

CBC/HB Model Explorer v1.1

Documentation

Introduction

- Do you want to search¹ for the prior variance and prior degrees of freedom settings that best fit the heterogeneity in your CBC or MaxDiff dataset?
- Do you want to explore all first-order (between two attributes at a time) interaction effects for CBC or ACBC data sets to see if any can provide a lift in predictive validity over your base main effects model?

At the 2016 SKIM/Sawtooth Software Conference, Bryan Orme and Walter Williams of Sawtooth Software presented evidence from their analysis of around 50 commercial data sets that the default priors (Prior Variance and Prior Degrees of Freedom) in CBC/HB software in place at that time² typically *overfit* the data for CBC studies and slightly *underfit* the data for MaxDiff studies. Orme and Williams also found that most CBC data sets do not benefit from additional interaction effects beyond main effects model specifications (though a couple data sets *did* benefit from specifying an additional interaction effect). They used this *CBC/HB Model Explorer* to conduct the research, which would have been extremely difficult and time-consuming to do with the commercial CBC/HB software alone.

What makes the *CBC/HB Model Explorer* unique is that it repeats the HB analysis multiple times (across multiple replicates) where for each replicate it jack-knife resamples across existing “random” (experimental design) choice tasks for individual-level hit rate validation rather than requiring “fixed” holdouts. Thus, any CBC or MaxDiff data set may be used, whether or not you included fixed holdouts.

The Model Explorer searches for optimal priors and interaction effects in an automated way that runs *much* faster than standard CBC/HB. It interacts with your existing CBC/HB v5 via the command interpreter to spin up multiple simultaneous HB runs across the available cores on your processor.

The CBC/HB Model Explorer is a “power tool” (less polished interface compared to our commercial applications) targeted for relatively advanced Sawtooth Software users. It uses a basic software interface and can leverage additional command interpreter instructions that are described in the current CBC/HB manual (Appendix N).

¹ We recognize that searching for priors using the data is not formerly acceptable among statistical purists. However, practitioners often are more pragmatic and the jack-knifing holdout procedure as describe herein seems a reasonable approach for obtaining better priors to improve modeling performance under HB.

² The default priors at that time were PriorVar=2, DF=5 for CBC data sets and PrioVar=1, DF=5 for MaxDiff data sets.

Using the Model Explorer

Installation:

- First, you must have a current copy of CBC/HB (v5.5.4 or later) installed on your Windows-enabled computer.
- Unzip the contents of the *CBC/HB Model Explorer* into its own directory on a convenient place on your hard drive.
- The Microsoft .Net framework v4.6.1 is required.

Running the software:

- Browse to the folder where you have unzipped the contents of the *CBC/HB Model Explorer*.
- Double-click the **ModelExplorer.exe** file and the dialog is displayed.
- When you have implemented all the settings as desired on the dialog, click the **Explore** button at the bottom of the dialog.

Notes: Typical runtime is 1 to 3 hours for most datasets for just the optimal priors search (but this depends a great deal on the number of cores of the processor you are using). You will probably want to run the Model Explorer program on a separate machine from where you are working since it spawns active console windows (one per core of your processor) for each replication that will interrupt your current work in other windows. There is no *Cancel* button. To cancel a search run, close the **ModelExplorer.exe** window by clicking the **X** at the upper-right, then shut down each console Window by clicking the **X**.

About Using the Model Explorer with ACBC Data

ACBC data may be exported to .CHO files and under proper conditions (scaling the Summed Pricing attribute to be in the single digits) may be analyzed using CBC/HB software. Thus, it is possible to use the *Model Explorer* to search for the proper priors or for interaction terms for ACBC data. ***We do think that it is useful to use the Model Explorer to search for significant interaction terms (that can lift hit rate) for ACBC data sets, but we caution against using it to search for the prior variance and degrees of freedom that maximize hit rates to randomly held out ACBC tasks.*** The reason is that many of the coded choice tasks in ACBC .cho files either involve or assume very low response error (e.g. BYO tasks, the often multiple inferred choice tasks that are coded as a result of the respondent selecting “must haves” or “unacceptables,” or augmented binary tasks that inform utility estimation that dropped levels are inferior to included levels in the case of constructed lists of levels). Because the *Model Explorer* randomly selects among *all* tasks to hold out for validation, the search procedure would favor HB solutions with relatively high prior variance and very little Bayesian smoothing to the hyperparameters. If we could isolate only the Tournament (CBC-looking) choice tasks for selection as holdout tasks, then the results of optimal priors searches should be reasonable. But, at this time, we haven’t developed that level of sophistication into the *Model Explorer*.

Settings in the Software Interface:

Input Files (*.cho)

☒ Explore a single data file

☐ Explore all data files within a folder

Browse...

☐ The data file contains MaxDiff data (the Prior Covariance Matrix will be set accordingly)

Input Files (*.cho): You can conduct your analysis on just a single .cho data file or across multiple .cho files (e.g. from different CBC and/or MaxDiff studies) if multiple .cho files are placed within the same folder.

Use the **Browse...** button to browse to the single .cho file or the folder containing multiple .cho files to be analyzed.

The data file contains MaxDiff data: If the data file(s) you are using have MaxDiff data, click this box. For MaxDiff data, the prior covariance matrix that is appropriate for dummy-coding will be set automatically by the software. The different prior variance values that the software searches for you are multiplied by the prior covariance matrix.

Sampling

☒ Use sample size same as the original

Sample size

☐ Use sampling with replacement

Holdout count

☒ Automatically determine replication count using ($r \times \text{respcount} \times \text{holdoutcount} \sim 20000$)

Sampling replication count

The **Sampling** settings control how the replicates are constructed (sample size to draw, whether to conduct sampling with or without replacement, how many of the experimental tasks to hold out for validation, and how many replicates to perform).

Use sample size same as the original: If you click this box, sample replicates are drawn with equal sample size as the original data file (recommended). For example, if you have 502 respondents in the .cho file, each replicate that the *Model Explorer* builds will also have 502 respondent records.

Use sampling with replacement: If you click this box, each replicate sample is built wherein respondents are sampled *with replacement*, meaning a given respondent could be represented multiple times in the sample replicate or no times at all. Sampling with replacement is often called *bootstrap resampling*. Using bootstrap samples leads to an additional source of sampling error in the replicates and thus requires that you conduct more replicates to obtain equal summary hit rate precision as when sampling without replacement. It might be argued that using bootstrap sampling can make the results of the search algorithms more generalizeable and predictive of out-of-sample observations. But, some researchers reacting to Orme and Williams work have suggested that bootstrap sampling (sampling with replacement) for the purposes of searching for optimal priors is an unnecessary procedure that just adds an additional source of error to the hit rates from each replicate.

Holdout count: This tells the software how many holdout tasks to randomly select and hold out per respondent. For example, if there are 12 total tasks in the .cho file and the holdout count is 2, then 2 of the 12 tasks are randomly held out per respondent for raw hit rate evaluation (a different 2 for each respondent). The other 10 tasks are used to estimate the model. We recommend you use the vast majority of the tasks to estimate the part-worth utilities and very few of the tasks to hold out each time for validation. We recommend generally that the holdout count be just 1 or 2. With fewer holdouts, more replicates need to be done to stabilize the hit rates, so there is a tradeoff between time to conduct the search procedure and the holdout count. The smaller the holdout count, the longer the run time.

Automatically determine replication count using ($r * \text{respcount} * \text{holdoutcount} \sim 20000$): This checkbox controls the number of replicates that are used for estimating HB utilities and computing raw hit rates on the jack-knifed holdout tasks. Each replicate leads to different part-worth utility estimates per individual and different hit rate scores for the held out tasks. One must repeat the analysis across multiple replicates and average across the resulting hit rates to stabilize the holdout hit rate statistics. We recommend selecting a number of replicates (r) such that $r * \text{respcount} * \text{holdoutcount}$ is approximately equal to 20,000 (where r is the number of replicates, *respcount* is the number of respondent records sampled in each replicate, and *holdoutcount* is the number of tasks held out for hit rate validation for each respondent in each replicate). This recommendation assumes you are sampling without replacement for the replicate data sets. However, if you check the box **Use sampling with replacement** this bootstrap resampling procedure adds an additional source of noise to the hit rate estimates and you should compensate by using even more iterations (for example, such that the $r * \text{respcount} * \text{holdoutcount} = 30000$).

You may override this default setting by specifying your own custom number of replicates in the *Sampling replication count* field.

Base Model

List of attributes to code as User-Specified (use numbers separated by commas)

List of interactions to use in base model (use NxN, with multiple separated by commas)

List of attributes to code as User-Specified: ACBC data sets can have “user-specified” attributes (for the “Summed Price” attribute and also the None parameter). Also, some researchers like to generate or modify their .cho files to include “user-specified” coding. To specify multiple attributes as user-specified, specify multiple integers here separated by commas. For example, if you have an ACBC data set where the 10th and 11th attributes are the Summed Price and None columns, then specify “10, 11” in this field.

List of interactions to use in base model: If you would like to specify certain interactions to be included in your base model, specify them in this field. For example, to include the interaction between attributes 1 and 4 in the base model, specify “1x4”. To specify the interactions between attributes 1&4 and 1&3, specify “1x4,1x3”.

Hierarchical Bayes Settings

Path to Command Interpreter (CBCHBCon.exe)

Browse...

The *Model Explorer* searches your hard drive for the **CBCHBCon.exe** file. If for some reason it cannot find it, you must **Browse...** and select the **CBCHBCon.exe** file.

Number of burn-in iterations

Number of saved iterations

The number of burn-in and saved iterations used to analyze each replicate is specified here. Sentis & Li (2000) found that very few iterations were needed to obtain near-optimal hit rates (though convergence has typically not yet occurred yet). Because our search procedure uses the goal of raw hit rates, we can take advantage of the findings of Sentis and Li to analyze many more replicates much faster. We’ve found that the hit rate will increase *slightly* by running 20K + 20K iterations vs. 1K + 5K iterations, but our experience so far is that the patterns of hit rate success pointing to the optimal priors or valuable interaction terms don’t change much if one uses much longer HB runs within each replicate (as this tool is quite new, we recommend you investigate your data set further with larger burn-in and saved iterations if in doubt). ***Of course, when you conduct your final model for delivery to the client, you should use many more burn-in and used iterations to obtain convergence!***

☒ Perform search for optimal prior variance and degrees of freedom

Candidate Prior Variances 0.2, 0.5, 1.0, 1.5

Candidate Prior Degrees Of Freedom 2, 5, 50

In the fields above, you specify the specific values to use in the grid search. In the example above, the candidate prior variances to test are 0.2, 0.5, 1.0, and 1.5. The candidate prior degrees of freedom to test are 2, 5, and 50 (note: degrees of freedom must be an integer at least 2 or larger). All combinations of these values ($4 \times 3 = 12$ combinations) are tested in the grid search. If you have a reasonably good idea about what prior settings are near-proper for your data set, you may be able to specify a tighter range and fewer increments to search (which saves you time).

We recommend you refer to the findings of Orme and Williams (2016) when setting the ranges and increments to search. Please refer to their PowerPoint slides (included in the .zip file for this software).

Below are a few key summary tables from Orme and Williams summarizing the findings of their research across ~50 commercial data sets:

► For full-profile CBC (showing 5 or fewer concepts per task):

	Prior Var
≥ 10 attributes	0.2 <i>Note: Set prior DF at least 10, even for small sample sizes</i>
7-9 attributes	0.3 <i>Note: Set prior DF at least 10, even for small sample sizes</i>
5-6 attributes	0.5 <i>Note: Set prior DF at least 5, even for small sample sizes</i>
3-4 attributes	0.8
2 attributes	1.5

For highly alternative-specific designs, consider the number of attributes applying to the typical concept, then follow recommendations as above.

► For shelf-display CBC with very few attributes but many levels of single attribute (e.g. ≥ 20) and showing 10 or more concepts per tasks:
set Prior Var=1.2

► If MaxDiff: Set Prior Var=1.3

► Set Degrees of Freedom depending on sample size:

Sample Size	DF
0-200	2
200-400	5
400-700	10
700-1200	30
1200-2400	50

☒ Perform interaction search

If you want to use the *Model Explorer* to search across all potential first-order interaction effects (interactions between attributes taken 2 at a time), click this box.

If you specified interaction effects already as part of the base model, these interaction effects will be left in the base model and skipped over when searching all potential *other* first-order interaction effects.

If you conduct both optimal priors and interaction searches (where you are searching across at least 3 values of prior variance and 3 values of degrees of freedom), the optimal priors search will be conducted first and then the interaction search will be performed *using the optimal prior variance and degrees of freedom found in the optimal priors search*.

If you want to perform an interactions search with a specific prior variance and degrees of freedom, you must specify it in the *Extra interpreter commands* field. To specify prior variance and degrees of freedom (in the example below 0.2 and 3, respectively), specify:

```
SetPriorVariance 0.2
SetPriorDegreesOfFreedom 3
```

Extra interpreter commands

Extra interpreter commands: If you want to include such things as covariates, you can specify these settings using the Extra Interpreter commands (described in Appendix N of the CBC/HB documentation).

☒ Clean up any HB estimation output files when finished

☐ Clean up replication sample files and interpreter scripts

Random Seed

Seed

Clean up any HB estimation output files when finished: By “Clean up” we mean to delete the files after the program has finished running the search. Running CBC/HB leads to many files written to your hard drive, some of which are very large (such as the covariances and alpha draws files) and very few people ever need to use them. By default, we automatically delete the files that we think most users will never refer to when running searches with this tool. However, if you want to include those files in the output folder, you can uncheck this box to choose to not clean them up.

Clean up replication sample files and interpreter scripts: By “Clean up” we mean to delete the files. For each replicate the Model Explorer creates a new data set (a .cho file) and file containing the held out tasks. There is usually no reason to look at these files, so the user may choose to have the program automatically delete them upon completion of the search.

Random Seed: You can specify a new random seed (an integer) to rerun the analysis and obtain a new result. The seed is used not only for seeding the HB algorithm for the different replicates but for randomization involving the jack-knife and bootstrap resampling.

Interpreting the Output of the Model Explorer

The *Model Explorer* spins up multiple command interpreter prompts to run CBC/HB across multiple cores of your processor. When it finishes, a series of files are written to the same folder containing the .cho file(s) that you used for estimation:

Optimalpriors.txt: *(This file only written if your grid search for optimal priors is at least 3x3 dimension.)* A summary report of the hit rates achieved (averaged across replicates) for each cell in the grid search is displayed. The optimal hit rate is reported based on your grid search (the prior variance and prior degrees of freedom combination that resulted in the highest hit rate). If you used the default number of replicates setting, these summary hit rates should have a good deal of stability and the optimal hit rate reported in your search grid should be nearly optimal. **Please note that if the estimated optimal priors lies on an boundary of your search grid, you’ll want to shift your search grid values (prior variances and prior degrees of freedom) and rerun the grid search.**

If you decided to use fewer replicates than the default (perhaps to speed up your run during initial investigations), the hit rates may have somewhat larger random error than desirable and a polynomial regression fit to the grid search results could provide a more accurate estimate of the optimal priors. Based on the search parameters you specified, the software writes out a regression estimate of the optimal prior variance and degrees of freedom to the **Optimalpriors.txt** file (if you specify at least a 3x3

dimensional grid search—3 values of prior variance and 3 values of degrees of freedom to search across). The regression is performed using the hit rate results found in the **studyname_hb_priors_finalreport.csv** file, where the hit rate for each replicate is the dependent variable and the prior variance and degrees of freedom settings are the independent variables. Because the relationship is not purely linear, we fit a polynomial regression. If you specify at least a 4x4 search grid, the regression with the following terms: $y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4y_1 + b_5y_2 + b_6y_3$, where x_1 is the prior variance, x_2 is the prior variance squared, and x_3 is the prior variance cubed, y_1 is the degrees of freedom, y_2 is the degrees of freedom squared, and y_3 is the degrees of freedom cubed. If you specify at least a 3x3 search grid (but less than a 4x4 grid), the regression uses the following terms: $y = a + b_1x_1 + b_2x_2 + b_3y_1 + b_4y_2$, where x_1 is the prior variance, x_2 is the prior variance squared, y_1 is the degrees of freedom, and y_2 is the degrees of freedom squared.

Using the regression equation, we compute the optimal prior variance and degrees of freedom (within the range you specified) that maximize the predicted hit rate and report this as the optimal solution.

If your grid search uses less than a 3x3 table (candidate prior variance x degrees of freedom), then a regression is not fit to the data to estimate the optimal priors.

studyname_hb_priors_finalreport.csv: This file contains the First Choice hit rate estimates for each replicate in the grid search. If you are interested in performing your own analysis on the results, we recommend you average the results across the rows (sample replicates) of the file to summarize the hit rates across multiple replicates.

studyname_hb_interactions_finalreport.csv: This file contains the First Choice hit rate estimates for each replicate in the grid search for each first-order interaction effect added to the base model. We recommend you average the results for each column across the replicate rows to see if any interaction terms added to the base model can significantly improve the hit rate.