Mixed Models for Clustered Ordinal Outcomes

Don Hedeker
Department of Public Health Sciences
Biological Sciences Division
University of Chicago

hedeker@uchicago.edu

Hedeker, D. (2014). Methods for multilevel ordinal data in prevention research. *Prevention Science.*

Hedeker, D. (2008). Multilevel models for ordinal and nominal variables. In J. de Leeuw & E. Meijer (Eds.), Handbook of Multilevel Analysis. Springer, New York.

Hedeker, D. (2005). Generalized linear mixed models. In B. Everitt & D. Howell (Eds.), Encyclopedia of Statistics in Behavioral Science. Wiley.

Why analyze as ordinal?

- Efficiency: Armstrong & Sloan (1989, Amer Jrn of Epid) and Strömberg (1996, Amer Jrn of Epid) report efficiency losses between 49% to 87% when dichotomizing an ordinal outcome with five categories.

- Bias: continuous model can yield correlated residuals and regressors when used for ordinal outcomes; continuous model does not take into account the ceiling and floor effects of the ordinal outcome. Results in biased estimates of regression coefficients and is most critical when the ordinal variables is highly skewed (see Bauer & Sterba, 2011, Psych Methods)

- Logic: continuous model can yield predicted values outside of the range of the ordinal variable.

Proportional Odds Model - McCullagh (1980)

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

$$\begin{aligned} c &= 1, \ldots, C-1 \text{ for the } C \text{ categories of the ordinal outcome} \\ \boldsymbol{x} &= \text{vector of explanatory variables (plus the intercept)} \\ \gamma_c &= \text{thresholds; reflect cumulative odds when } \boldsymbol{x} = 0 \text{ (for identification: } \gamma_1 = 0 \text{ or } \beta_0 = 0) \end{aligned}$$

- positive association between x and Y is reflected by $\beta > 0$
- the effect of x is assumed to be the same for each cumulative odds ratio
- odds that the response is greater than or equal to c (for fixed c) is multiplied by e^{β} for every unit change in x:

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

Ordinal Model for Dichotomous Response: same as it ever was!

$$\log\left[\frac{P(Y=0)}{1-P(Y=0)}\right] = 0 - \boldsymbol{x}'\boldsymbol{\beta}$$

$$\frac{P(Y=0)}{1-P(Y=0)} = \exp(0 - \boldsymbol{x}'\boldsymbol{\beta})$$

$$\frac{1-P(Y=0)}{P(Y=0)} = \left[\exp(0 - \boldsymbol{x}'\boldsymbol{\beta})\right]^{-1}$$

$$\frac{1-P(Y=0)}{P(Y=0)} = \exp(\boldsymbol{x}'\boldsymbol{\beta})$$

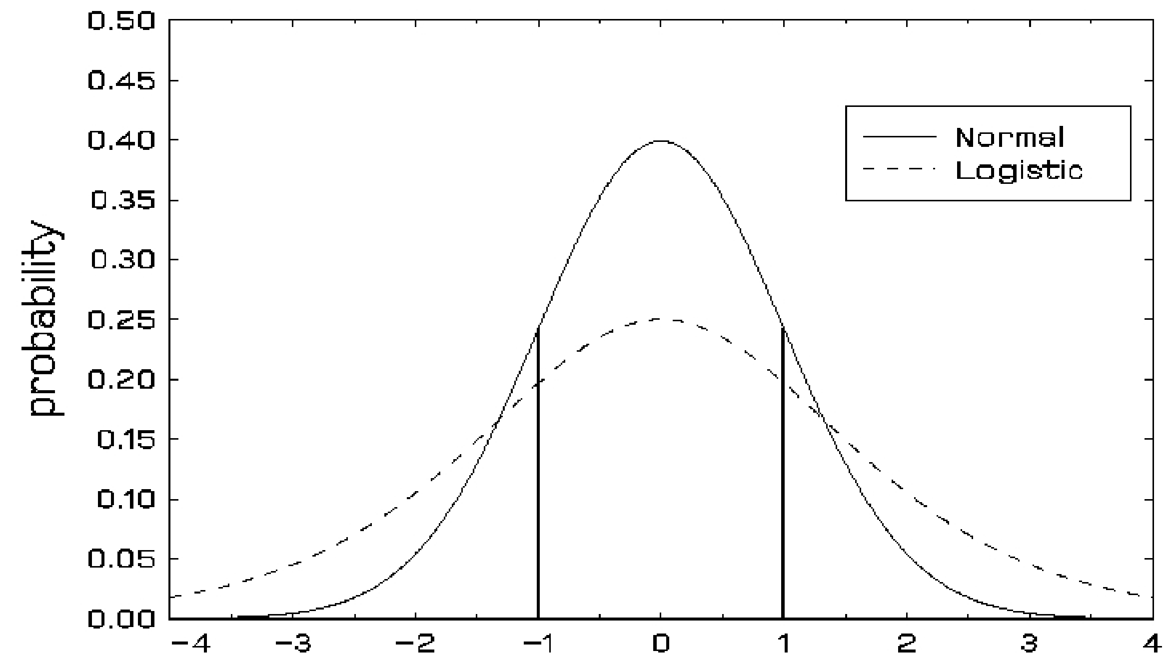
$$\log\left[\frac{P(Y=1)}{1-P(Y=1)}\right] = \boldsymbol{x}'\boldsymbol{\beta}$$

Ordinal Response and Threshold Concept

Continuous y_i - unobservable latent variable - related to ordinal response Y_i via "threshold concept"

- threshold values $\gamma_1, \gamma_2, \ldots, \gamma_{C-1}$ ($\gamma_0 = -\infty$ and $\gamma_C = \infty$)
- C = number of ordered categories

Response occurs in category c, $\quad Y_i = c \quad$ if $\quad \gamma_{c-1} < y_i < \gamma_c$



The Threshold Concept in Practice

"How was your day?"
(what is your level of satisfaction today?)

- Satisfaction may be continuous, but we sometimes emit an ordinal response:

Model for Latent Continuous Responses

Consider the model with p covariates for the latent response strength y_i $(i = 1, 2, \ldots, N)$:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i$$

- probit: $\varepsilon_i \sim$ standard normal (mean=0, variance=1)
- logistic: $\varepsilon_i \sim$ standard logistic (mean=0, variance=$\pi^2/3$)

$\Rightarrow \boldsymbol{\beta}$ estimates from logistic regression are larger (in abs. value) than from probit regression by approximately $\sqrt{\pi^2/3} = 1.8$

Underlying latent variable

- useful way of thinking of the problem
- not an essential assumption of the model

Mixed-effects ordinal logistic regression model

(Hedeker & Gibbons, 1994, 1996)

- $i = 1, \ldots N$ level-2 units (clusters or subjects)
- $j = 1, \ldots, n_i$ level-1 units (subjects or repeated observations)
- $c = 1, 2, \ldots, C$ response categories
- Y_{ij} = ordinal response of level-2 unit i and level-1 unit j

How was your day? (asked repeatedly each day for a week)

Random-intercept Ordinal Logistic Regression Model

$$\lambda_{ijc} = \log\left[\frac{P_{ijc}}{(1 - P_{ijc})}\right] = \gamma_c - (\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \upsilon_{0i})$$

- $P_{ijc} = \Pr\left(Y_{ij} \leq c \mid \boldsymbol{\upsilon}\, ;\, \gamma_c, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\upsilon\right) = \frac{1}{1+\exp(-\lambda_{ijc})}$
- $p_{ijc} = \Pr\left(Y_{ij} = c \mid \boldsymbol{\upsilon}\, ;\, \gamma_c, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\upsilon\right) = P_{ijc} - P_{ijc-1}$

- $C - 1$ strictly increasing model thresholds γ_c
- $\boldsymbol{x}_{ij} = p \times 1$ covariate vector
- $\boldsymbol{\beta} = p \times 1$ fixed regression parameters
- υ_{0i} = cluster effects distributed $\sim N(0, \sigma_\upsilon^2)$

Model for Latent Continuous Responses

Model with p covariates for the latent response strength y_{ij}:

$$y_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \upsilon_{0i} + \varepsilon_{ij}$$

where $\upsilon_{0i} \sim N(0, \sigma^2_\upsilon)$, and assuming

- $\varepsilon_{ij} \sim$ standard normal (mean 0 and $\sigma^2 = 1$) leads to mixed-effects ordinal probit regression
- $\varepsilon_{ij} \sim$ standard logistic (mean 0 and $\sigma^2 = \pi^2/3$) leads to mixed-effects ordinal logistic regression

Underlying latent variable

- not an essential assumption of the model
- useful for obtaining intra-class correlation (r)

$$r = \frac{\sigma_v^2}{\sigma_v^2 + \sigma^2}$$

and for design effect (d)

$$d = \frac{\sigma_v^2 + \sigma^2}{\sigma^2} = 1/(1 - r)$$

ratio of actual variance to the variance that would be obtained by simple random sampling (holding sample size constant)

Scaling of regression coefficients

Fixed-effects model

$\boldsymbol{\beta}$ estimates from logistic regression are larger (in abs. value) than from probit regression by approximately

$$\sqrt{\frac{\pi^2/3}{1}} = 1.8$$

because

- $V(y) = \sigma^2 = \pi^2/3$ for logistic
- $V(y) = \sigma^2 = 1$ for probit

Mixed-effects model

$\boldsymbol{\beta}$ estimates from mixed-effects (random intercepts) model are larger (in abs. value) than from fixed-effects model by approximately

$$\sqrt{d} = \sqrt{\frac{\sigma_v^2 + \sigma^2}{\sigma^2}}$$

because

- $V(y) = \sigma_v^2 + \sigma^2$ in mixed-effects (random intercepts) model
- $V(y) = \sigma^2$ in fixed-effects model

- difference depends on size of random-effects variance σ_v^2
- more complex for models with multiple random effects

Numerical Quadrature: integration over random effect distribution

- method to numerically perform an integration

$$\int_{v} f(\mathsf{y}_i \mid v) g(v) dv \approx \sum_{q=1}^{Q} f(\mathsf{y}_i \mid B_q) A(B_q)$$

where B_q $(q = 1, \ldots, Q)$ are the quadrature nodes or points
$A(B_q)$ $(q = 1, \ldots, Q)$ are the weights (sum = 1)

- More points, more accurate the approximation, but more time

- For standard normal distribution, Gauss-Hermite quadrature

- Yields a likelihood value that can be used for LR tests

- Full-likelihood approach found in STATA, SUPERMIX, MIXOR, SAS PROC NLMIXED & GLIMMIX

Other methods for integration of θ

Methods based on first- or second-order Taylor series expansions

- Marginal quasi-likelihood (MQL) involves expansion around the fixed part of the model
- Penalized or predictive quasi-likelihood (PQL) also includes the random part in its expansion
- fast, but doesn't yield a likelihood for LR tests
- can yield downwardly biased estimates in certain situations (if N and/or n is small, or ICC is high), especially for MQL
- Not available in Supermix, but other software programs use these (e.g., SPSS, some SAS PROCs, MLwiN)

Laplace approximation - Raudenbush et. al., (2000)

- a combination of a fully multivariate Taylor series expansion and Laplace approximation
- fast and computationally accurate, though some bias for variance parameters
- yields a likelihood for LR tests
- available in Stata, also in HLM (though not for all models)

Other methods

- Markov Chain Monte Carlo (MCMC) Bayesian approach (in BUGS)
- Maximum Simulated Likelihood (in some STATA programs) in econometric, transportation, political science literatures

Effects of a School-based Intervention

The Television School and Family Smoking Prevention and Cessation Project (Flay, *et al.*, 1988); a subsample:

- *sample* - 1600 7th-graders - 135 classes - 28 schools
  - 1 to 13 classes per school, 2 to 28 students per class
- *outcome* - knowledge of the effects of tobacco use
- *timing* - students tested at pre and post-intervention
- *design* - schools exposed to
  - a social-resistance classroom curriculum (CC)
  - a media (television) intervention (TV)
  - CC combined with TV
  - a no-treatment control group

Main question of interest:

- Influence of the intervention on the tobacco health knowledge scores (THKS) ?

Challenges in the analysis:

- outcome variable (THKS) is number correct of 7 items
- controlling for intra-school and intra-class variability
- potential explanatory variables are at different levels

Tobacco and Health Knowledge Scale
Post-Intervention Scores - Frequencies (percentages)

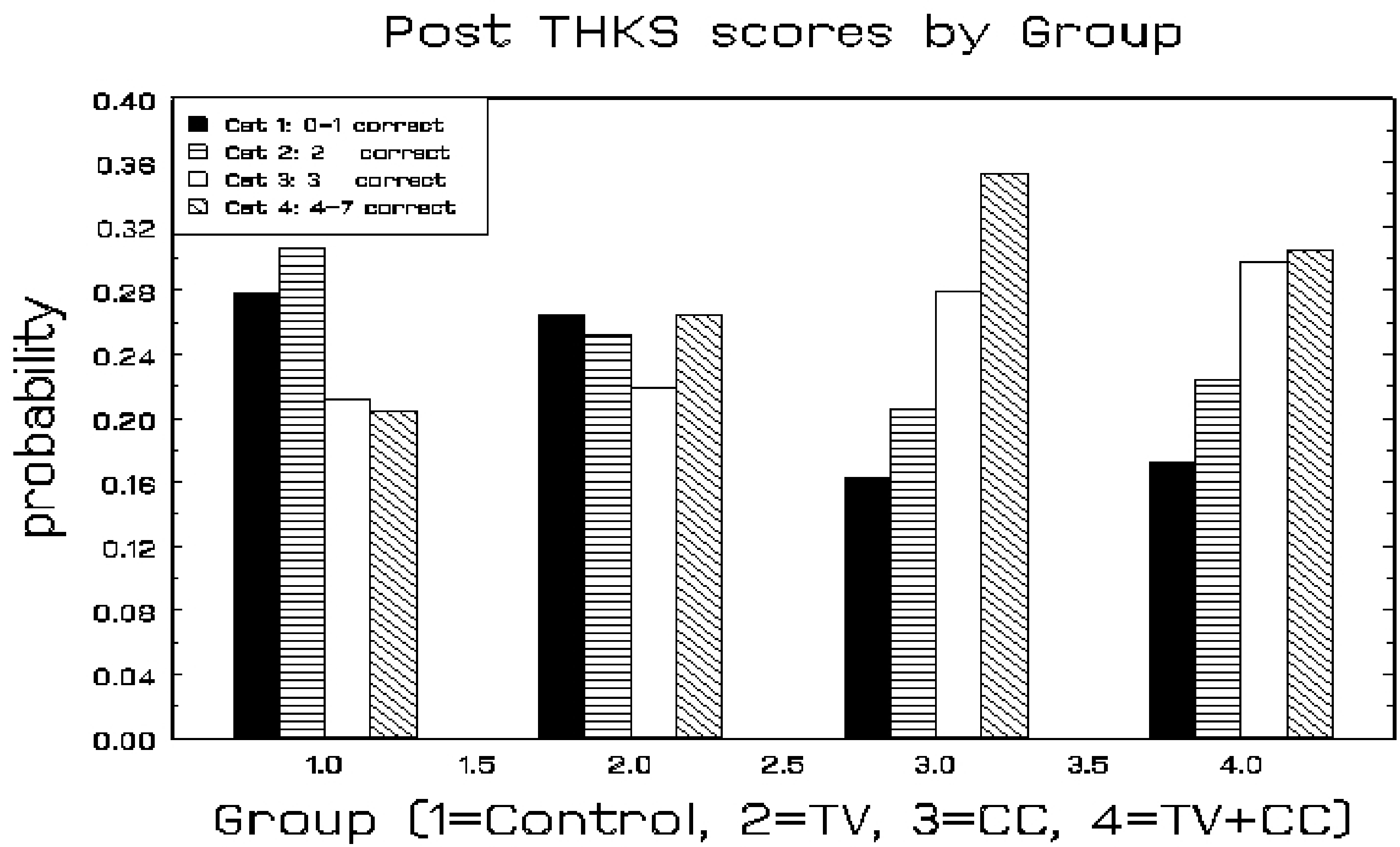
subgroup		*THKS score*				
CC	TV	0-1	2	3	4-7	*total*
no	no	117 (27.8)	129 (30.6)	89 (21.1)	86 (20.4)	421
no	yes	110 (26.4)	105 (25.2)	91 (21.9)	110 (26.4)	416
yes	no	62 (16.3)	78 (20.5)	106 (27.9)	134 (35.3)	380
yes	yes	66 (17.2)	86 (22.5)	114 (29.8)	117 (30.5)	383
	total	355 (22.2)	398 (24.9)	400 (25.0)	447 (27.9)	1600

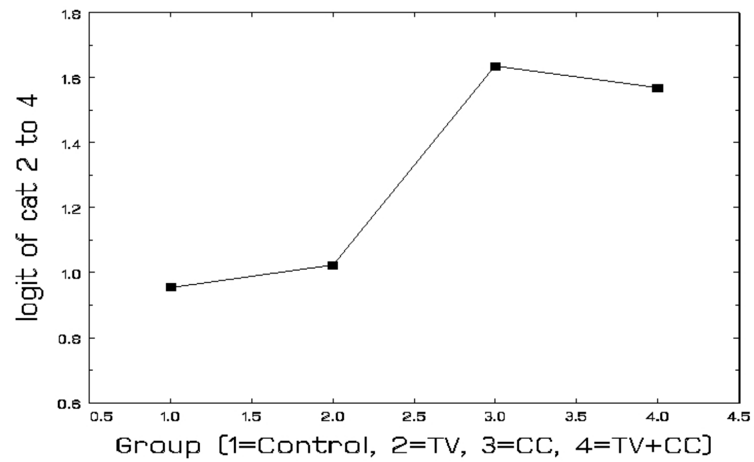
THKS Post-Intervention Scores - Proportions, Odds, Logits

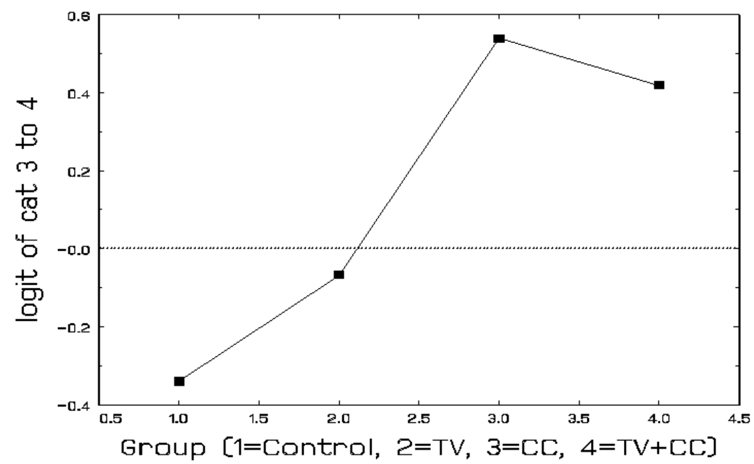
subgroup		*proportions*				*cumulative prop*		
CC	TV	1	2	3	4	2-4	3-4	4
no	no	.278	.306	.211	.204	.722	.416	.204
no	yes	.264	.252	.219	.264	.736	.483	.264
yes	no	.163	.205	.279	.353	.837	.632	.353
yes	yes	.172	.225	.298	.305	.826	.603	.305

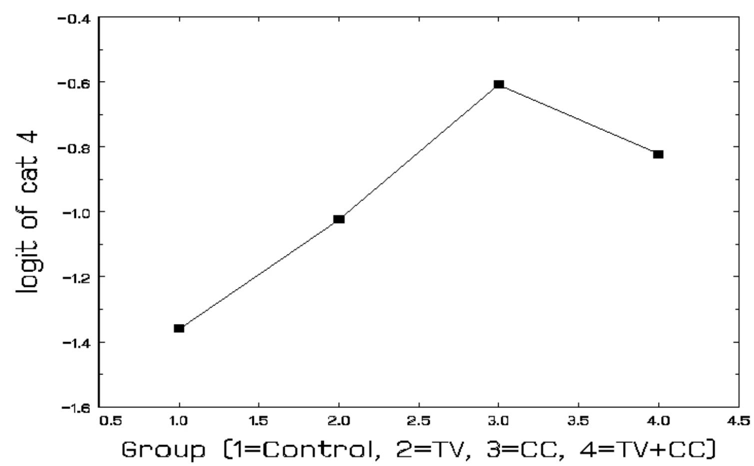
subgroup		*odds*			*logits*		
CC	TV	2-4 vs 1	3-4 vs 1-2	4 vs 1-3	2-4 vs 1	3-4 vs 1-2	4 vs 1-3
no	no	2.598	.711	.257	.955	-.341	-1.360
no	yes	2.782	.935	.359	1.023	-.067	-1.023
yes	no	5.129	1.714	.545	1.635	.539	-.607
yes	yes	4.803	1.520	.440	1.569	.419	-.821

Observed Proportions by Group




logit of cat 2 to 4
Group (1=Control, 2=TV, 3=CC, 4=TV+CC)


logit of cat 3 to 4
Group (1=Control, 2=TV, 3=CC, 4=TV+CC)


logit of cat 4
Group (1=Control, 2=TV, 3=CC, 4=TV+CC)

Within-Clusters / Between-Clusters components

Within-clusters model - level 1 $\quad (j = 1, \ldots, n_i \text{ subjects})$

$$logit_{ijc} = b_{0ic}$$

Between-clusters model - level 2 $\quad (i = 1, \ldots, N \text{ clusters})$

$$b_{0ic} \;=\; \gamma_c \;-\; [\beta_1 CC_i \;+\; \beta_2 TV_i \;+\; \beta_3 (CC_i \times TV_i) \;+\; \upsilon_{0i}]$$

$$\upsilon_{0i} \sim \mathcal{NID}(0, \sigma_\upsilon^2)$$

$$\gamma_c \;=\; (C-1) \text{ THKS logits for CC=no TV=no subgroup}$$

$$\beta_1 \;=\; \text{logit diff. between CC=yes vs CC=no (for TV=no)}$$

$$b_{0ic} \;=\; \gamma_c \;-\; [(\beta_1 + \beta_3 TV_i)CC_i \;+\; \beta_2 TV_i \;+\; \upsilon_{0i}]$$

$$\beta_2 \;=\; \text{logit diff. between TV=yes vs TV=no (for CC=no)}$$

$$b_{0ic} \;=\; \gamma_c \;-\; [(\beta_2 + \beta_3 CC_i)TV_i \;+\; \beta_1 CC_i \;+\; \upsilon_{0i}]$$

$$\beta_3 \;=\; \text{difference in logit attributable to interaction}$$

$$\upsilon_{0i} \;=\; \text{random cluster deviation}$$

note: interpretation depends on coding of variables, and βs are adjusted for the cluster effects (cluster-specific effects)

3-level model

Within-classrooms (and schools) model - *level 1* ($k = 1, \ldots, n_{ij}$ students)

$$logit_{ijkc} = b_{0ijc}$$

Between-classrooms (within-schools) model - *level 2* ($j = 1, \ldots, n_i$ classrooms)

$$b_{0ijc} \;=\; b_{0ic} \;+\; \upsilon_{0ij}$$

Between-schools model - *level 3* ($i = 1, \ldots, N$ schools)

$$b_{0ic} \;=\; \gamma_c \;-\; [\beta_1 CC_i \;+\; \beta_2 TV_i \;+\; \beta_3 (CC_i \times TV_i) \;+\; \upsilon_{0i}]$$

$$\upsilon_{0ij} \sim \mathcal{NID}(0, \sigma^2_{\upsilon(2)}) \quad \text{and} \quad \upsilon_{0i} \sim \mathcal{NID}(0, \sigma^2_{\upsilon(3)})$$

$$
\begin{aligned}
\gamma_c &= \text{(C-1) THKS logits for CC=no TV=no subgroup} \\
\beta_1 &= \text{logit diff. between CC=yes vs CC=no (for TV=no)} \\
\beta_2 &= \text{logit diff. between TV=yes vs TV=no (for CC=no)} \\
\beta_3 &= \text{difference in logit attributable to interaction} \\
\upsilon_{0ij} &= \text{random classroom deviation} \\
\upsilon_{0i} &= \text{random school deviation}
\end{aligned}
$$

THKS Post-Int (ordinal) Scores - LR Estimates (std errs)

	Fixed		*Multilevel* *2-level*		*Multilevel* *3-level*	
cut 1	-.889	***	-.919	***	-.925	***
	(.093)		(.132)		(.180)	
cut 2	.275	***	.309	**	.302	*
	(.090)		(.130)		(.178)	
cut 3	1.366	***	1.459	***	1.452	***
	(.096)		(.136)		(.182)	
CC	.777	***	.764	***	.823	***
	(.128)		(.186)		(.254)	
TV	.224	*	.151		.236	
	(.125)		(.183)		(.249)	
CC× TV	-.372	**	-.269		-.431	
	(.180)		(.263)		(.356)	
class var			.260		.161	
			(.074)		(.067)	
school var					.106	
					(.061)	
-2 log L	4377.98		4345.36		4339.31	

***$p < .01$ **$p < .05$ *$p < .10$ (Wald-tests not done for vars)

Calculation of ICC - 2 level model

$$r = \frac{\sigma_\upsilon^2}{\sigma_\upsilon^2 + \sigma^2}$$

Random classrooms model

$$r = \frac{.260}{.260 + \pi^2/3} = .073$$

$\Rightarrow$ 7.3% of the unexplained variation is at the classroom level

Calculation of ICC - 3 level model

Level-3 (likeness of students in the same school)

$$r = \frac{\sigma^2_{\upsilon(3)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2} = \frac{.106}{.106 + .161 + \pi^2/3} = .030$$

Level-2 (likeness of students in same classroom & school)

$$r = \frac{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2} = \frac{.106 + .161}{.106 + .161 + \pi^2/3} = .075$$

Level-2 (likeness of classes in the same school)

$$r = \frac{\sigma^2_{\upsilon(3)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)}} = \frac{.106}{.106 + .161} = .397$$

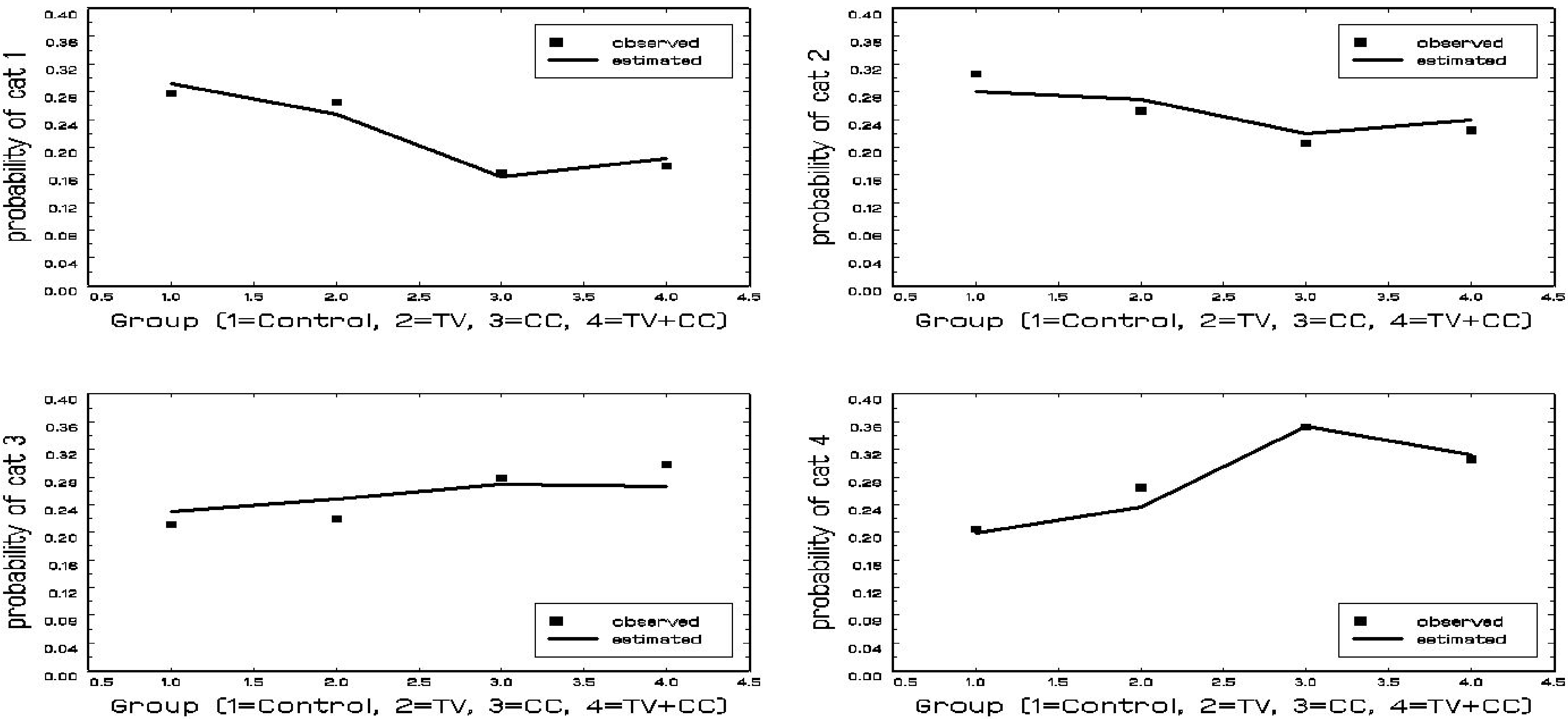
- $r < .5$: the school level contributes slightly less to variability than the class level
- average classroom post THKS scores are moderately similar within schools

Model fit of proportions: 3-level model

CC	TV	logistic $\Psi(z) = \frac{1}{1+\exp(-z)}$	estimate	observed
		Probability of Category 1 response		
0	0	$\Psi(-.925/\sqrt{\hat{d}})$	.291	.278
0	1	$\Psi((-.925 + .236)/\sqrt{\hat{d}})$	.247	.264
1	0	$\Psi((-.925 + .823)/\sqrt{\hat{d}})$	.157	.163
1	1	$\Psi((-.925 - (.236 + .823 - .431))/\sqrt{\hat{d}})$	.183	.172
		Probability of Category 1 or 2 response		
0	0	$\Psi(.302/\sqrt{\hat{d}})$	.572	.584
0	1	$\Psi((.302 + .236)/\sqrt{\hat{d}})$	.516	.517
1	0	$\Psi((.302 + .823)/\sqrt{\hat{d}})$	.377	.368
1	1	$\Psi((.302 - (.236 + .823 - .431))/\sqrt{\hat{d}})$	.422	.397
		Probability of Category 1, 2, or 3 response		
0	0	$\Psi(1.453/\sqrt{\hat{d}})$	.802	.796
0	1	$\Psi((1.453 - .236)/\sqrt{\hat{d}})$	.763	.736
1	0	$\Psi((1.453 - .823)/\sqrt{\hat{d}})$	.647	.647
1	1	$\Psi((1.453 - (.236 + .823 - .431))/\sqrt{\hat{d}})$	.689	.695

d = design effect = $(\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2)/\sigma^2$ $\quad \hat{d} = (.106 + .161 + \pi^2/3)/(\pi^2/3)$

Model Fit of Observed Proportions: 3-level model



Proportional and Non-proportional Odds

Proportional Odds model

$$\log\left[\frac{P(Y_{ij} \le c)}{1 - P(Y_{ij} \le c)}\right] = \gamma_c - \left[\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \upsilon_{0i}\right]$$

$$\text{with } \upsilon_{0i} \sim N(0, \sigma_\upsilon^2)$$

- relationship between the explanatory variables and the cumulative logits does not depend on c
- effects of $\boldsymbol{x}$ variables DO NOT vary across the $C - 1$ cumulative logits

Hedeker & Mermelstein (1998, *Mult Behav Res*) extension:

$$\log\left[\frac{P(Y_{ij} \leq c)}{1 - P(Y_{ij} \leq c)}\right] = \gamma_{c(0)} - \left[\boldsymbol{u}'_{ij}\boldsymbol{\gamma}_c + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \upsilon_{0i}\right]$$

$\boldsymbol{u}_{ij} = h \times 1$ vector for the set of h covariates for which proportional odds is not assumed

- effects of $\boldsymbol{u}$ variables DO vary across the $C - 1$ cumulative logits
- more flexible model for ordinal response relations
- can't estimate this model in most software programs, but is available in Supermix

Proportional Odds Assumption: covariate effects are the same across all cumulative logits

group	Absent	Response Mild	Severe	total
control	27	46	27	100
cumulative odds	$\frac{27}{73} = .37$	$\frac{73}{27} = 2.7$		
logit	*-1*	*1*		
treatment	38	44	18	100
cumulative odds	$\frac{38}{62} = .61$	$\frac{82}{18} = 4.6$		
logit	*-.5*	*1.5*		

$\Rightarrow$ *group difference = .5 for both cumulative logits*

Non-Proportional Odds: covariate effects vary across the cumulative logits

group	Response Absent	Mild	Severe	total
control	27	46	27	100
cumulative odds	$\frac{27}{73} = .37$	$\frac{73}{27} = 2.7$		
logit	*-1*	*1*		
treatment	28	60	12	100
cumulative odds	$\frac{28}{72} = .39$	$\frac{88}{12} = 7.3$		
logit	*-.95*	*2*		

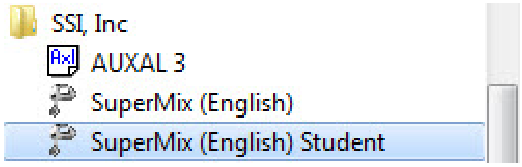
$\Rightarrow$ *UNEQUAL group difference across cumulative logits*

TVSFP Study: Post-Intervention THKS ($N = 1600$)
Ordinal LR Estimates (se) - *3-level model*

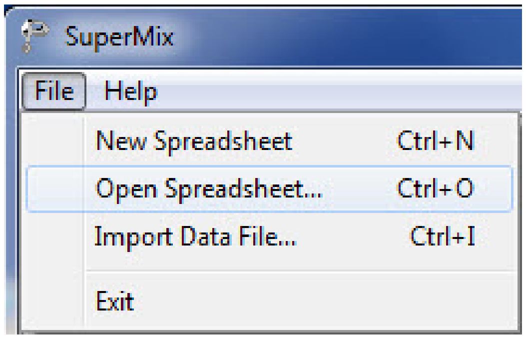
	Proportional Odds Model	Non-Proportional Odds Model		
		2-4 vs 1	3,4 vs 1,2	4 vs 1-3
CC	.823	.727	.928	.780
	(.254)	(.281)	(.262)	(.272)
TV	.236	.109	.281	.310
	(.249)	(.266)	(.256)	(.271)
CC by TV	-.431	-.205	-.444	-.584
	(.356)	(.396)	(.368)	(.381)
$-2 \log L$	4339.31	4332.42		

- Proportional Odds accepted ($\chi_6^2 = 4339.31 - 4332.42 = 6.89$)

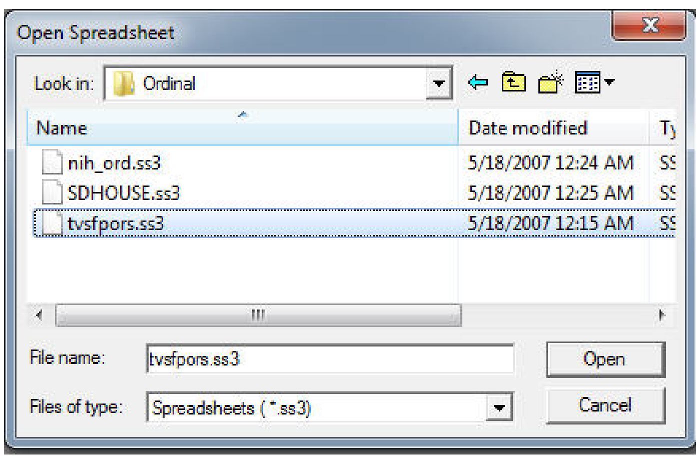
- Under SSI, Inc > "SuperMix (English)" or "SuperMix (English) Student"



- Under "File" click on "Open Spreadsheet"



- Open C:\SuperMixEn Examples\Workshop\Ordinal\tvsfpors.ss3 (or C:\SuperMixEn Student Examples\Workshop\Ordinal\tvsfpors.ss3)



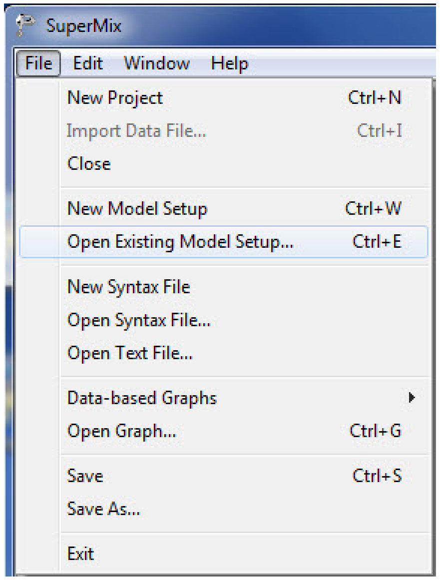
C:\SuperMixEn Examples\Workshop\Ordinal\tvsfpors.ss3

tvsfpors.ss3

Apply

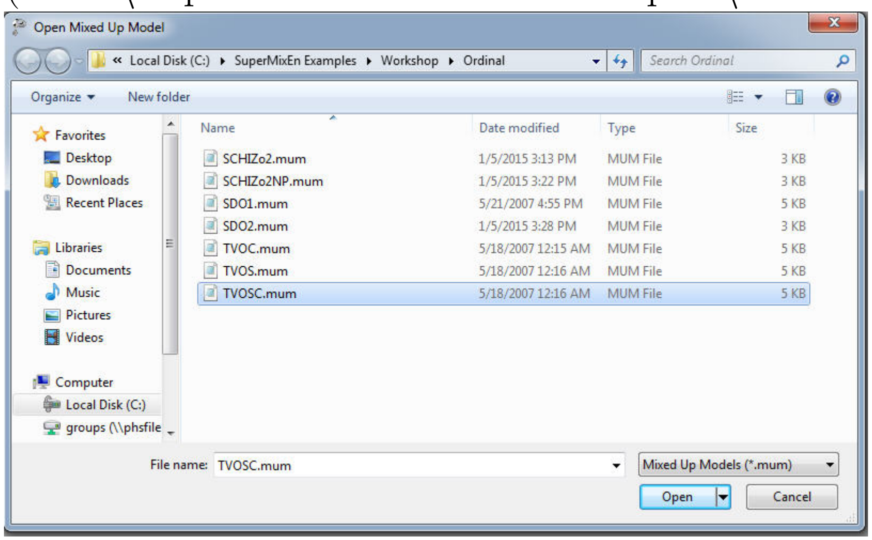
	(A)_School	(B)_Class	(C)_THKSor	(D)_THKSbi	(E)_Intrcpt	(F)_PreTHK	(G)_CC	(H)_TV	(I)_CC*TV
1	403	403101	3	1	1	2	1	0	0
2	403	403101	4	1	1	4	1	0	0
3	403	403101	3	1	1	4	1	0	0
4	403	403101	4	1	1	3	1	0	0
5	403	403101	4	1	1	3	1	0	0
6	403	403101	3	1	1	4	1	0	0
7	403	403101	2	0	1	2	1	0	0
8	403	403101	4	1	1	4	1	0	0
9	403	403101	4	1	1	5	1	0	0
10	403	403101	4	1	1	3	1	0	0
11	403	403101	3	1	1	3	1	0	0
12	403	403101	4	1	1	3	1	0	0
13	403	403101	3	1	1	1	1	0	0
14	403	403101	4	1	1	2	1	0	0
15	403	403101	2	0	1	2	1	0	0
16	403	403101	4	1	1	1	1	0	0
17	403	403101	4	1	1	4	1	0	0
18	403	403101	3	1	1	3	1	0	0
19	403	403101	3	1	1	0	1	0	0
20	403	403101	4	1	1	3	1	0	0
21	403	403102	2	0	1	0	1	0	0
22	403	403102	4	1	1	1	1	0	0
23	403	403102	3	1	1	5	1	0	0
24	404	404101	3	1	1	1	1	1	1
25	404	404101	4	1	1	2	1	1	1
26	404	404101	2	0	1	4	1	1	1
27	404	404101	3	1	1	3	1	1	1
28	404	404101	2	0	1	1	1	1	1

Under “File” click on “Open Existing Model Setup”

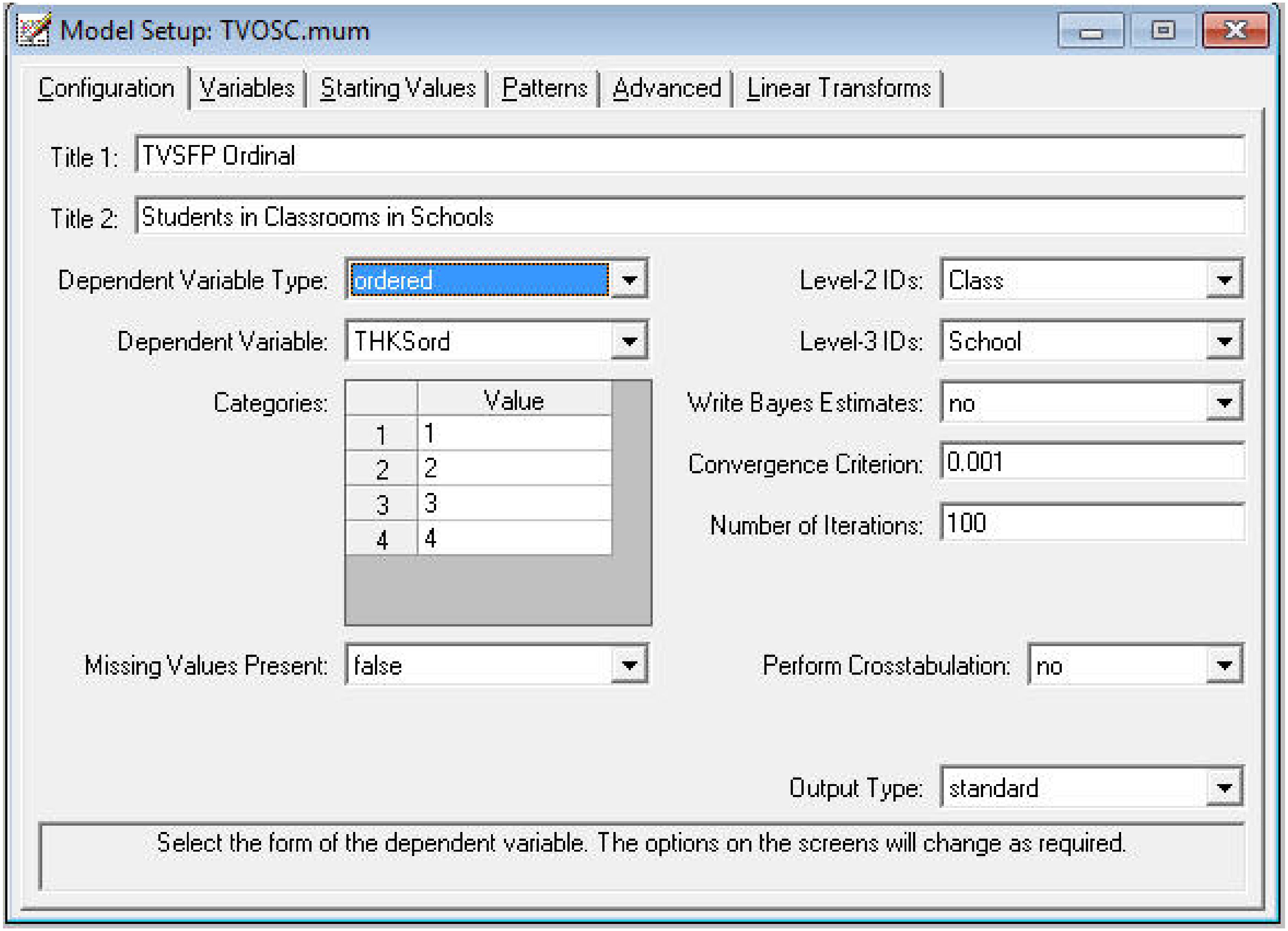


Open C:\SuperMixEn Examples\Workshop\Ordinal\tvosc.mum

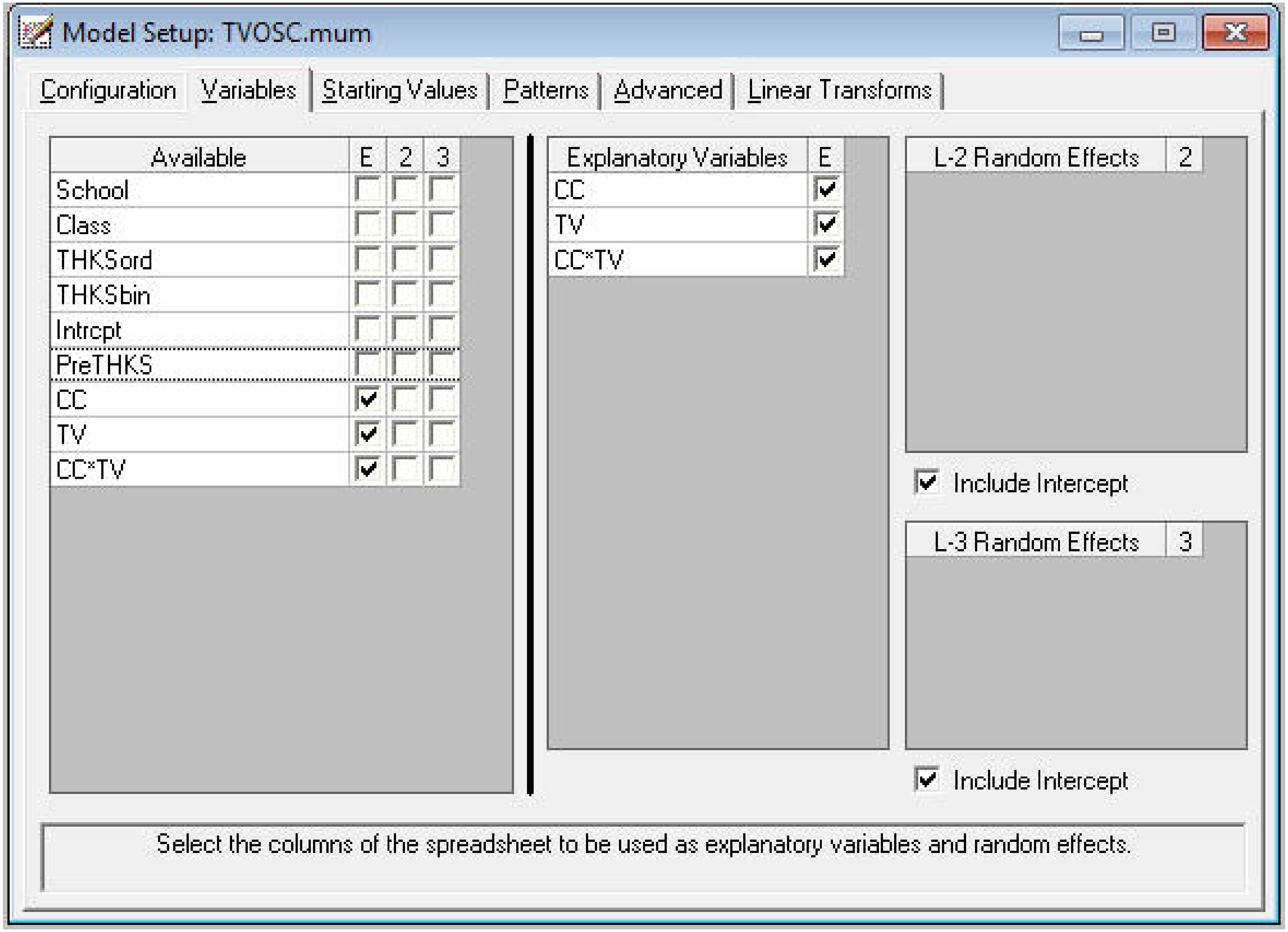
(or C:\SuperMixEn Student Examples\Workshop\Ordinal\tvosc.mum)



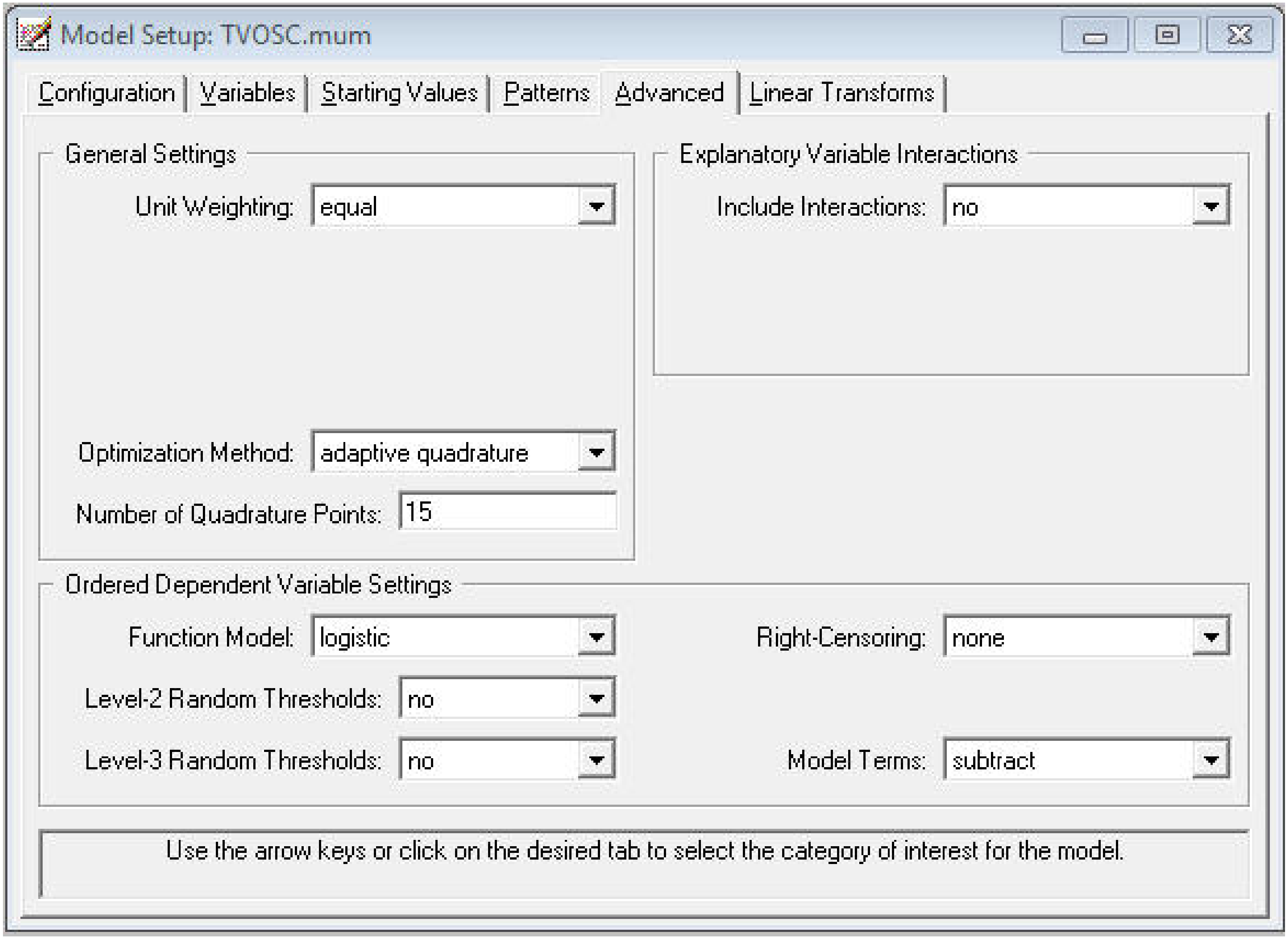
Note “Dependent Variable Type” is “ordered”



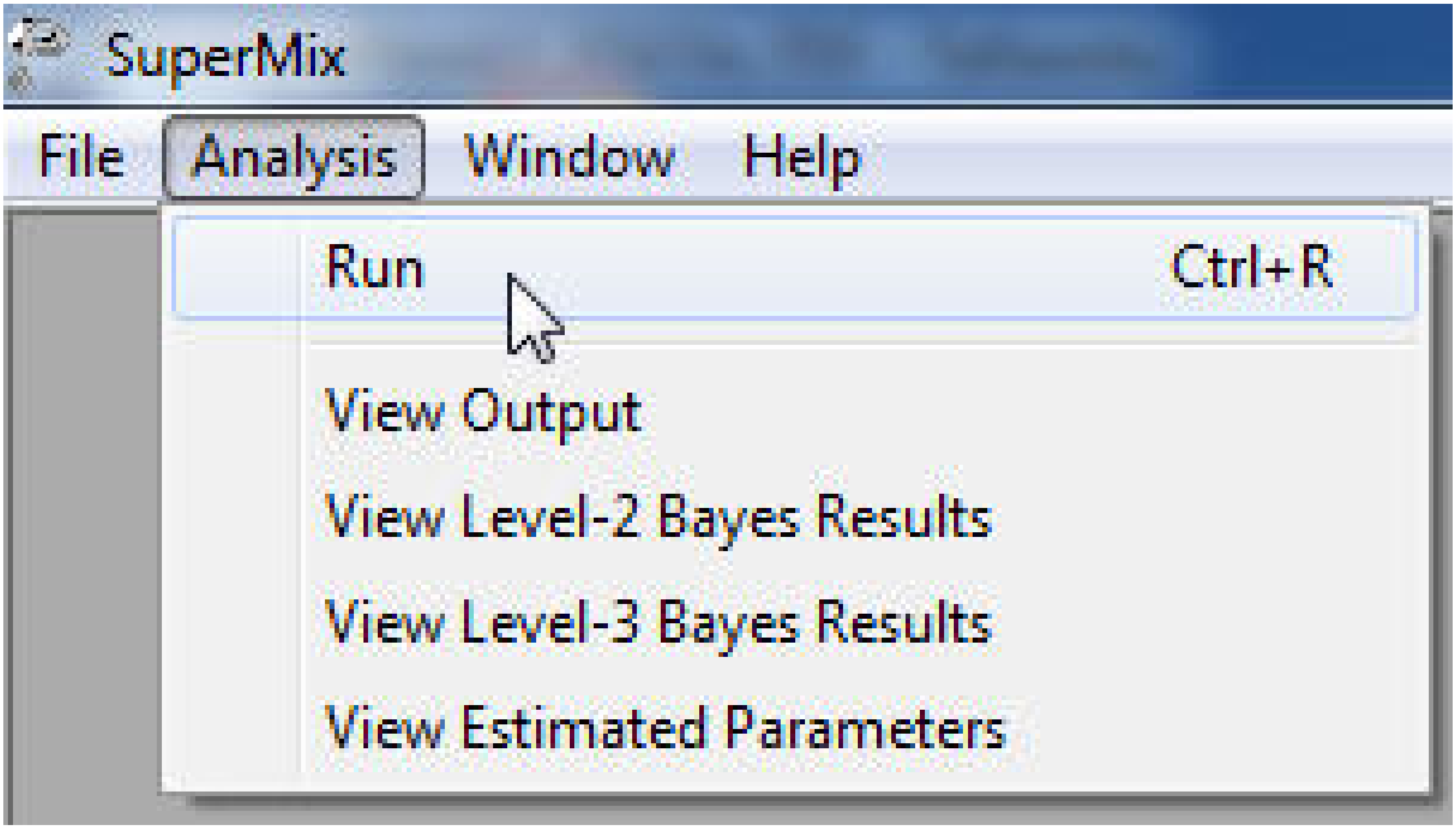
For the moment, unselect **PreTHKS** as an explanatory variable



Note "Optimization Method" is "adaptive quadrature"



Under "Analysis" click on "Run"



TVOSC.out

```
o==================================o
| TVSFP Ordinal                    |
| Students in Classrooms in Schools |
o==================================o



           Model and Data Descriptions

   Sampling Distribution                   = Multinomial
   Link Function                           = Cumulative Logit
   Number of Level-3 Units                 = 28
   Number of Level-2 Units                 = 135
   Number of Level-1 Units                 = 1600
   Number of Level-2 Units per Level-3 Unit =
     2     3     1     6     2     4     3     6     2     2     3     3
     4     4     4     2     6     5     5     7    11     7     4     8
     7     4     7    13
   Number of level-1 units for the first (level-3, level-2) unit combination =
    20     3



    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
   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
```

Save As... Close

TVOSC.out

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

Number of quadrature points =              15
Number of free parameters =                 8
Number of iterations used =                13

-2lnL (deviance statistic) =       4339.31017
Akaike Information Criterion       4355.31017
Schwarz Criterion                  4398.33224
```

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
Threshold1	-0.9254	0.1799	-5.1454	0.0000
Threshold2	0.3019	0.1782	1.6946	0.0901
Threshold3	1.4529	0.1822	7.9753	0.0000
CC	0.8232	0.2537	3.2453	0.0012
TV	0.2357	0.2485	0.9482	0.3430
CC*TV	-0.4310	0.3563	-1.2098	0.2264

Odds Ratio and 95% Odds Ratio Confidence Intervals

Parameter	Estimate	Odds Ratio	Bounds Lower	Bounds Upper
Threshold1	-0.9254	0.3964	0.2786	0.5639
Threshold2	0.3019	1.3524	0.9538	1.9176
Threshold3	1.4529	4.2754	2.9916	6.1100
CC	0.8232	2.2778	1.3855	3.7449
TV	0.2357	1.2657	0.7777	2.0601
CC*TV	-0.4310	0.6499	0.3233	1.3064

Save As... Close

TVOSC.out

Estimated level 2 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intercept/intercept	0.1612	0.0665	2.4233	0.0154

Level 2 covariance matrix

	intercept
intercept	0.161160

Level 2 correlation matrix

	intercept
intercept	1.000000

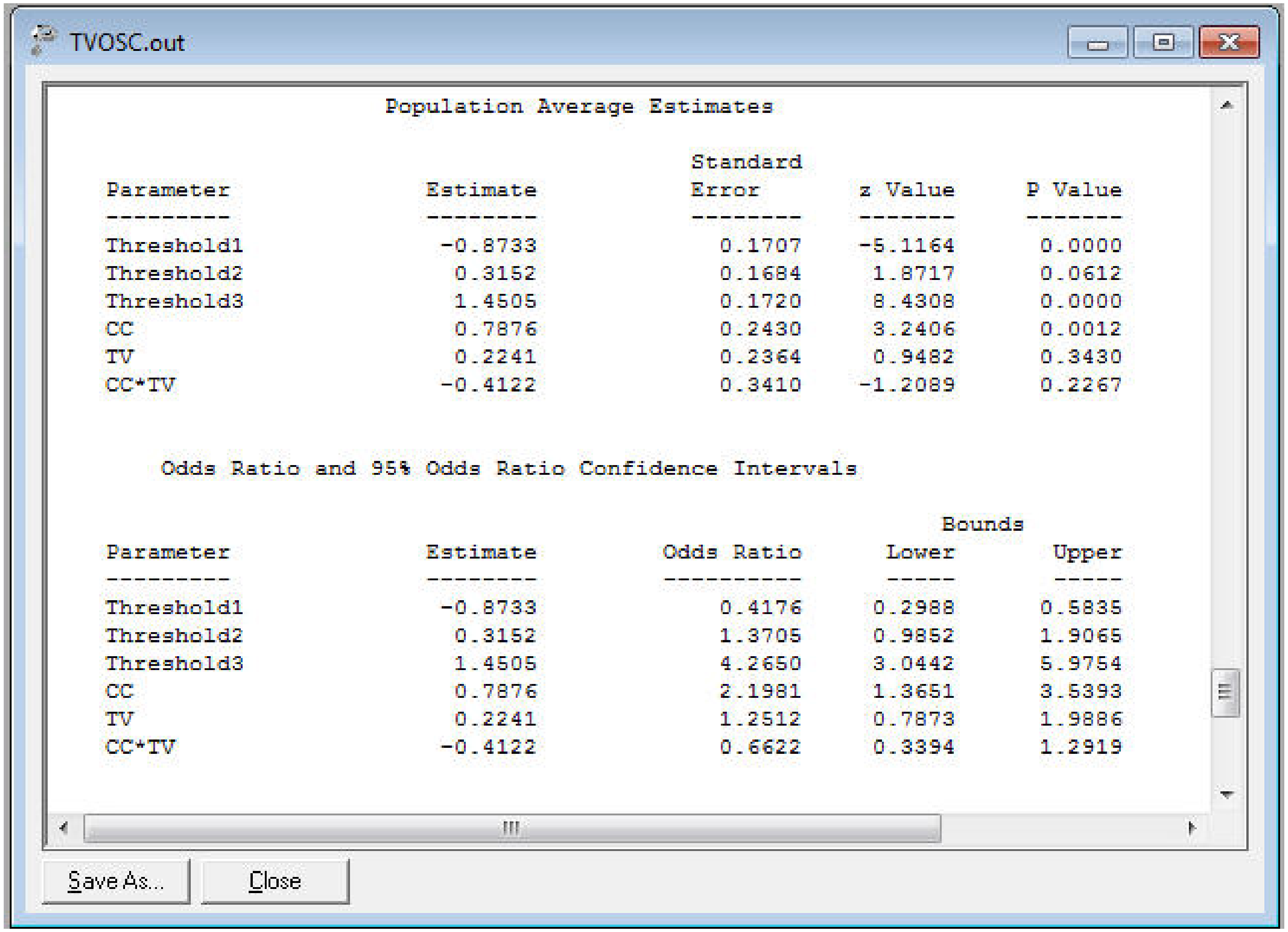
Estimated level 3 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intercept/intercept	0.1057	0.0605	1.7463	0.0808

Level 3 covariance matrix

	intercept
intercept	0.105677

Save As... Close



TVOSC.out

Population Average Estimates

Parameter	Estimate	Standard Error	z Value	P Value
Threshold1	-0.8733	0.1707	-5.1164	0.0000
Threshold2	0.3152	0.1684	1.8717	0.0612
Threshold3	1.4505	0.1720	8.4308	0.0000
CC	0.7876	0.2430	3.2406	0.0012
TV	0.2241	0.2364	0.9482	0.3430
CC*TV	-0.4122	0.3410	-1.2089	0.2267

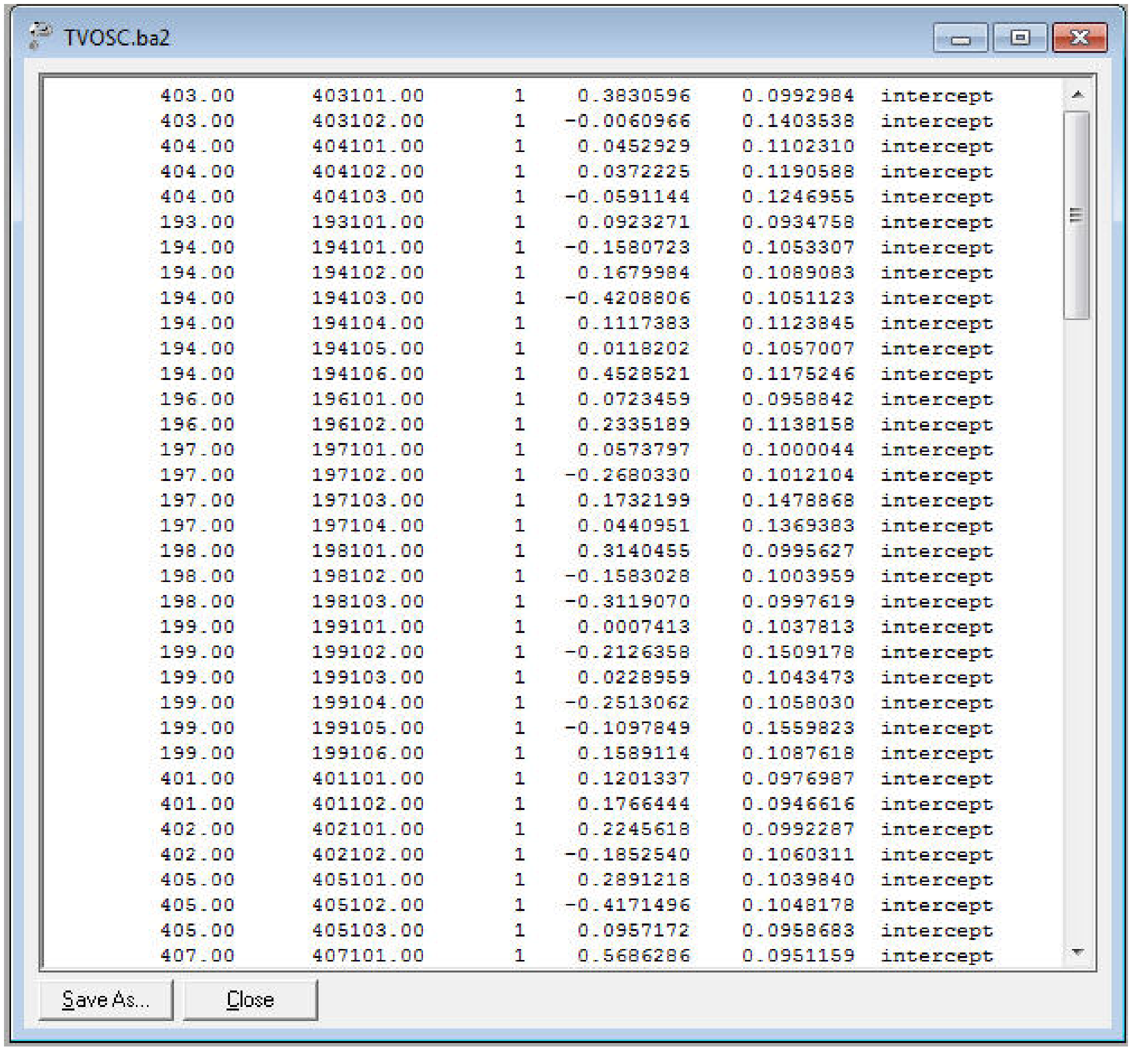
Odds Ratio and 95% Odds Ratio Confidence Intervals

Parameter	Estimate	Odds Ratio	Bounds Lower	Bounds Upper
Threshold1	-0.8733	0.4176	0.2988	0.5835
Threshold2	0.3152	1.3705	0.9852	1.9065
Threshold3	1.4505	4.2650	3.0442	5.9754
CC	0.7876	2.1981	1.3651	3.5393
TV	0.2241	1.2512	0.7873	1.9886
CC*TV	-0.4122	0.6622	0.3394	1.2919

Save As... Close

Empirical Bayes Estimates of Random Effects

Select "Analysis" > "View Level-2 Bayes Results"



403.00	403101.00	1	0.3830596	0.0992984	intercept
403.00	403102.00	1	-0.0060966	0.1403538	intercept
404.00	404101.00	1	0.0452929	0.1102310	intercept
404.00	404102.00	1	0.0372225	0.1190588	intercept
404.00	404103.00	1	-0.0591144	0.1246955	intercept
193.00	193101.00	1	0.0923271	0.0934758	intercept
194.00	194101.00	1	-0.1580723	0.1053307	intercept
194.00	194102.00	1	0.1679984	0.1089083	intercept
194.00	194103.00	1	-0.4208806	0.1051123	intercept
194.00	194104.00	1	0.1117383	0.1123845	intercept
194.00	194105.00	1	0.0118202	0.1057007	intercept
194.00	194106.00	1	0.4528521	0.1175246	intercept
196.00	196101.00	1	0.0723459	0.0958842	intercept
196.00	196102.00	1	0.2335189	0.1138158	intercept
197.00	197101.00	1	0.0573797	0.1000044	intercept
197.00	197102.00	1	-0.2680330	0.1012104	intercept
197.00	197103.00	1	0.1732199	0.1478868	intercept
197.00	197104.00	1	0.0440951	0.1369383	intercept
198.00	198101.00	1	0.3140455	0.0995627	intercept
198.00	198102.00	1	-0.1583028	0.1003959	intercept
198.00	198103.00	1	-0.3119070	0.0997619	intercept
199.00	199101.00	1	0.0007413	0.1037813	intercept
199.00	199102.00	1	-0.2126358	0.1509178	intercept
199.00	199103.00	1	0.0228959	0.1043473	intercept
199.00	199104.00	1	-0.2513062	0.1058030	intercept
199.00	199105.00	1	-0.1097849	0.1559823	intercept
199.00	199106.00	1	0.1589114	0.1087618	intercept
401.00	401101.00	1	0.1201337	0.0976987	intercept
401.00	401102.00	1	0.1766444	0.0946616	intercept
402.00	402101.00	1	0.2245618	0.0992287	intercept
402.00	402102.00	1	-0.1852540	0.1060311	intercept
405.00	405101.00	1	0.2891218	0.1039840	intercept
405.00	405102.00	1	-0.4171496	0.1048178	intercept
405.00	405103.00	1	0.0957172	0.0958683	intercept
407.00	407101.00	1	0.5686286	0.0951159	intercept

School ID, Class ID, random effect number, estimate, variance, name

Empirical Bayes Estimates of Random Effects

Select “Analysis” > “View Level-3 Bayes Results”

TVOSC.ba3

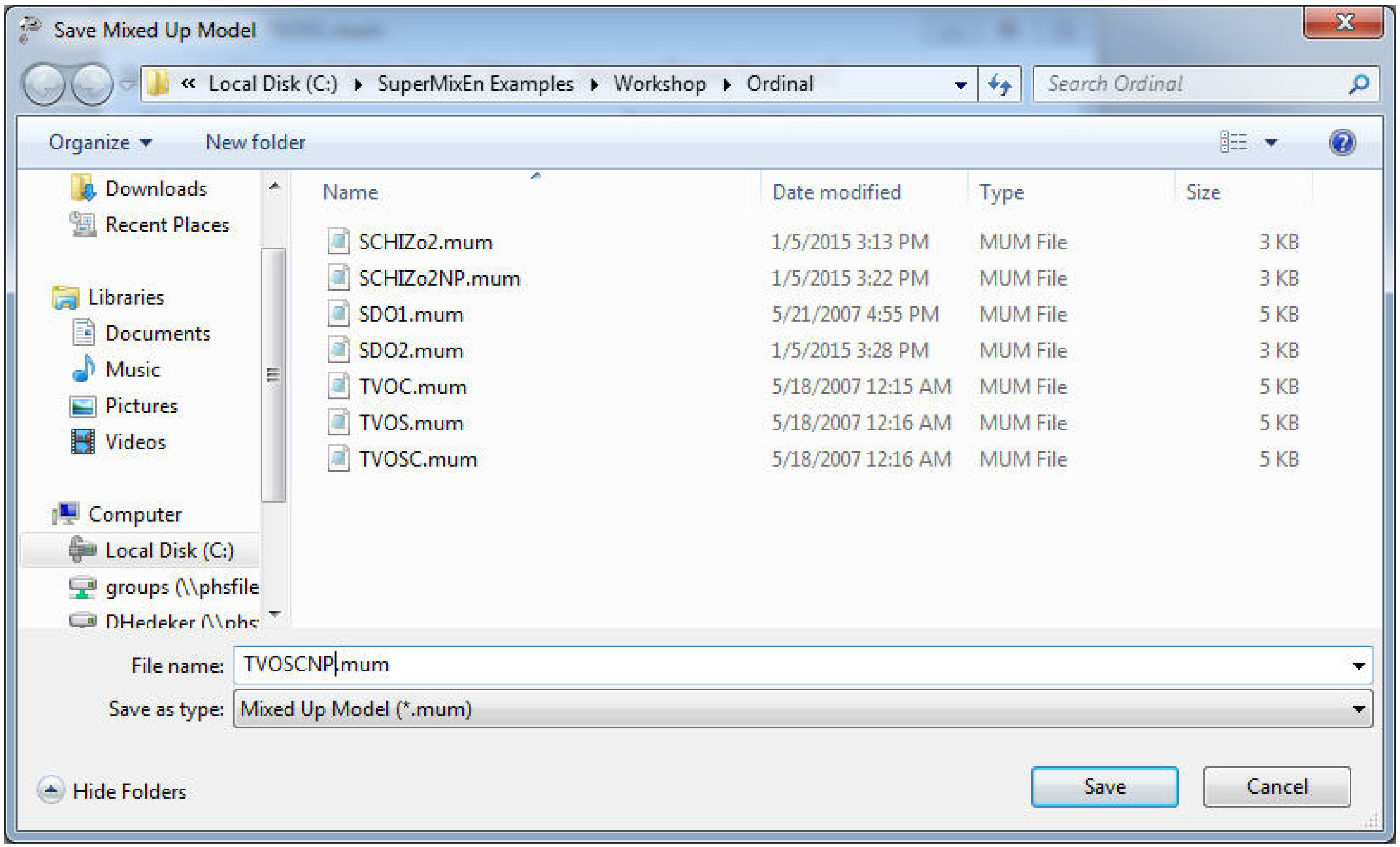
403.00	1	0.2471881	0.0745439	intercept
404.00	1	0.0153458	0.0658483	intercept
193.00	1	0.0605439	0.0765745	intercept
194.00	1	0.1084941	0.0438868	intercept
196.00	1	0.2005668	0.0675637	intercept
197.00	1	0.0043688	0.0602617	intercept
198.00	1	-0.1024008	0.0560158	intercept
199.00	1	-0.2565086	0.0504731	intercept
401.00	1	0.1946084	0.0636548	intercept
402.00	1	0.0257770	0.0666819	intercept
405.00	1	-0.0211861	0.0567750	intercept
407.00	1	0.3132113	0.0531603	intercept
408.00	1	-0.1369094	0.0684003	intercept
409.00	1	0.1033115	0.0462808	intercept
410.00	1	-0.4393079	0.0606901	intercept
411.00	1	0.3263763	0.0755432	intercept
412.00	1	-0.1111421	0.0569602	intercept
414.00	1	-0.0262411	0.0560123	intercept
415.00	1	0.4152628	0.0470290	intercept
505.00	1	-0.1847633	0.0410649	intercept
506.00	1	-0.4638935	0.0413189	intercept
507.00	1	-0.2733055	0.0434032	intercept
508.00	1	0.3752750	0.0455138	intercept
509.00	1	-0.2404592	0.0338281	intercept
510.00	1	0.0669854	0.0372791	intercept
513.00	1	-0.2471750	0.0593768	intercept
514.00	1	0.1012731	0.0371836	intercept
515.00	1	-0.0552958	0.0272297	intercept

Save As... Close

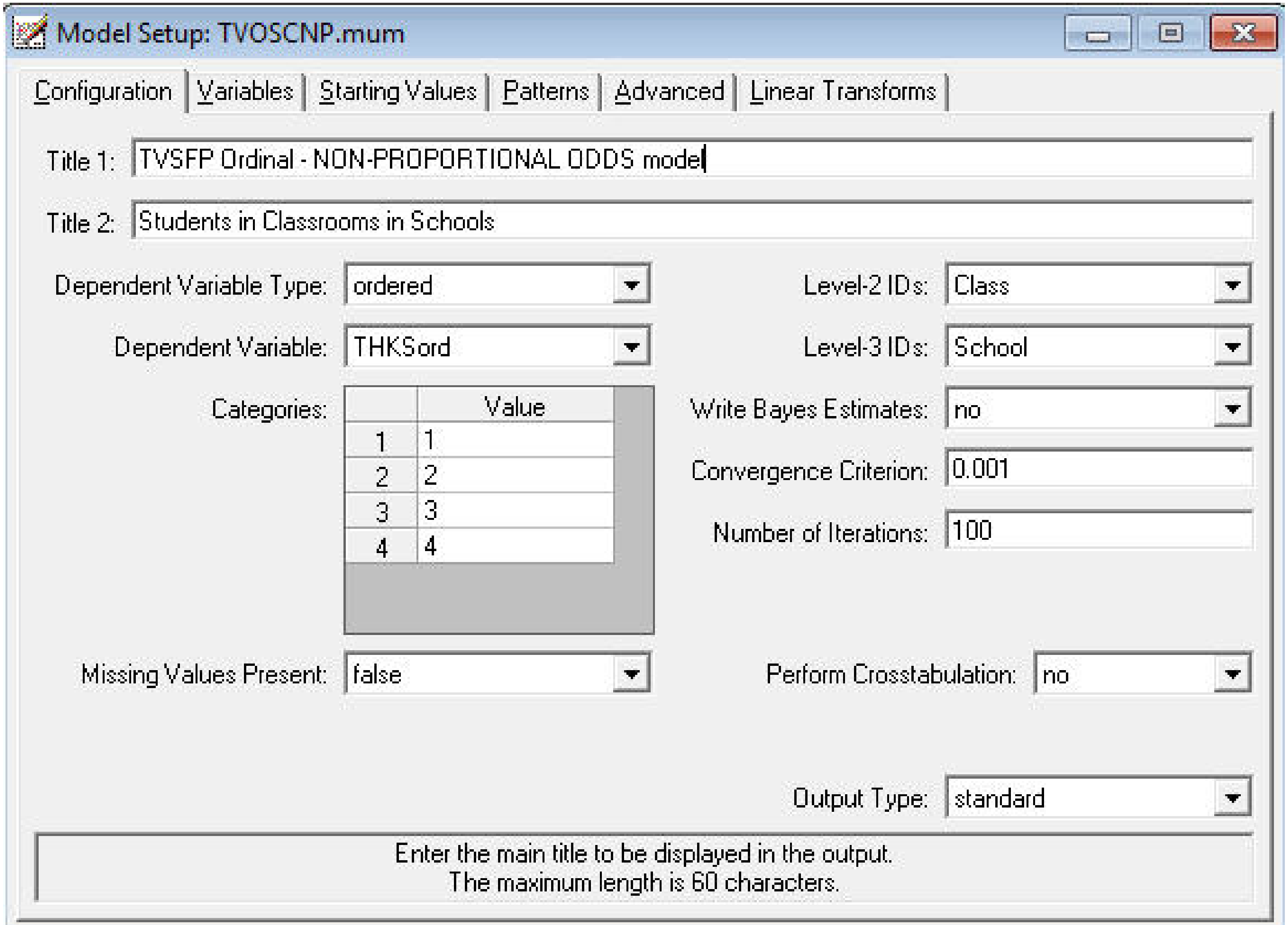
School ID, random effect number, estimate, variance, name

To estimate the non-proportional odds model, we'll make some modifications to the mum file. First, let's save the mum file to a new name so that our previous results are not written over.

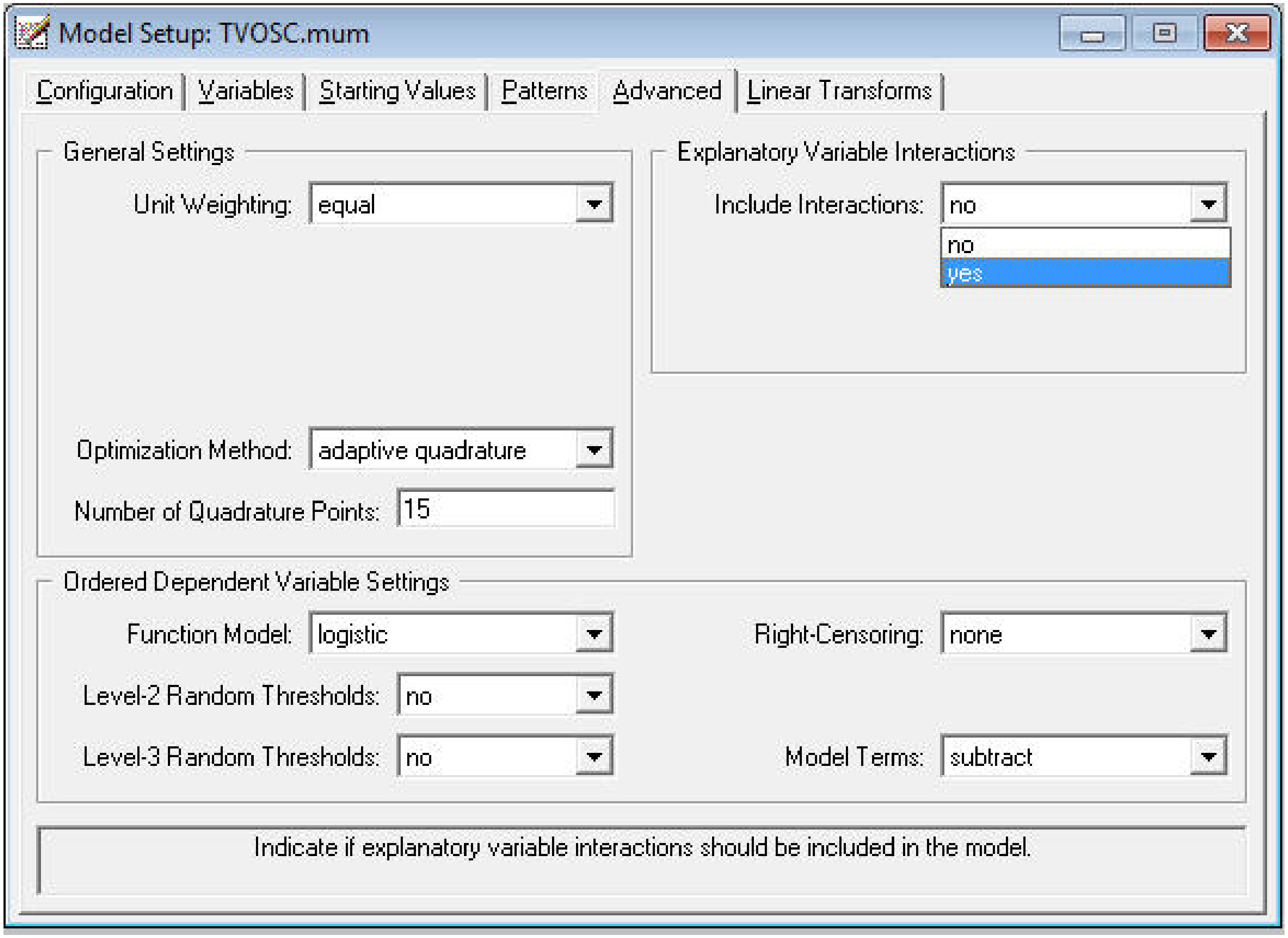
Under "File" > "Save as" type in "TVOSCNP.mum"



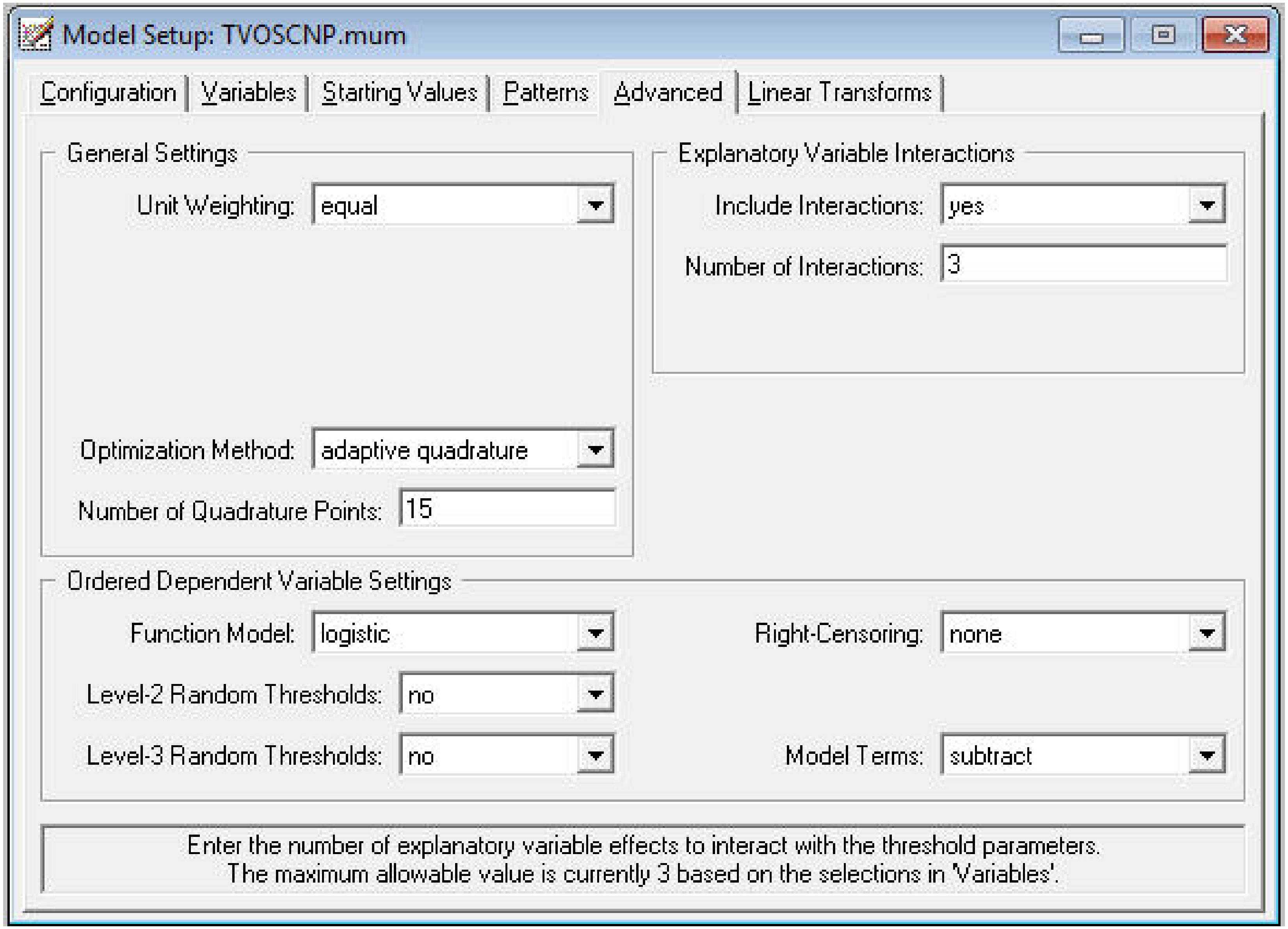
On the Configuration Card, modify the title



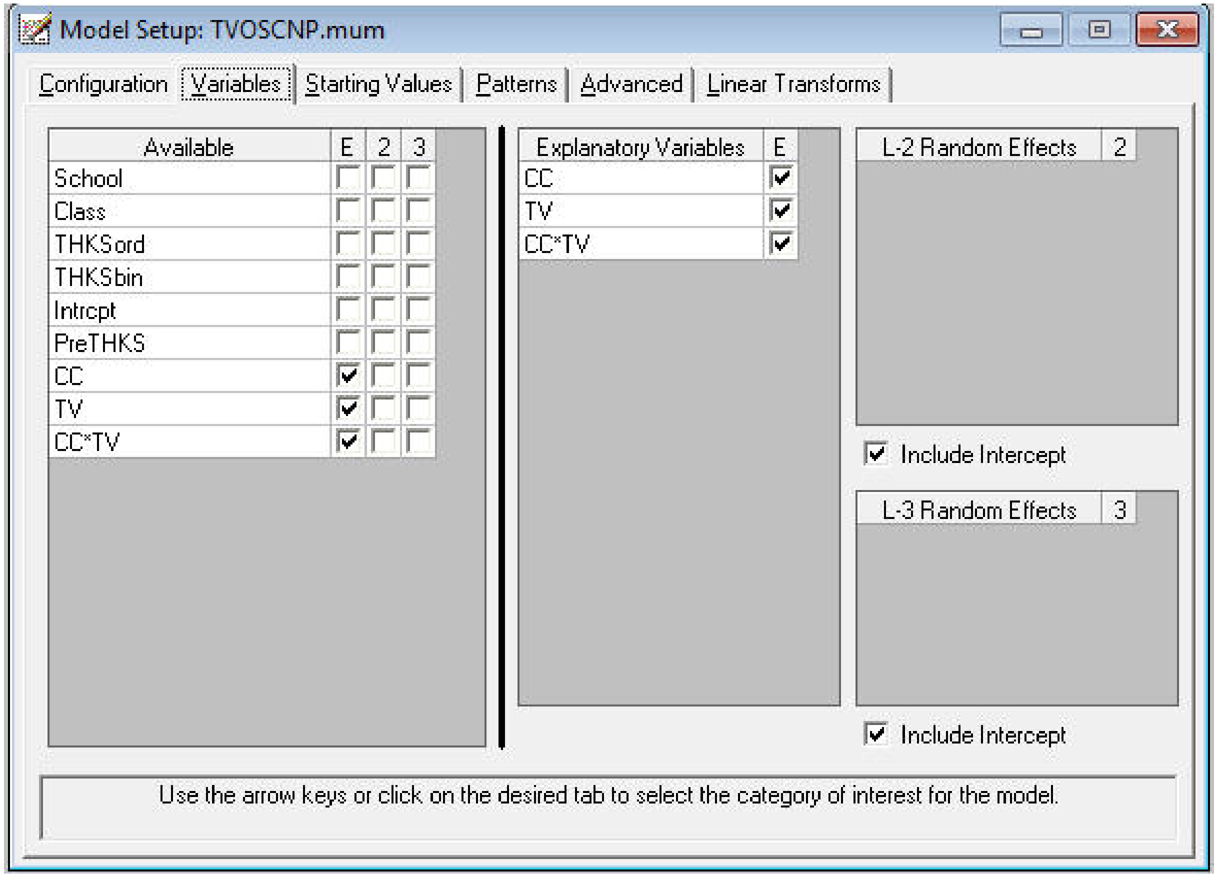
On the Advanced Card, select "yes" for Explanatory Variable Interactions



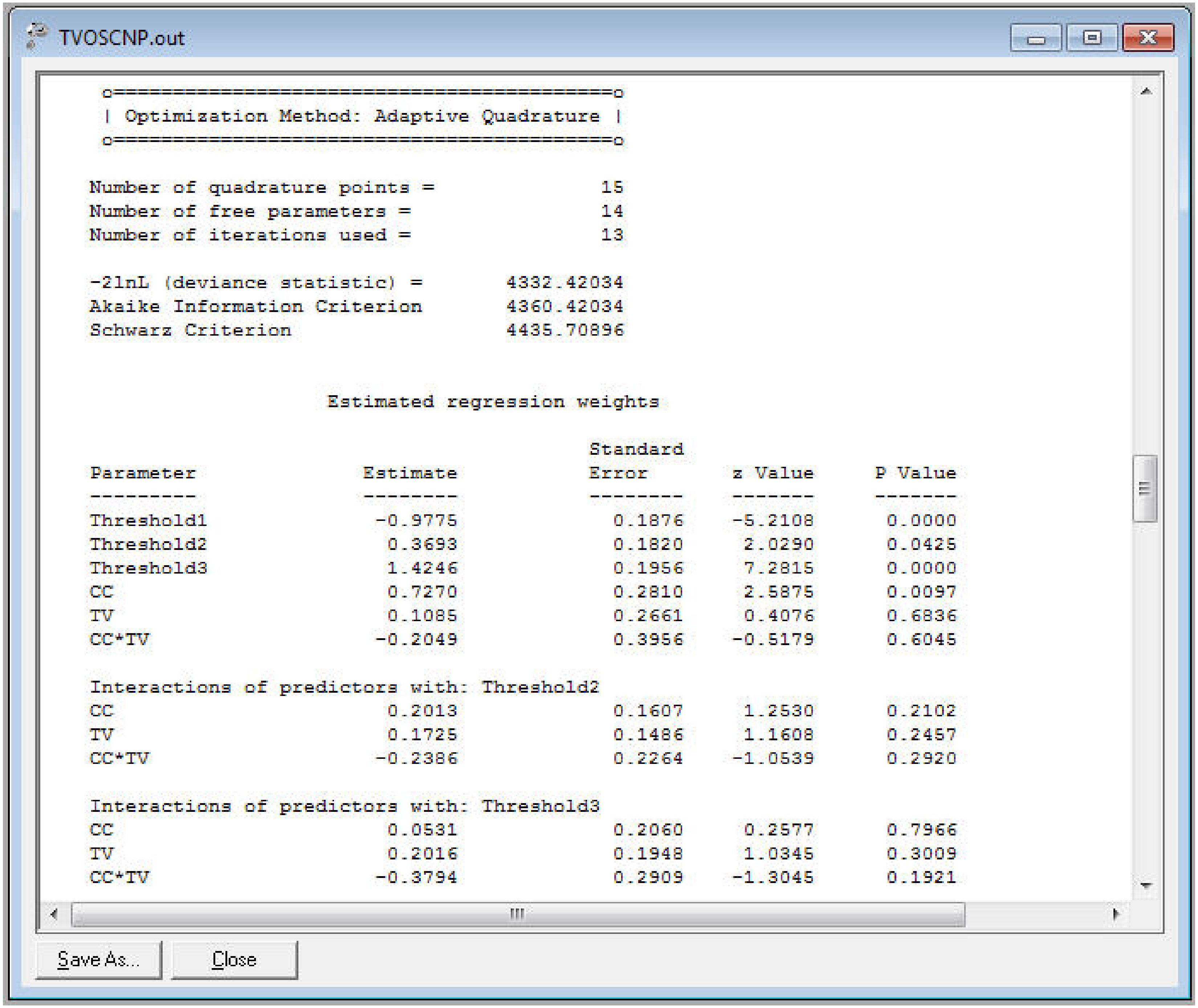
Type in “3” for the Number of Interactions (all explanatory variables will have non-proportional effects)



If fewer than 3 is specified, then which explanatory variables have non-proportional effects depends on the order that the variables were selected on the Variables card



If 1 was specified for the number of interactions, then CC would have non-proportional effects (but TV and CC*TV would have proportional effects). Can unselect and reselect variables if different ordering is desired.


TVOSCNP.out

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

Number of quadrature points =              15
Number of free parameters =                14
Number of iterations used =                13

-2lnL (deviance statistic) =       4332.42034
Akaike Information Criterion       4360.42034
Schwarz Criterion                  4435.70896
```

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
Threshold1	-0.9775	0.1876	-5.2108	0.0000
Threshold2	0.3693	0.1820	2.0290	0.0425
Threshold3	1.4246	0.1956	7.2815	0.0000
CC	0.7270	0.2810	2.5875	0.0097
TV	0.1085	0.2661	0.4076	0.6836
CC*TV	-0.2049	0.3956	-0.5179	0.6045
Interactions of predictors with: Threshold2				
CC	0.2013	0.1607	1.2530	0.2102
TV	0.1725	0.1486	1.1608	0.2457
CC*TV	-0.2386	0.2264	-1.0539	0.2920
Interactions of predictors with: Threshold3				
CC	0.0531	0.2060	0.2577	0.7966
TV	0.2016	0.1948	1.0345	0.3009
CC*TV	-0.3794	0.2909	-1.3045	0.1921

Save As... Close

- First set of estimates (under Estimated regression weights) are for the effects on the first cumulative logit
  `CC = 0.7270, TV = 0.1085, CC*TV = -0.2049`

- Next estimates (under Interactions of predictors with: Threshold2) indicate how the effects are DIFFERENT on the second cumulative logit, relative to the first
  `CC = 0.2013, TV = 0.1725, CC*TV = -0.2386`

- Next estimates under Interactions of predictors with: Threshold3) indicate how the effects are DIFFERENT on the third cumulative logit, relative to the first
  `CC = 0.0531, TV = 0.2016, CC*TV = -0.3794`

$\Rightarrow$ note that none of these six interactions are significant; in agreement with overall LR test ($\chi^2_6 = 4339.31 - 4332.42 = 6.89$)
$\Rightarrow$ proportional odds assumption not rejected

Effects on the cumulative logits

- First cumulative logit:
  CC = 0.7270, TV = 0.1085, CC*TV = -0.2049

- Second cumulative logit:
  CC = 0.7270 + 0.2013 = 0.9283
  TV = 0.1085 + 0.1725 = 0.2810
  CC*TV = -0.2049 + (-0.2386) = -0.4435

- Third cumulative logit:
  CC = 0.7270 + 0.0531 = 0.7801
  TV = 0.1085 + 0.2016 = 0.3101
  CC*TV = -0.2049 + (-0.3794) = -0.5843

Linear Transforms

Fixed part of model:

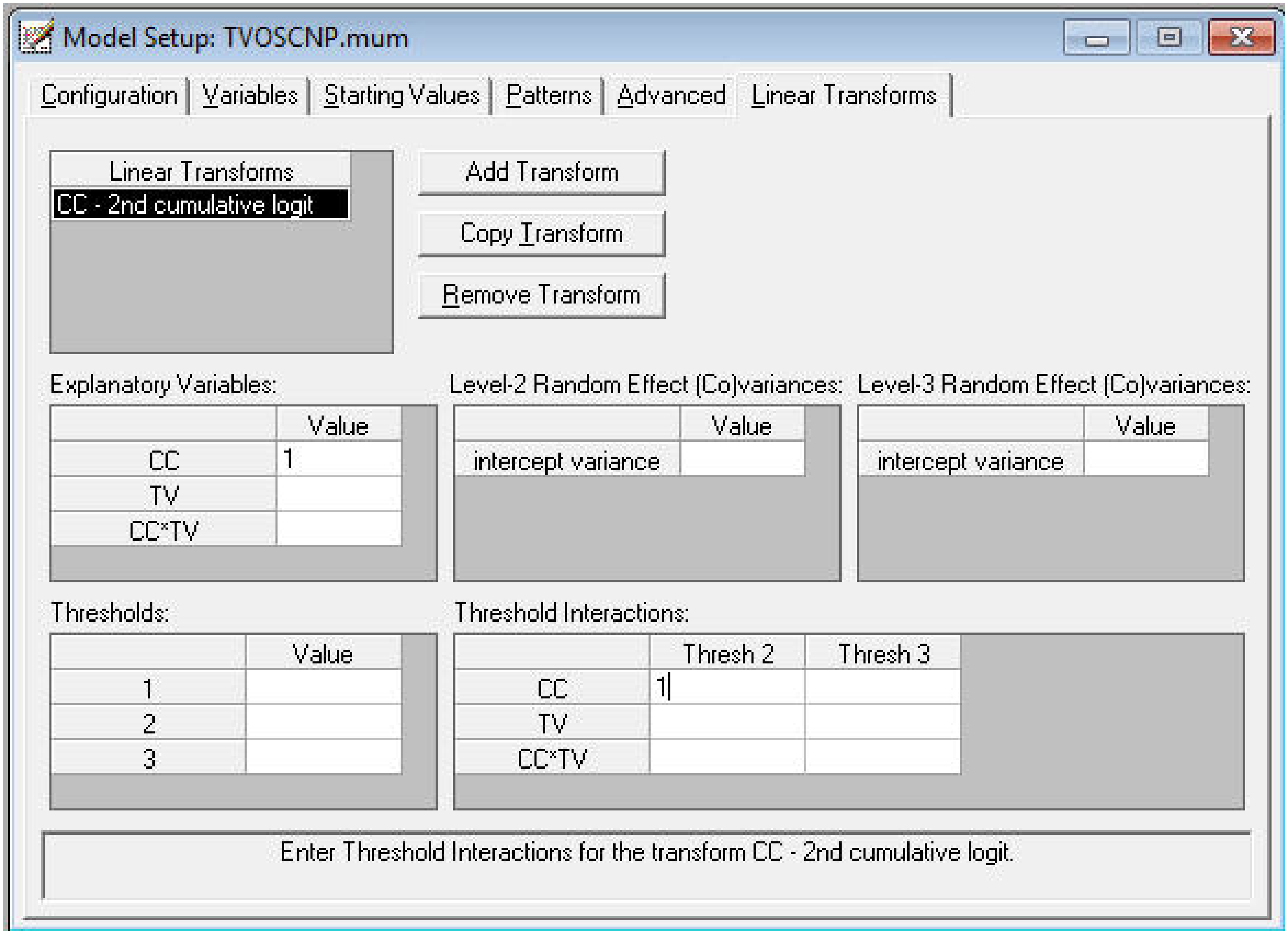
$$\lambda_c = \hat{\gamma}_{0c} - \Big[\hat{\beta}_1\texttt{CC} + \hat{\beta}_2\texttt{TV} + \hat{\beta}_3\texttt{CC*TV} \\ + \; \hat{\gamma}_{1c}\texttt{CC} + \hat{\gamma}_{2c}\texttt{TV} + \hat{\gamma}_{3c}\texttt{CC*TV}\Big]$$

	cumulative logit		
variable	1 vs 2,3,4	1,2 vs 3,4	1,2,3, vs 4
`CC`	$\hat{\beta}_1$	$\hat{\beta}_1 + \hat{\gamma}_{12}$	$\hat{\beta}_1 + \hat{\gamma}_{13}$
`TV`	$\hat{\beta}_2$	$\hat{\beta}_2 + \hat{\gamma}_{22}$	$\hat{\beta}_2 + \hat{\gamma}_{23}$
`CC*TV`	$\hat{\beta}_3$	$\hat{\beta}_3 + \hat{\gamma}_{32}$	$\hat{\beta}_3 + \hat{\gamma}_{33}$

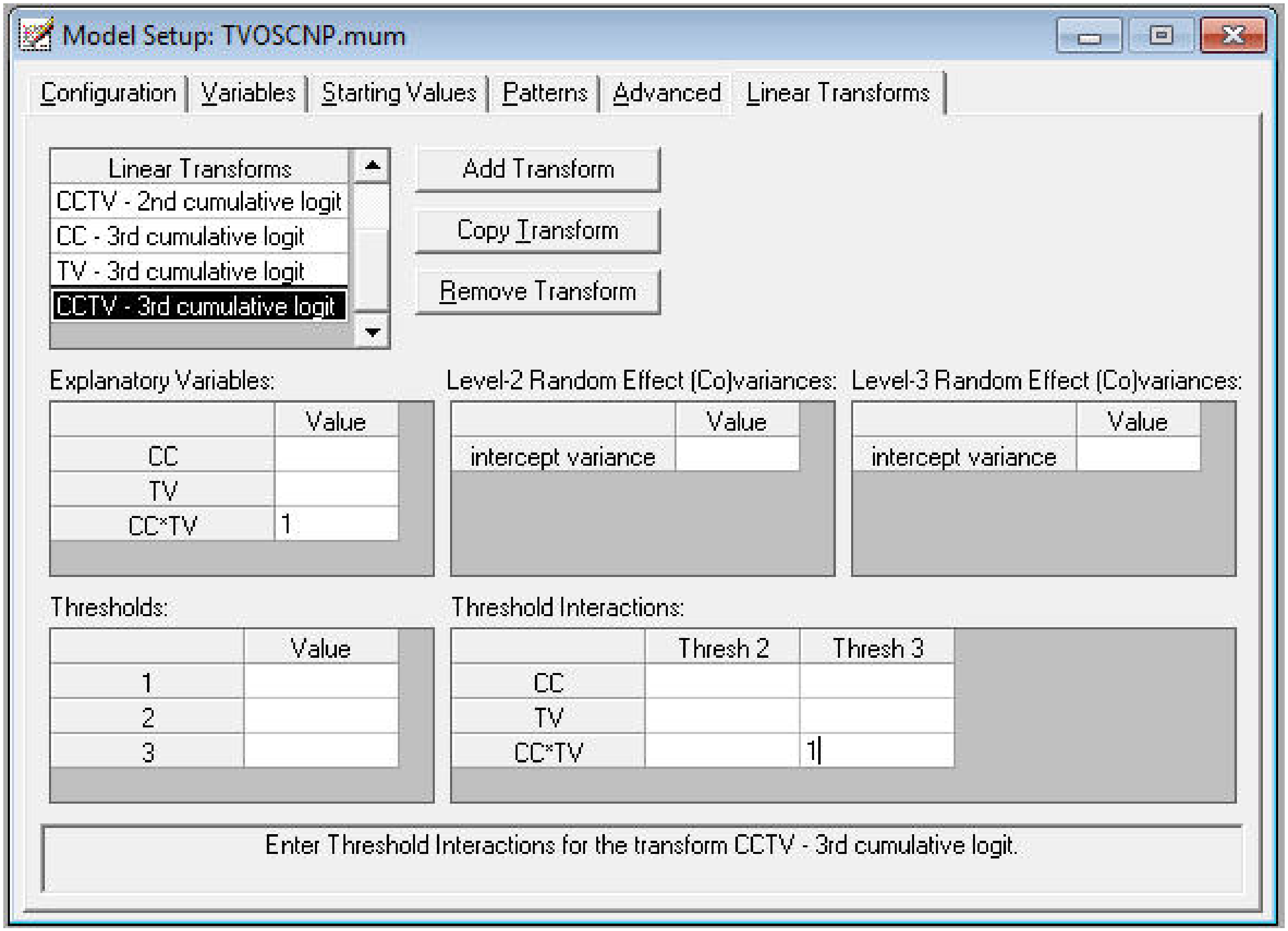
$H_0 : \beta_1 + \gamma_{12} = 0$; `CC` effect is 0 on the 2nd cumulative logit

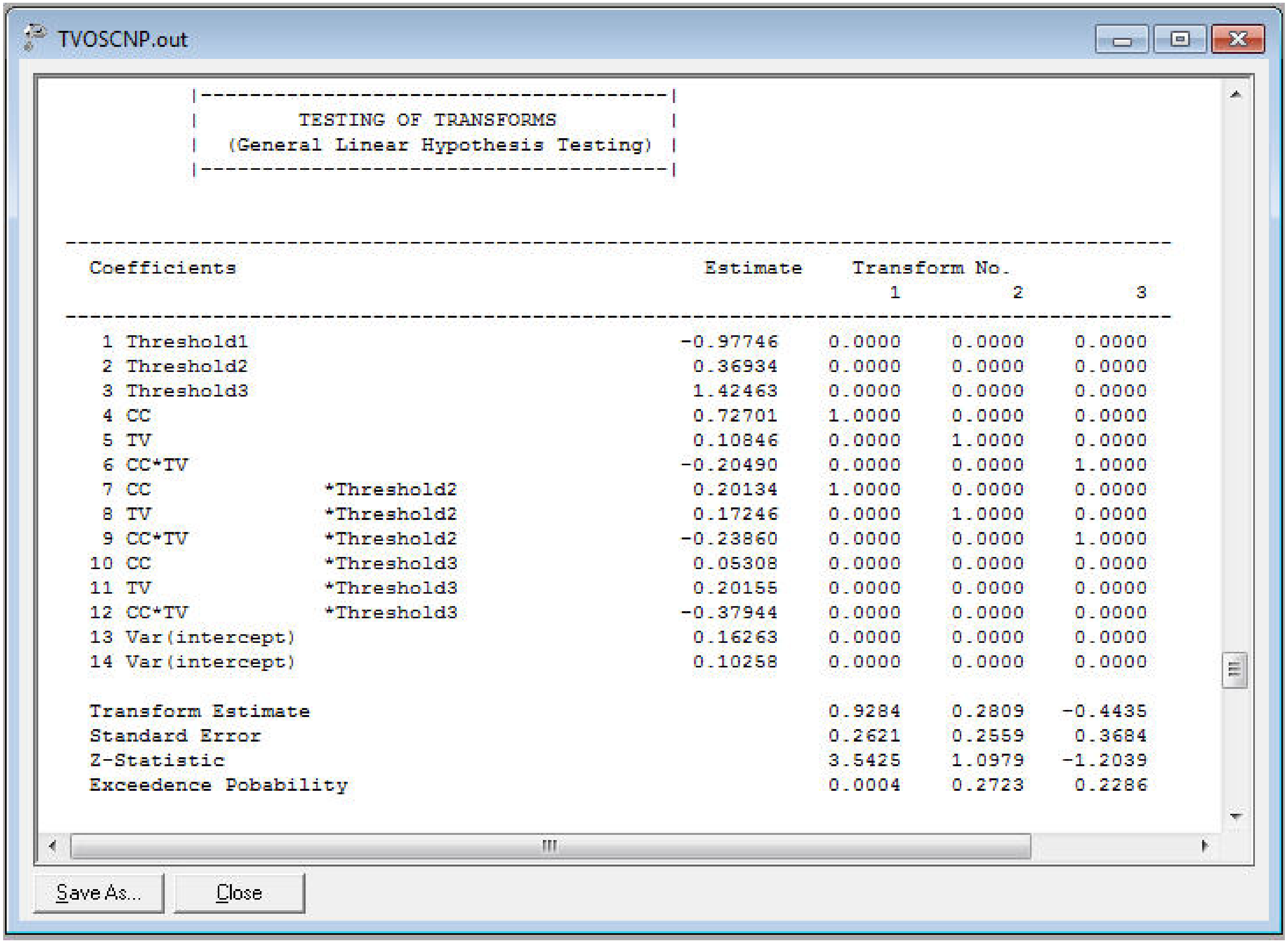
$$z = \frac{\hat{\beta}_1 + \hat{\gamma}_{12}}{SE(\hat{\beta}_1 + \hat{\gamma}_{12})}$$

Linear Transforms: for estimate, std error, p-value $\beta_1 + \gamma_{12}$



Linear Transforms: for estimate, std error, p-value $\beta_3 + \gamma_{33}$





TVOSCNP.out

TESTING OF TRANSFORMS
(General Linear Hypothesis Testing)

Coefficients		Estimate	Transform No. 1	2	3
1 Threshold1		-0.97746	0.0000	0.0000	0.0000
2 Threshold2		0.36934	0.0000	0.0000	0.0000
3 Threshold3		1.42463	0.0000	0.0000	0.0000
4 CC		0.72701	1.0000	0.0000	0.0000
5 TV		0.10846	0.0000	1.0000	0.0000
6 CC*TV		-0.20490	0.0000	0.0000	1.0000
7 CC	*Threshold2	0.20134	1.0000	0.0000	0.0000
8 TV	*Threshold2	0.17246	0.0000	1.0000	0.0000
9 CC*TV	*Threshold2	-0.23860	0.0000	0.0000	1.0000
10 CC	*Threshold3	0.05308	0.0000	0.0000	0.0000
11 TV	*Threshold3	0.20155	0.0000	0.0000	0.0000
12 CC*TV	*Threshold3	-0.37944	0.0000	0.0000	0.0000
13 Var(intercept)		0.16263	0.0000	0.0000	0.0000
14 Var(intercept)		0.10258	0.0000	0.0000	0.0000
Transform Estimate			0.9284	0.2809	-0.4435
Standard Error			0.2621	0.2559	0.3684
Z-Statistic			3.5425	1.0979	-1.2039
Exceedence Pobability			0.0004	0.2723	0.2286

Save As... Close

TVOSCNP.out

LINEAR TRANSFORMS (continued)

Coefficients		Estimate	Transform No. 4	5	6
1 Threshold1		-0.97746	0.0000	0.0000	0.0000
2 Threshold2		0.36934	0.0000	0.0000	0.0000
3 Threshold3		1.42463	0.0000	0.0000	0.0000
4 CC		0.72701	1.0000	0.0000	0.0000
5 TV		0.10846	0.0000	1.0000	0.0000
6 CC*TV		-0.20490	0.0000	0.0000	1.0000
7 CC	*Threshold2	0.20134	0.0000	0.0000	0.0000
8 TV	*Threshold2	0.17246	0.0000	0.0000	0.0000
9 CC*TV	*Threshold2	-0.23860	0.0000	0.0000	0.0000
10 CC	*Threshold3	0.05308	1.0000	0.0000	0.0000
11 TV	*Threshold3	0.20155	0.0000	1.0000	0.0000
12 CC*TV	*Threshold3	-0.37944	0.0000	0.0000	1.0000
13 Var(intercept)		0.16263	0.0000	0.0000	0.0000
14 Var(intercept)		0.10258	0.0000	0.0000	0.0000
Transform Estimate			0.7801	0.3100	-0.5843
Standard Error			0.2719	0.2705	0.3813
Z-Statistic			2.8692	1.1460	-1.5323
Exceedence Pobability			0.0041	0.2518	0.1254

Save As... Close

Summary: models for clustered ordinal data as developed as models for continuous data

- Proportional odds models
  - covariate effects are equal across $C - 1$ cumulative logits
- Non and partial proportional odds models
  - covariate effects are all unequal across $C - 1$ cumulative logits (non proportional odds); or some covariate effects are unequal and some are equal across $C - 1$ cumulative logits (partial proportional odds)
- Scaling models (Hedeker, Berbaum, & Mermelstein, 2006; Hedeker, Demirtas, & Mermelstein, 2009; not yet in Supermix)
  - Dispersion across the ordinal categories can depend on covariates; examination of extreme response styles
- Can be recast as ordinal probit mixed models by selecting probit link in Supermix