

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

**DemoData = ‘data3.sav’ (or ‘data3\_copy.sav’)**

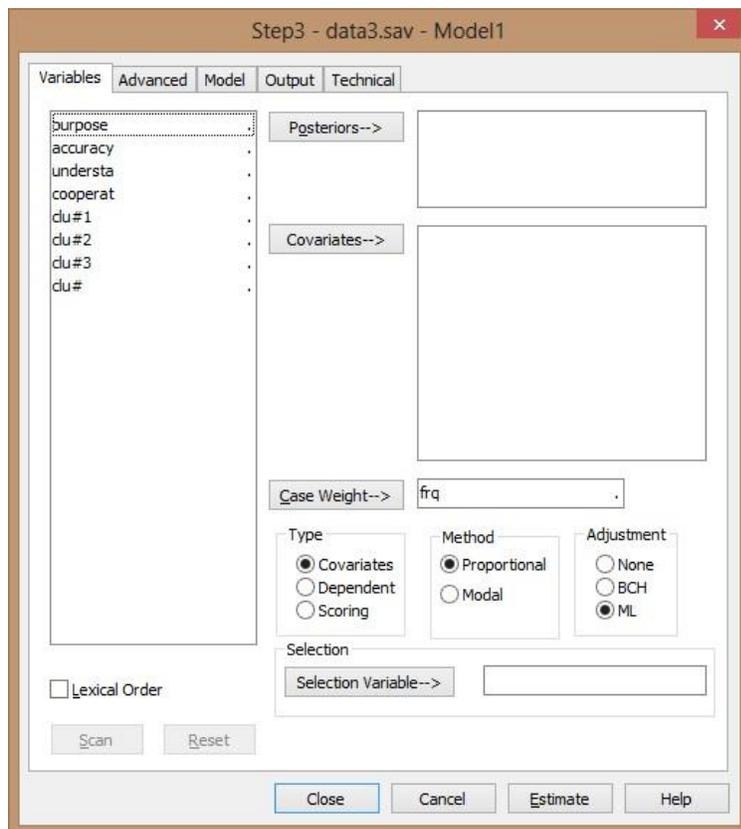
In this tutorial, we will show how to use the Step3 module in Latent GOLD to obtain equations for a scoring algorithm and related SPSS syntax for scoring new cases as a function of the indicators used to develop the latent classes. That is, the equations show how the indicators can be used to *predict* latent class membership.

### Obtaining scoring equations from Step3 Module

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

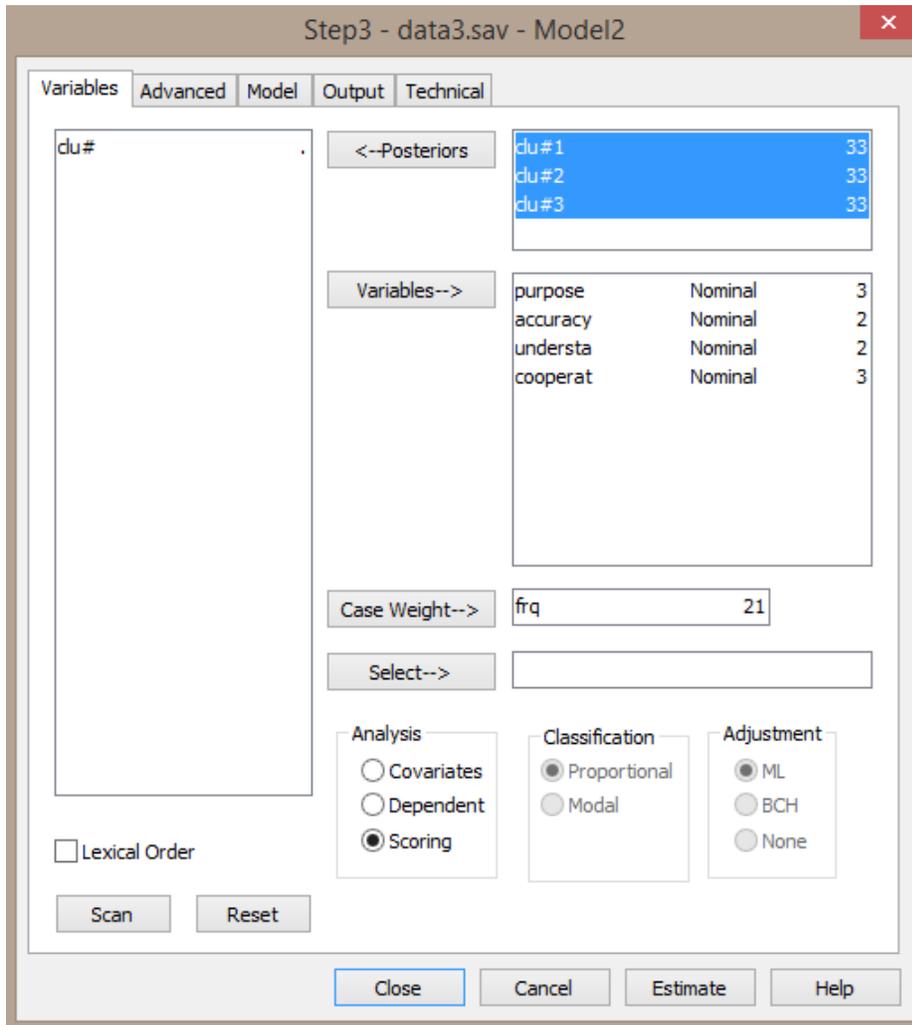
- Open data3.sav using “File→Open” (‘data3\_copy.sav’ works with the LG demo version)
- From the menu bar, click ‘Model’ and select ‘Step3’.

The following dialog box appears:



**Figure 1.** Variables Tab of Step3 Module

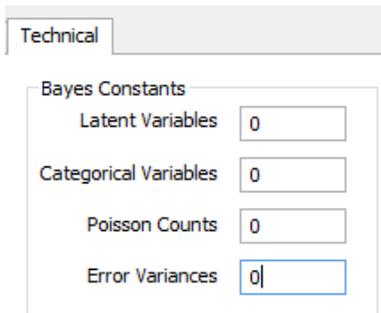
- For Type, select 'Scoring'
- Click on the 4 indicators and click 'Covariates' to move them to the Covariates box
- Right click on the 4 predictors and select 'Nominal'
- Click on clu#1, clu#2, clu#3 and click 'Posteriors' to move them to the Posteriors box
- Click 'Scan'



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

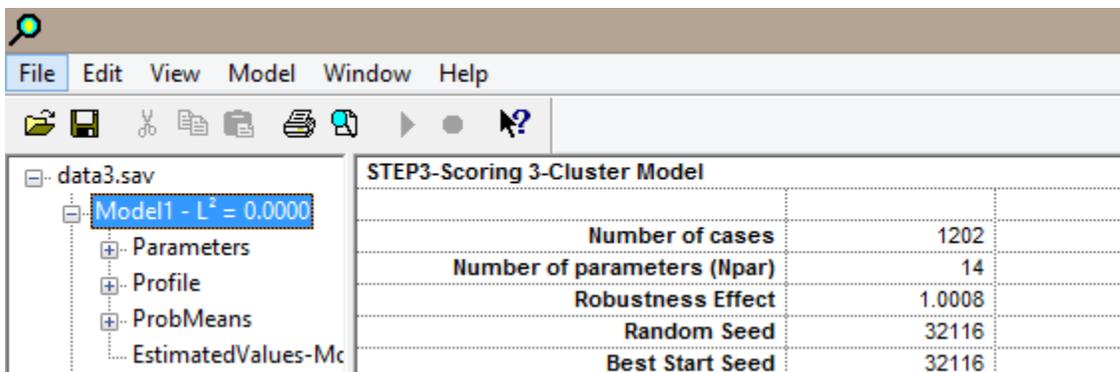
Your screen should now look like Figure 2.

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



- 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).

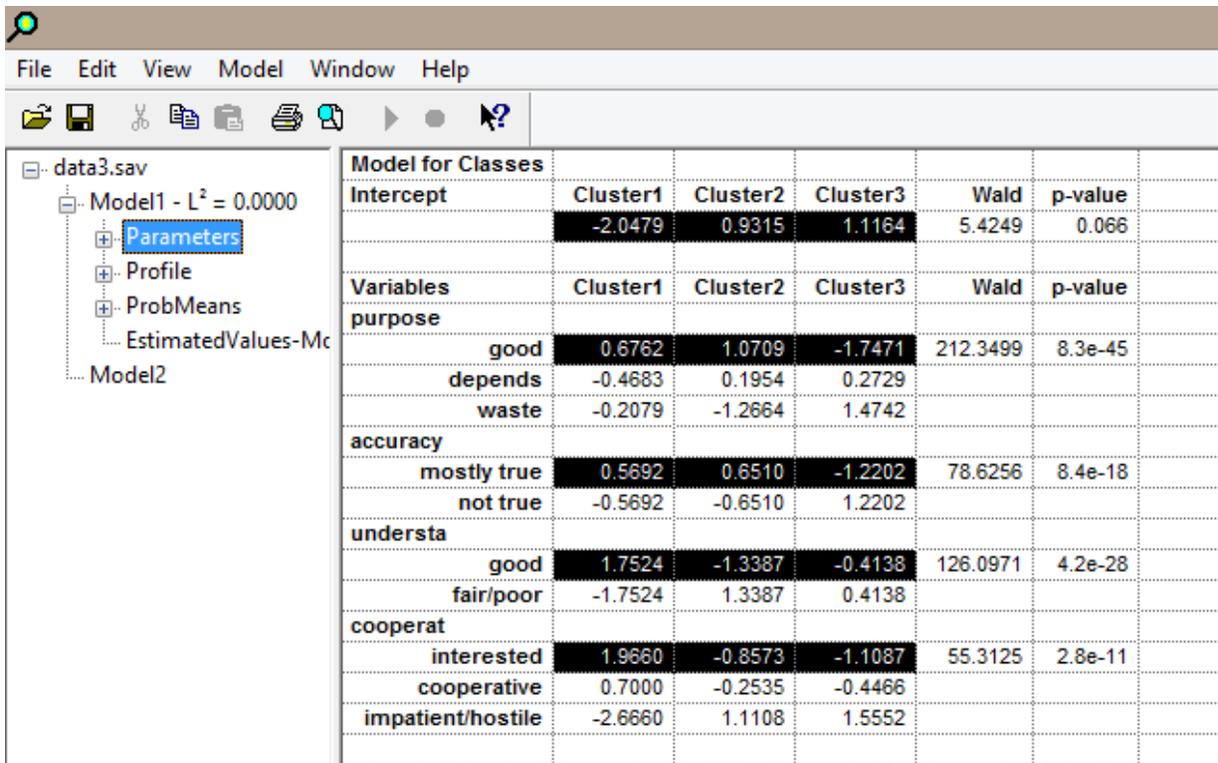


**Figure 3.** Step 3 Model output

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



**Figure 4.** Step 3 Parameters 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 translate these logit scores to posterior membership probabilities using the formula provided below. This yields the following posterior membership probabilities associated with this response pattern:

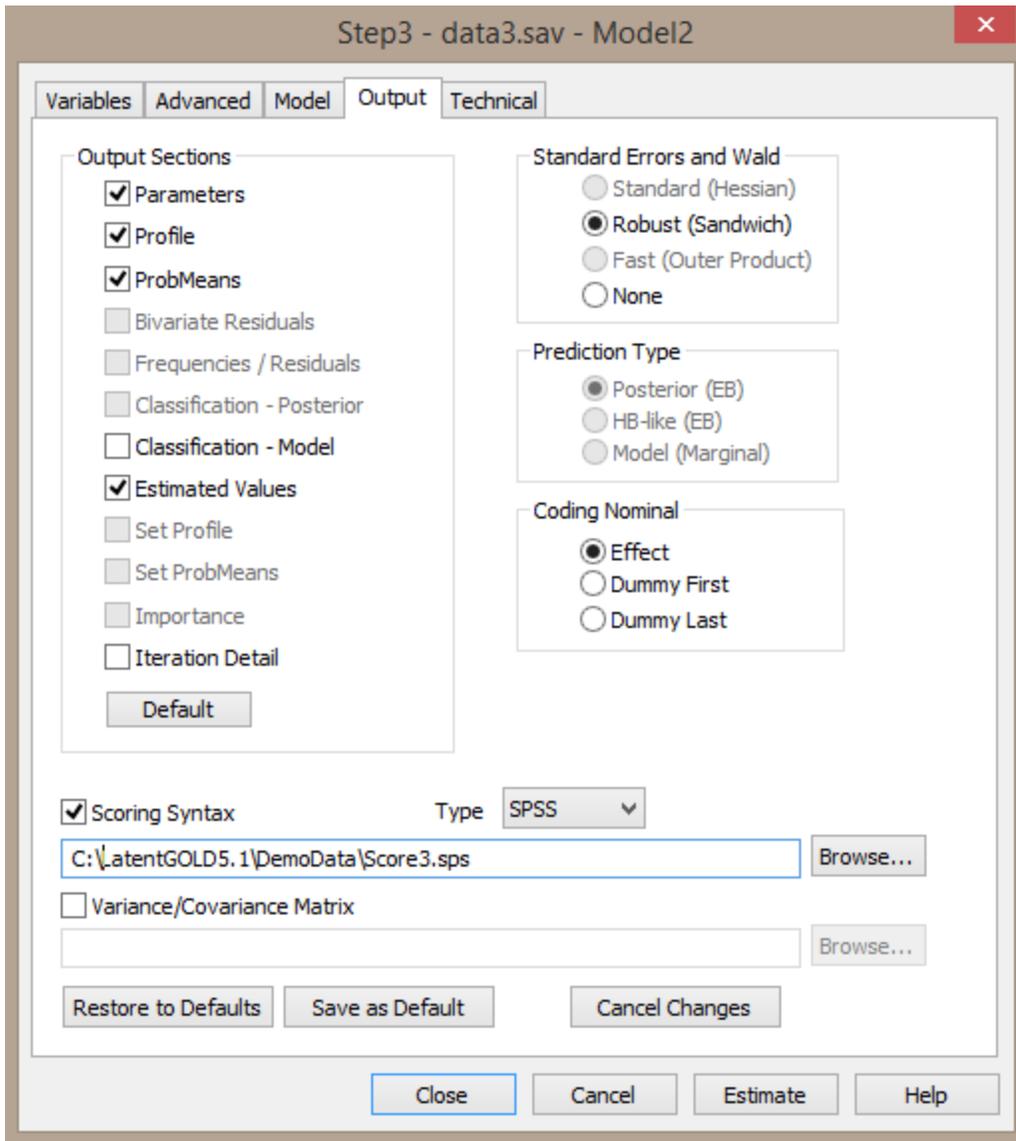
Posterior 1 = 0.9196, Posterior 2 = 0.0787, Posterior 3 = 0.0017

Here is the formula to obtain the posterior probabilities:

$$\text{Posterior}(k) = \frac{\exp[\text{score}(k)]}{\exp(\text{score}1) + \exp(\text{score}2) + \exp(\text{score}3)}, k = 1, 2, 3.$$

It is also possible to obtain SPSS syntax equations which compute the posteriors:

- 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).

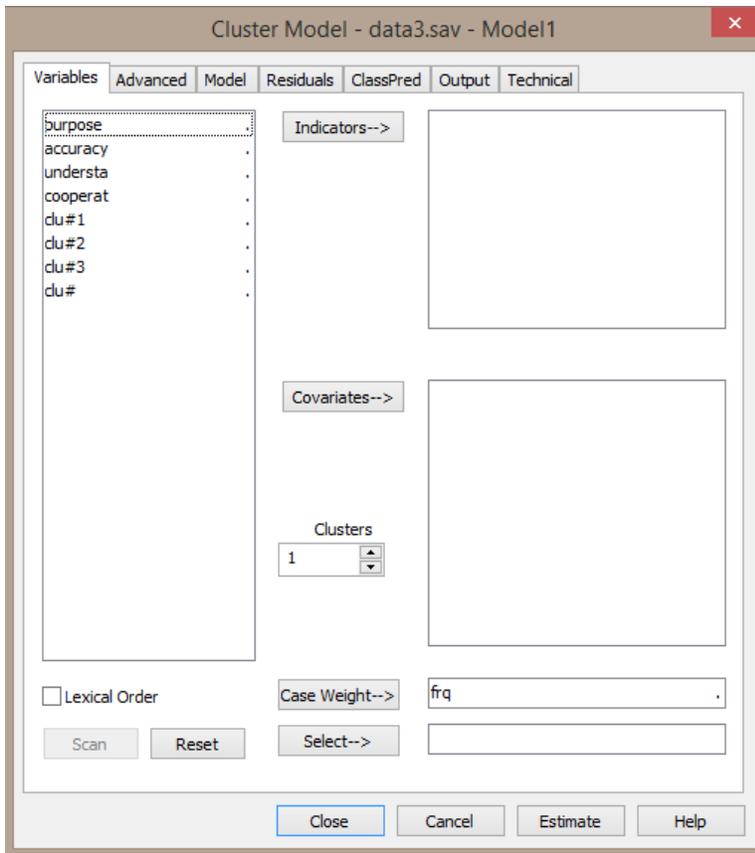


**Figure 5.** Step 3 Output tab: Requesting SPSS Syntax

## Using 1-Step Scoring to obtain Scoring Equations from Cluster Module

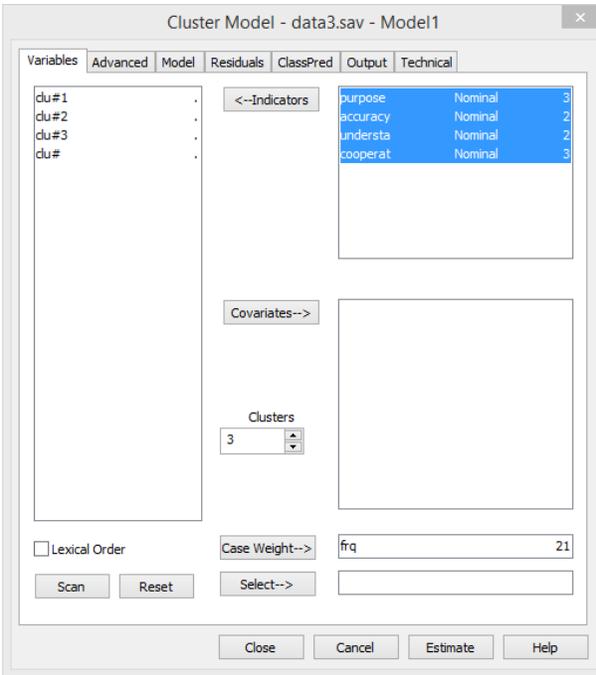
In addition to providing the scoring equations, the Step 3 module can also be used to relate the latent classes to other variables not used in the latent class model. A faster way to obtain these scoring equations is to use the Output option for LC Cluster models implemented in LG 5.1:

- Open data3.sav using “File→Open”
- From the menu bar, click ‘Model’ and select ‘Cluster’. A dialog box will pop up (Figure 6)



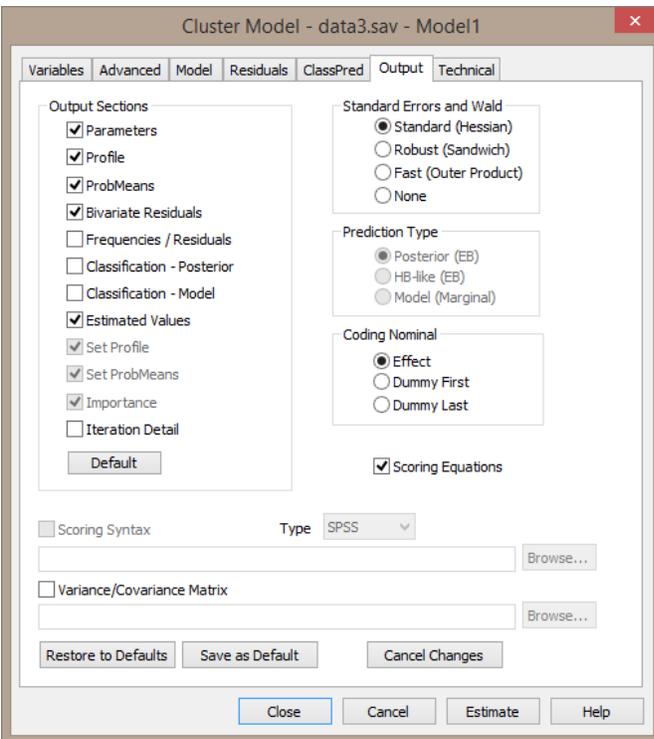
**Figure 6.** Variables tab from Cluster Module

- Move the 4 indicators to the Indicators box
- Right click on the 4 indicators and select ‘Nominal’
- Change the number of Clusters from ‘1’ to ‘3’
- Click ‘Scan’



**Figure 7.** Variables tab from Cluster Module

- Open the Output tab and select ‘Scoring Equations’ (see Figure 8)



**Figure 8.** Output tab from Cluster Module

- Click Estimate

- Expand the Parameters output to display the Scoring Equations

	Cluster1	Cluster2	Cluster3
<b>Intercept</b>			
	-2.0479	0.9315	1.1164
<b>purpose</b>			
good	0.6762	1.0709	-1.7471
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
not true	-0.5692	-0.6510	1.2202
<b>understa</b>			
good	1.7524	-1.3387	-0.4138
fair/poor	-1.7524	1.3387	0.4138
<b>cooperat</b>			
interested	1.9660	-0.8573	-1.1087
cooperative	0.7000	-0.2535	-0.4466
impatient/hostile	-2.6660	1.1108	1.5552

**Figure 9. Scoring Equations**

Confirm that these coefficients are identical to those obtained using the Step3 Module (recall Figure 4).