Proceedings of the Sawtooth Software Conference

September 2016

SIMULATING FROM HB UPPER LEVEL MODEL

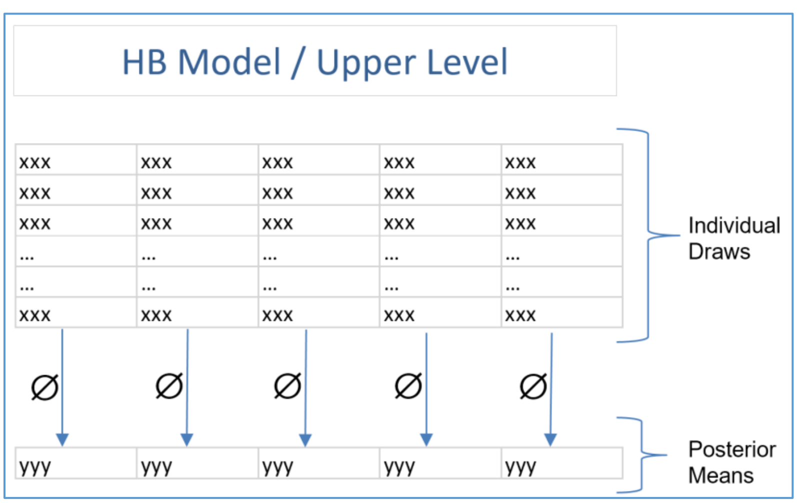
PETER KURZ
TNS INFRATEST
STEFAN BINNER
BMS MARKETING RESEARCH + STRATEGY

Motivation for this Paper

Today many practitioners in market research conduct conjoint analysis or discrete choice modeling (DCM) studies in their day-to-day research work. Since HB became available many years ago to the research community, simulations are quite accurate and therefore many researchers simply base their estimations on standard HB settings and use standard simulation tools. This means that "point estimates," also known as **"posterior means,"** are used in most simulation models. Posterior means are the individual respondents' part-worth utilities, calculated by taking the mean value for each parameter from a certain number of random draws from the posterior distribution after the convergence phase of the HB estimation process. Using posterior means is the standard in Sawtooth Software simulation tools. The disadvantage of this popular simulation method is the risk that due to the averaging of the draws, distribution uncertainty information gets lost. Therefore, simulations based on posterior means might calculate artificial or too simplified preference shares.

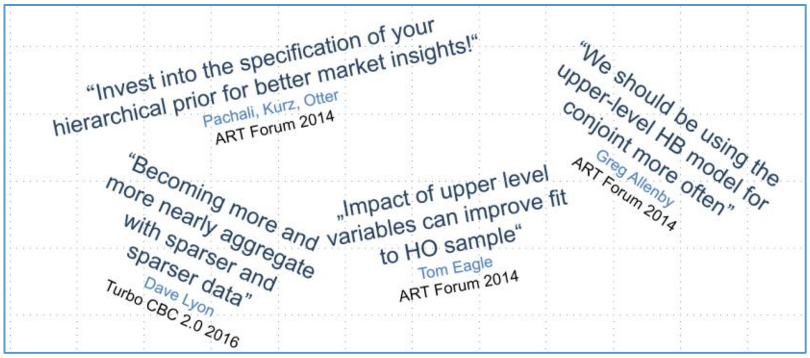
In order to account for both heterogeneity and uncertainty at the individual level, some researchers use individual **random draws** from the lower level model of the HB estimation after convergence is reached and apply those in simulations. Random draws should provide more insights into the uncertainty of respondents' choice behavior and therefore provide more accurate preference shares. One disadvantage of using random draws is that most current standard simulation tools do not support random draw simulations and therefore such simulation tools need to be created individually (e.g., in Excel). Furthermore, the large number of random draws in such tools (e.g., 100 draws for each respondent) leads to very large data sets which might make the simulation tools slow, difficult to operate and sometimes even hard to distribute (e.g., to clients, due to their size).

Figure 1. HB Model



Both simulation methods, posterior means and random draws, simulate from the lower level model of the Hierarchical Bayes (HB) model, thus neglecting the upper level model. However, some well-known scientists and research experts emphasize the relevance and impact of the upper level model of HB.

Figure 2. Expert Quotes

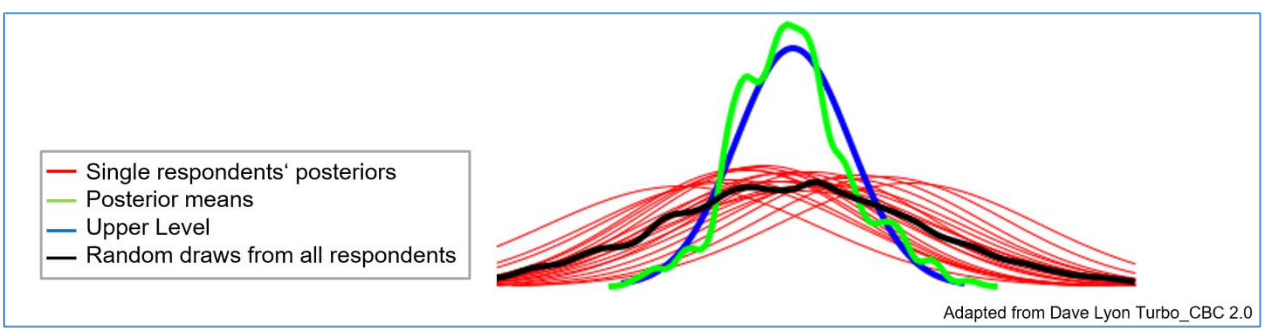


As there are concerns about losing uncertainty at the individual data, the motivation of this paper is to understand whether simulation results can be improved by using the **upper level model** as the simulation method.

How HB Works

HB "shrinks" individual-level utilities towards the means of all respondents. This is necessary because often it is impossible to estimate individual respondents—simply because there is not enough information in the data for each single respondent. It is the hierarchical "prior" of HB that pools information across respondents at the population level and allows the calculation of pseudo-individual values and simulations. Therefore, the weaker the individual data, the stronger is the resulting "shrinkage" or smoothing effect of the population level and the results are actually more based on the prior (see Figure 3). The use of posterior means—aggregated mean values of the draws—further strengthens this "shrinkage" effect as it ignores the uncertainty information within individual-level posterior draws (green line in Figure 3 misses the greater variance in the black line).

Figure 3. HB Shrinkage Effect

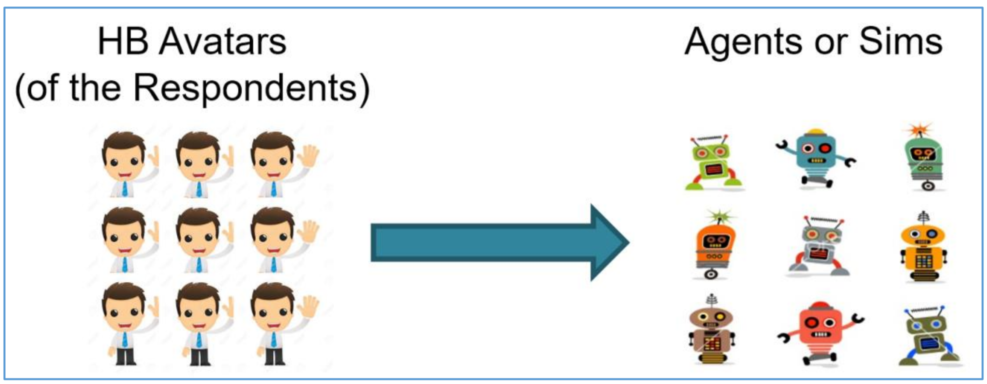


At the upper level, we assume individuals are distributed in some specified way, usually as multivariate normal, with means and covariances to be estimated. In the lower level, we assume that each individual's answers conform to a separate model, such as logit or regression. Hierarchical Bayes determines the optimal degree to which the upper level model and the lower level model influence the parameters for each individual. The lower level model only dominates if a lot of information per respondent is available.

By applying these hierarchical models, we estimate the population means and covariances at the upper level as well as the part-worths (betas) of each individual at the lower level. The information about the population means and covariances strengthens our estimation of individual results for each respondent. The payoff is that HB permits more precise estimation for each individual, often permitting individual-level estimation where previously only aggregated or segment-level estimation (such as latent class) was possible.

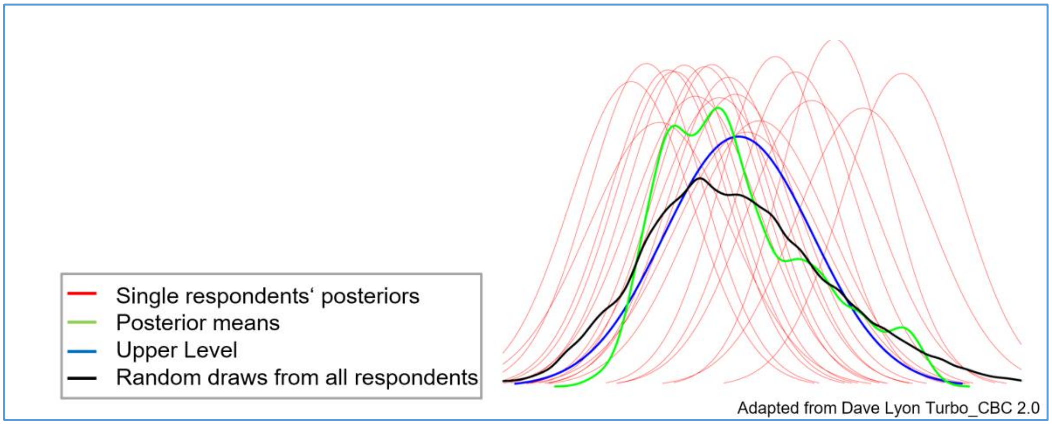
Due to the lack of individual information and the subsequent shrinkage effect, the individual estimates of the lower level model represent "avatars" rather than real respondents. On the other hand, the upper level model allows us to create "agents" or "sims" based on the aggregated functional form which is derived from the respondents in the lower level. The mean values over the avatars and the mean values over the agents are more or less the same.

Figure 4. Model Characteristics



The upper level model usually makes the invariant assumption that the data follows a multivariate normal (MVN) distribution (see CBC/HB technical paper, Sawtooth Software 2009). The upper level population means and variance-covariance of the estimates follow that multivariate normal distribution. These parameters are updated in every iteration of the sampler based on draws. The upper level captures the variance as well as the correlation structure in draws at an aggregate level. It is sensitive to the assumption about the functional form (MVN). The covariance matrix characterizes the extent of unobserved heterogeneity. Large diagonal elements, for instance, indicate more (preference) heterogeneity across consumers. Off-diagonal elements indicate patterns in the evaluation of attribute levels (the covariance structure of the part-worth coefficients). For example, positive covariations indicate pairs of attribute levels which tend to be evaluated similarly across respondents. The off-diagonal values can be translated into correlation coefficients. Figure 5 illustrates a model with good representation of the individual heterogeneity by draws and only a small shrinkage effect through the posterior means (the green and blue lines are relatively similar to the black).

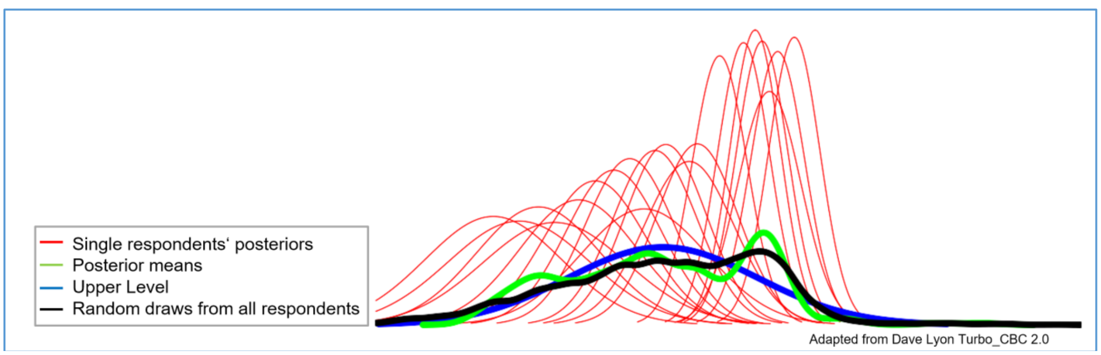
Figure 5. Good Representation of the Individual Heterogeneity



It is the combination of upper level and lower level models that allows us to estimate the part-worth values. In HB, a part-worth (beta) follows a distribution over respondents. The upper level model contains this full distribution over all respondents. The lower level model identifies the best spot for each individual within that distribution.

Figure 6 shows a different picture: Posterior means, upper level model and draws show similar results which indicates a good overall model fit. However, the plots of single respondents' posteriors show a rather poor representation of individual heterogeneity, especially on the right side of the distribution. This could lead to misinterpretation of the results, if we were—for example—looking at niche segments which are actually not sufficiently covered by the model.

Figure 6. Shrinkage of Individual Heterogeneity



UPPER LEVEL MODEL—RESULTS

When estimated with the HB sampler, the upper level model has the following aggregated results:

- The "mean value of alphas"—these values are often called the Bayesian logit model, because the alpha values usually come very close to the aggregate logit model. The alphas are the mean values of the population for each attribute level. Mathematically speaking, they are the mean values of the normal distribution of the upper level model.
- The variance and covariance structures—these structures describe the captured heterogeneity and the correlation between the different attribute levels.

Using the above measures offers an adequate representation of the underlying normal distribution of the parameter estimates. (See the blue lines in Figures 3, 5 and 6.)

The described results are part the summary file of CBC/HB:

Figure 7. Example of a Summary File with the Estimation Results

```
Point Estimate of Alpha
    -0.66685    1.22433   -0.43454   -0.89084    -1.0014
    1.04983   -0.12098   -1.56340   -3.31921   -0.84321
   0.43398   -1.26168   -0.60543   -0.31794    0.59853
-0.14683    0.85830    0.20906   -1.06737    1.04106


Average of Mean Betas
    -0.78214    1.23609   -0.41435   -0.91609    -1.1344
    0.94834   -0.15127   -1.63700   -3.38361   -0.97048
   0.41417   -1.24698   -0.59032   -0.24783    0.95484
-0.15116    0.90683    0.17477   -1.08160    1.11679

Estimated Covariances
    25.67529    7.02041   10.23446    7.53555    8.1666
   15.18765   13.39615    8.71322    8.37642    4.69160
-17.36621  -20.35964  -17.61392  -13.10647  -31.12503
  1.63775   -4.28386    0.02830    4.25556   -5.46803
```

By extracting the estimated variance-covariance matrix from the summary file of the estimation results, one could analyze the information contained in the matrix much better.

Figure 8. Upper Level Variance-Covariance Matrix

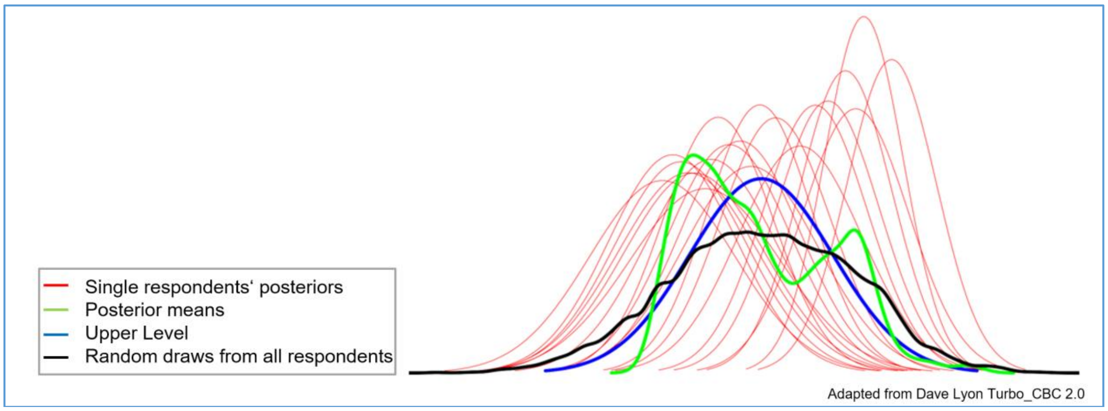
25,67529	7,02041	10,23446	7,53555	8,16667	17,54559	8,21049	-1,2395	
7,02041	23,98516	9,58513	12,61428	2,12021	-1,51361	-5,91708	-3,29358	
10,23446	9,58513	27,60617	10,33315	8,39988	4,5102	-0,99454	-9,63422	-
7,53555	12,61428	10,33315	22,13781	3,92074	3,3956	-2,75615	-21,1134	-1
8,16667	2,12021	8,39988	3,92074	24,02689	23,92051	16,25889	-1,37769	-
17,54559	-1,51361	4,5102	3,3956	23,92051	47,66646	27,53991	0,02311	-
8,21049	-5,91708	-0,99454	-2,75615	16,25889	27,53991	25,45274	5,99608	
-1,2395	-3,29358	-9,63422	-21,1134	-1,37769	0,02311	5,99608	46,39125	2
-0,5207	-9,6907	-2,57188	-16,63492	-5,59265	-2,99503	3,20832	24,55607	3
-4,58831	-12,88398	0,444	-12,12195	-0,67386	-7,66235	0,41442	16,72954	1
16,9943	0,95741	4,30712	5,50545	8,8315	26,34463	15,03351	-2,26342	-
11,90522	11,12359	8,29718	11,00259	-0,79709	-0,46767	-4,84708	-7,35196	-
15,18765	12,98496	9,40361	9,42139	7,10223	16,67777	5,12112	-8,01273	-1
13,39615	9,01569	9,64405	12,12484	5,0648	9,69743	4,38628	-11,37429	-

The variances of the attribute level alphas are represented in the main diagonal and explain the heterogeneity of the attribute level, which is the spread around the mean value of the underlying normal distribution. The off-diagonal values reflect the correlation between the attribute levels. For example, the value "7.0241" in above table indicates that Level 3 has a higher preference together with Level 1 compared with for example Level 2. This means that Level 2 is less frequently chosen when Level 1 appears than Level 3.

The hierarchical prior aggregates the information provided by individual-level draws. This "aggregation," however, is sensitive to the specification of the hierarchical prior unless flexible semi-parametric models were considered. This means we capture the overall heterogeneity in general, but not necessarily the functional form of its distribution. As long as we capture most of the heterogeneity, this is only a small loss. However, if the individual draws capture a lot of uncertainty for individual respondents, we lose some of that information about the parameter uncertainty in our model.

Individual draws can be one little step away from a mis-specified model even though they strongly depend on the hierarchical prior if the data is sparse. Among the three, posterior means are least sensitive to the functional form assumed in the hierarchical prior and may produce better aggregate results too, as the following Figure shows:

Figure 9. Capturing Individual Uncertainty through the Lower Level



Simulating from the Upper Level

Under the assumption that our upper level model parameters are a good representation of the data, simulation from the upper level model would miss a small extent of uncertainty—at the respondent level only. This could be easily compensated for by adding a small extreme value distributed error term. To build a simulator one can simply apply the alphas and the variance/covariance structure, together with this extreme value distributed error term, in order to create a certain number of new "respondents" (or better, "agents") based on the resulting normal distribution. Based on these new created agents, logit or first choice simulations can be performed with the model. Each agent is used for simulation in the same way as posterior means or random draws would be used.

If the model is based on sparse data and therefore doesn't capture enough individual respondent behavior (as illustrated in Figure 9), one can add the average within-respondent standard deviation of the draws, instead of the extreme value distributed error term, in order to restore the individual uncertainty—at least approximately.

Remember that each agent comes from the same prior, same scale (variance), same shape (covariance) and same functional form (assumed MVN). However, each agent has a different combination of those—a different random location under the normal distribution.

In contrast to posterior means, both random draws and upper level model agents account for parameter uncertainty caused by sparse data on the individual level. Both depend on the assumption about the functional form of the hierarchical prior. A better understanding of the problem-specific combination of these parameters therefore improves both. The only difference is that random draws are relatively less dependent on the functional form and if data is sparse they show a lot of spread (uncertainty) in the data. Sometimes this information could improve the results, sometimes it could make them worse.

EMPIRICAL COMPARISON OF SIMULATION METHODS

In order to understand the impact of the different simulation techniques we analyzed six real studies (all estimated with standard settings via Sawtooth Software's CBC/HB) and compared the preference shares resulting from these different simulation methods:

1. Posterior Means
2. Random Draws
3. Upper Level Model

For this purpose we randomly selected six real market studies. These empirical studies differ in their complexity (e.g., number of parameters, tasks, concepts per task) as well as sample size, so that one can assume different degrees of sparseness in these datasets.

Furthermore, these studies are different in their research objectives—which could also have an impact on the complexity of the choice tasks. Our sample of studies consisted of:

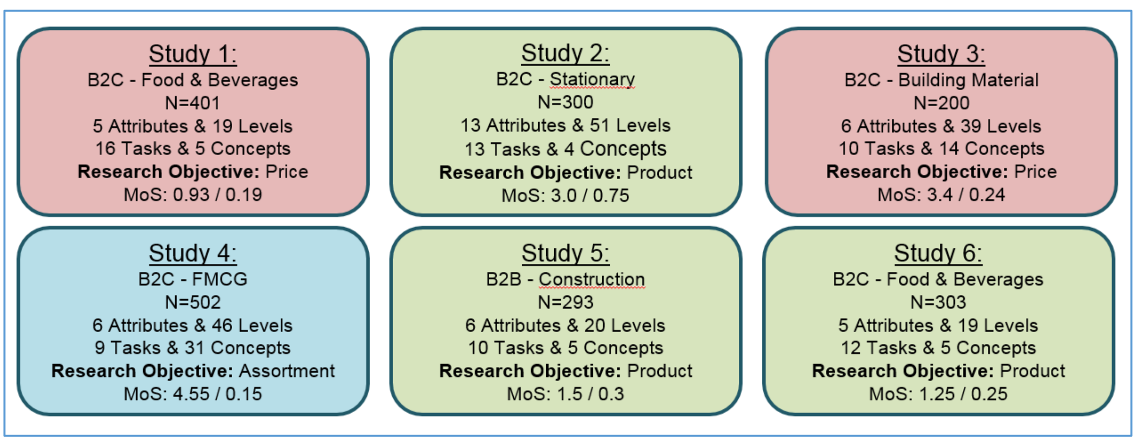
- 3 product configuration studies
- 2 pricing studies
- 1 assortment study

As an indication for expected complexity and sparseness of data we used two measures:

1. Measure of Sparseness 1 (MoS1): $$\text{MoS1}=\frac{\#\text{ of Levels} - \#\text{ of Attributes}) + 1}{\#\text{ of tasks}}$$
2. Measure of Sparseness 2 (MoS2): $$\text{MoS2}=\frac{MoS1}{\#\text{ of concepts}}$$

The MoS1 of the six studies ranged between 0.93 (Study 1) and 4.55 (Study 4); MoS2 between 0.15 (Study 4) and 0.75 (Study 2).

Figure 10. Overview of Six Empirical Studies



First we compared the within-sample prediction performance of the three simulation methods. As benchmarks for model performance, we used aggregated hit rates and RLH.

In order to calculate comparable hit rates in each study we used random tasks, which were not included in the estimations.

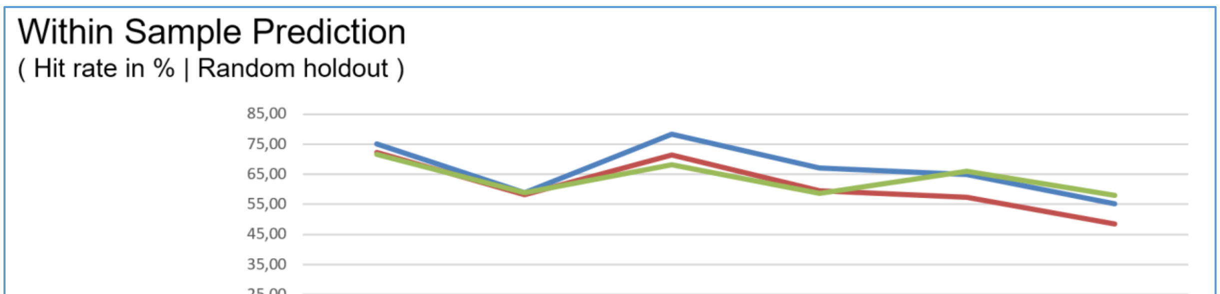
In regard to random draws, we considered two ways to derive hit rates:

> "0/1 Method": If more than 50% of the draws of one respondent fit with the random holdout task, this respondent is counted as a hit (if </=50% not).

> "100% Method": Each draw is individually counted as a hit if there is a fit with the random holdout, so one respondent's hit rate might be 0.6 if, for example, 60 out of 100 draws fit with the random holdout.

As the 100% method provides a higher likelihood to reach hits, hit rates for Random Draws are usually higher when the 100% method is applied. Therefore, we used the 0/1 approach to derive hit rates from random draws simulations.

Figure 11. Within Sample Prediction—Hit Rates

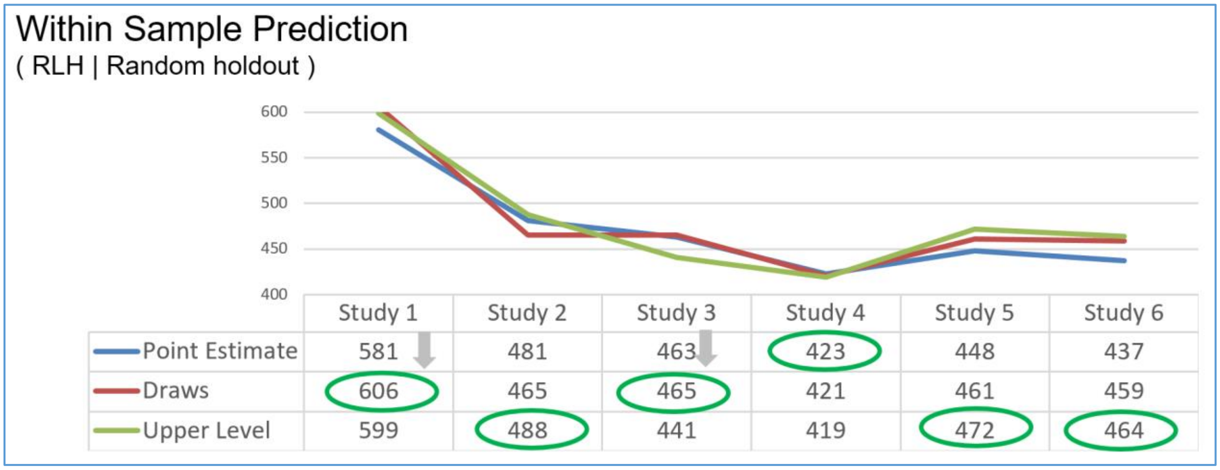


	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Point Estimate	75,06	58,8	78,25	67,03	65,02	55,28
Draws	72,36	58,25	71,40	59,61	57,28	48,57
Upper Level	71,54	58,92	68,27	58,76	65,95	57,93

As Figure 11 shows, posterior means and upper level model simulations often lead to very similar results. In the three complex studies focusing on product features (studies 2, 5 and 6), the upper level model simulation performed slightly better than the other two. However, these improvements are quite marginal (and in case of studies 3 and 4, certainly not significant). Nevertheless, a first hypothesis is that the upper level model might improve in-sample predictions compared to posterior means if there is a lot of individual level uncertainty. The rather poor performance of random draws might be caused by the 0/1 holdout method.

The comparison of RLH results shows a quite similar performance between the different simulation methods:

Figure 12. Within Sample Prediction—RLHs

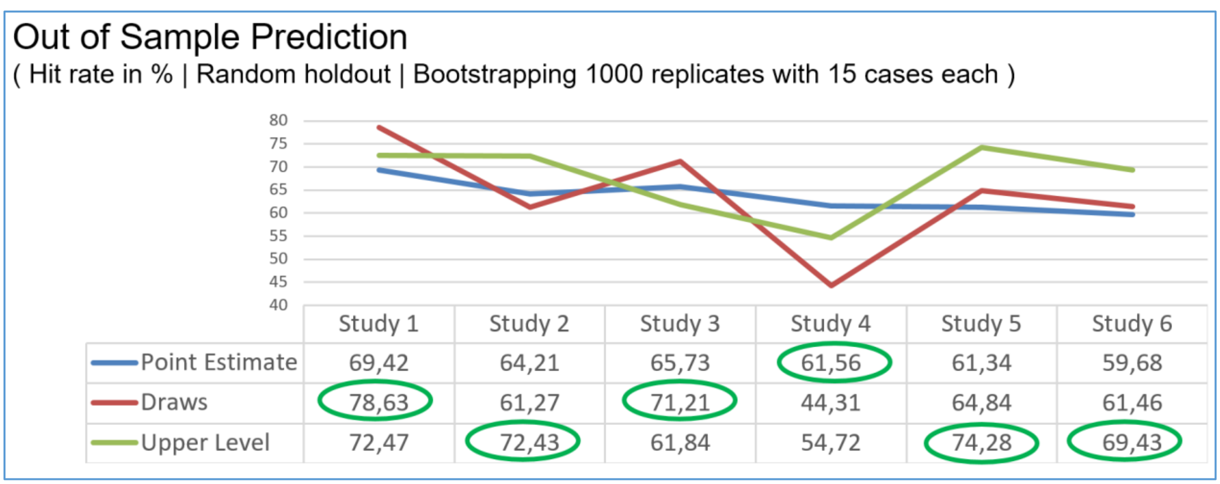


	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Point Estimate	581	481	463	423	448	437
Draws	606	465	465	421	461	459
Upper Level	599	488	441	419	472	464

In contrast to the holdout results, the RLH scores showed that the random draws perform slightly better than posterior means in Studies 1 and 3, which were both pricing studies. This is in line with the practical perception of random draws being the simulation model of choice for price-only discrete choice models.

How do the three simulation methods perform in regard to the more relevant measure, the ability of the models to predict *out-of-sample* choice behavior? To provide an indication for out-of-sample validity we used a bootstrapping process: For each of the 6 studies we drew 1,000 different samples, where in each sample 15 respondents were randomly excluded from the estimation and used as holdout respondents. The simulation of these holdout respondents allows us to use both out of sample hit rates and out-of-sample RLH as measures of performance.

Figure 13: Out of Sample Prediction—Hit Rates

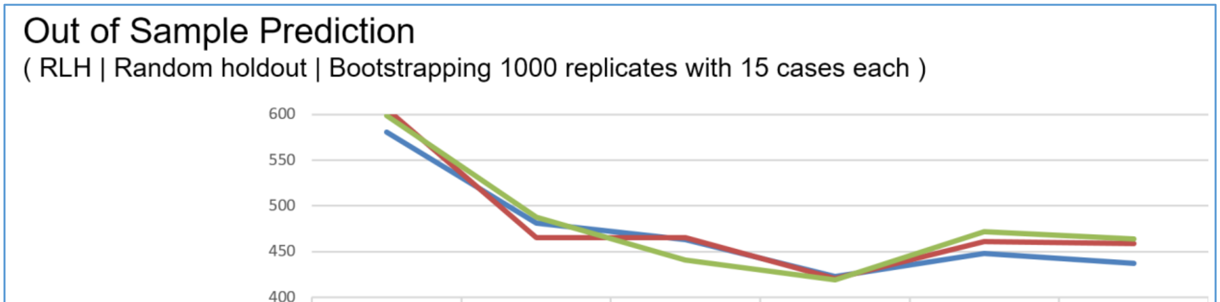


	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Point Estimate	69,42	64,21	65,73	61,56	61,34	59,68
Draws	78,63	61,27	71,21	44,31	64,84	61,46
Upper Level	72,47	72,43	61,84	54,72	74,28	69,43

In the out-of-sample holdout prediction the posterior means performed best only in Study 4 (the one with smallest number of choice tasks and an assortment objective). As expected, random draws performed best in the two pricing studies. The upper level model clearly produced the best results for the product configuration studies and outperformed the other two simulation methods.

Similar results can be observed looking at the RLH values:

Figure 14: Out of Sample Prediction—RLHs



	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Point Estimate	581	481	463	423	448	437
Draws	606	465	465	421	461	459
Upper Level	599	488	441	419	472	464

Again, RLH results are quite close between simulation methods. As in the hit rate results, posterior means performed best in the assortment study, random draws performed best in the two pricing studies and the upper level model clearly produced the best results for the product configuration studies.

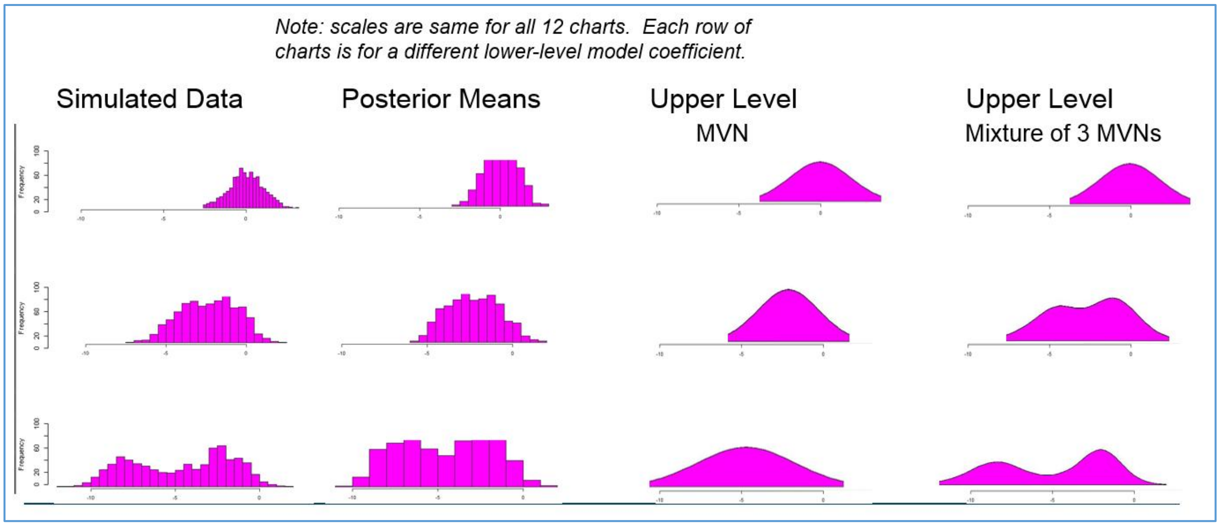
To summarize the performance of different simulation methods in out-of-sample prediction:

1. **Posterior means** only performed best in the assortment study, which had the lowest number of choice tasks and highest number of levels (thus, the highest sparsity of data, with MoS1 = 4.55). Our hypothesis is that there is lot of uncertainty in the lower level which is not real heterogeneity, but rather a reflection of how little we know about the individual respondent. Therefore, the upper level model has no chance to learn much from the respondents and is up to that point mis-specified.
2. **Random draws** performed best in the two pricing studies (even using the 0/1 method). Our hypothesis is that posterior means throw out too much uncertainty here. In the price-only studies we have enough individual information so that the lower level can capture real respondents' heterogeneity and therefore performs better. The simple upper level model we used in this paper was not able to represent the information we captured with the draws.
3. **The upper level model** showed the best results for the studies dealing with product configuration. Here our hypothesis is that there is a lot of uncertainty on the individual level due to the large number of parameters (complex choice tasks). Therefore, posterior means performance is inferior compared to the upper level model. The performance of random draws depends on the way the hit rate is computed (0/1 vs. 100%), but the upper level at least represents the heterogeneity of the respondents in a better way than 0/1 draws.

Difference between Posterior Means and Upper Level Model

The main difference between posterior means and the upper level model is that the upper level model describes a functional form (the model) of the data while posterior means describe the data by mean values created from a statistical model of the data. The upper level model is therefore more flexible and can be used to sample new agents whenever needed. The only problem is that the upper level model has to be specified more carefully if one wants to use it to build a simulator. Usually the standard HB approach (as in Sawtooth Software's CBC HB) assumes only a single multivariate normal distribution on the upper level model. This could be a significant limitation in estimating the best upper level model, as Figure 15 illustrates:

Figure 15. Model Comparison



The three rows show three different lower level model coefficients, the first column the real data (simulated), the second column the posterior means and the third and fourth column the upper level model with two different functional forms. The upper level model is determined by the functional form we have decided to use. As one can clearly see if we use a single MVN for the upper level model, we can capture the normally distributed beta coefficent (first row) very well. The beta coefficient representing a bi-modal distribution with two small peaks (second row) is not recovered very well with the a single MVN upper level model, but a mixture of normals does better. The beta coefficient with a sttrongly bi-modal distribution (row 3) is captured very poorly. Compared to the single MVN upper level model, the posterior means fit the the two bi-modal beta coefficients much better. If we use a more sophisticated upper level model—in this example a mixture of three MVN distributions—the simulation fits very well to all three shapes of beta coefficients.

This underlines that it's really worthwhile to invest more in the proper selection of the *right* upper level model if one attempts to build a simulator based on it. A proper specification of the upper level model is not just a question of the right functional form. Including meaningful covariates could make the upper level model even more powerful. But, we must use really *meaningful* covariates, otherwise the model could be distorted more than improved. Useful covariates are usually exogenous variables, which are related to the attributes and level.

CONCLUSIONS

Upper level model simulators performed very well—in some cases better than random draws or posterior means, especially when the functional form fits well. In our practical studies this was most often the case in product optimization studies with more attributes when we do not have clear compromise alternative in the choice tasks (Dhar & Simonson 2003). Additionally, in complex studies some attributes or attribute levels may not be taken into account by the respondents. In these cases we assume that the uncertainty captured by the individual respondent draws is due to "attribute non-attendance" (Chrzan & White 2016) and is not useful for predicting the respondents' choice.

Posterior means ignore parameter uncertainty and are shrunk towards the sample mean. This especially leads to poor out-of-sample predictions. In the case of a high "parameters-to-tasks ratio" (sparse data) the upper level model and random draws are superior.

In two-attribute cases (e.g., price-only DCMs), where we usually have clear compromise alternative in our choice tasks, the random draws and upper level model are superior.

The upper level model could be a solution for complex objectives or small sample sizes such as we often find in small markets (e.g., B2B). In such cases the upper level model, which only specifies aggregate parameters and a functional form, often results in more stable and meaningful estimates. An additional advantage of the upper level model approach for such small samples is the ability to resample and rebalance the simulated agents. When it's not possible to have samples that are representative for the market, as is often true in B2B contexts, one can use the upper level model to generate agents that represent the market structure in order to build a better simulator.

In order to cope with high model complexity, investment in the upper level model could be a better solution than longer interviews (Individual Choice Task Threshold, Kurz and Binner 2012). Furthermore the application of meaningful covariates and more complex distributions (such as mixtures of normals) instead of using a standard MVN MNL-HB program could minimize the risk of simulating from a mis-specified hierarchical prior. Working with the upper level model will create a need for HB estimations for choice based conjoint with more customizable computer programs that can vary for each new study.

Although HB methods do not converge to a closed form solution—in the way most of our classical statistical methods do—we should be comfortable with the fact that the variance stabilizes after a few thousand iterations, but there will be still considerable variation in the averages of the parameter estimates. This means that we end up with a distribution of estimates for each individual rather than a single point estimate for each part-worth value.

While the random draws (representing this distribution) are powerful in terms of understanding uncertainty, they add complexity to the analysis. Random draws adequately account for parameter uncertainty but may become impractical for large N data files because we multiply the size of our data files by 100 or even 1000. A well specified upper level model can represent the functional form nearly as accurately as random draws, and is relatively easy to handle because one needs only the aggregate parameters and can easily sample as many agents as necessary on the fly.

SUGGESTIONS FOR FUTURE RESEARCH:

We often talk about "sparse data" but without a clear understanding of how to define sparseness. Some researchers even talk about sparse data when simulating 2 parameters with 12 choice tasks and 500 respondents, which would mean a MoS1 index of 0.25 (Pachali, Kurz, Otter 2014). It is important to develop a formula or heuristic in order to determine sparse/not-sparse data (similar to the MoS indices) that could be used as a common indicator in the research community. Such a measure could also include parameter/information ratio which would facilitate comparisons among different studies.

Most practitioners use experimental designs which are tested with aggregate logit calculations for d-efficiency only. However, most studies are estimated with HB techniques that take individual respondents into account. Especially if we want to use the upper level model, we should pay more attention to how to optimize the experimental designs for HB techniques.

Should the upper level model be used more often, covariates will become more important for model estimations. Therefore we need good guidance on how to determine meaningful covariates in practice. Up to now, there have been very controversial discussions as to which covariates are meaningful and which ones are not, or could even harm the estimations.

Up to now simulators based on upper level models have been used mainly in the academic world and only a few practitioners have tried to work with them so far. Therefore, we believe that more research has to be done in developing sophisticated but practical upper level model simulators that also take advantage of the resampling and rebalancing capabilities.


Peter Kurz


Stefan Binner

References

Allenby, G. M.; Rossi, P. E. (2006): "Hierarchical Bayes Models," in Grover, R.; Vriens, M. (Eds.), *The Handbook of Marketing Research: Uses, Misuses, and Future Advances*, pp. 418–440, SAGE Publications Inc., Thousand Oaks.

Dhar, R., & Simonson, I. (2003): "The effect of forced choice on choice," *Journal of Marketing Research*, 40 (May), 146–160.

Chrzan, K; White, J. (2016): "Mapping Attribute Non-Attendance," *Proceedings of the 2016 Sawtooth Software Conference*.

Hein, M.; Kurz, P.; Steiner, W.(2013): "Limits for Parameter Estimation in Choice-Based Conjoint Analysis: A Simulation Study," presentation to European Conference on Data Analysis 2013.

Johnson, R. M. (2000): "Understanding HB: An Intuitive Approach," Sawtooth Software Research Paper Series.

Kurz, P; Binner, S.(2011): "Added Value through Covariates in HB Modeling?," *Proceedings of the 2011 Sawtooth Software Conference*.

Kurz, P.; Binner, S. (2012): "'The Individual Choice Task Threshold' Need for Variable Number of Choice Tasks," *Proceedings of the 2012 Sawtooth Software Conference*.

Kurz, P; Binner, S.(2015): "Capturing Individual Level Behavior in DCM," *Proceedings of the 2015 Sawtooth Software Conference*.

Liakhovitski, D.; Shmulyian, F. (2011): "Covariates in Discrete Choice Models: Are They Worth the Trouble?," 2011 Advanced Research Techniques (ART) Forum presentation.

Lyon, D. W. (2016): "Hierarchical Bayes: Poking Around Under the Hood," presentation at Sawtooth Software Turbo CBC, Captiva Island.

Pachali, M.; Kurz, P.; Otter, T. (2014): "How to Generalize from a Hierarchical Model," 2014 Advanced Research Techniques (ART) Forum presentation

Sawtooth Software (2009): "The CBC/HB System for Hierarchical Bayes Estimation Version 5.0 Technical Paper," Sawtooth Software Technical Paper Series.

Sentis, K. and Li, L. (2001): "One Size Fits All or Custom Tailored: Which HB Fits Better?," *Proceedings of the 2001 Sawtooth Software Conference*.

Sentis, K.; Geller, V. (2011): "The Impact of Covariates on HB Estimates," *Proceedings of the 2011 Sawtooth Software Conference*.