Models for proportional and non-proportional odds

Contents

1 Introduction

The term "ordinal" is applied to variables where the response measure of interest is measured in a series of ordered categories. Examples of such variables include Likert scales and psychiatric ratings of severity. Nominal and ordinal outcome models can be seen as generalizations of the binary outcome model. The ordinal model becomes important when the outcome variable is not dichotomous, or not truly continuous. If an ordinal outcome is analyzed within a continuous model, such a model can yield predicted values outside the range of the ordinal variable. As with binary data, some transformation or link function becomes necessary to prevent this from happening. The continuous model can also yield correlated residuals and regressors when applied to ordinal outcomes because the continuous model does not take the ceiling and floor effects of the ordinal outcome into account. This can then result in biased estimates of regression coefficients, and is most critical when the ordinal variable in question is highly skewed. Armstrong & Sloan (1989) also report efficiency losses between 89% and 99% when comparing an ordinal to a continuous outcome, depending on the number of categories and distribution within the ordinal categories.

Extensive work on the development of methods for the analysis of ordinal response data has been undertaken by numerous researchers, including Hedeker & Gibbons (1994). These developments have focused on the extension of methods for dichotomous variables to ordinal response data, and have been mainly in terms of logistic and probit regression models. The proportional odds model proposed by McCullagh (1980) is a common choice for analysis of ordinal data. This model, which is described in detail in Section 2.2, is based on the logistic regression formulation.

In this chapter we will now build on the dichotomous model discussed earlier and introduce the ordinal model, illustrating the use of this model using the TVSFP (Flay, *et. al.*, 1988) data previously used in this manual.

2 Two-level ordinal analysis of TVSFP data

2.1 The data

The study was designed to test independent and combined effects of a school-based social-resistance curriculum and a television-based program in terms of tobacco use and cessation.

The structure of this study indicates a three-level hierarchical structure. However, for illustration purposes in this chapter we will consider a two-level structure in which students are nested within schools. Data for the first 10 participants on most of the variables used in this section are shown below in the form of a LISREL spreadsheet file, named **tvsfpors.lsf**, located in the **Multilevel Generalized Linear Model examples** subfolder.

tvsfpors.lsf

	School	Class	THKSord	THKSbin	PreTHKS	CC	TV	CC*TV
1	403.0	403101.0	3.0	1.0	2.0	1.0	0.0	0.0
2	403.0	403101.0	4.0	1.0	4.0	1.0	0.0	0.0
3	403.0	403101.0	3.0	1.0	4.0	1.0	0.0	0.0
4	403.0	403101.0	4.0	1.0	3.0	1.0	0.0	0.0
5	403.0	403101.0	4.0	1.0	3.0	1.0	0.0	0.0
6	403.0	403101.0	3.0	1.0	4.0	1.0	0.0	0.0
7	403.0	403101.0	2.0	0.0	2.0	1.0	0.0	0.0
8	403.0	403101.0	4.0	1.0	4.0	1.0	0.0	0.0
9	403.0	403101.0	4.0	1.0	5.0	1.0	0.0	0.0
10	403.0	403101.0	4.0	1.0	3.0	1.0	0.0	0.0

The variables of interest are:

- School indicates the school a student is from (28 schools in total).
- Class identifies the classroom (135 classrooms in total).
- THKSord represents the post-intervention tobacco and health knowledge scaled score, with 4 categories ranging between 1 and 4. The frequency distribution of the post-intervention THKS scores indicated that approximately half the students had scores of 2 or less, and half of 3 or greater. In terms of quartiles, four ordinal classifications were suggested corresponding to 0 – 1, 2, 3, and 4 – 7 correct responses.
- PreTHKS indicates a student's score prior to intervention, *i.e.* the number correct of 7 items.
- CC is a binary variable indicating whether a social-resistance classroom curriculum was introduced, where 0 indicates "no" and 1 "yes."
- TV is an indicator variable for the use of media (television) intervention, with a "1" indicating the use of media intervention, and "0" the absence thereof.
- CC*TV was constructed by multiplying the variables TV and CC, and represents the CC by TV interaction.

In this example we will explore a random intercept model using the ordinal variable THKSord as outcome. The original post-intervention score was assumed to be a continuous variable. In contrast, here categories are created and the implied data collapse may lead to a loss of information and thus results may differ from those obtained previously.

2.1.1 Exploring the data

The focus in this chapter is on the influence of the intervention on the tobacco health knowledge scores of the students, as represented by the ordinal outcome variable THKSord. A cross-tabulation of the variables CC, TV, and THKSord are given in Table 1.1 below.

In general, students not exposed to the social-resistance classroom curriculum (CC = 0) seem to have less health knowledge than those students exposed to the social-resistance classroom curriculum (CC = 1), regardless of their exposure to media intervention. The opposite is true for students from groups assigned the social-resistance classroom curriculum (CC = 1).

Table 1.1: Crosstabulation of CC, TV and THKSord

TV			CC 0	CC 1	Total
0	THKSord	1	117	62	179
		2	129	78	207
		3	89	106	195
		4	86	134	220
	Total		421	380	801
1	THKSord	1	110	66	176
		2	105	86	191
		3	91	114	205
		4	110	117	227
	Total		416	383	799

The trend is also apparent when the post-intervention scores are expressed as proportions (see Table 1.2 below).

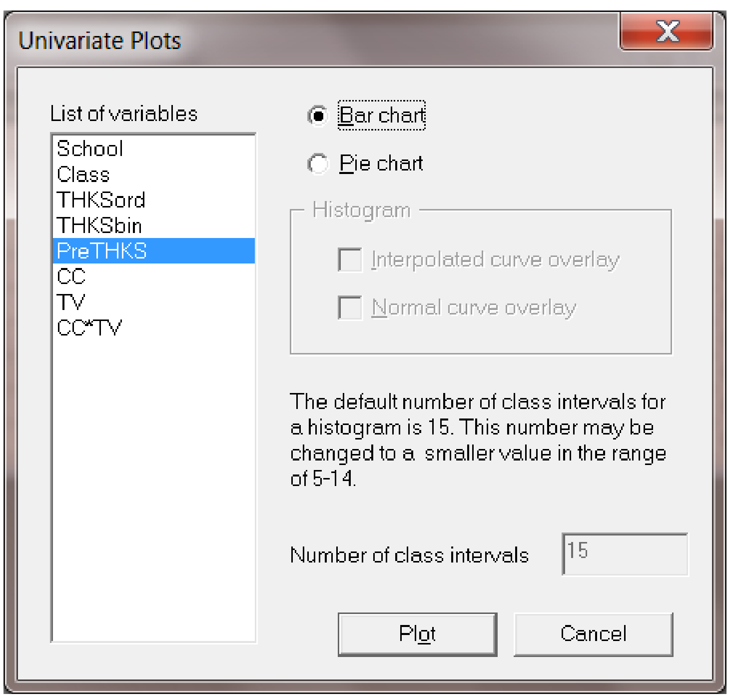
First, notice that the outcome variable THKSord has a skewed distribution. By combining the proportions per category over interventions, we find that 0.2219 of the 1600 students had a value of 1 for THKSord, 0.2488 had a value of 2, 0.25 had a value of 3, and 0.2794 a value of 4 for THKSord. The monotonic increase in the proportion observed in each category of THKSord indicates that it would be inappropriate to try to fit a continuous model to the data.

The pre-intervention scores of the students may be used as a covariate in the analysis. To get some idea of the relationship between the scale score PreTHKS and the post-intervention score THKSord, an exploratory graph may be useful.

Table 1.2: Observed proportion of high post–intervention scores

TV			CC 0	CC 1	Total
0	THKSord	1	0.0731	0.0388	0.1119
		2	0.0806	0.0488	0.1294
		3	0.0556	0.0663	0.1219
		4	0.0538	0.0838	0.1375
	Total		0.2631	0.2375	0.5006
1	THKSord	1	0.0688	0.0413	0.1100
		2	0.0656	0.0538	0.1194
		3	0.0569	0.0713	0.1281
		4	0.0688	0.0731	0.1419
	Total		0.2600	0.2394	0.4994

To take a closer look at the distribution of PreTHKS, select the **Graphs**, **Univariate…** option from the **File** menu after opening the Lisrel spreadsheet **tvsfpors.lsf**.



The **Univariate plots** dialog box is activated. Select the variable PreTHKS, and request a **Bar Chart**. Click **Plot**.

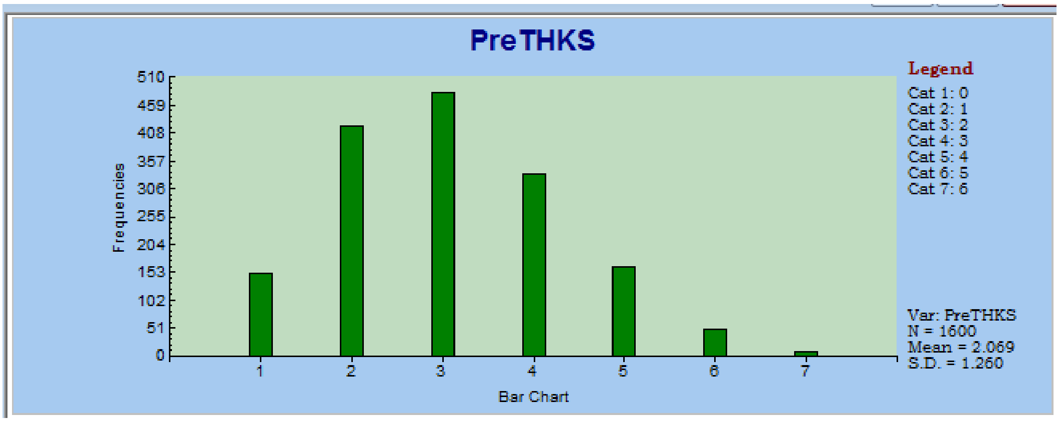
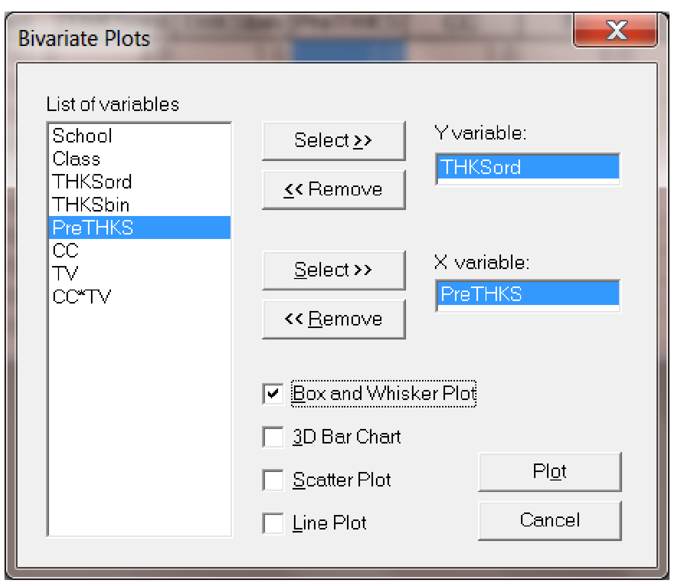


Figure 1.1: Distribution of the PreTHKS scores

Figure 1.1 is obtained. In contrast to the outcome variable THKSord, the distribution of the PreTHKS score has a lower mean, with very few students exhibiting extensive knowledge on the subject matter (PreTHKS = 5 or PreTHKS = 6).

We now take a closer look at the distribution of the outcome variable at each distinct pre-intervention score value by utilizing the **Graphs**, **Bivariate** option on the **File** menu. By default, a bar chart will be produced. Select the variable THKSord in the **Y** column and the variable PreTHKS in the **X** column, and request a **Box and Whiske**r plot before clicking the **Plot** button.



The figure below shows a reasonably steady increase in the mean THKSord with increasing PreTHKS scores. This seems to be expected: students with more initial knowledge ending up having higher post-intervention scores as well. Note that only 55 of the 1600 observations showed a score of 5 or higher on the pre-intervention score, and that no student obtained a post-intervention score of 7 out of 7.

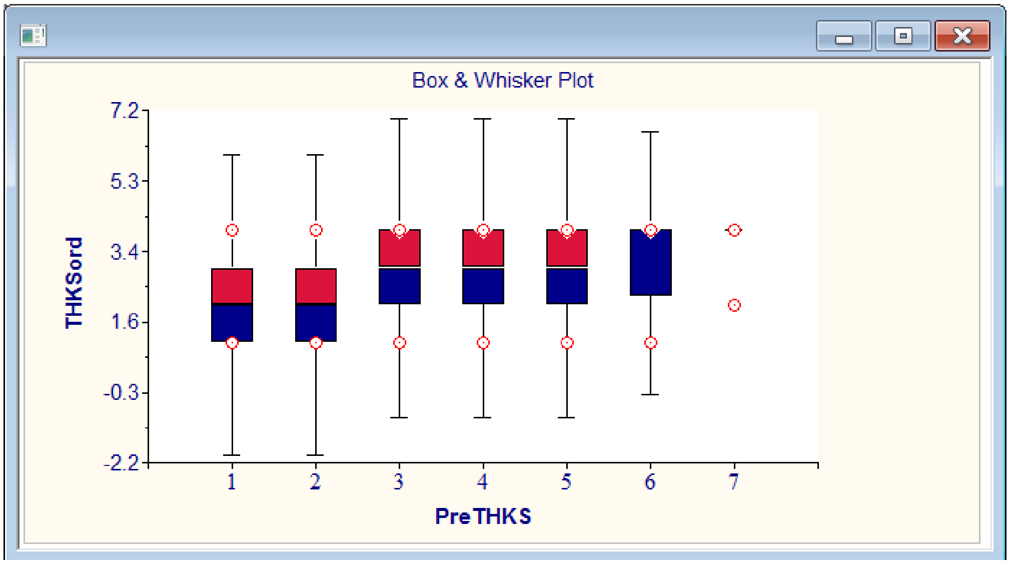


Figure 1.2: Box-and-whisker plot of THKSord for values of PreTHKS

Finally, we also take a look at the mean pre-intervention scores of the students for each of the four subgroups. These are summarized in Table 1.3 below, and show that the mean pre-intervention scores do not differ much.

Table 1.3: Mean pre-intervention scores

Study condition	Mean
CC = 0, TV = 0	2.152
CC = 0, TV = 1	2.087
CC = 1, TV = 0	2.050
CC = 1, TV = 1	1.979

2.2 A multilevel ordinal model with logistic link function

2.2.1 The proportional odds model

The model we use for the analysis of ordinal data is based on McCullagh's (1980) proportional odds model, which characterizes the ordinal responses in C categories in terms of $C-1$ cumulative category comparisons, specifically $C-1$ cumulative logits. The McCullagh model can be written as

$$\log\left[\frac{P(y \le c)}{1-P(y \le c)}\right] = \gamma_c - \mathbf{x}'\boldsymbol{\beta}$$

where

- $c = 1,...,C-1$ for the C categories of the ordinal outcome
- $\mathbf{x}$ is the vector of explanatory variables, plus the intercept
- γ_c represent the threshold parameter(s); and reflect the cumulative odds when $\mathbf{x} = 0$.

The positive association between a predictor variable x and the ordinal outcome variable y is reflected by β. It is assumed that the effect of x is the same for each of the cumulative odds ratios.

To illustrate, consider a model with a single predictor x. The odds that the response is less than or equal to c (for any fixed c) is divided by e^{β} for every unit change in x, as shown below:

$$\left[\frac{P(y \le c)}{1-P(y \le c)}\right] = \exp\left(\gamma_c - x\beta\right) = \frac{e^{\gamma_c}}{\left(e^{\beta}\right)^{x}}.$$

On the other hand, the odds that the response is greater than or equal to c (again for a fixed c) is multiplied by e^{β} for every unit change in x:

$$\left[\frac{1-P(y \le c)}{P(y \le c)}\right] = e^{-\gamma_c} \times \left(e^{\beta}\right)^{x}.$$

To motivate the ordinal regression model, it is often assumed that there is an unobservable latent variable ($y*$) which is related to the actual response through the "threshold concept." An example of this is when respondents are asked to rate their agreement with a given statement using the categories "Disagree," "Neutral," "Agree." These three options leave no room for any other response, though one can argue that these are three possibilities along a continuous scale of agreement that would also make provision for "Strongly Agree" and

"Disagree somewhat." The ordinal responses capture in y and the latent continuous variable $y*$ are linked through some fixed, but unknown, thresholds.

For the dichotomous model, one threshold value is assumed, while for the ordinal model, a series of threshold values $\gamma_0, \gamma_1, \gamma_2, ..., \gamma_C$, where C equals the number of ordered categories, $\gamma_0 = -\infty$, and $\gamma_C = \infty$, is assumed. Here, a response occurs in category c $(Y = c)$ if the latent response process y exceeds the threshold value γ_{C-1}, but does not exceed the threshold value γ_c. The cumulative probabilities are given in terms of the cumulative logits with $C-1$ strictly increasing model thresholds $\gamma_1, \gamma_2, ..., \gamma_{C-1}$. In the current case, we will thus have $C-1$ = 3 cumulative probabilities, given in terms of 3 thresholds γ_1, γ_2 and γ_3. The thresholds represent the marginal response probabilities in the C categories. We will illustrate the use of the logistic link function in this example.

To set the location of the latent variable, it is common to set a threshold to zero. Usually, the first of the threshold parameters (γ_1) is set to zero. Alternatively, the model intercept (β_0) is set to zero and $C-1$ thresholds are estimated.

2.2.2 The mixed-effect proportional model

The mixed-effect proportional odds model can be formulated as:

$$\log\left[\frac{P(y \le c)}{1-P(y \le c)}\right] = \gamma_c - \left[\mathbf{x}_{ij}^{'}\boldsymbol{\beta} + \mathbf{z}_{ij}^{'}\mathbf{v}_i\right].$$

In this model, as in the proportional odds model, the origin of the latent variable y is set by setting the first threshold, γ_1, equal to zero. It is assumed that $\mathbf{v}_i \sim NID\left(\mathbf{0}, \boldsymbol{\Sigma}_v\right)$. The unit of measurement is $\sigma = \pi / \sqrt{3}$.

For this model, the category probabilities are defined as

$$P(y_{ij} \le c) = \psi\left(\gamma_c - \left(\mathbf{x}_{ij}^{'}\boldsymbol{\beta} + \mathbf{z}_{ij}^{'}\mathbf{v}_i\right)\right)$$

and

$$P\left(y_{ij} = c\right) = \psi\left(\gamma_c - \left(\mathbf{x}_{ij}^{'}\boldsymbol{\beta} + \mathbf{z}_{ij}^{'}\mathbf{v}_i\right)\right) - \psi\left(\gamma_{c-1} - \left(\mathbf{x}_{ij}^{'}\boldsymbol{\beta} + \mathbf{z}_{ij}^{'}\mathbf{v}_i\right)\right)$$

where the cumulative standard logistic distribution function is

$$\psi\left(\gamma_c-\left(\mathbf{x}_{ij}^{'}\boldsymbol{\beta}+\mathbf{z}_{ij}^{'}\mathbf{v}_i\right)\right)=\frac{1}{1+\exp\left[-\gamma_c-\left(\mathbf{x}_{ij}^{'}\boldsymbol{\beta}+\mathbf{z}_{ij}^{'}\mathbf{v}_i\right)\right]}.$$

Various link functions may be used with this model. If we define $G^{-1}\left[P\left(y_{ij}\leq c\right)\right]$ as

$$G^{-1}\left[P\left(y_{ij}\leq c\right)\right]=\gamma_c-\left(\mathbf{x}_{ij}^{'}\boldsymbol{\beta}+\mathbf{z}_{ij}^{'}\mathbf{v}_i\right),$$

or, equivalently,

$$P\left(y_{ij}\leq c\right)=G\left[\gamma_c-\left(\mathbf{x}_{ij}^{'}\boldsymbol{\beta}+\mathbf{z}_{ij}^{'}\mathbf{v}_i\right)\right],$$

three types of models can easily be fitted:

- Using $G^{-1}\left(P\right)=\log\left[P/\left(1-P\right)\right]$ will give a cumulative logit model, *i.e.* a proportional odds model,
- using $G^{-1}\left(P\right)=\boldsymbol{\Phi}^{-1}\left[P/\left(1-P\right)\right]$ will produce a cumulative probit model, and
- using $G^{-1}\left(P\right)=\log\left[-\log\left(1-P\right)\right]$, the so-called complementary log-log link, will give a proportional hazards model.

2.2.3 The mixed-effect non-proportional logistic regression model

A limitation of the model specified in the previous section is that it is assumed that the effect of covariates is the same across the cumulative logits. To overcome this limitation, an extension of the mixed-effects ordinal logistic regression model to allow for nonproportional odds for a set of regressors was developed by Hedeker & Mermelstein (1998). This generalization of the proportional odds model can be formulated as

$$\log\left[\frac{P(y\leq c)}{1-P(y\leq c)}\right]=\gamma_c-\left[\mathbf{u}_{\mathbf{ij}}^{'}\boldsymbol{\theta}_{\mathbf{c}}+\mathbf{x}_{ij}^{'}\boldsymbol{\beta}+\mathbf{z}_{ij}^{'}\mathbf{v}_i\right].$$

In this model, as in the proportional odds model, the origin of the latent variable y is set by setting the first threshold, γ_1, equal to zero. It is assumed that $\mathbf{v}_i\sim NID\left(\mathbf{0},\boldsymbol{\Sigma}_v\right)$. The unit of measurement is $\sigma=\pi/\sqrt{3}$.

$\mathbf{u}_{\mathbf{ij}}^{'}$ = hx1 vector for the set of h covariates for which the proportional odds is not assumed.

For this model, the category probabilities are defined as

$$P(y_{ij}\leq c)=\psi\left(\gamma_c-\left(\mathbf{u}_{\mathbf{ij}}^{'}\boldsymbol{\theta}_{\mathbf{c}}+\mathbf{x}_{ij}^{'}\boldsymbol{\beta}+\mathbf{z}_{ij}^{'}\mathbf{v}_i\right)\right)$$

- The effects of the **u** variables do vary across the c-1 cumulative logits

- The non-proportional odds model is a more flexible model for ordinal response relations

2.2.4 A general multilevel ordinal model

The multilevel ordinal model is defined in terms of the cumulative probability $P\left(y_{ij} \le c\right)$ where c denotes the category of interest. The level-1 model is written in terms of the cumulative logits, as shown below.

Level-1 model:

$$\log\left[\frac{P\left(y_{ij} \le c\right)}{1-P\left(y_{ij} \le c\right)}\right] = \gamma_c - \left[\mathbf{x}_{ij}^{'}\mathbf{b}_i\right].$$

where $\mathbf{x}_{ij}$ represent the values of the covariates corresponding to level-1 unit j nested within level-2 unit i.

Level-2 model:

If all the elements of the coefficient vector $\mathbf{b}_i$ are allowed to vary randomly across level-2 units, then

$$\mathbf{b}_i = \boldsymbol{\beta} + \mathbf{v}_i,$$

which models the level-2 effects as a function of an overall mean $\boldsymbol{\beta}$ and a unique random component $\mathbf{v}_i \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_v)$. The latter is also referred to as the level-2 residuals and indicates the extent to which a given level-2 unit differs from the average, as estimated by the first part of the level-2 model.

Note that the level-2 model does not depend on the response variable. As the regression coefficients $\beta_0, \beta_1, \beta_2$ and β_3 are without subscript, it is assumed that they do not vary across the categories and hence that the relationship between the predictor variables and the cumulative logits is not dependent on c. McCullagh (1980) referred to this as the assumption of identical odds ratios across the $C-1$ categories.

In practice, a subset of the coefficients $\mathbf{b}_i$ are assumed to have fixed, but unknown, values. For example, a random intercept-and-slope model with 2 predictors of which the first has a random slope would have a level-2 model of the form

$$b_{0i} = \beta_0 + v_{0i}$$
$$b_{1i} = \beta_1 + v_{1i}$$
$$b_{2i} = \beta_2$$

In this model, only the first two coefficients are assumed to vary randomly across the level-2 units.

Another characteristic of the current model is that a positive coefficient for a regressor indicates that the odds that the response is greater than or equal to c increases with an increase in regressor values. However, another formulation as shown below, in which the regression parameters $\boldsymbol{\beta}$ are identical but of opposite sign, is commonly used in survival analysis models (see Chapter **Error! Reference source not found.**):

$$\log\left[\frac{P_{ijc}}{1-P_{ijc}}\right] = \gamma_c + \left[\mathbf{x}_i^{'}\mathbf{b}_i\right] \qquad (c = 1,...,C-1).$$

2.2.5 An ordinal model with 2 covariates and an interaction term

As in the case of the binary variable THKSbin, we intend to explore the relationship between the type of intervention, the pre-intervention scores of students and the ordinal outcome variable THKSord. We do so using a 2-level model, with students nested within schools.

Level-1 model:

At the first level, the pre-intervention score is used as predictor.

$$\log\left[\frac{P\left(\text{THKSord}_{ij} \le c\right)}{1-P\left(\text{THKSord}_{ij} \le c\right)}\right] = \gamma_c - \left[b_{0i} + b_{1i}\text{PreTHKS}_{ij}\right] \qquad (j = 1,...,n_i \text{ subjects})$$

Level-2 model:

At the school level, the types of intervention (represented by the dummy variables CC and TV) are used to explain differences in the intercepts of the groups. In addition, the interaction between CC and TV is included in the model.

$$b_{0i} = \beta_0 + \beta_2\text{CC}_i + \beta_3\text{TV}_i + \beta_4\left(\text{CC*TV}\right)_i + v_{0i} \quad (i = 1,...,N \text{ groups})$$

$$b_{1i} = \beta_1$$

It is assumed that $v_{0i} \sim NID(0, \sigma_v^2)$.

The model can also be formulated in a single expression as:

$$\log\left[\frac{P\left(\text{THKSord}_{ij} \le c\right)}{1-P\left(\text{THKSord}_{ij} \le c\right)}\right]$$

$$=\gamma_c-[\beta_0+\beta_1\text{PreTHKS}_{ij}+\beta_2\text{CC}_i+\beta_3\text{TV}_i+\beta_4\left(\text{CC*TV}\right)_i+v_{0i}]$$

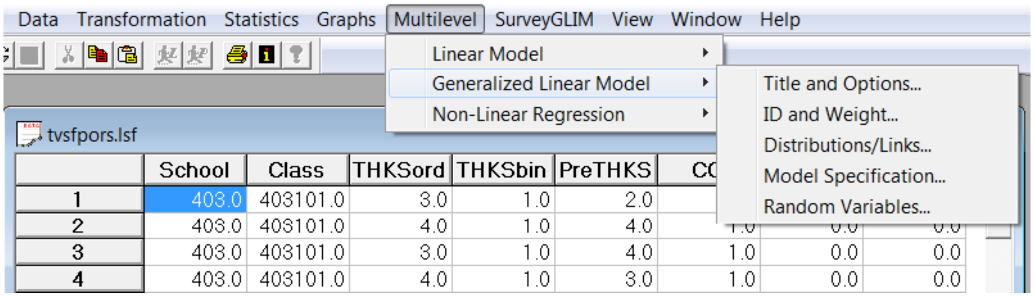
Recall that the outcome variable has 4 categories. There are thus 3 thresholds. In this model

- $0-\beta_0$ (remember that $\gamma_1=0$ for identification purposes) is the first logit (category 1 vs. categories 2 to 4) for groups with no intervention (CC = TV = 0). This logit is adjusted for the effect of PreTHKS.
- $\gamma_2-\beta_0$ is the second logit, representing categories 1 and 2 vs. categories 3 and 4, for groups with no intervention (CC = TV = 0). This logit is also adjusted for the effect of PreTHKS.
- $\gamma_3-\beta_0$ is the third logit, representing categories 1 to 3 vs. category 4, for the same groups and again adjusted for the effect of PreTHKS.
- The coefficient β_1 represents the effect of PreTHKS on THKSord.
- The coefficient β_2 denotes the PreTHKS adjusted logit differences between CC = 1 and CC = 0 (for TV = 0).
- The coefficient β_3 denotes the PreTHKS adjusted logit differences between TV = yes and TV =no (for CC = 0).
- The coefficient β_4 is the adjusted difference in logit attributable to interaction between CC and TV (CC * TV).
- The random school deviation is represented by v_{0i}. Note that we assume a single, fixed and thus common PreTHKS slope over the level-2 units.
- The interpretation of the coefficients is dependent on the coding of the variables used in the model.

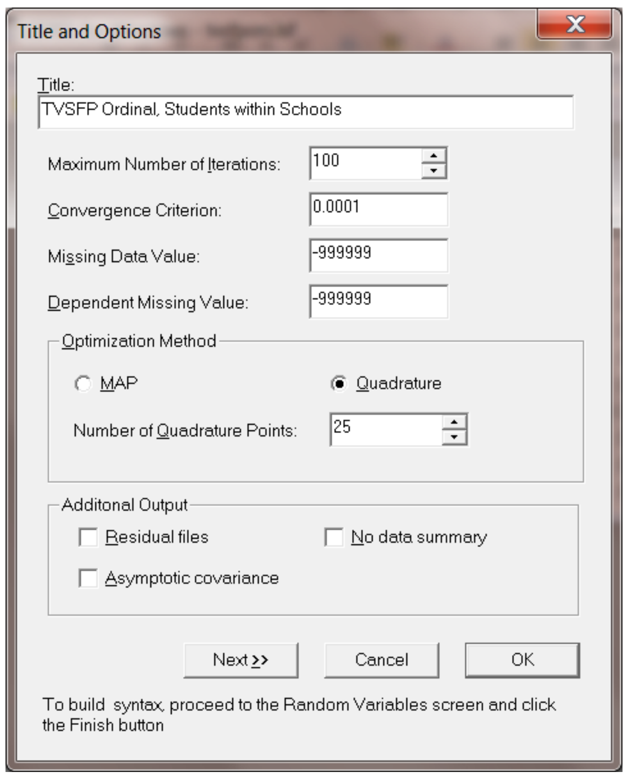
2.2.6 Setting up the analysis: Proportional Odds Model

Using the data in **tvsfpors.lsf**, we consider the situation where students are nested within schools and fit a two-level model with the ordinal variable THKSord as outcome. We wish to examine the relationships between the outcome and the two intervention methods employed, simultaneously taking students' pre-intervention scores into account. To do so, we use the model described above with schools as the level-2 units.

Use the **File**, **Open Spreadsheet** option to activate the display of an **Open** dialog box. Browse for the file **tvsfpors.lsf**. Select the **Multilevel, Generalized Linear Model** option from the main menu.

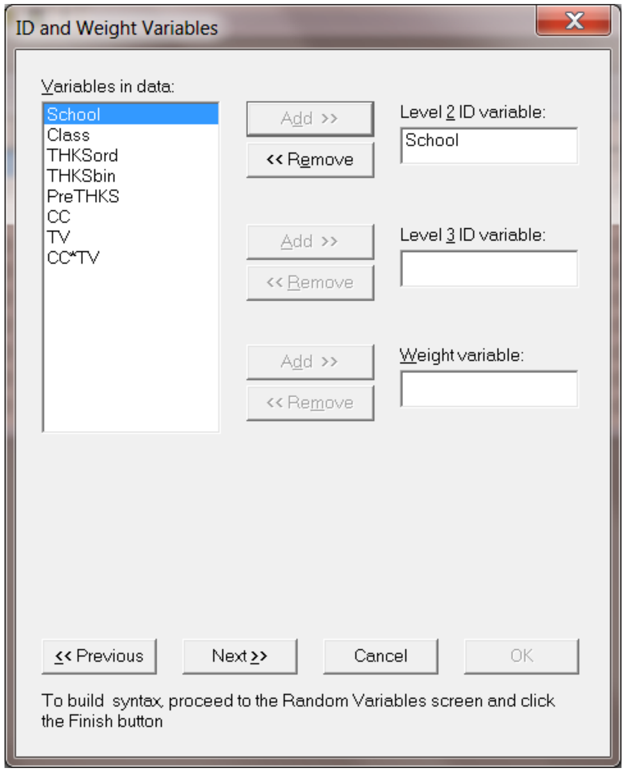


We are now ready to provide model specifications by clicking on the **Title and Options…** menu item. We opt to increase the number of quadrature points to be used during estimation. to 25 (default is 10).

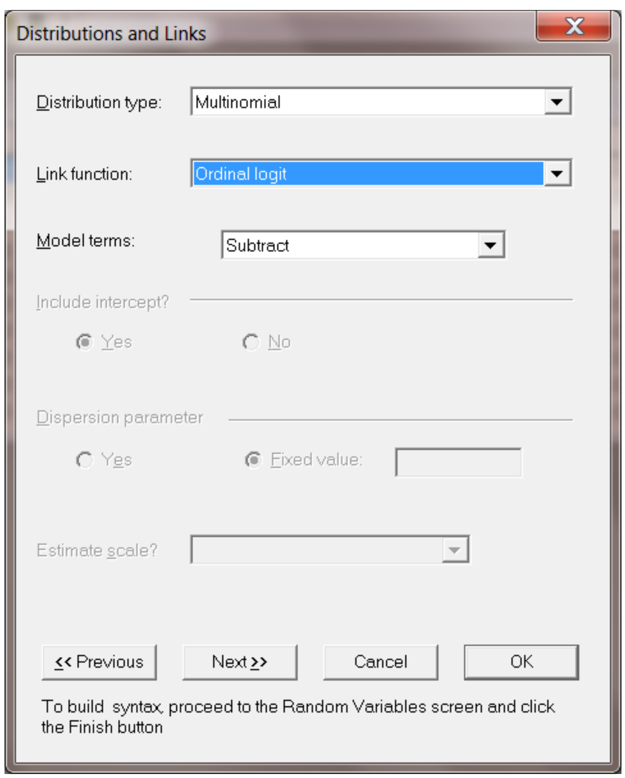


Click the **Next** button to obtain the **ID and Weight Variables** dialog.

The School identification variable is used to define the hierarchical structure of the data, and is selected as the **Level-2 ID Variable** from the **Variables in data** list.



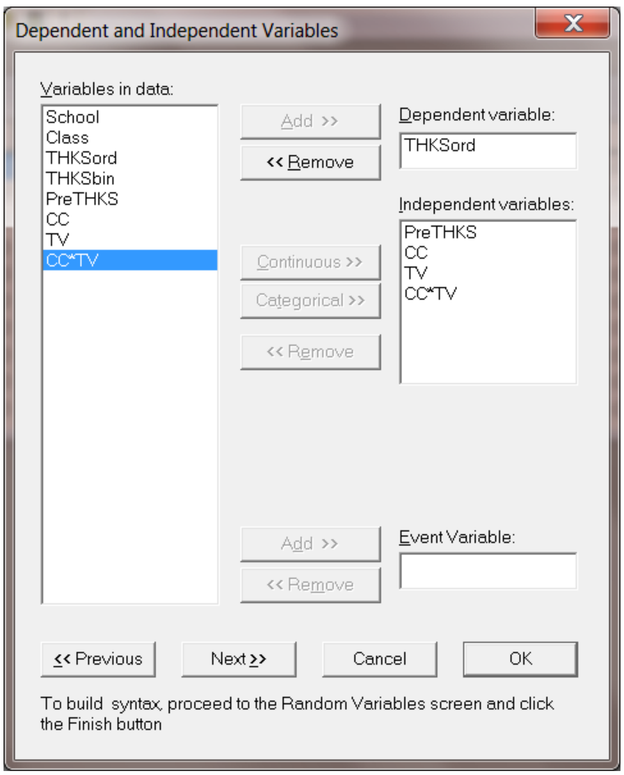
To proceed to the **Distribution and Links dialog** click the Next button and select **Multinomial** as the distribution type and **Ordinal Logit** as the link function type as shown below:



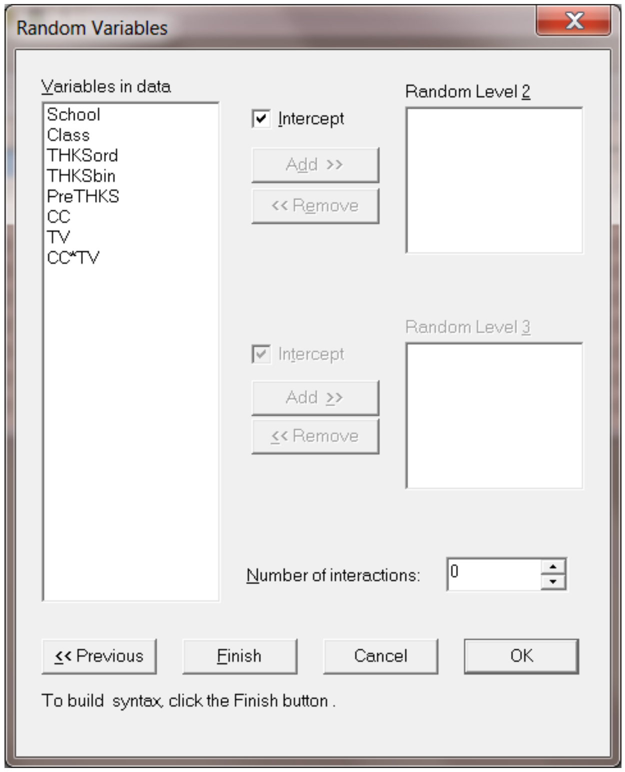
Note (see the dialog above) that covariate and random effect means are subtracted from the thresholds, implying that a positive coefficient indicates a positive association between the outcome and the predictor in question. To add the covariate and random effect means instead of using the default subtract setting, the **add** option must be selected in the **Model Terms** field on the **Advanced** tab of the **Model Setup** dialog box.

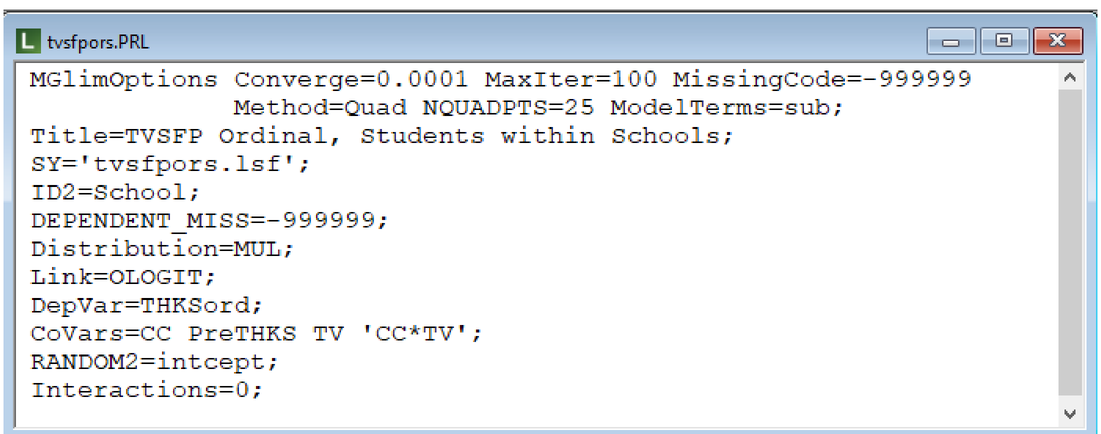
By clicking the **Next** button on the dialog shown above, the **Dependent and Independent Variables** dialog is displayed.

Start by selecting the ordinal outcome variable THKSord from the **Variables in data** list. Complete the model specifications by selecting PreTHKS, CC, TV and CC*TV as the independent variables (predictors).



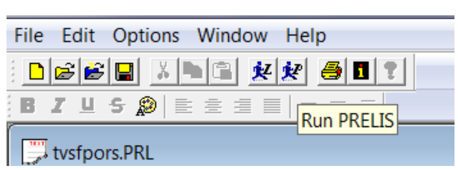
After selecting all the independent (explanatory) variables, the random effect(s) at level 2 must be selected. This is accomplished by again clicking **Next** to proceed to the last dialog. In this case, we wish to allow only the intercept to vary randomly over the schools. By default, the intercept is assumed to vary randomly over higher levels of the hierarchy as indicated by the checked boxes for the **Include Intercept** options.






```
MGlimOptions Converge=0.0001 MaxIter=100 MissingCode=-999999
             Method=Quad NQUADPTS=25 ModelTerms=sub;
Title=TVSFP Ordinal, Students within Schools;
SY='tvsfpors.lsf';
ID2=School;
DEPENDENT_MISS=-999999;
Distribution=MUL;
Link=OLOGIT;
DepVar=THKSord;
CoVars=CC PreTHKS TV 'CC*TV';
RANDOM2=intcept;
Interactions=0;
```

Run the analysis by clicking the **Run Prelis** icon on the main menu bar.



2.2.7 Discussion of results

Portions of the output file **tvsfpors.out** are shown below.

The section shown below contains a description of the model specifications and of the number of level-2 and level-1 units in the data.

```
tvsfpors.OUT

o=========================================o
| TVSFP Ordinal, Students within Schools |
|                                         |
o=========================================o


          Model and Data Descriptions

   Sampling Distribution                   = Multinomial
   Link Function                           = Cumulative Logit
   Number of Level-2 Units                   28
   Number of Level-1 Units                   1600
   Number of Level-1 Units per Level-2 Unit =
    23     25     26     70     31     42     52     55     39     33     52     65
    27     80     33     18     34     38     67     73     70     74     82    114
   113     33     94    137
```

Descriptive statistics

After the observation counts, descriptive statistics for all variables included in the model are given.

```
o=============================================================o
| Descriptive statistics for all the variables in the model |
o=============================================================o
                                                        Standard
   Variable       Minimum      Maximum        Mean     Deviation
   --------       -------      -------        ----     ---------
   THKSord1        0.0000       1.0000       0.2219       0.4156
   THKSord2        0.0000       1.0000       0.2487       0.4324
   THKSord3        0.0000       1.0000       0.2500       0.4331
   THKSord4        0.0000       1.0000       0.2794       0.4488
   PreTHKS         0.0000       6.0000       2.0694       1.2602
   CC              0.0000       1.0000       0.4769       0.4996
   TV              0.0000       1.0000       0.4994       0.5002
   CC*TV           0.0000       1.0000       0.2394       0.4268
```

Fixed effects results

The output describing the estimated parameters after convergence is shown next. Two iterations were required to obtain convergence, using 25 quadrature points per dimension. The likelihood function value at convergence as well as the deviance are also given, and may be used to compare a set of nested models. The estimates are shown in the column with heading Estimate and correspond to the coefficients $\beta_0, \beta_1, \ldots, \beta_4$ in the model specification. Significant effects of PreTHKS and CC are observed. With the exception of the CC *TV interaction term, positive relationships between the predictors and the ordinal outcome variable are indicated by these results. We also note that the coefficient associated with the curriculum-based intervention (CC) is almost three times the size of the estimated coefficient for media intervention (TV).

```
o==========================================o
| Optimization Method: Adaptive Quadrature |
o==========================================o

   Number of quadrature points =              25
   Number of free parameters =                 8
   Number of iterations used =                 2

   -2lnL (deviance statistic) =        4239.48553
   Akaike Information Criterion        4255.48553
   Schwarz Criterion                   4298.50760
```

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
Thresh1	-0.0885	0.1641	-0.5390	0.5899
Thresh2	1.1534	0.1656	6.9639	0.0000
Thresh3	2.3319	0.1734	13.4466	0.0000
PreTHKS	0.4033	0.0389	10.3780	0.0000
CC	0.9238	0.2041	4.5266	0.0000
TV	0.2750	0.1977	1.3906	0.1643
CC*TV	-0.4659	0.2846	-1.6371	0.1016

Odds Ratio and 95% Odds Ratio Confidence Intervals

Parameter	Estimate	Odds Ratio	Bounds Lower	Bounds Upper
Thresh1	-0.0885	0.9153	0.6636	1.2626
Thresh2	1.1534	3.1688	2.2904	4.3840
Thresh3	2.3319	10.2979	7.3304	14.4666
PreTHKS	0.4033	1.4967	1.3870	1.6152
CC	0.9238	2.5188	1.6884	3.7576
TV	0.2750	1.3165	0.8935	1.9398
CC*TV	-0.4659	0.6276	0.3592	1.0963

The alternative parameterization, setting threshold = 0 is shown next. The estimates of γ_2 and γ_3 are 1.242 and 2.420 respectively – recall that for identification purposes γ_1 was set to zero.

Alternative Parameterization, setting Threshold1= 0

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
intcept	0.0885	0.1641	0.5390	0.5899
Thresh2	1.2418	0.0571	21.7663	0.0000
Thresh3	2.4204	0.0748	32.3401	0.0000

Odds Ratio and 95% Odds Ratio Confidence Intervals

Parameter	Estimate	Odds Ratio	Bounds Lower	Upper
intcept	0.0885	1.0925	0.7920	1.5070
Thresh2	1.2418	3.4619	3.0956	3.8715
Thresh3	2.4204	11.2502	9.7153	13.0277

Random effects results

The last part of the output file contains information on the random effects and calculation of the intra-cluster correlation coefficient. The variation in intercept over schools is estimated at 0.0735, with the associated *p*-value of 0.055 indicating its statistical significance.

Estimated level 2 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intcept/intcept	0.0735	0.0383	1.9189	0.0550

Level 2 covariance matrix

	intcept
intcept	0.073516

Level 2 correlation matrix

	intcept
intcept	1.000000

```
Calculation of the intracluster correlation
-------------------------------------------
residual variance = pi*pi / 3 (assumed)
cluster  variance =      0.0735

intracluster correlation =     0.0735 / (     0.0735 + (pi*pi/3)) = 0.022
```

In the case of the fixed effects, a 2-tailed p-value is used, as the alternative hypothesis considered here is of the form $H_1 : \beta \neq 0$. As variances are constrained to be elements of the interval $[0, +\infty)$, the p-values used for these effects are 1-tailed. If the model is true, it is assumed that the level-1 error variance, σ_e^2, is equal to $\pi^2 / 3$ for the logistic link function, where π represents the constant 3.141592654 (see, *e.g.*, Hedeker & Gibbons (2006), p. 157).

Finally, the calculation of the intra-cluster correlation is shown. The value of 0.022 indicates that almost all variation is attributable to students, rather than to the schools.

2.2.8 Setting up the analysis: Non-Proportional Odds Model

Use the **File**, **Open** and browse for the syntax file **tvsfpors.prl** (see Section 1.2.2.6):

tvsfpors.PRL

```
MGlimOptions Converge=0.0001 MaxIter=100 MissingCode=-999999
             Method=Quad NQUADPTS=25 ModelTerms=sub;
Title=TVSFP Ordinal, Students within Schools;
SY='C:\LISREL Examples\MGLIMEX\tvsfpors.lsf';
ID2=School;
DEPENDENT_MISS=-999999;
Distribution=MUL;
Link=OLOGIT;
DepVar=THKSord;
CoVars=PreTHKS CC TV 'CC*TV';
RANDOM2=intcept;
Interactions=0;
```

Suppose that we want to check the non-proportional assumption for the variables CC and TV. In this case we need to reorder the predictors (CoVars) so that the variables CC and TV are listed first. Since there are two predictors of interest, we need to change **Interactions=0** to **Interactions=2**. The revised syntax file (saved as **tvsfpors1.prl**) is shown below:

tvsfpors.PRL

```
MGlimOptions Converge=0.0001 MaxIter=100 MissingCode=-999999
             Method=Quad NQUADPTS=25 ModelTerms=sub;
Title=TVSFP Ordinal, Students within Schools;
SY='tvsfpors.lsf';
ID2=School;
DEPENDENT_MISS=-999999;
Distribution=MUL;
Link=OLOGIT;
DepVar=THKSord;
CoVars=CC TV 'CC*TV' PreTHKS;
RANDOM2=intcept;
Interactions=2;
```

Run the analysis by clicking the **Run Prelis** icon on the main menu bar.

2.2.9 Discussion of results

Portions of the output file **Ttvsfpors1.prl** are shown below.

Fixed effects results

The output describing the estimated parameters after convergence is shown next. Two iterations were required to obtain convergence, using 25 quadrature points per dimension. The likelihood function value at convergence as well as the deviance are also given, and may be used to compare a set of nested models.

tvsfpors1.OUT

```
Number of quadrature points =              25
Number of free parameters =                12
Number of iterations used =                 2

-2lnL (deviance statistic) =       4234.01997
Akaike Information Criterion       4258.01997
Schwarz Criterion                  4322.55308
```

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
Thresh1	-0.1041	0.1694	-0.6144	0.5389
Thresh2	1.2148	0.1689	7.1919	0.0000
Thresh3	2.2643	0.1826	12.3976	0.0000
CC	0.9369	0.2224	4.2128	0.0000
TV	0.2286	0.2106	1.0857	0.2776
CC*TV	-0.4704	0.2841	-1.6556	0.0978
PreTHKS	0.4045	0.0389	10.4033	0.0000
Interactions of predictors with: Thresh2				
CC	0.0952	0.1141	0.8341	0.4043
TV	0.0716	0.1133	0.6323	0.5272
Interactions of predictors with: Thresh3				
CC	-0.1325	0.1460	-0.9073	0.3643
TV	0.0444	0.1458	0.3045	0.7607

Random effects results

The last part of the output file contains information on the random effects and calculation of the intra-cluster correlation coefficient. The variation in intercept over schools is estimated at 0.0721, with the associated *p*-value of 0.058 indicating its statistical significance.

```
                    Estimated level 2 variances and covariances

                                                Standard
  Parameter                  Estimate            Error      z Value    P Value
  ---------                  --------           --------    -------    -------
  intcept/intcept              0.0721             0.0381     1.8933     0.0583

 Level 2 covariance matrix

              intcept
 intcept     0.072087


 Level 2 correlation matrix

              intcept
 intcept     1.000000


Calculation of the intracluster correlation
-------------------------------------------
residual variance = pi*pi / 3 (assumed)
cluster  variance =     0.0721

intracluster correlation =     0.0721 / (     0.0721 + (pi*pi/3)) = 0.021
```

Testing the proportional odds assumption

For a pair of nested models, the difference in $-2\ln L$ values has a χ^2 distribution, with degrees of freedom equal to the difference in number of parameters estimated in the models compared. From the information contained in **tvsfpors.out** (proportional odds model) and **tvsfpors1.out** (non-proportional odds model) it follows that χ^2 =4239.49 – 4234.02 = 5.47 with degrees of freedom 12 – 8 =4. Since this is a non-significant result, we conclude that the proportional odds assumption cannot be rejected.

3 References

Hedeker, D. (2008). Multilevel Models for Ordinal and Nominal Variables. in J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 237-274). New York: Springer.

Hedeker, D. & Gibbons, R.D. (2006). *Longitudinal Data Analysis*. New York: Wiley.

Hedeker, D. & Gibbons, R.D. (1994). A random-effects ordinal regression model for multilevel analysis, *Biometrics*, **50**, 933-944.

Hedeker D., Mermelstein R.J. (1998) A multilevel thresholds of change model for analysis of stages of change data. *Multivariate Behavioral Research*, **33**: 427-55.