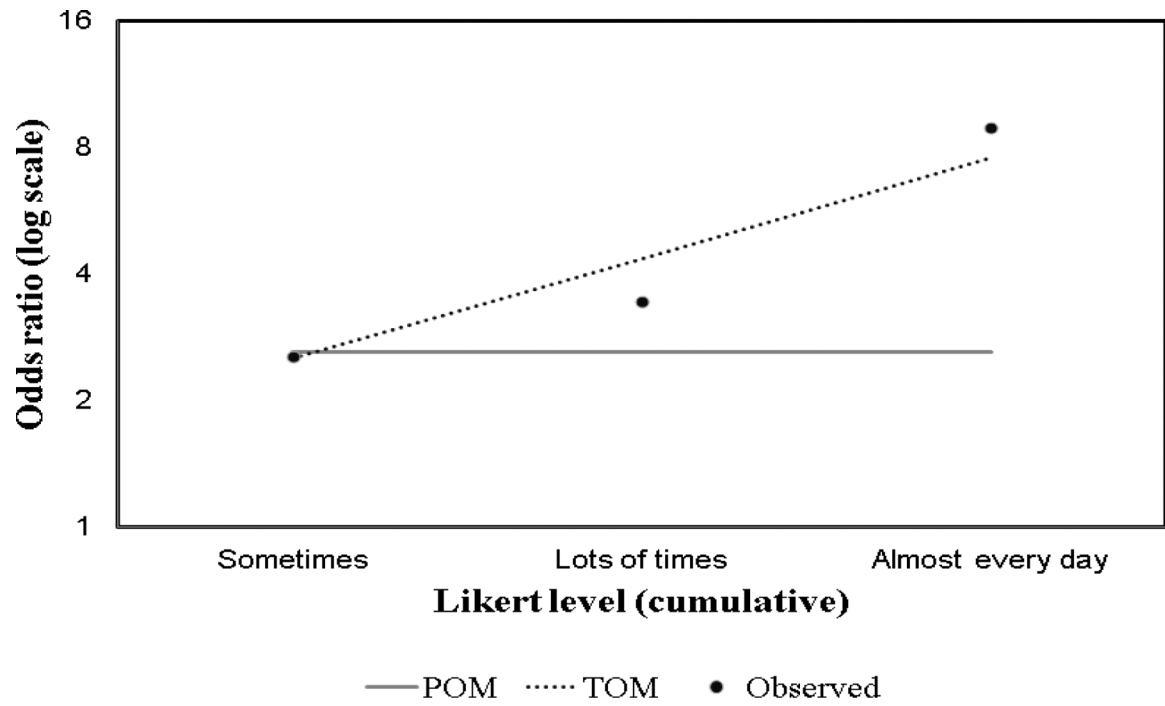


Figure 2. Unadjusted cumulative log odds ratio of severe violence for children with disability compared to others.

Table 1.

Example I, multivariable analyses of frequencies of perceived dishonesty with Proportional-Trend Odds model

Variable	Multivariable Proportional-Trend Odds model					
	β	Odds Ratio (e^β)	p value for β	γ	Trend Multiplicative Effect (e^γ)	p value for γ
Sex						
Male	0.47	1.61	.029	0.51	1.66	.047
Female	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Age at baseline	-0.06	0.94	.001	---	---	---

Notes. $LRT=0.2$, $p = .90$, compared to unconstrained model for sex.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Example II, multivariable analyses of level of severe violence victimization (victim beat school) with Proportional and Trend Odds model, accounting for school random effect

Variable	Multivariable Proportional-Trend Odds model					
	β	Odds Ratio (e^β)	p value for β	γ	Trend Multiplicative Effect (e^γ)	p value for γ
Sex						
Male	1.03	2.81	<.0001	---	---	---
Female	Ref.	Ref.	Ref.	---	---	---
Student Ethnicity						
All others	0.46	1.58	<.01	---	---	---
Asian	0.45	1.57	<.01	---	---	---
Black/African-American	Ref.	Ref.	Ref.	---	---	---
Latino	0.46	1.58	<.0001	---	---	---
White	0.35	1.42	<.01	---	---	---
Disability						
Not in special education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
In special education	1.16	3.18	<.0001	0.52	1.68	.056
Past Violence exposure subscales						
Victimized	0.65	1.91	<.0001	---	---	---
Witnessed	-0.02	0.98	.05	---	---	---

Notes. School random effect = 2.3947 ($SE=0.544$, $p<.0001$). $LRT=2$, $p= .37$, compared to unconstrained model disability.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Example II, multivariable analyses of level of severe violence victimization (victim beat school) with Proportional and Trend Odds model, without a random effect for school.

Variable	Multivariable Proportional-Trend Odds model					
	β	Odds Ratio (e^β)	p value for β	γ	Trend Multiplicative Effect (e^γ)	p value for γ
Sex						
Male	0.85	2.34	<.0001	---	---	---
Female	Ref.	Ref.	Ref.	---	---	---
Student Ethnicity						
All others	0.39	1.48	<.01	---	---	---
Asian	0.45	1.57	<.01	---	---	---
Black/African-American	Ref.	Ref.	Ref.	---	---	---
Latino	0.40	1.49	<.0001	---	---	---
White	0.32	1.38	<.01	---	---	---
Disability						
Not in special education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
In special education	0.94	2.56	<.0001	0.51	1.66	.03
Past Violence exposure subscales						
Victimized	0.48	1.61	<.0001	0.06	1.06	<.0001
Witnessed	---	---	---	---	---	---

Notes. $LRT=16$, $p<0.01$, compared to unconstrained model for disability and past violence victimized.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript