

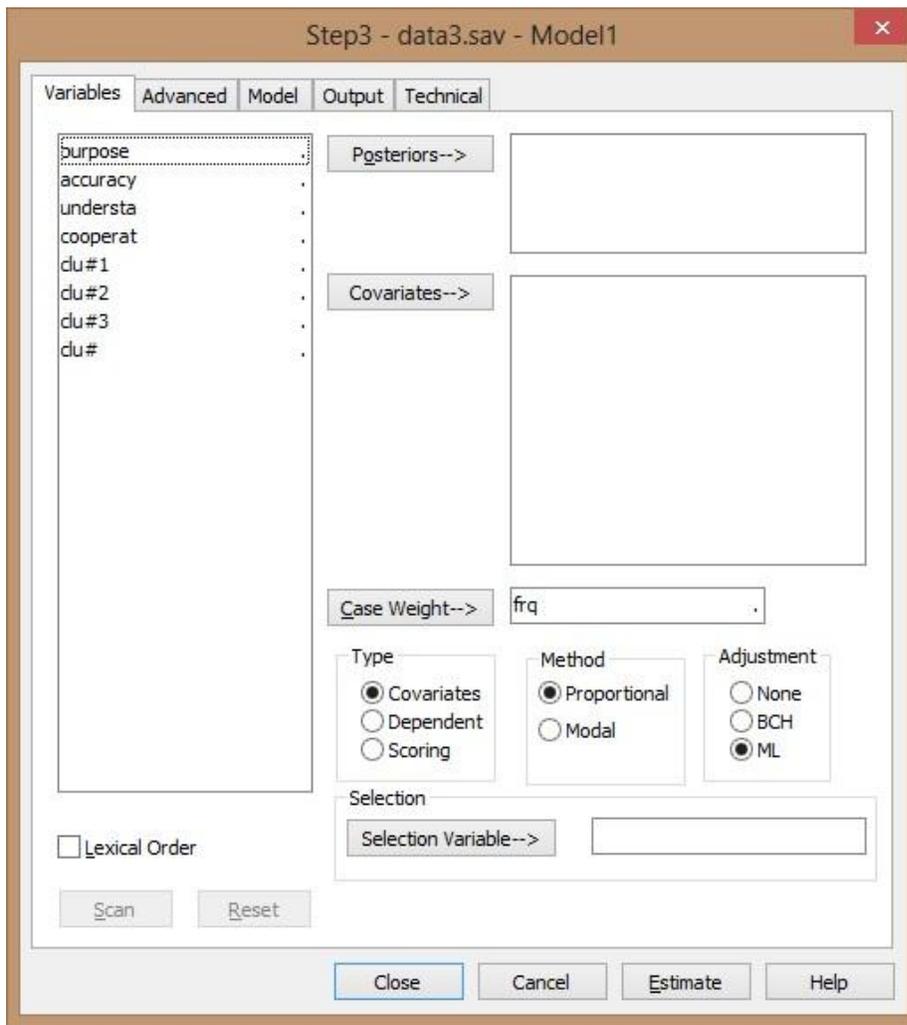
## Step 3 Tutorial #2: Obtaining equations for scoring new cases in a basic example with main affects

**DemoData = 'data3.sav'**

In this tutorial, we will show how to use the Step3 module in Latent GOLD to obtain an algorithm (equations) and related SPSS syntax for scoring new cases as a function of those same indicators. That is, the indicators will now be used as predictors in the regression equation.

Recall the output file, data3.sav, created in Latent GOLD Tutorial 1.

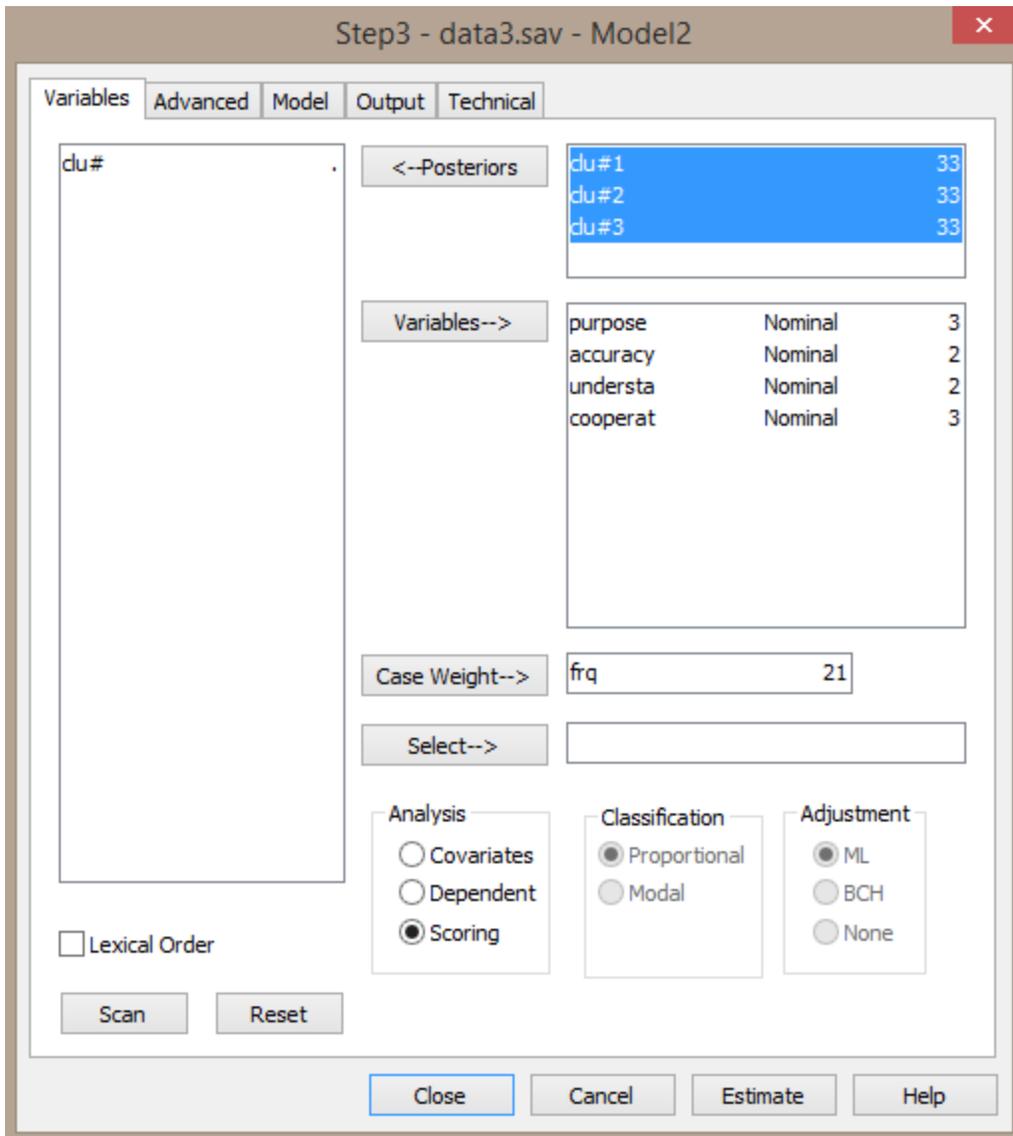
- Open data3.sav using “File→Open”
- From the menu bar, click ‘Model’ and select ‘Step3’. A dialog box will pop up (Figure 1)



**Figure 1.** Step3 Module

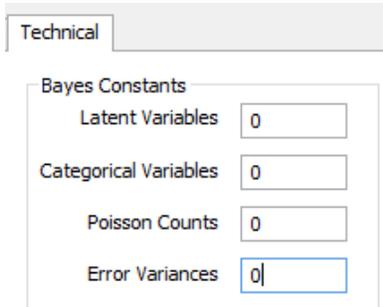
- For Type, select ‘Scoring’
- Click on the 4 indicators and click ‘Covariates’ to move them to the Covariates box

- Right click on the 4 predictors and select 'Nominal'
- Click on clu#1, clu#2, clu#3 and click 'Posteriors' to move them to the Posteriors box
- Click 'Scan'



**Figure 2.** Step 3 Variables Tab: Selecting the variables and the analysis type

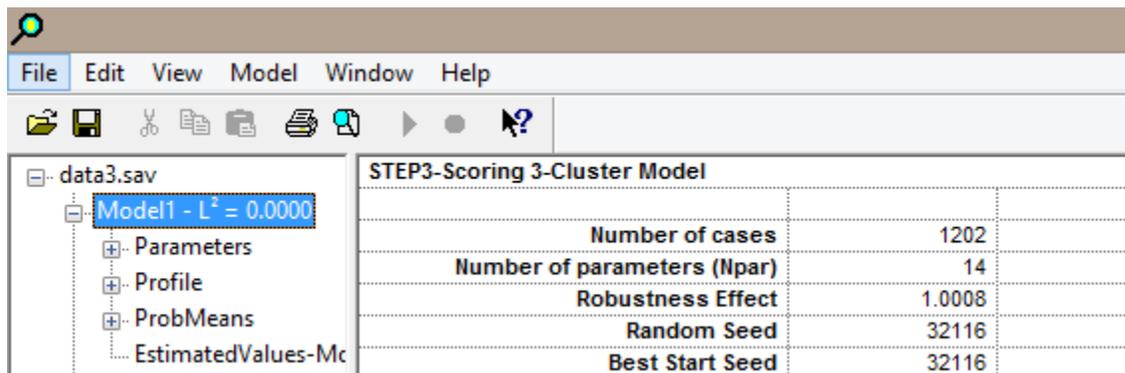
- Click on the 'Technical Tab' and set all Bayes constants to '0'.



The window should now look like Figure 2.

- Click 'Estimate'

Confirm that the  $L^2 = 0$  (Figure 3), which means that the posterior probabilities are reproduced perfectly as a function of the 4 predictors. The Profile and Probmeans output obtained here also reproduce the Profile and Probmeans output from the original 3-class model developed in Latent GOLD Tutorial 1. (Note that since we used the default Bayes=1 technical setting in the original 3-class model, the parameters output may differ in the 4<sup>th</sup> decimal place).



**Figure 3.** Step 3 Model output

The equations are provided in the Parameters output.

- Click Parameters to view the coefficients in the Parameters output (Figure 4)

Note: Ignore the other statistics as they are not relevant for the scoring procedure

Model for Classes						
Intercept	Cluster1	Cluster2	Cluster3	Wald	p-value	
	-2.0479	0.9315	1.1164	5.4249	0.066	
Variables	Cluster1	Cluster2	Cluster3	Wald	p-value	
<b>purpose</b>						
good	0.6762	1.0709	-1.7471	212.3499	8.3e-45	
depends	-0.4683	0.1954	0.2729			
waste	-0.2079	-1.2664	1.4742			
<b>accuracy</b>						
mostly true	0.5692	0.6510	-1.2202	78.6256	8.4e-18	
not true	-0.5692	-0.6510	1.2202			
<b>understa</b>						
good	1.7524	-1.3387	-0.4138	126.0971	4.2e-28	
fair/poor	-1.7524	1.3387	0.4138			
<b>cooperat</b>						
interested	1.9660	-0.8573	-1.1087	55.3125	2.8e-11	
cooperative	0.7000	-0.2535	-0.4466			
impatient/hostile	-2.6660	1.1108	1.5552			

**Figure 4.** Step 3 Parameter output

Each response pattern is scored on each cluster, and is assigned to the cluster with the highest score. For example, the response pattern:

Purpose = Good, Accuracy = Mostly True, Understanding = good, Cooperation = Interested

is scored based on the coefficients highlighted above. This results in the following scores:

Cluster 1 score = 2.916, Cluster 2 score = 0.457, Cluster 3 score = -3.373.

Thus, this response pattern is assigned to Cluster 1. To obtain more meaningful scores, we can generate the posterior membership probabilities using the formula provided below. This yields the following posterior membership probabilities associated with this response pattern:

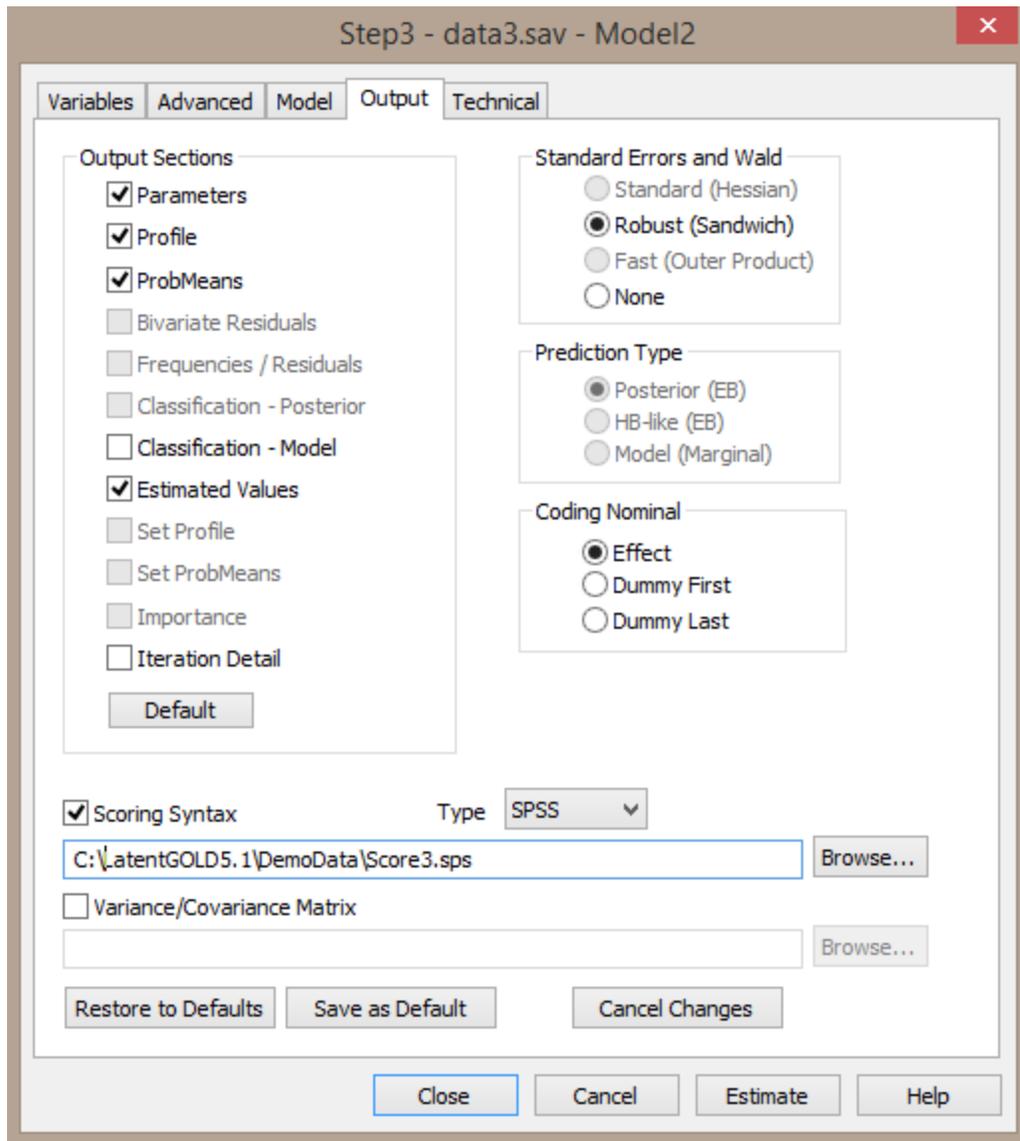
Posterior 1 = 0.9196, Posterior 2 = 0.0787, Posterior 3 = 0.0017

To generate the corresponding posteriors, use the following equation:

Posterior(k) =  $\exp[\text{score}(k)] / [\exp(\text{score}1) + \exp(\text{score}2) + \exp(\text{score}3)]$   
 k=1,2,3, which can be obtained in an SPSS syntax.

In order to get the SPSS syntax equations:

- Double click on Model 2
- Click on the Output tab
- At the bottom of the tab, check the box 'Scoring Syntax'.  
By default, the syntax type will be SPSS syntax (see Figure 5).



**Figure 5.** Step 3 Output tab: Requesting SPSS Syntax