Latent GOLD® Choice 5.0 tutorial #11A (3-file format)

How to Estimate Scale-Adjusted Latent Class (SALC) Models and Obtain Better Segments with Discrete Choice Data

Introduction and Goal of this tutorial

Researchers frequently use Latent Class (LC) choice models for strategic segmentation and targeting purposes to

1) find meaningful segments of respondents having different preferences, and
2) estimate part-worth utilities for these segments.

However, LC models can form spurious segments (such as Class 4 in **Figure 1** below) that differ primarily in terms of scale (response error) but don't differ much in terms of real preference patterns.
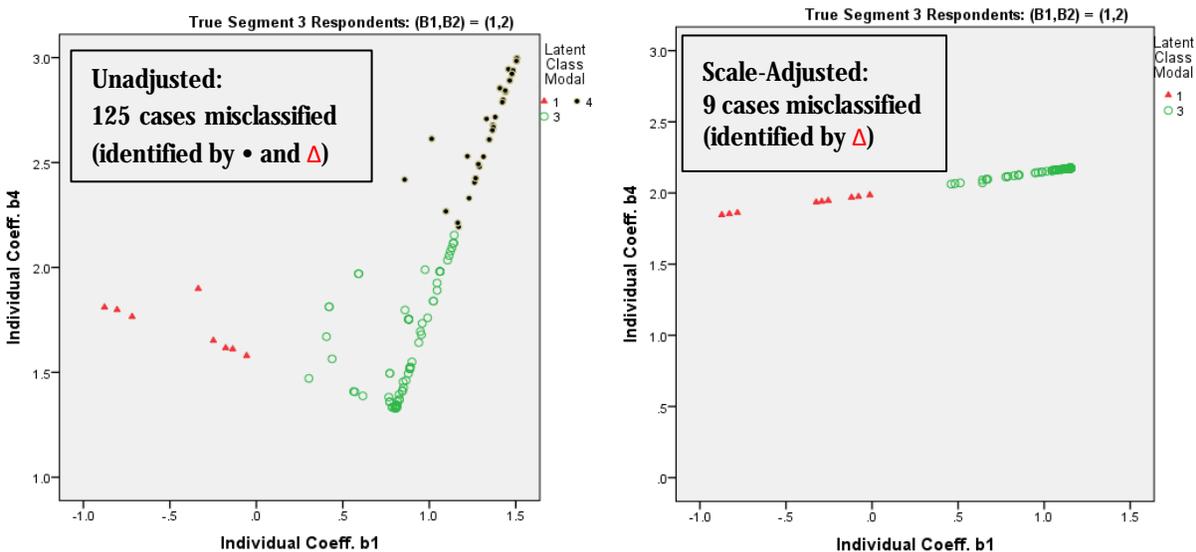

Figure 1. For a true homogeneous population (Segment 3) with preference part-worth utilities B1=1 and B4=2, the standard solution ("Unadjusted") shows much more error variation and greater misclassification than the "Scale-Adjusted" solution.

In this tutorial we illustrate a simple model-fitting strategy that can be used with the syntax module of Latent GOLD® Choice 5.0 to estimate continuous scale Scale-Adjusted LC (SALC) models which adjust for differences in scale, resulting in more meaningful segments that differ only in preference.

The approach is illustrated using responses simulated based on 3 homogeneous segments (3 latent classes) that differ according to their preference part-worth utilities, and also in their error variance (simulated with a normally distributed latent factor to account for continuous scale heterogeneity). That is, each respondent has a unique error variance (unique scale factor) and belongs to 1 of 3 LC segments.

We will also see that an almost identical result can be obtained with the point-and-click interface of Latent GOLD 5.0 Advanced with a SALC model that assumes that respondents have either low or high scale factors (i.e., using 2 latent scale classes rather than a continuous scale factor).

SALC models can be estimated in LG Choice 5.0 with observed and/or unobserved (latent) scale heterogeneity (Magidson and Vermunt 2007). The former can be done by allowing the scale factor to depend on observed variables (e.g., time to complete task, stated vs. revealed preference), the latter by including latent scale classes or latent continuous factors in the model. In this tutorial we illustrate the latter situation (latent continuous scale heterogeneity).

For the example used in this tutorial, we assume that there are three latent class segments differing in their part-worth utilities (see Table 1) but the responses from some individuals exhibit less consistency about their preferences than others (higher response error).

Table 1. Part-worth Utilities for Low Error Cases (Reference Scale Factor = 1).

| | Reference (Scale Factor = 1) | | |
	Segment 1	Segment 2	Segment 3
Attribute 1, Level 1 (b1)	-1	0	1
Attribute 1, Level 2 (b2)	0	-3	0
Attribute 1, Level 3 (b3)	1	3	-1
Attribute 2, Level 1 (b4)	2	-1.5	2
Attribute 2, Level 2 (b5)	0	0	0
Attribute 2, Level 3 (b6)	-2	1.5	-2

All respondents plotted in the left-most plot shown in Figure 1 (True Segment 3 population respondents) have true preference part-worth utilities of 1 for b1 and 2 of b2 as shown in Table 1. As we will show in this tutorial, non-SALC models result in a 4-class solution, with only 110 of the 200 true Segment 3 respondents assigned to the same class (Class 3), the remaining 90 cases being assigned to class 1 or class 4 as shown in Figure 1.

Background

As explained by Louviere and Eagle (2006), standard discrete choice models suffer from a potential confound between scale factors and utilities -- the larger the response error, the smaller the magnitude of the part-worth utilities; the smaller the response error, the larger the magnitude of the part-worth

utilities. As a consequence, two segments of respondents could have essentially the same preferences, but appear to have different utilities due to a stretching or shrinking of the utilities (the scale factor).

SALC models were first introduced by Magidson and Vermunt (2007) and implemented in the LG Choice Syntax version 4.5 in a somewhat complex manner. Since that time, several applications have appeared in the literature (see e.g., Burke et al., 2010). Version 5.0 of LG Choice simplifies the earlier implementation, allowing basic SALC models to be setup and estimated with the point-and-click interface. Moreover, the current implementation allows greater user control via the new log-linear scale model. This model estimates the log-scale factor rather than the scale factor directly, thus guaranteeing that the scale factor is always non-negative.

For further details, see the LG Choice 5.0 Upgrade Manual and references cited there.

Simulated Data

Simulated Response Data: 'SALC_Sim1sCFac_Resp.sav' (see Figure 4).

Note: The design consists of 9 alternatives which differ on 2 attributes, each having 3 categories. Using these 9 alternatives, 12 choice sets are created, each consisting of 3 alternatives: See 'SALC_Alt.sav' and 'SALC_Set.sav'.
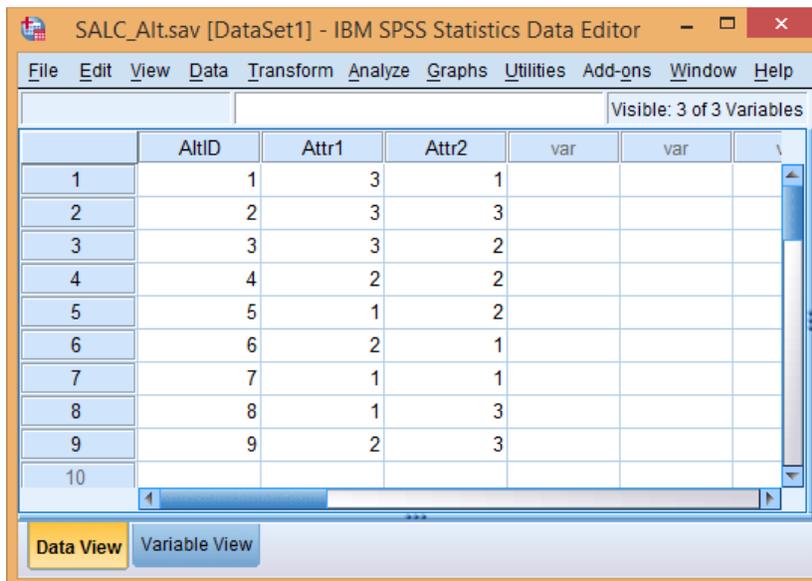
Figure 2. Alternatives file showing how the 9 alternatives differ on the attributes Attr1 and Attr2.

Figure 3. Sets file: Each of 12 choice sets contain 3 of the 9 alternatives.

This simulation ('sim#'=1), was performed using the parameters shown in Table 1). The response file below contains the responses of 600 individuals ('Choice') for each set, as well as their true segment ('True') and scale class ('sclss').

Figure 4. Response File showing the selected response (Choice) for case ID=1 on each of the 12 choice sets.

Estimating Models with LG Choice Syntax Module

We will begin by opening previously saved sCFac SALC models.[1]

In LG Choice:

∅ **Click on File à Open.**

∅ **In the Open window, click on the 'Files of type' drop down menu and select 'LG Syntax files(*.lgs)' (see Figure 5).**

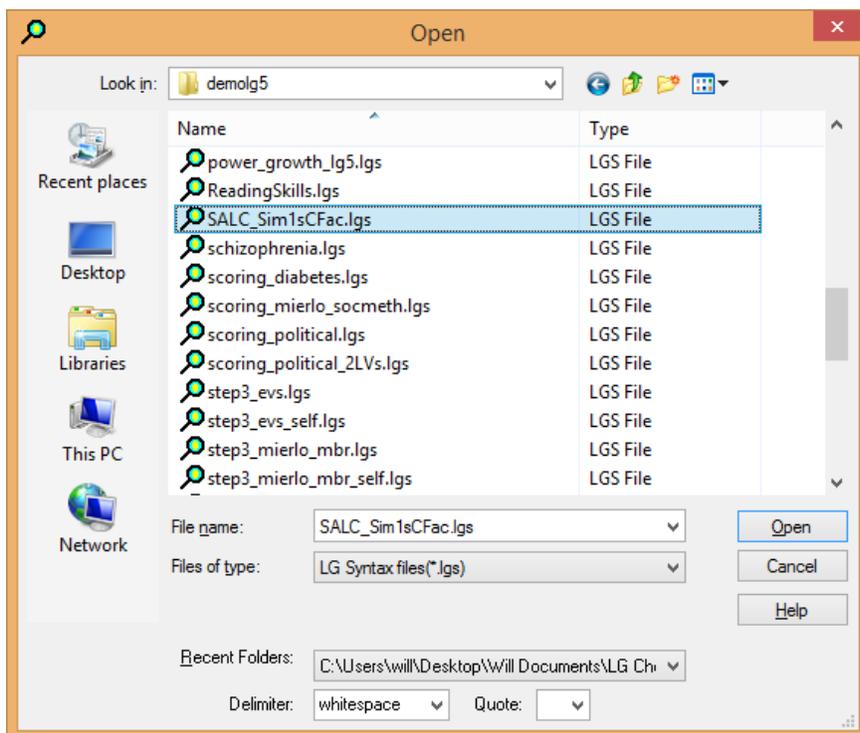∅ **Click on the file 'SALC_Sim1sCFac.lgs' and then click 'Open'.**

Figure 5. Opening previously saved models.

Our 11 previously saved models appear in the left pane (see Figure 6).

---

[1] For a step-by-step tutorial on how to 1) set up a model in LG Choice using the 3-file format, 2) determine the number of classes, and 3) save model definitions in a .lgf file, please see Tutorial 1.
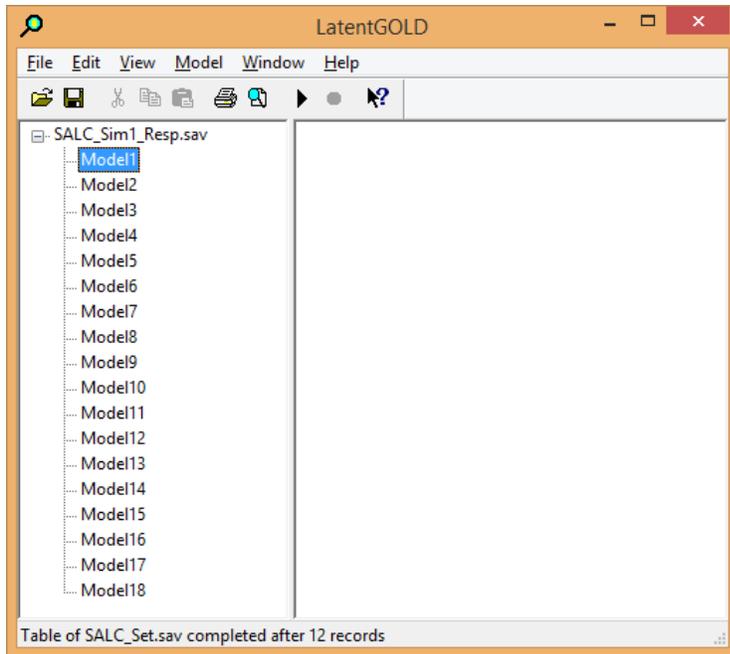
Figure 6. Previously saved models.

⊘ **Left click on 'SALC_Sim1sCFac_Resp1.sav' in the Outline pane.**

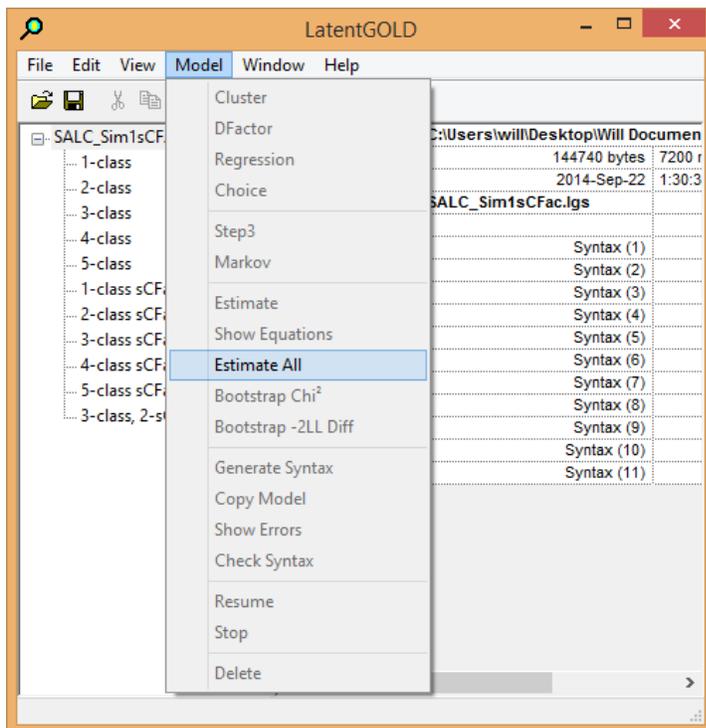⊘ **From the Toolbar, click 'Model' and then click 'Estimate All' (Figure 7).**

Figure 7. Estimating all saved models.

After estimation has completed, results for all estimated models appear in the Outline pane.

From the model labels (Figure 8) it can be seen that we estimated SALC models with 1-5 classes that have a continuous scale class (Models 6-10), 1 3-class, 2-sClass model (Model), and 5 non-SALC models that are homogeneous with respect to scale (Models 1-6).

∅   Click on the data file name 'SALC_Sim1sCFact_Resp.sav' in the Outline (left hand) pane to compare the fit of these models by viewing the Data File Summary output (see Figure 8).
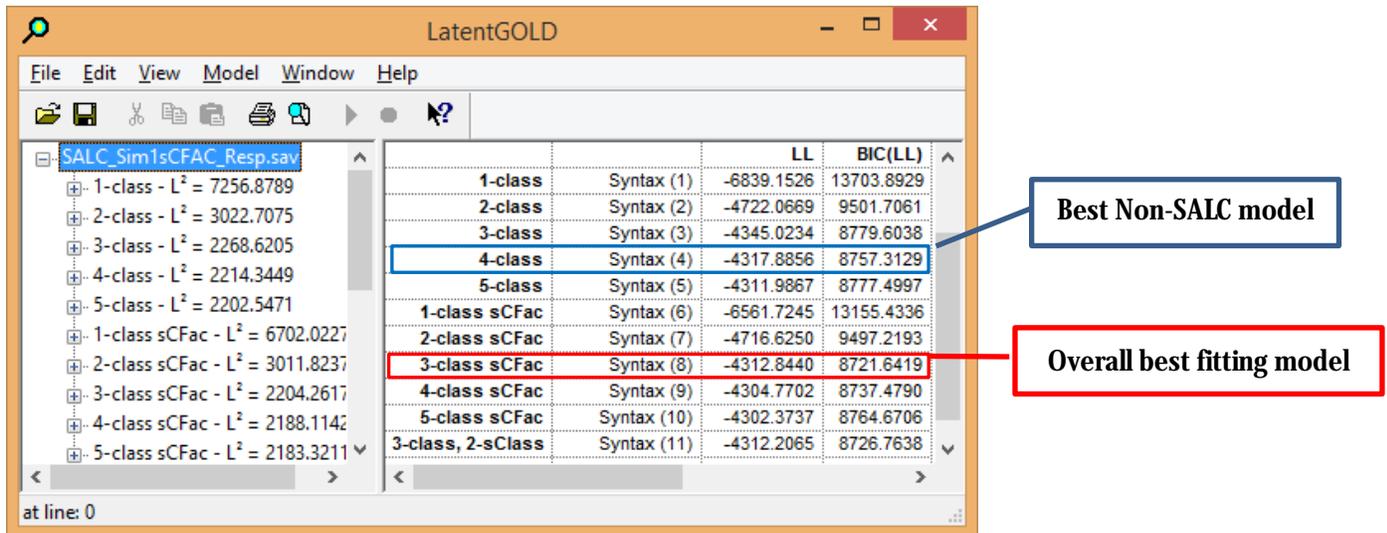

Figure 8. Data File Summary Output.

From this summary, we obtain the following results:

- Overall, the 3-class sCFac SALC model ('Model8') fits best (lowest BIC).
- Not accounting for scale heterogeneity (non-SALC Models 1-6), results in the 4-class model as best (Model 4).

∅   Click on '3-class sCFac' to view the setup for this model (Figure 9).
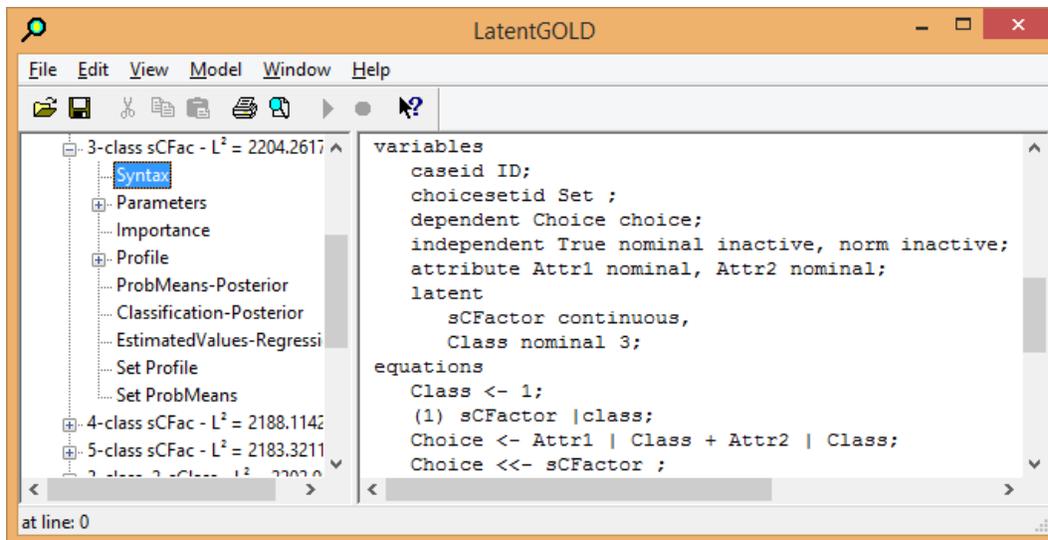
Figure 9. Model Setup for the 3-Class, sCFac model.

∅ Click on 'ProbMeans-Posterior' to view the Probmeans output for the '3-class sCFac' model to view the setup for this model (Figure 10).

Figure 10. Probmeans output for the '3-class sCFac' model.

References

Burke, P., C. Burton, T. Huybers, T. Islam, J. Louviere, and C.Wise (2010). The scale-adjusted latent class model: application to museum visitation. Tourism Analysis, Vol.15, pp.147-165.

Louviere, J. J., and T. Eagle (2006). Confound it! That pesky little scale constant messes up our convenient assumptions. Proceedings, 2006 Sawtooth Software Conference, Sawtooth Software, pp.211-228.

Magidson, J., and J.K. Vermunt (2007). Removing the scale factor confound in multinomial logit choice models to obtain better estimates of preference. Proceedings, 2007 Sawtooth Software Conference, Sawtooth Software, pp.139-154.

Magidson, J., T. Eagle, and J.K. Vermunt (2003). New Developments in Latent Class Choice Models. Proceedings, 2003 Sawtooth Software Conference, Sawtooth Software, pp. 89-112.