Mixed Models for Clustered Binary Outcomes

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

hedeker@uchicago.edu

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

This work was supported by National Institute of Mental Health Contract N44MH32056.

Multilevel models for categorical outcomes

- dichotomous outcomes
  - mixed-effects logistic regression

- ordinal outcomes
  - mixed-effects ordinal logistic regression
    - proportional odds model
    - partial or non-proportional odds model

- discrete or grouped time-to-event data
  - mixed-effects dichotomous or ordinal regression
  - replace logistic link with complementary log-log link to yield proportional (and non-proportional) hazards models

Logistic Regression Model

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

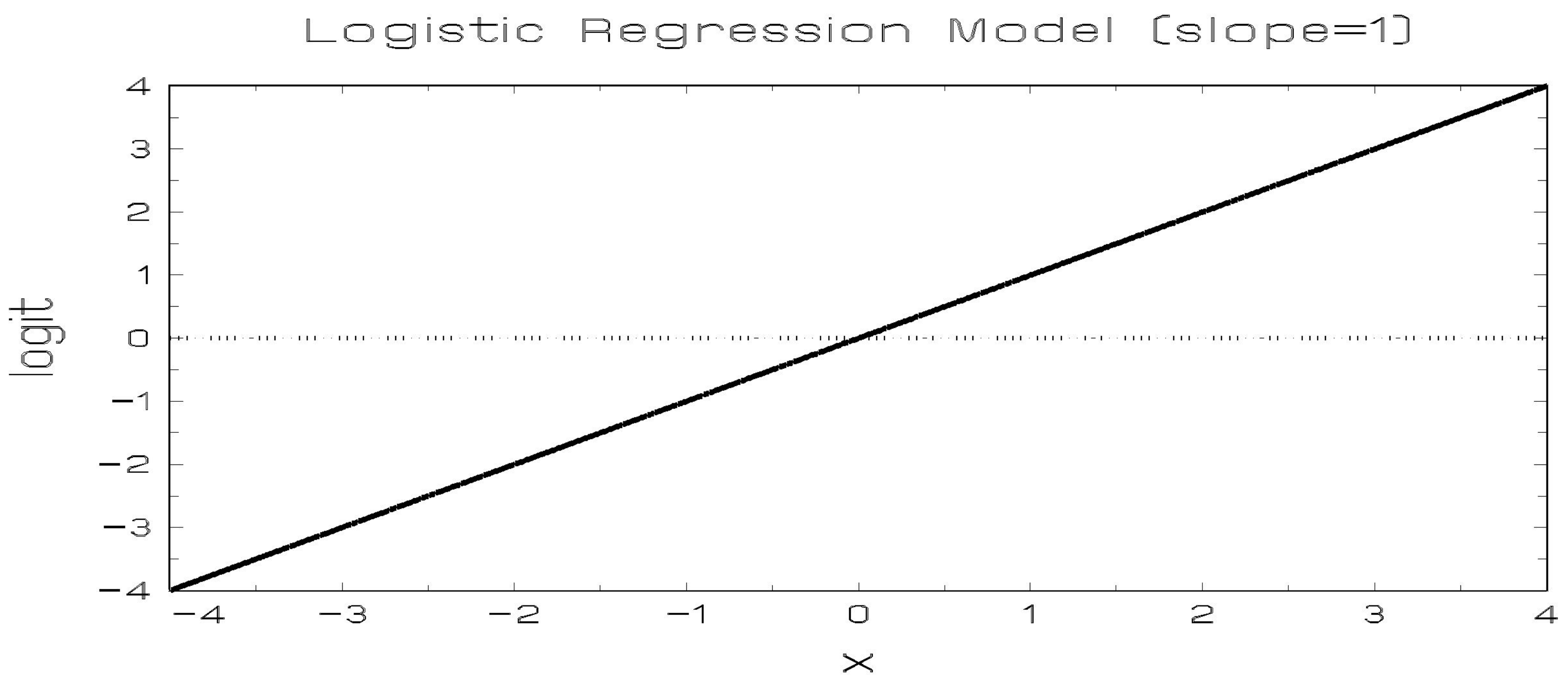
- Dichotomous outcome ($Y = 0$ absence, $Y = 1$ presence).
- Function that links probabilities to regressors is the logit (or log odds) function $\log\left[P/(1 - P\right]$. Logit is called the link function.

The model can be written in terms of probabilities:

$$P(Y_i = 1) = \frac{1}{1 + \exp(-\boldsymbol{x}_i'\boldsymbol{\beta})}$$

- Model is a linear model for the logits, not for the probabilities. Logits can take on any values between negative and positive infinity, probabilities can only take on values between 0 and 1.


Logistic Regression Model (slope=1)
probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−4
−3
−2
−1
0
1
2
3
4
X


Logistic Regression Model (slope=1)
logit
4
3
2
1
0
−1
−2
−3
−4
−4
−3
−2
−1
0
1
2
3
4
X

The model can also be written in terms of the odds:

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

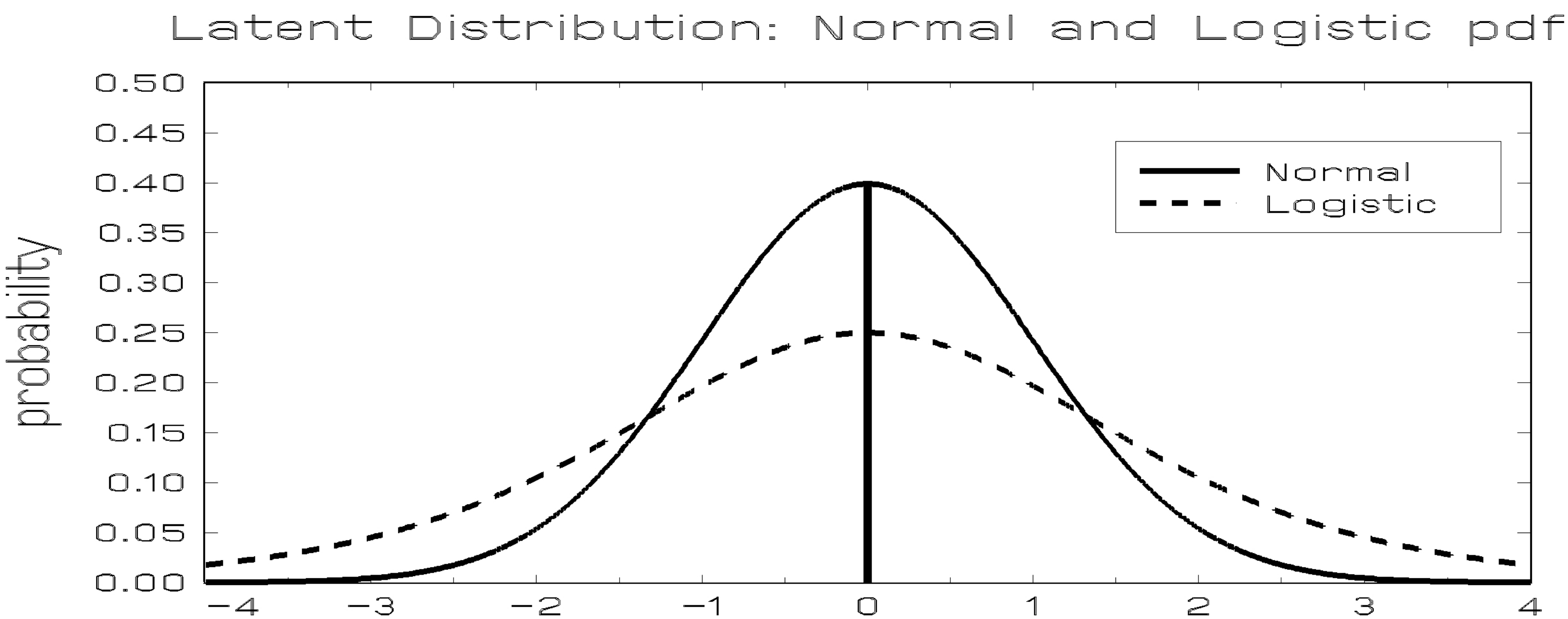
$\exp \beta$ = change in odds for Y per unit change of x

- $\beta = 0$ yields no effect on the odds
- $\beta > 0$ increases odds Y is present with increasing x
- $\beta < 0$ decreases odds Y is present with increasing x

Dichotomous Response and Threshold Concept

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

Response occurs $(Y_i = 1)$ if $\gamma < y_i$
otherwise, a response does not occur $(Y_i = 0)$



The Threshold Concept in Practice

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

- Satisfaction may be continuous, but we usually emit a dichotomous 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

Random-intercept Logistic Regression Model

Consider the model with p covariates for the response Y_{ij} for subject j ($j = 1, 2, \ldots, n_i$) in cluster i ($i = 1, 2, \ldots, N$):

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

where

$$\begin{aligned}
Y_{ij} &= \text{ dichotomous response for subject } j \text{ in cluster } i \\
\boldsymbol{x}_{ij} &= (p+1) \times 1 \text{ covariate vector (includes 1 for intercept)} \\
\boldsymbol{\beta} &= (p+1) \times 1 \text{ vector of unknown parameters} \\
\upsilon_{0i} &= \text{ cluster effects distributed } \mathcal{NID}(0, \sigma_\upsilon^2)
\end{aligned}$$

Characteristics of $v_{0i} \sim \mathcal{NID}(0, \sigma_v^2)$

- separates model from ususal (fixed-effects) multiple logistic regression model
- takes on $i = 1, 2, \ldots, N$ values
- assess impact of cluster i on individual outcome $logit_{ij}$, represents effect of subject clustering
- common for each cluster member, but changes for each cluster
- if $v_{0i} = 0$, then cluster has no effect for cluster i
- if $v_{0i} = 0$ for all clusters, cluster structure has no impact on individual data ($\sigma_v^2 = 0$)
  - no need for multilevel approach
  - ordinary logistic regression is OK
- if subject clustering has strong effect, estimates of $v_{0i} \neq 0$ and σ_v^2 will increase from 0

Model for Latent Continuous Responses

Consider the model with p covariates for the $n_i \times 1$ latent response strength y_{ij}:

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

where assuming

- $\varepsilon_{ij} \sim$ standard normal (mean 0 and $\sigma^2 = 1$) leads to multilevel probit regression
- $\varepsilon_{ij} \sim$ standard logistic (mean 0 and $\sigma^2 = \pi^2/3$) leads to multilevel 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 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 model
- $V(y) = \sigma^2$ in fixed-effects model

difference depends on size of random-effects variance σ_v^2

Within-Clusters / Between-Clusters models

Within-clusters model - *level 1* $\quad (j = 1, \ldots, n_i)$

observed response

$$\log\left[\frac{P(Y_{ij} = 1 \mid \upsilon_{0i})}{1 - P(Y_{ij} = 1 \mid \upsilon_{0i})}\right] = b_{0i} \; + \; b_{1i}\, Sex_{ij}$$

latent response

$$y_{ij} \; = \; b_{0i} \; + \; b_{1i}\, Sex_{ij} \; + \; \varepsilon_{ij}$$

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

$$\begin{aligned} b_{0i} &= \beta_0 \; + \; \beta_2\, Grp_i \; + \; \upsilon_{0i} \\ b_{1i} &= \beta_1 + \; \beta_3\, Grp_i \end{aligned}$$

$$\upsilon_{0i} \sim \mathcal{NID}(0, \sigma_\upsilon^2) \quad \text{and} \quad \varepsilon_{ij} \sim \mathcal{LID}(0, \pi^2/3)$$

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 ≥ 3 (out of 7)
Subgroup Descriptive Statistics

	CC = no		CC = yes	
	TV=no	TV=yes	TV=no	TV=yes
n	421	416	380	383
proportions	.416	.483	.632	.603
odds	.711	.935	1.714	1.520
logits	-.341	-.067	.539	.419

Within-Clusters / Between-Clusters components

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

$$logit_{ij} = b_{0i}$$

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

$$b_{0i} \ = \ \beta_0 \ + \ \beta_1 CC_i \ + \ \beta_2 TV_i \ + \ \beta_3 (CC_i \times TV_i) \ + \ \upsilon_{0i}$$

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

$$\beta_0 = \text{THKS logit for CC=no TV=no subgroup}$$

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

$$b_{0i} = \beta_0 + (\beta_1 + \beta_3 TV_i) CC_i + \beta_2 TV_i + \upsilon_{0i}$$

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

$$b_{0i} = \beta_0 + (\beta_2 + \beta_3 CC_i) TV_i + \beta_1 CC_i + \upsilon_{0i}$$

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

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

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)

$$logit_{ijk} = b_{0ij}$$

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

$$b_{0ij} \; = \; b_{0i} \; + \; \upsilon_{0ij}$$

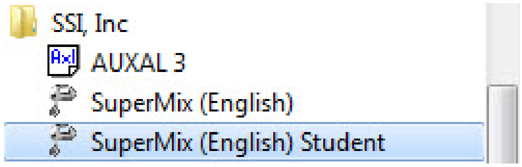
Between-schools model - *level 3* $\quad$ ($i = 1, \ldots, N$ schools)

$$b_{0i} \; = \; \beta_0 \; + \; \beta_1 CC_i \; + \; \beta_2 TV_i \; + \; \beta_3 (CC_i \times TV_i) \; + \; \upsilon_{0i}$$

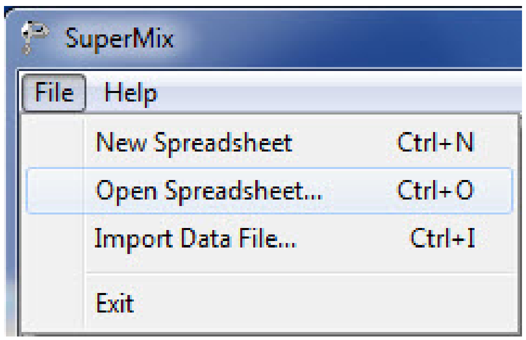
$$\upsilon_{0ij} \sim \mathcal{NID}(0, \sigma^2_{\upsilon(2)}) \quad \text{and} \quad \upsilon_{0i} \sim \mathcal{NID}(0, \sigma^2_{\upsilon(3)})$$

$$
\begin{aligned}
\beta_0 &= \text{THKS logit for CC=no TV=no subgroup} \\
\beta_1 &= \text{logit diff. between CC=yes vs CC=no (for TV=no)} \\
\beta_2 &= \text{logit diff. between TV=yes vs TV=no (for CC=no)} \\
\beta_3 &= \text{difference in logit attributable to interaction} \\
\upsilon_{0ij} &= \text{random classroom deviation} \\
\upsilon_{0i} &= \text{random school deviation}
\end{aligned}
$$

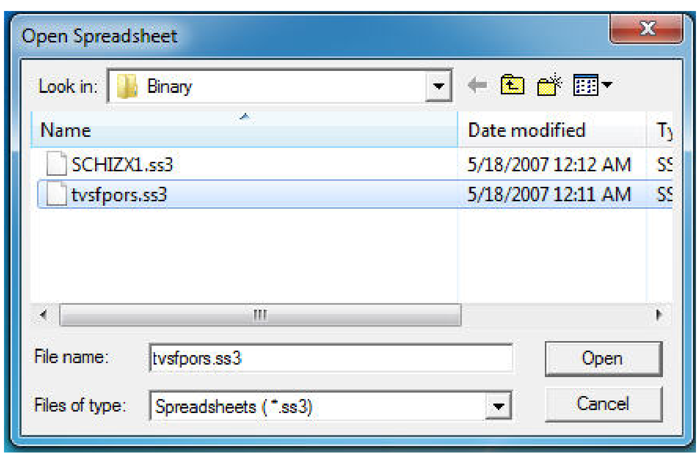
- Under SSI, Inc > "SuperMix (English)" or "SuperMix (English) Student"



- Under "File" click on "Open Spreadsheet"



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



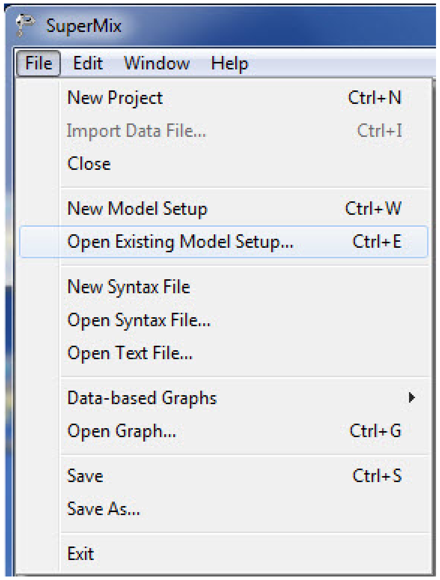
C:\SuperMixEn Examples\Workshop\Binary\tvsfpors.ss3

tvsfpors.ss3

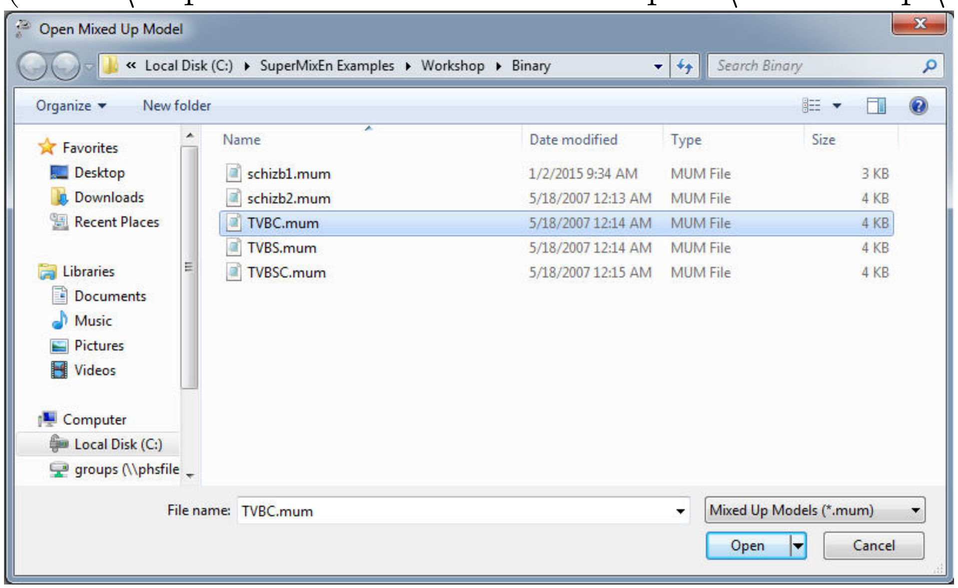
Apply

	(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

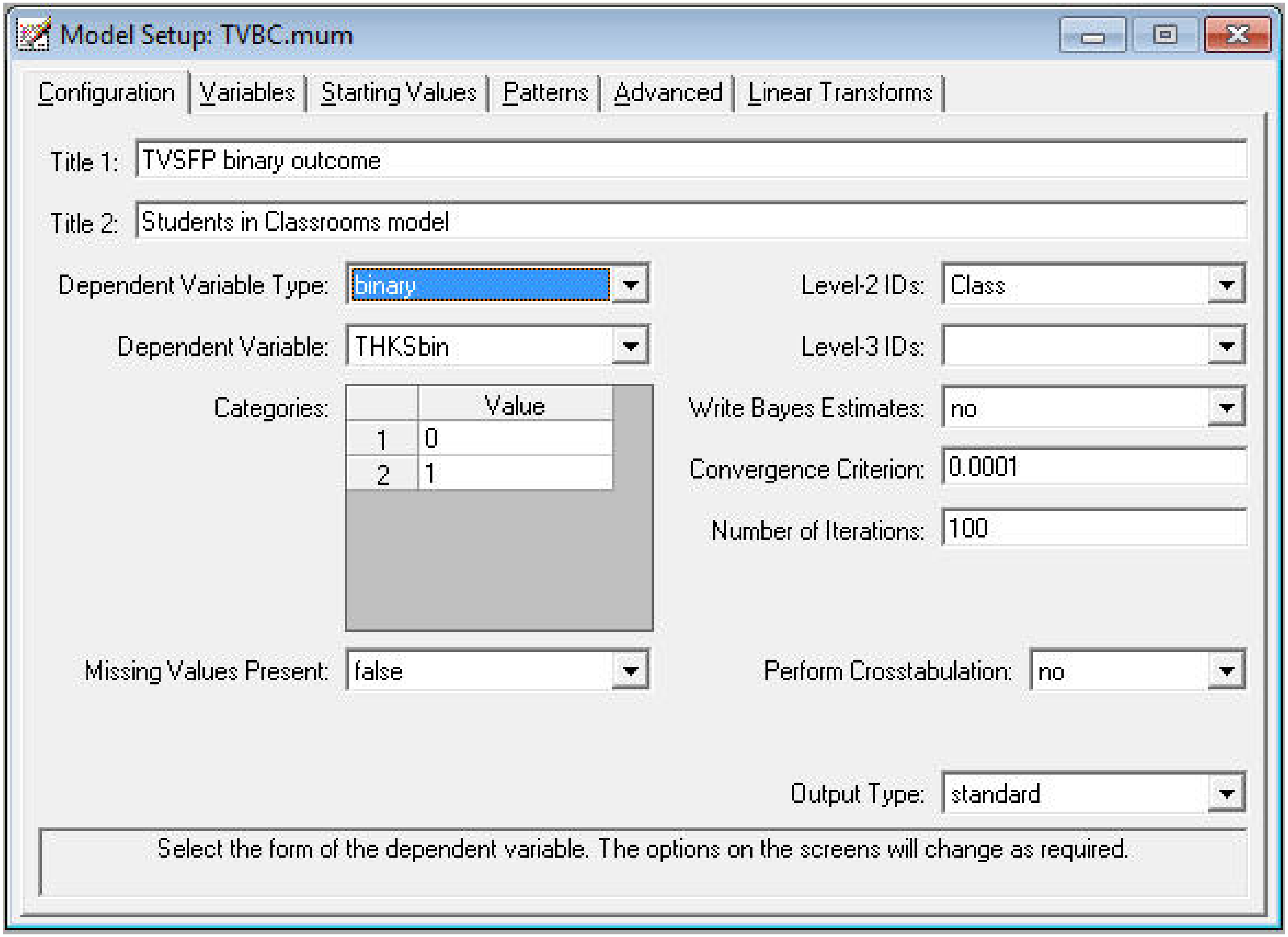
Under "File" click on "Open Existing Model Setup"



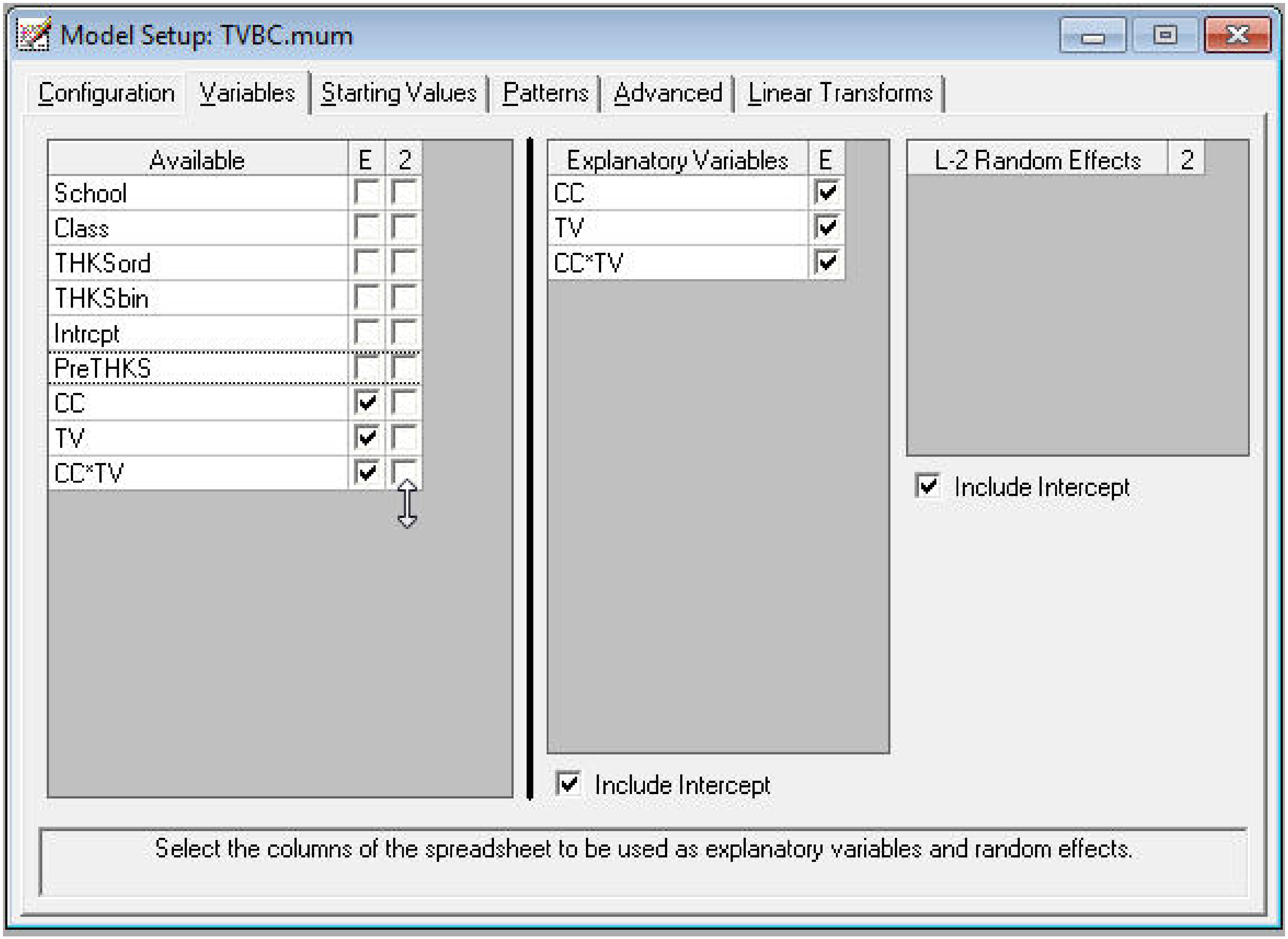
Open C:\SuperMixEn Examples\Workshop\Binary\tvbc.mum
(or C:\SuperMixEn Student Examples\Workshop\Binary\tvbc.mum)



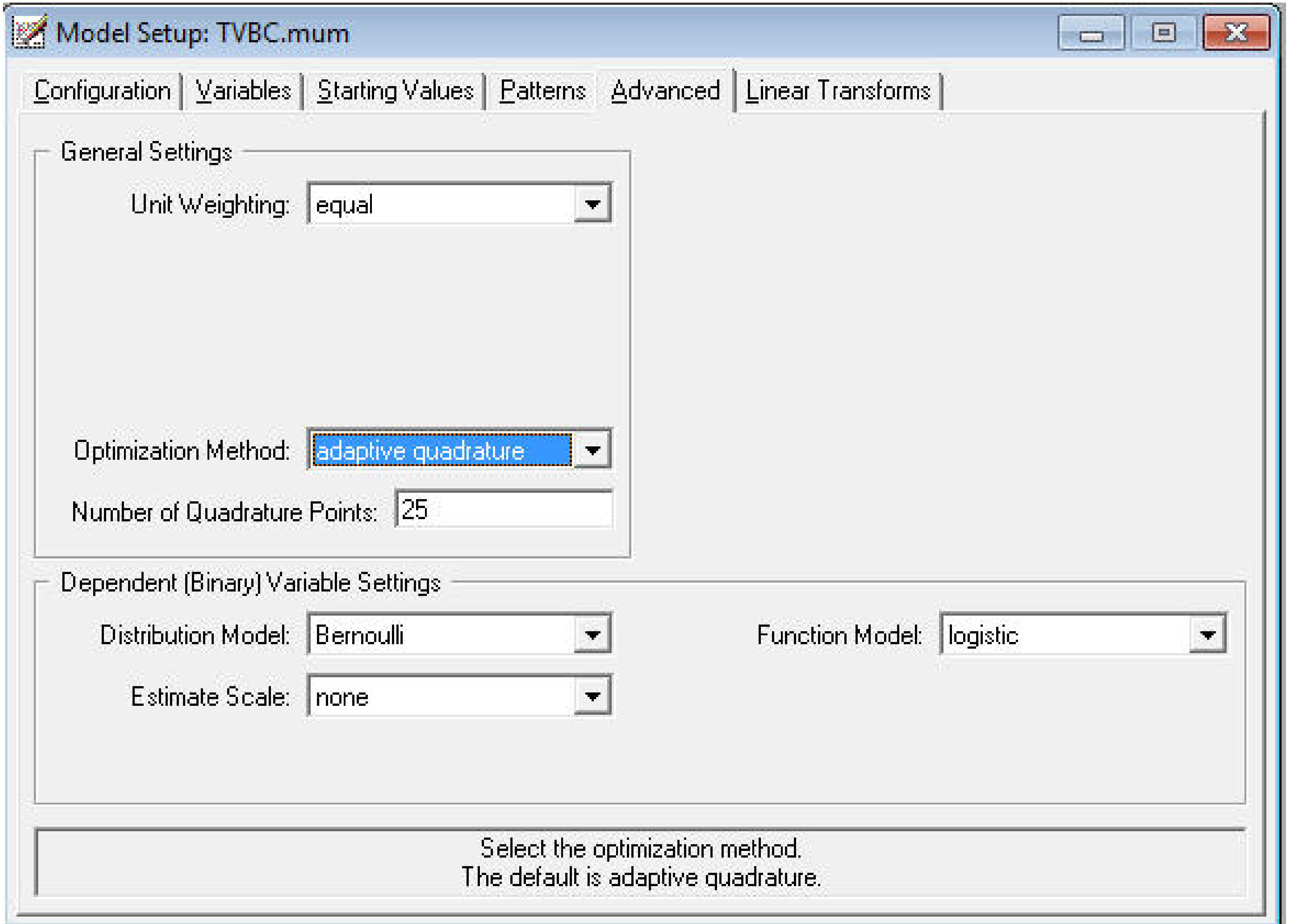
Note “Dependent Variable Type” should be “binary”



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



Note “Optimization Method” should be “adaptive quadrature”



TVBC.out

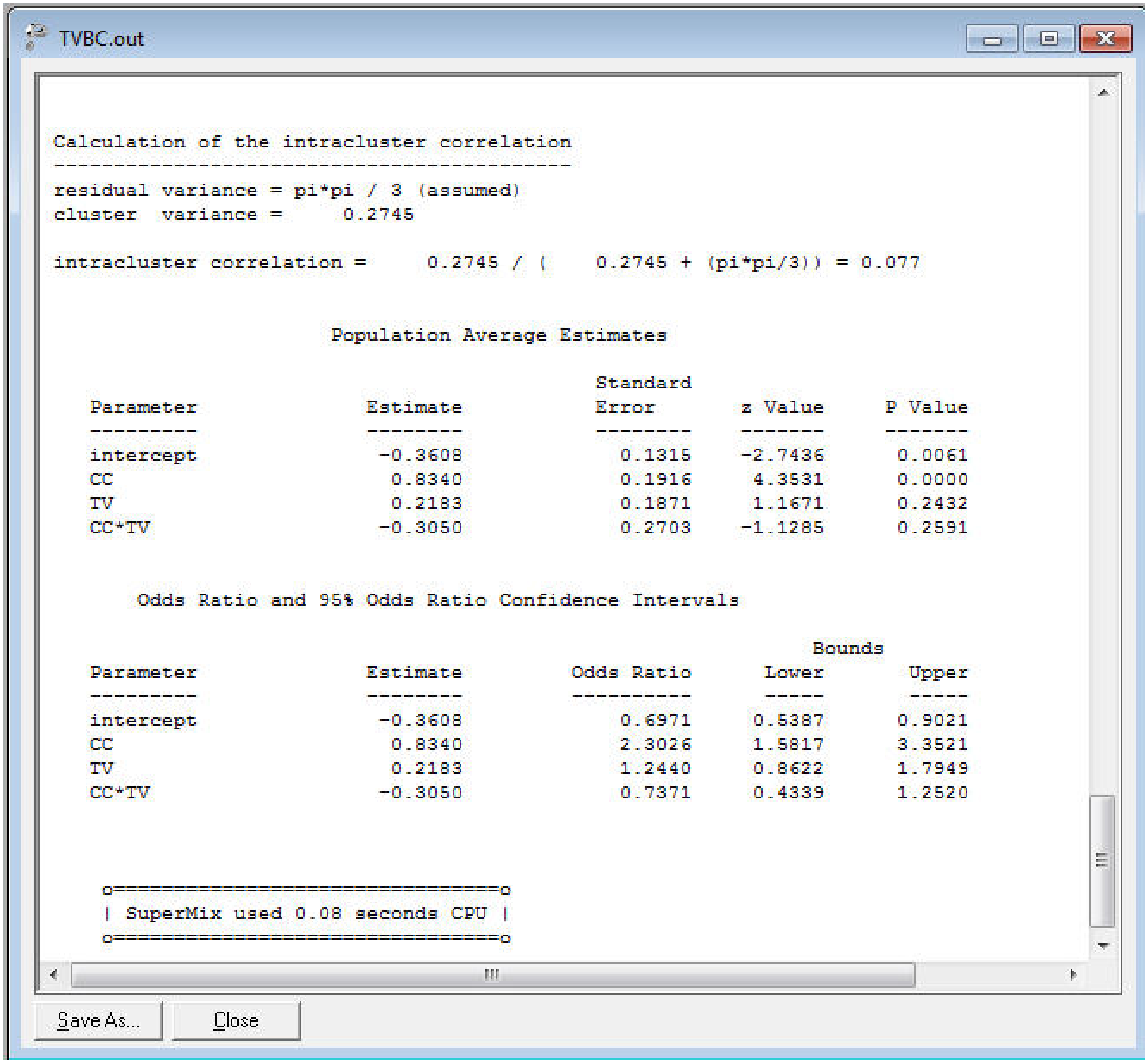
```
o==============================o
| TVSFP binary outcome         |
| Students in Classrooms model |
o==============================o



          Model and Data Descriptions

   Sampling Distribution                    = Bernoulli
   Link Function                            = Logistic
   PROB(Success)= 1.0/[1.0+EXP(-ETA)]

   Number of Level-2 Units                    135
   Number of Level-1 Units                    1600
   Number of Level-1 Units per Level-2 Unit =
    20     3    11     9     5    26    11    10    15    12    12    10
    21    10    17    19     2     4    21    16    15    13     2    14
    13     1    12    18    21    17    16    15    16    21    21    27
    17     3     2    15     7    24    22    15    19     7    12     8
     6    11     7     7     8     3     5     8     3     8     9     8
     2    11     9    21    13    12    12    14     9     6    11    10
    12    11     6     6    14    10    14     2     3     2     4     3
     6    10    14    11     6    22     4     7    22    18    23    19
    14     5    14    28    15    15    11    12    11    11    15    17
    24    20    15     6     8    14     5    11     9    17    14    11
    17    15     6     7    14    10    14    18     4     9     7    12
    15    11    10



    o============================================================o
    | Descriptive statistics for all the variables in the model |
    o============================================================o
                                                                Standard
   Variable               Minimum      Maximum          Mean   Deviation
   --------               -------      -------          ----   ---------
   THKSbin1                0.0000       1.0000        0.4706      0.4993
   THKSbin2                0.0000       1.0000        0.5294      0.4993
   intercept               1.0000       1.0000        1.0000      0.0000
   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

TVBC.out

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

Number of quadrature points =              25
Number of free parameters =                 5
Number of iterations used =                 4

-2lnL (deviance statistic) =       2138.14940
Akaike Information Criterion       2148.14940
Schwarz Criterion                  2175.03819


                    Estimated regression weights

                                           Standard
Parameter              Estimate            Error         z Value     P Value
---------              --------            --------      -------     -------
intercept               -0.3839              0.1400      -2.7418      0.0061
CC                       0.8872              0.2032       4.3657      0.0000
TV                       0.2322              0.1987       1.1686      0.2426
CC*TV                   -0.3243              0.2870      -1.1302      0.2584


    Odds Ratio and 95% Odds Ratio Confidence Intervals

                                                             Bounds
Parameter              Estimate          Odds Ratio      Lower       Upper
---------              --------          ----------      -----       -----
intercept               -0.3839              0.6812      0.5177      0.8963
CC                       0.8872              2.4284      1.6305      3.6167
TV                       0.2322              1.2614      0.8545      1.8621
CC*TV                   -0.3243              0.7230      0.4120      1.2689


               Estimated level 2 variances and covariances

                                                Standard
Parameter                       Estimate        Error          z Value    P Value
---------                       --------        --------       -------    -------
intercept/intercept               0.2745          0.0871        3.1499     0.0016
```

Save As... Close



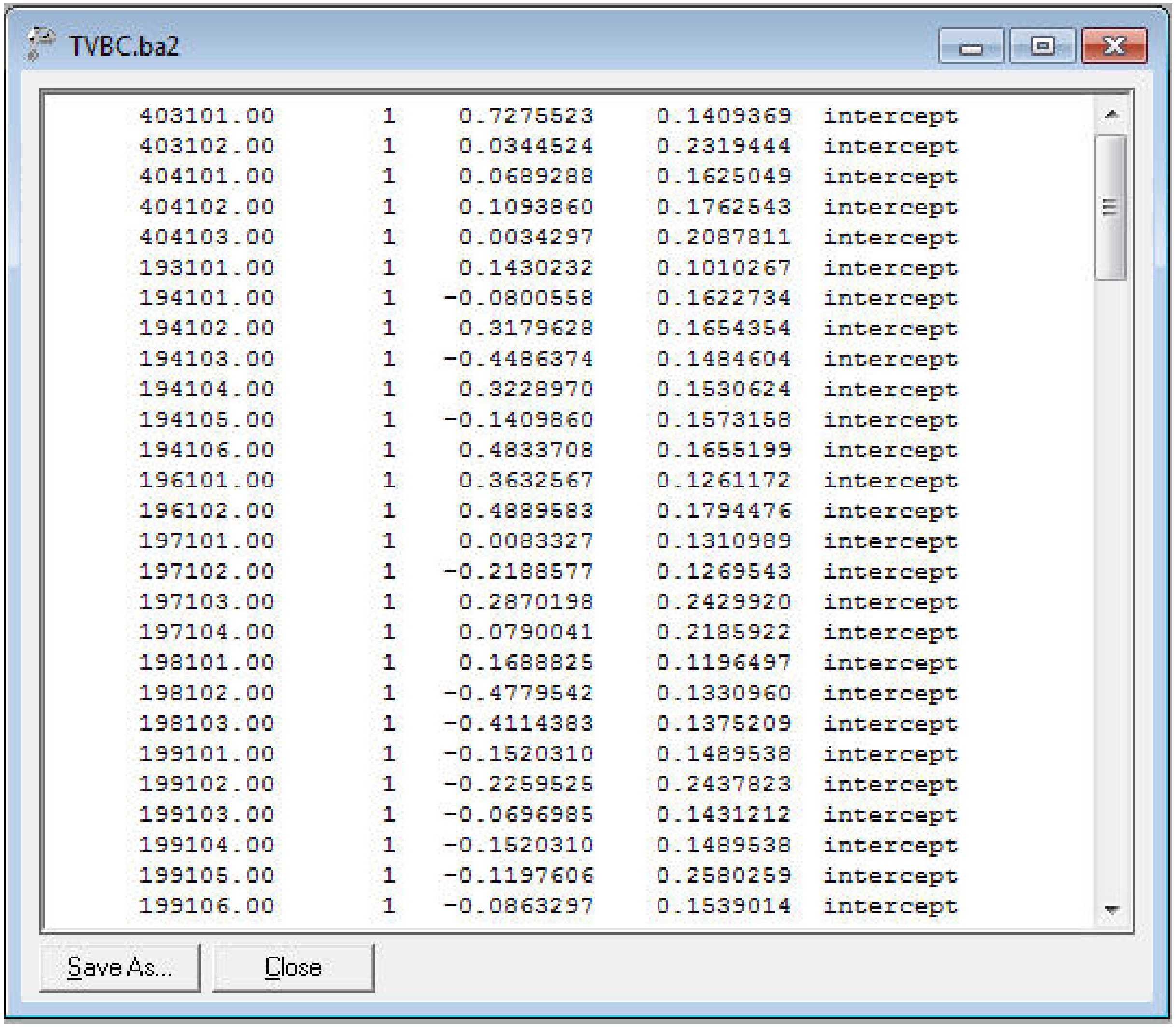
TVBC.out

```
Calculation of the intracluster correlation
-------------------------------------------
residual variance = pi*pi / 3 (assumed)
cluster  variance =     0.2745

intracluster correlation =     0.2745 / (     0.2745 + (pi*pi/3)) = 0.077


                       Population Average Estimates

                                                   Standard
   Parameter                Estimate               Error       z Value      P Value
   ---------                --------               --------    -------      -------
   intercept                 -0.3608                 0.1315    -2.7436       0.0061
   CC                         0.8340                 0.1916     4.3531       0.0000
   TV                         0.2183                 0.1871     1.1671       0.2432
   CC*TV                     -0.3050                 0.2703    -1.1285       0.2591


       Odds Ratio and 95% Odds Ratio Confidence Intervals

                                                                     Bounds
   Parameter                Estimate             Odds Ratio     Lower        Upper
   ---------                --------             ----------     -----        -----
   intercept                 -0.3608                 0.6971     0.5387       0.9021
   CC                         0.8340                 2.3026     1.5817       3.3521
   TV                         0.2183                 1.2440     0.8622       1.7949
   CC*TV                     -0.3050                 0.7371     0.4339       1.2520


       o=================================o
       | SuperMix used 0.08 seconds CPU |
       o=================================o
```

Save As... Close

Empirical Bayes Estimates of Random Effects

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



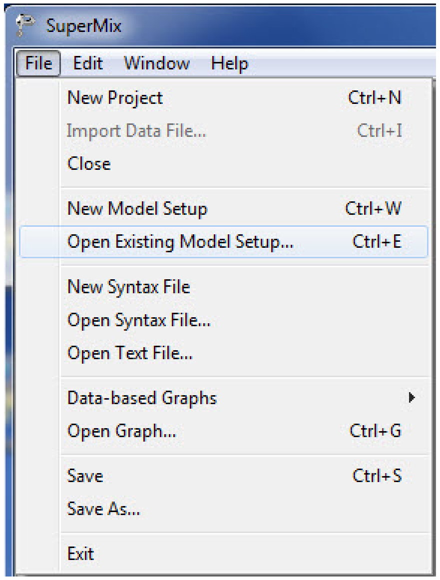
TVBC.ba2

403101.00	1	0.7275523	0.1409369	intercept
403102.00	1	0.0344524	0.2319444	intercept
404101.00	1	0.0689288	0.1625049	intercept
404102.00	1	0.1093860	0.1762543	intercept
404103.00	1	0.0034297	0.2087811	intercept
193101.00	1	0.1430232	0.1010267	intercept
194101.00	1	-0.0800558	0.1622734	intercept
194102.00	1	0.3179628	0.1654354	intercept
194103.00	1	-0.4486374	0.1484604	intercept
194104.00	1	0.3228970	0.1530624	intercept
194105.00	1	-0.1409860	0.1573158	intercept
194106.00	1	0.4833708	0.1655199	intercept
196101.00	1	0.3632567	0.1261172	intercept
196102.00	1	0.4889583	0.1794476	intercept
197101.00	1	0.0083327	0.1310989	intercept
197102.00	1	-0.2188577	0.1269543	intercept
197103.00	1	0.2870198	0.2429920	intercept
197104.00	1	0.0790041	0.2185922	intercept
198101.00	1	0.1688825	0.1196497	intercept
198102.00	1	-0.4779542	0.1330960	intercept
198103.00	1	-0.4114383	0.1375209	intercept
199101.00	1	-0.1520310	0.1489538	intercept
199102.00	1	-0.2259525	0.2437823	intercept
199103.00	1	-0.0696985	0.1431212	intercept
199104.00	1	-0.1520310	0.1489538	intercept
199105.00	1	-0.1197606	0.2580259	intercept
199106.00	1	-0.0863297	0.1539014	intercept

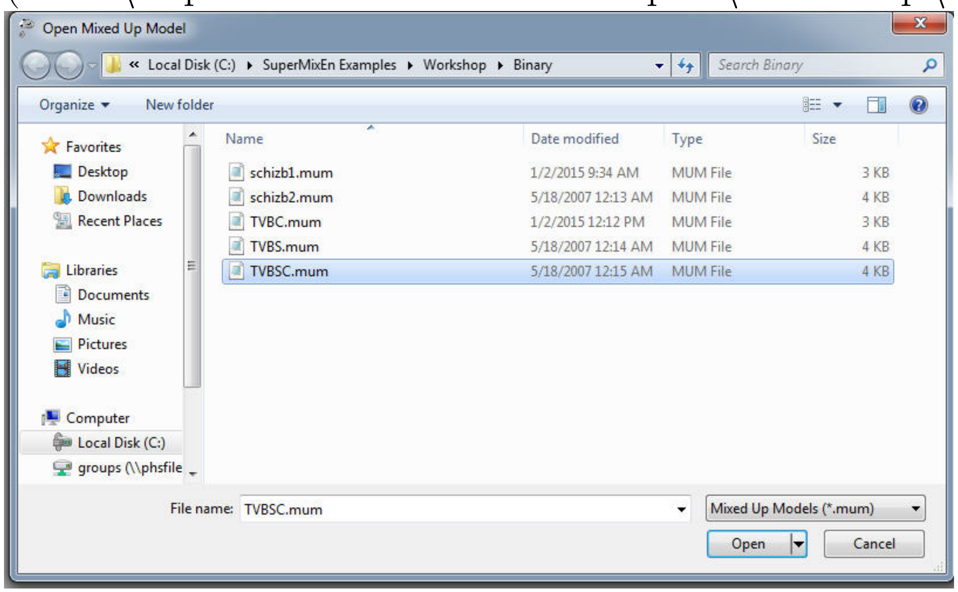
Save As... Close

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

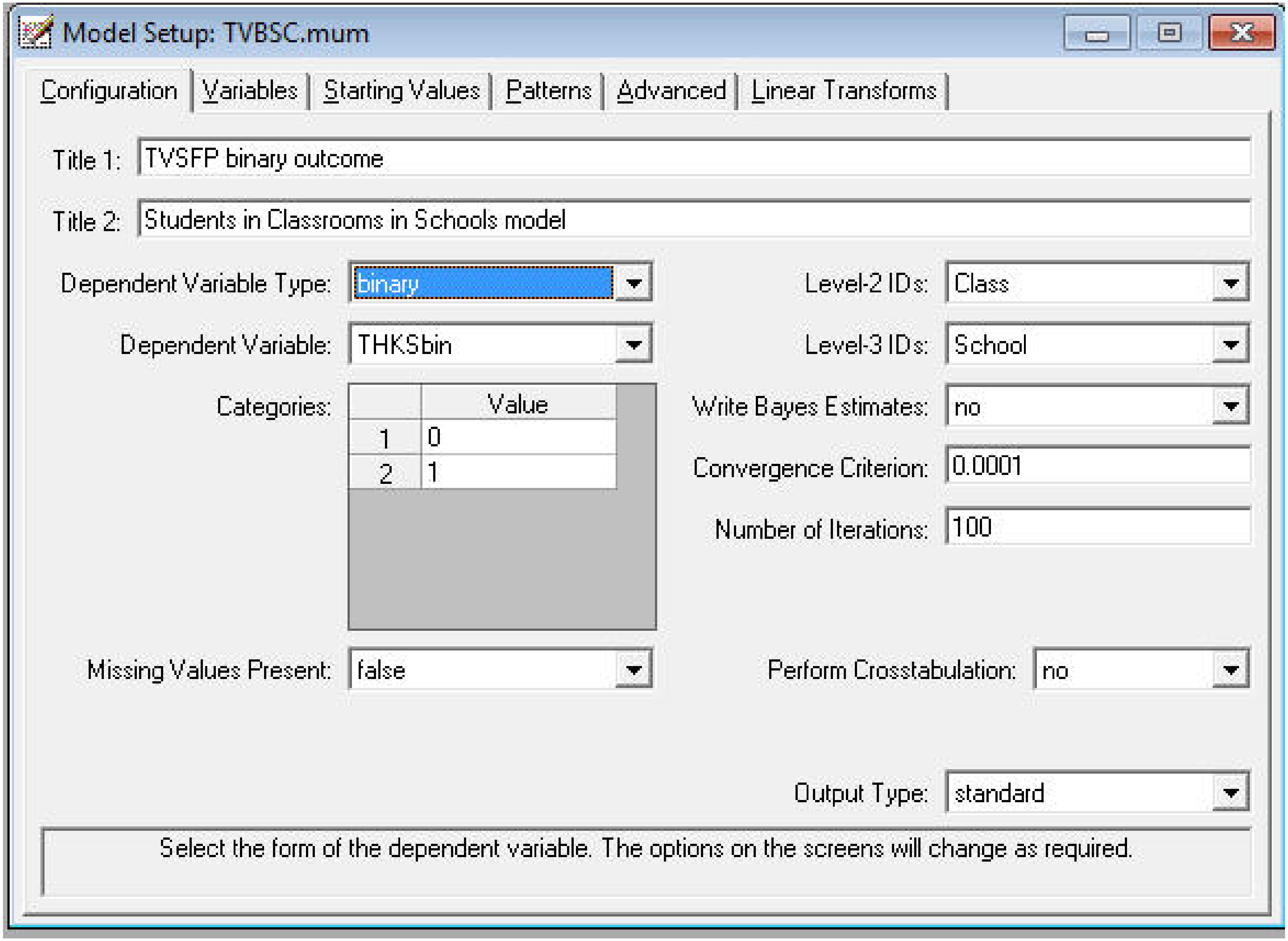
Under "File" click on "Open Existing Model Setup"



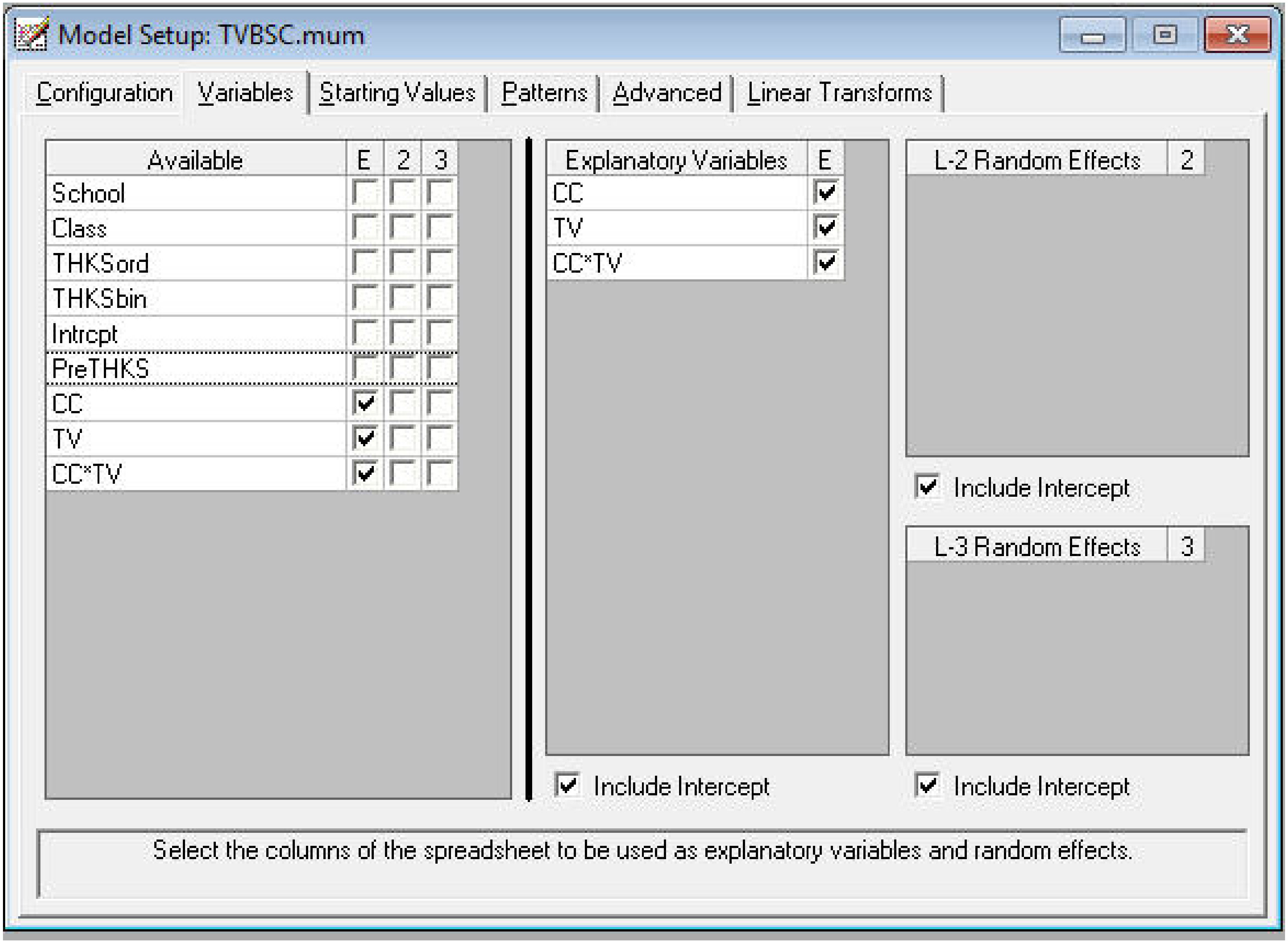
Open C:\SuperMixEn Examples\Workshop\Binary\tvbsc.mum
(or C:\SuperMixEn Student Examples\Workshop\Binary\tvbsc.mum)



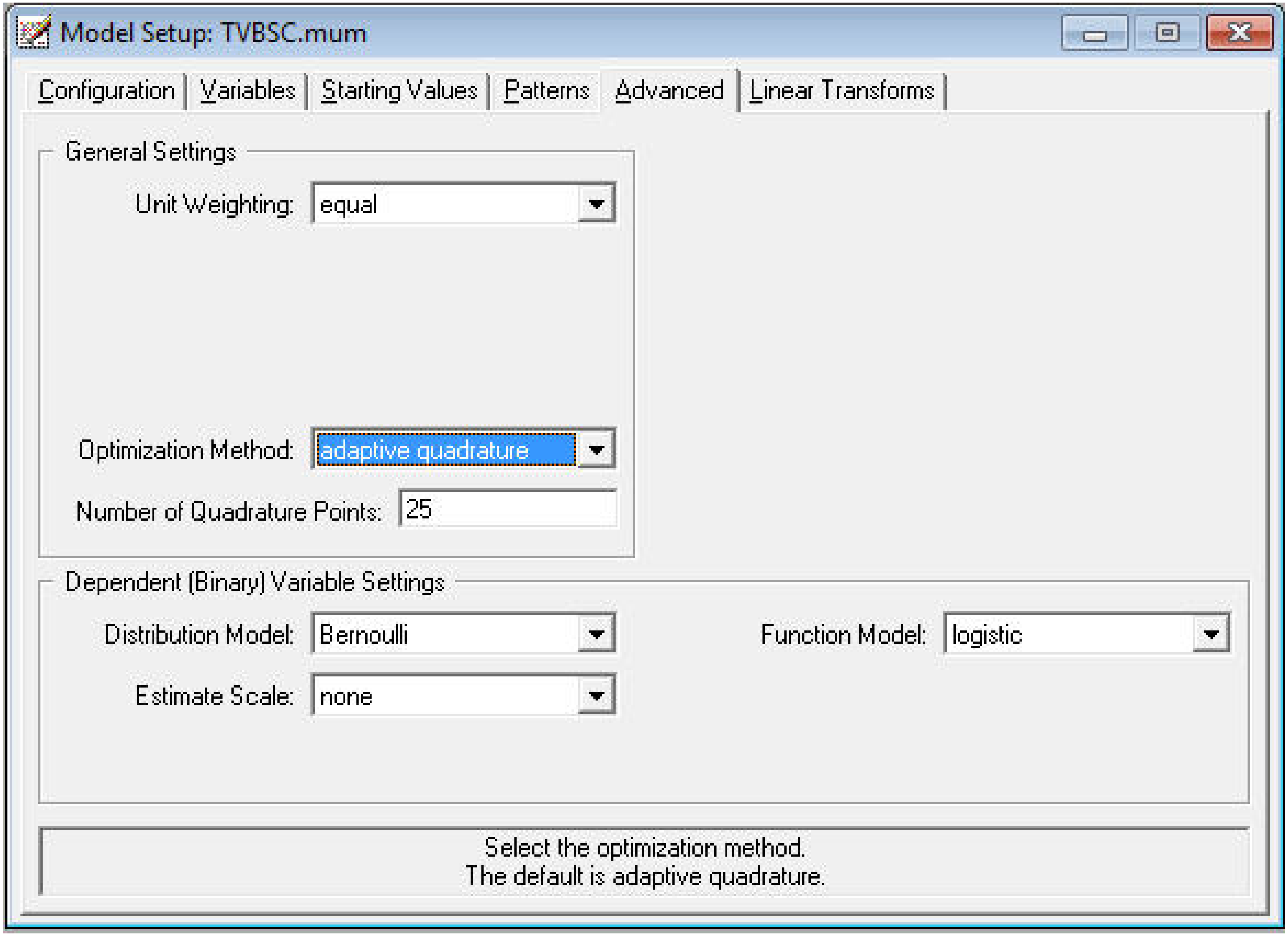
Note “Dependent Variable Type” should be “binary”



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



Note “Optimization Method” should be “adaptive quadrature”



TVBSC.out

```
o========================================o
| TVSFP binary outcome                   |
| Students in Classrooms in Schools model |
o========================================o


          Model and Data Descriptions

   Sampling Distribution                  = Bernoulli
   Link Function                          = Logistic
   PROB(Success)= 1.0/[1.0+EXP(-ETA)]

   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
   --------             -------     -------        ----   ---------
   THKSbin1              0.0000      1.0000      0.4706      0.4993
   THKSbin2              0.0000      1.0000      0.5294      0.4993
   intercept             1.0000      1.0000      1.0000      0.0000
   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

TVBSC.out

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

Number of quadrature points =              25
Number of free parameters =                 6
Number of iterations used =                10

-2lnL (deviance statistic) =       2133.70171
Akaike Information Criterion       2145.70171
Schwarz Criterion                  2177.96826


                    Estimated regression weights

                                           Standard
Parameter              Estimate            Error       z Value     P Value
---------              --------            --------    -------     -------
intercept               -0.3914              0.1918    -2.0407      0.0413
CC                       0.9794              0.2788     3.5128      0.0004
TV                       0.3230              0.2700     1.1967      0.2314
CC*TV                   -0.5013              0.3901    -1.2851      0.1987


    Odds Ratio and 95% Odds Ratio Confidence Intervals

                                                           Bounds
Parameter              Estimate          Odds Ratio      Lower       Upper
---------              --------          ----------      -----       -----
intercept               -0.3914              0.6761     0.4643      0.9846
CC                       0.9794              2.6628     1.5418      4.5989
TV                       0.3230              1.3813     0.8138      2.3447
CC*TV                   -0.5013              0.6057     0.2820      1.3011


               Estimated level 2 variances and covariances

                                                  Standard
Parameter                       Estimate          Error        z Value    P Value
---------                       --------          --------     -------    -------
intercept/intercept               0.1687            0.0805      2.0963     0.0361
```

Save As... Close

TVBSC.out

Estimated level 3 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intercept/intercept	0.1209	0.0771	1.5677	0.1169

Level 3 covariance matrix

	intercept
intercept	0.120890

Level 3 correlation matrix

	intercept
intercept	1.000000

Population Average Estimates

Parameter	Estimate	Standard Error	z Value	P Value
intercept	-0.3667	0.1798	-2.0399	0.0414
CC	0.9180	0.2614	3.5116	0.0004
TV	0.3027	0.2528	1.1973	0.2312
CC*TV	-0.4701	0.3656	-1.2857	0.1986

Odds Ratio and 95% Odds Ratio Confidence Intervals

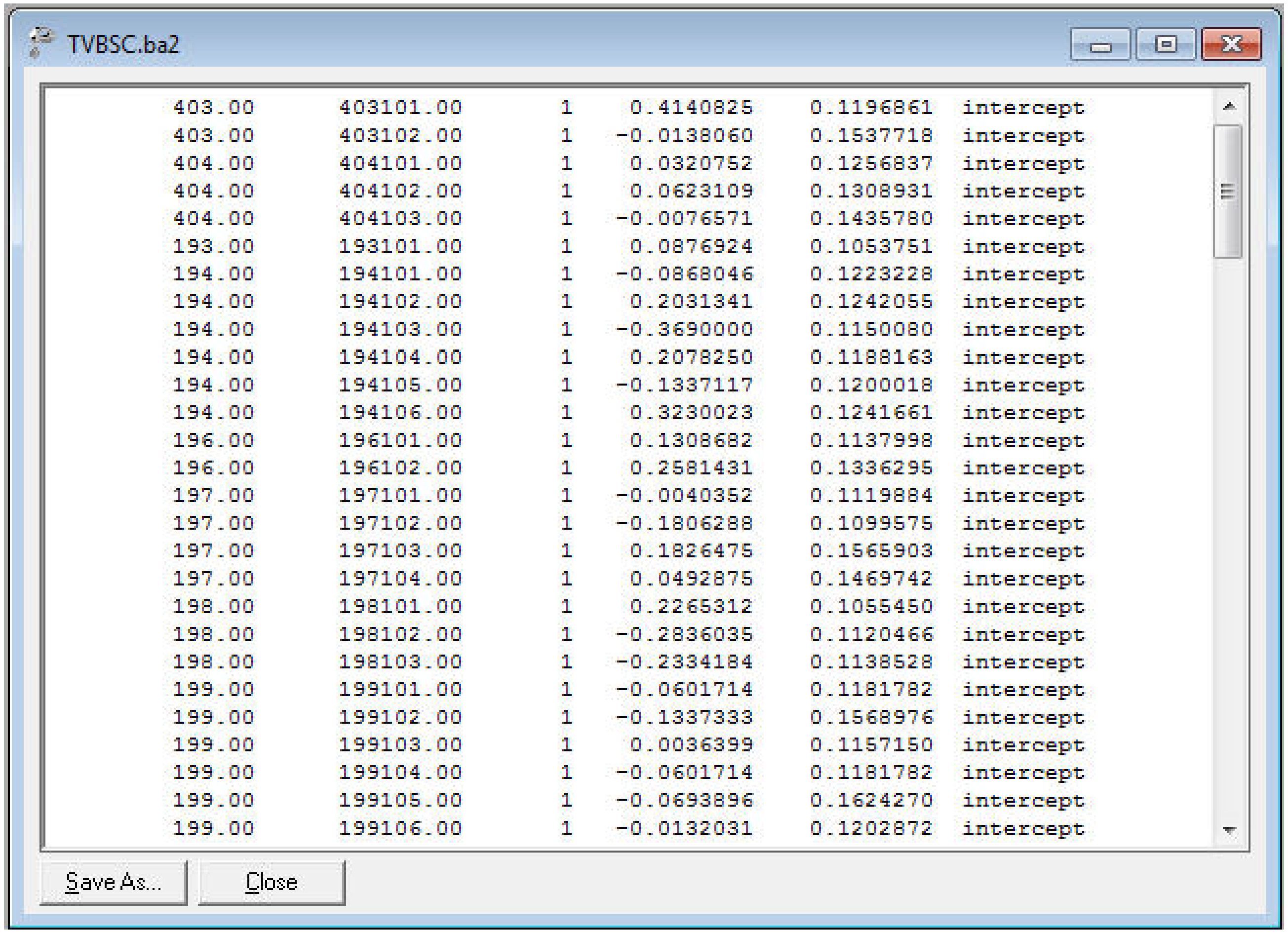
Parameter	Estimate	Odds Ratio	Bounds Lower	Upper
intercept	-0.3667	0.6930	0.4872	0.9857
CC	0.9180	2.5043	1.5002	4.1803
TV	0.3027	1.3535	0.8246	2.2217
CC*TV	-0.4701	0.6249	0.3052	1.2796

```
o=================================o
| SuperMix used 3.45 seconds CPU |
o=================================o
```

Save As... Close

Empirical Bayes Estimates of Random Class Effects

Select “Analysis” > “View Level-2 Bayes Results”



403.00	403101.00	1	0.4140825	0.1196861	intercept
403.00	403102.00	1	-0.0138060	0.1537718	intercept
404.00	404101.00	1	0.0320752	0.1256837	intercept
404.00	404102.00	1	0.0623109	0.1308931	intercept
404.00	404103.00	1	-0.0076571	0.1435780	intercept
193.00	193101.00	1	0.0876924	0.1053751	intercept
194.00	194101.00	1	-0.0868046	0.1223228	intercept
194.00	194102.00	1	0.2031341	0.1242055	intercept
194.00	194103.00	1	-0.3690000	0.1150080	intercept
194.00	194104.00	1	0.2078250	0.1188163	intercept
194.00	194105.00	1	-0.1337117	0.1200018	intercept
194.00	194106.00	1	0.3230023	0.1241661	intercept
196.00	196101.00	1	0.1308682	0.1137998	intercept
196.00	196102.00	1	0.2581431	0.1336295	intercept
197.00	197101.00	1	-0.0040352	0.1119884	intercept
197.00	197102.00	1	-0.1806288	0.1099575	intercept
197.00	197103.00	1	0.1826475	0.1565903	intercept
197.00	197104.00	1	0.0492875	0.1469742	intercept
198.00	198101.00	1	0.2265312	0.1055450	intercept
198.00	198102.00	1	-0.2836035	0.1120466	intercept
198.00	198103.00	1	-0.2334184	0.1138528	intercept
199.00	199101.00	1	-0.0601714	0.1181782	intercept
199.00	199102.00	1	-0.1337333	0.1568976	intercept
199.00	199103.00	1	0.0036399	0.1157150	intercept
199.00	199104.00	1	-0.0601714	0.1181782	intercept
199.00	199105.00	1	-0.0693896	0.1624270	intercept
199.00	199106.00	1	-0.0132031	0.1202872	intercept

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

Empirical Bayes Estimates of Random School Effects

Select "Analysis" > "View Level-3 Bayes Results"

TVBSC.ba3

403.00	1	0.2869095	0.0912322	intercept
404.00	1	0.0621644	0.0800387	intercept
193.00	1	0.0628534	0.0883777	intercept
194.00	1	0.1035370	0.0514053	intercept
196.00	1	0.2788348	0.0828379	intercept
197.00	1	0.0338824	0.0694394	intercept
198.00	1	-0.2082206	0.0643113	intercept
199.00	1	-0.2387087	0.0588660	intercept
401.00	1	0.2933296	0.0770402	intercept
402.00	1	-0.0194551	0.0773640	intercept
405.00	1	-0.0126517	0.0643529	intercept
407.00	1	0.2971076	0.0612066	intercept
408.00	1	-0.1337075	0.0789259	intercept
409.00	1	0.1783368	0.0532141	intercept
410.00	1	-0.4240002	0.0709826	intercept
411.00	1	0.3365628	0.0887505	intercept
412.00	1	-0.0279489	0.0694564	intercept
414.00	1	-0.1228827	0.0676033	intercept
415.00	1	0.3624831	0.0574220	intercept
505.00	1	-0.1754494	0.0492031	intercept
506.00	1	-0.5082910	0.0505349	intercept
507.00	1	-0.2349807	0.0505248	intercept
508.00	1	0.3608973	0.0531524	intercept
509.00	1	-0.2026640	0.0407207	intercept
510.00	1	-0.0783592	0.0419983	intercept
513.00	1	-0.3368798	0.0751064	intercept
514.00	1	0.0896937	0.0450339	intercept
515.00	1	-0.0223928	0.0332637	intercept

Save As... Close

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

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

	Fixed		Multilevel 2-level		Multilevel 3-level	
intercept	-.341	***	-.384	***	-.391	***
	(.099)		(.140)		(.192)	
CC	.880	***	.887	***	.979	**
	(.145)		(.203)		(.278)	
TV	.273	**	.232		.323	
	(.139)		(.199)		(.270)	
CC× TV	-.394	*	-.324		-.501	
	(.204)		(.287)		(.390)	
class var			.275		.170	
			(.087)		(.081)	
school var					.120	
					(.077)	
-2 log L	2162.53		2138.15		2133.70	

***$p < .01$ **$p < .05$ *$p < .10$ (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 $(\pi^2/3 = 3.2897)$

$$r = \frac{.275}{.275 + \pi^2/3} = .077$$

$\Rightarrow$ 7.7% of the unexplained variation is at the classroom level

Calculation of ICC - 3 level model

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

$$r = \frac{\sigma^2_{\upsilon(3)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2} = \frac{.121}{.121 + .169 + \pi^2/3} = .034$$

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

$$r = \frac{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2} = \frac{.121 + .169}{.121 + .169 + \pi^2/3} = .081$$

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

$$r = \frac{\sigma^2_{\upsilon(3)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)}} = \frac{.121}{.121 + .169} = .415$$

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

CC	TV	logistic $\Psi(z) = [1 + \exp(-z)]^{-1}$	estimate
Fixed-effects model			
0	0	$\Psi(-.341)$	.416
0	1	$\Psi(-.341 + .273)$	.483
1	0	$\Psi(-.341 + .880)$	.632
1	1	$\Psi(-.341 + .273 + .880 - .394)$	.603
Random-classrooms model $\hat{d} = (.2745 + \pi^2/3)/(\pi^2/3)$			
0	0	$\Psi((-.384)/\sqrt{\hat{d}})$	.409
0	1	$\Psi((-.384 + .232)/\sqrt{\hat{d}})$	.464
1	0	$\Psi((-.384 + .887)/\sqrt{\hat{d}})$	.619
1	1	$\Psi((-.384 + .232 + .887 - .324)/\sqrt{\hat{d}})$	.597
Random-classrooms model using Population Average Estimates			
0	0	$\Psi(-.361)$	.411
0	1	$\Psi(-.361 + .218)$	.464
1	0	$\Psi(-.361 + .834)$	.616
1	1	$\Psi(-.361 + .218 + .834 - .305)$	.595

d = design effect = $(\sigma_v^2 + \sigma^2)/\sigma^2$

CC	TV	logistic $\Psi(z) = [1+\exp(-z)]^{-1}$	estimate
3-level model		$\hat{d} = (.121 + .169 + \pi^2/3)/(\pi^2/3)$	
0	0	$\Psi((-.391)/\sqrt{\hat{d}})$	.407
0	1	$\Psi((-.391 + .323)/\sqrt{\hat{d}})$	.484
1	0	$\Psi((-.391 + .979)/\sqrt{\hat{d}})$	.638
1	1	$\Psi((-.391 + .323 + .979 - .501)/\sqrt{\hat{d}})$	.597
3-level model using Population Average Estimates			
0	0	$\Psi(-.367)$	.409
0	1	$\Psi(-.367 + .303)$	.484
1	0	$\Psi(-.367 + .918)$	.634
1	1	$\Psi(-.367 + .303 + .918 - .470)$	.595

d = design effect = $(\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2)/\sigma^2$

Within-Clusters / Between-Clusters components

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

$$logit_{ij} = b_{0i} + b_{1i} PRETHKS_{ij}$$

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

$$\begin{aligned} b_{0i} &= \beta_0 \ + \ \beta_2 CC_i \ + \ \beta_3 TV_i \ + \ \beta_4 (CC_i \times TV_i) \ + \ \upsilon_{0i} \\ b_{1i} &= \beta_1 \end{aligned}$$

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

β_0 = (PRETHKS adjusted) logit for CC=no TV=no subgroup

β_1 = effect of PRETHKS on POSTTHKS

β_2 = (PRETHKS adjusted) logit diff. between CC=yes vs CC=no (for TV=no)

β_3 = (PRETHKS adjusted) logit diff. between TV=yes vs TV=no (for CC=no)

β_4 = (PRETHKS adjusted) difference in logit attributable to interaction

v_{0i} = random cluster deviation

3-level model

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

$$logit_{ijk} = b_{0ij} + b_{1ij} PRETHKS_{ijk}$$

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

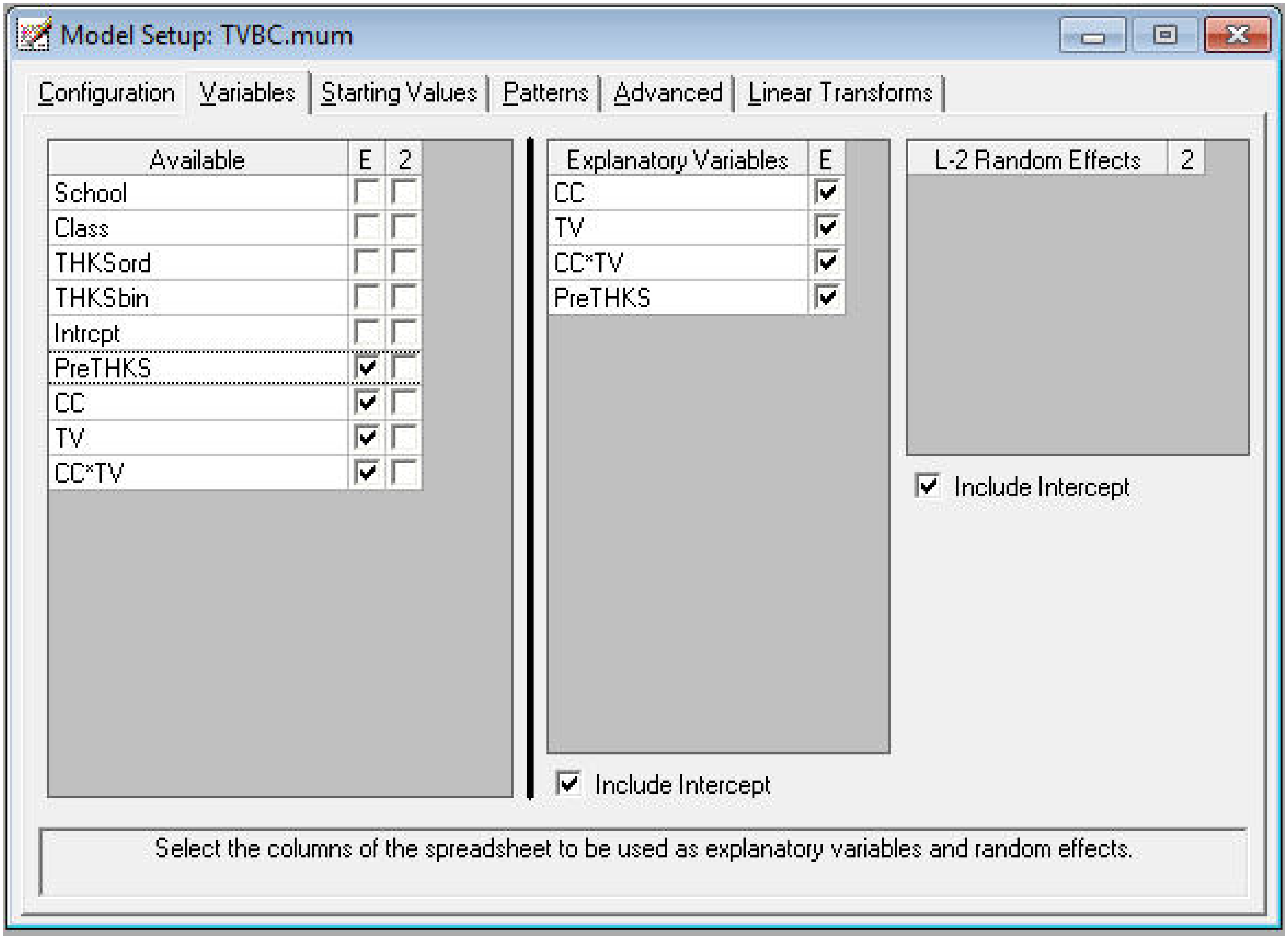
$$\begin{aligned} b_{0ij} &= b_{0i} + \upsilon_{0ij} \\ b_{1ij} &= b_{1i} \end{aligned}$$

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

$$\begin{aligned} b_{0i} &= \beta_0 + \beta_2 CC_i + \beta_3 TV_i + \beta_4 (CC_i \times TV_i) + \upsilon_{0i} \\ b_{1i} &= \beta_1 \end{aligned}$$

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

Reopening **TVBC.mum** and selecting **PreTHKS** as an explanatory variable



TVBC.out

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

Number of quadrature points =            25
Number of free parameters =               6
Number of iterations used =               4

-2lnL (deviance statistic) =       2057.17849
Akaike Information Criterion       2069.17849
Schwarz Criterion                  2101.44504


                    Estimated regression weights

                                          Standard
Parameter              Estimate           Error       z Value      P Value
---------              --------           --------    -------      -------
intercept               -1.2534             0.1695    -7.3936       0.0000
CC                       0.9883             0.1973     5.0094       0.0000
TV                       0.2870             0.1920     1.4949       0.1349
CC*TV                   -0.3691             0.2774    -1.3306       0.1833
PreTHKS                  0.4010             0.0461     8.6971       0.0000


    Odds Ratio and 95% Odds Ratio Confidence Intervals

                                                            Bounds
Parameter              Estimate         Odds Ratio      Lower        Upper
---------              --------         ----------      -----        -----
intercept               -1.2534             0.2855     0.2048       0.3981
CC                       0.9883             2.6865     1.8250       3.9548
TV                       0.2870             1.3325     0.9146       1.9412
CC*TV                   -0.3691             0.6914     0.4014       1.1908
PreTHKS                  0.4010             1.4933     1.3643       1.6345


                   Estimated level 2 variances and covariances

                                                 Standard
Parameter                        Estimate        Error         z Value    P Value
---------                        --------        --------      -------    -------
intercept/intercept                0.2192          0.0803       2.7308     0.0063
```

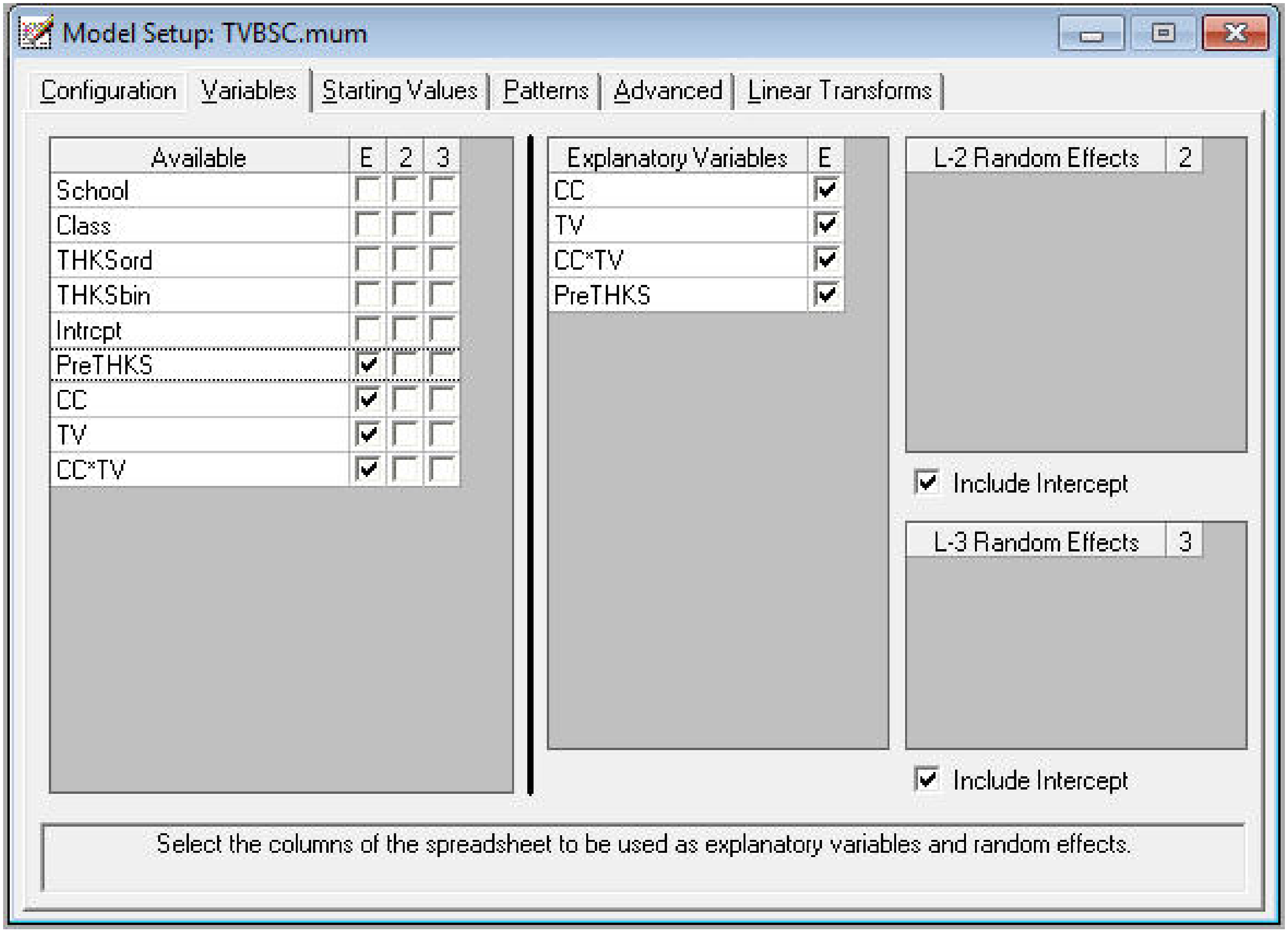
Save As... Close

TVBC.out

```
Calculation of the intracluster correlation
-------------------------------------------
residual variance = pi*pi / 3 (assumed)
cluster  variance =     0.2192

intracluster correlation =     0.2192 / (     0.2192 + (pi*pi/3)) = 0.062


                       Population Average Estimates

                                                Standard
   Parameter                Estimate            Error       z Value     P Value
   ---------                --------            --------    -------     -------
   intercept                 -1.1953              0.1621    -7.3754      0.0000
   CC                         0.9421              0.1885     4.9981      0.0000
   TV                         0.2736              0.1833     1.4928      0.1355
   CC*TV                     -0.3518              0.2649    -1.3283      0.1841
   PreTHKS                    0.3825              0.0444     8.6121      0.0000


      Odds Ratio and 95% Odds Ratio Confidence Intervals

                                                                  Bounds
   Parameter                Estimate          Odds Ratio    Lower        Upper
   ---------                --------          ----------    -----        -----
   intercept                 -1.1953              0.3026    0.2203       0.4158
   CC                         0.9421              2.5654    1.7730       3.7119
   TV                         0.2736              1.3147    0.9179       1.8829
   CC*TV                     -0.3518              0.7034    0.4185       1.1821
   PreTHKS                    0.3825              1.4660    1.3438       1.5994


    o=================================o
    | SuperMix used 0.09 seconds CPU |
    o=================================o
```

Save As...   Close

Reopening **TVBSC.mum** and selecting **PreTHKS** as an explanatory variable



TVBSC.out

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

Number of quadrature points =              25
Number of free parameters =                 7
Number of iterations used =                11

-2lnL (deviance statistic) =       2055.70206
Akaike Information Criterion       2069.70206
Schwarz Criterion                  2107.34637
```

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
intercept	-1.2465	0.1957	-6.3687	0.0000
CC	1.0383	0.2448	4.2416	0.0000
TV	0.3325	0.2358	1.4104	0.1584
CC*TV	-0.4644	0.3427	-1.3552	0.1754
PreTHKS	0.3954	0.0463	8.5332	0.0000

Odds Ratio and 95% Odds Ratio Confidence Intervals

Parameter	Estimate	Odds Ratio	Bounds Lower	Bounds Upper
intercept	-1.2465	0.2875	0.1959	0.4220
CC	1.0383	2.8245	1.7481	4.5636
TV	0.3325	1.3945	0.8785	2.2137
CC*TV	-0.4644	0.6285	0.3211	1.2303
PreTHKS	0.3954	1.4849	1.3560	1.6261

Estimated level 2 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intercept/intercept	0.1649	0.0813	2.0277	0.0426

Save As... Close

TVBSC.out

Estimated level 3 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intercept/intercept	0.0630	0.0616	1.0213	0.3071

Level 3 covariance matrix

	intercept
intercept	0.062954

Level 3 correlation matrix

	intercept
intercept	1.000000

Population Average Estimates

Parameter	Estimate	Standard Error	z Value	P Value
intercept	-1.1866	0.1869	-6.3478	0.0000
CC	0.9881	0.2330	4.2416	0.0000
TV	0.3164	0.2243	1.4106	0.1584
CC*TV	-0.4420	0.3262	-1.3551	0.1754
PreTHKS	0.3766	0.0446	8.4335	0.0000

Odds Ratio and 95% Odds Ratio Confidence Intervals

Parameter	Estimate	Odds Ratio	Bounds Lower	Bounds Upper
intercept	-1.1866	0.3053	0.2116	0.4403
CC	0.9881	2.6861	1.7015	4.2405
TV	0.3164	1.3722	0.8841	2.1298
CC*TV	-0.4420	0.6427	0.3391	1.2181
PreTHKS	0.3766	1.4572	1.3351	1.5905

Save As... Close

THKS Post-Int (dichotomized) Scores - LR Estimates (std err)

			Multilevel			
	Fixed		*2-level*		*3-level*	
intercept	-1.217	***	-1.253	***	-1.246	***
	(.141)		(.170)		(.196)	
PRETHKS	.400	***	.401	***	.395	***
	(.044)		(.046)		(.046)	
CC	.973	***	.988	***	1.038	***
	(.150)		(.197)		(.245)	
TV	.316	**	.287		.333	
	(.143)		(.192)		(.236)	
CC$\times$ TV	-.413	**	-.369		-.464	
	(.210)		(.277)		(.343)	
class var			.219		.165	
			(.080)		(.081)	
school var					.063	
					(.062)	
-2 log L	2073.3		2057.18		2055.70	

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

Calculation of ICC - 2 level models

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

Random classrooms model

$$r = \frac{.219}{.219 + \pi^2/3} = .062$$

$\Rightarrow$ 6.2% of the unexplained variation is at the classroom level

Calculation of ICC - 3 level model

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

$$r = \frac{\sigma^2_{\upsilon(3)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2} = \frac{.063}{.063 + .165 + \pi^2/3} = .018$$

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

$$r = \frac{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2} = \frac{.063 + .165}{.063 + .165 + \pi^2/3} = .063$$

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

$$r = \frac{\sigma^2_{\upsilon(3)}}{\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)}} = \frac{.063}{.063 + .165} = .276$$

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

CC	TV	logistic $\Psi(z) = [1+\exp(-z)]^{-1}$	estimate
Fixed-effects model			
0	0	$\Psi(-1.217 + 2.152 \times .400)$	.412
0	1	$\Psi(-1.217 + .316 + 2.087 \times .400)$	.483
1	0	$\Psi(-1.217 + .973 + 2.050 \times .400)$	.640
1	1	$\Psi(-1.217 + .316 + .973 - .413 + 1.979 \times .400)$	.610
Random-classrooms model $\hat{d} = (.219 + \pi^2/3)/(\pi^2/3)$			
0	0	$\Psi((-1.253 + 2.152 \times .401)/\sqrt{\hat{d}})$	.407
0	1	$\Psi((-1.253 + .287 + 2.087 \times .401)/\sqrt{\hat{d}})$	.469
1	0	$\Psi((-1.253 + .988 + 2.050 \times .401)/\sqrt{\hat{d}})$	.632
1	1	$\Psi((-1.253 + .287 + .988 - .369 + 1.979 \times .401)/\sqrt{\hat{d}})$	.606
Random-classrooms model using Population Average Estimates			
0	0	$\Psi(-1.195 + 2.152 \times .383)$	.408
0	1	$\Psi(-1.195 + .287 + 2.087 \times .383)$	.469
1	0	$\Psi(-1.195 + .988 + 2.050 \times .383)$	.630
1	1	$\Psi(-1.195 + .287 + .988 - .369 + 1.979 \times .383)$	.605
d = design effect = $(\sigma_v^2 + \sigma^2)/\sigma^2$			

CC	TV	logistic $\Psi(z) = [1 + \exp(-z)]^{-1}$	estimate
3-level model		$\hat{d} = (.063 + .165 + \pi^2/3)/(\pi^2/3)$	
0	0	$\Psi((-1.246 + 2.152 \times .395)/\sqrt{\hat{d}})$	.405
0	1	$\Psi((-1.246 + .333 + 2.087 \times .395)/\sqrt{\hat{d}})$	.479
1	0	$\Psi((-1.246 + 1.038 + 2.050 \times .395)/\sqrt{\hat{d}})$	.642
1	1	$\Psi((-1.246 + .333 + 1.038 - .464 + 1.979 \times .395)/\sqrt{\hat{d}})$	.605
3-level model using Population Average Estimates			
0	0	$\Psi(-1.187 + 2.152 \times .377)$	.407
0	1	$\Psi(-1.187 + .316 + 2.087 \times .377)$	.479
1	0	$\Psi(-1.187 + .988 + 2.050 \times .377)$	.640
1	1	$\Psi(-1.187 + .316 + .988 - .442 + 1.979 \times .377)$	.604

d = design effect = $(\sigma^2_{\upsilon(3)} + \sigma^2_{\upsilon(2)} + \sigma^2)/\sigma^2$

Summary

- Mixed logistic regression model is direct extension of ordinary logistic regression
- Useful approach for multilevel data
- Software is available in Supermix (and other programs)
- (Extended) methods are available for ordinal, nominal, count outcomes
- Similar models can be used for longitudinal, albeit more issues
  - more random effects are typically necessary
  - missing data and attrition