PROCEEDINGS OF THE SAWTOOTH SOFTWARE CONFERENCE

March 2015

CAPTURING INDIVIDUAL LEVEL BEHAVIOR IN DCM

PETER KURZ[1]
TNS INFRATEST
STEFAN BINNER[2]
BMS MARKETING RESEARCH + STRATEGY

PROLOGUE: SOMEWHERE IN A CENTRAL TEST LOCATION . . .

At the beginning of this paper we want to take you to a real market research situation. Imagine you are with your client in a "central test location," a research studio where consumers are invited to participate in a survey. As such studios are usually set up for qualitative research purposes, there are observation rooms where clients or researchers can observe group discussions or in-depth interviews. When conjoint studies, a quantitative research method, are not conducted online, such test locations are often used in order to screen for the right target group, to use stimuli (such as dummies) and to conduct interviews in front of a computer. Of course such a set-up represents a great opportunity to observe interviews: either by the researcher, e.g., to conduct pre-tests, or by the client who wants to understand how consumers react to the choice tasks and to gain some insights as to their preferences.

We are now in such a central test location with our client for a large and important study. The respondents are instructed by the interviewers to always comment on what they select and to explain the motivation for their choice decisions. We are in the observation room listening to a respondent explaining her preferences while clicking through the choice tasks. As she goes through the choice tasks our client is impressed with how consistently she is making her choices. Our client is especially pleased that she has a high preference for a particular product feature as he is convinced that this feature is quite important for many of his customers.

A few weeks later we present the results of the study to the client and his team. To their general surprise (our client having already reported his experience during the interview to his colleagues), the product feature in question came out as being not desirable at all. Another feature was clearly the winner. Our client is irritated and tries to understand why the results do not fit the observations he made during the interviews. He asks about the interview with the respondent he found so interesting and wants to see her individual results. As we are prepared for all types of discussions, we have the individual results of all respondents on hand and we look for the individual results of this specific consumer. To our surprise, her individual part-worth utility for the feature she liked so much is negative, while the utility value of the alternative feature is positive. Our client raises his eyebrows and asks for an explanation. Didn't we promise him that we can derive individual utility values instead of an aggregated result? Didn't we tell him that by using HB we are able to deliver best results even if the model is quite complex and we cannot ask each respondent to complete very many choice tasks? Did something go wrong during the estimation process? How can we explain this result to our client without losing his trust in discrete choice modeling, and us?

---

[1] Head of Research & Development TNS Infratest (Peter.Kurz@tns-global.com)
[2] Managing Director, bms marketing research + strategy (s.binner@bms-net.de)

MOTIVATION FOR THIS PAPER

In the last few years many papers presented at market research conferences in regard to conjoint analysis or DCM focused on such topics as how (if at all) to apply covariates, how many choice tasks can be asked or how parameters (e.g., priors) should be set in the Hierarchical Bayes estimation and how relevant these parameters are for researchers.

Although these papers were sometimes controversial (especially since academics and practitioners often came up with different conclusions) some of the common conclusions from these discussions are:

1. In complex DCM designs one usually does not derive pure individual utilities, but kind of artificial or "pseudo individual" utilities which more-or-less represent the sample.
2. As long as one uses the resulting part-worth utility values for market simulations and not for segmentation or other additional analysis procedures it is believed that "enough" heterogeneity is captured and the simulation models work fine.
3. Even if there is not "enough" heterogeneity captured, it is certainly more than with aggregate logit models or other "practical" alternative approaches such as latent class analysis.

On the other hand when we are forced to use these "pseudo individual" utilities for segmentation or when we dig deeper into the data structure we sometimes find these individual estimates as being not really intuitive: We might not find the segments clients expect or we have observed while attending real interviews (as in our prologue). In the worst case, these "pseudo-individual" results, sold to our clients as real individual results, can lead to distrust in the simulation results and the value of the whole study.

Therefore we need a deeper understanding of how much heterogeneity we really capture with our DCMs using hierarchical Bayes techniques. This understanding will help us to further improve our research designs (e.g., sample structure) and estimation processes and guide us in how to interpret the resulting estimates and results.

QUESTION 1: CAN WE CAPTURE MORE HETEROGENEITY BY APPLYING COVARIATES?

Let us start with a short introduction to covariates and how they can be applied. If one uses "standard HB" as it is implemented in Sawtooth Software packages, default settings are applied and there are no covariates defined and one single multivariate normal prior is assumed. This leads to a "shrinkage" of respondents preferences towards the population mean and this effect is sometimes quite significant. This shrinkage tends to reduce differences between individual respondents and thus, between customer segments. The effect becomes more noticeable for small customer segments, unless they are boosted by using disproportionate sample structures. When such segment differences are reduced due to shrinkage, their chance of being identified and acted upon goes down as significance testing of segment level differences is based on the shrunken, less-different values.

The purpose of the application of covariates is to allow segments to keep their characteristics within a total sample estimation model. Instead of estimating disaggregate part-worth utilities with hierarchical Bayes based on one sample mean, which might reduce heterogeneity greatly compared to the underlying reality (Figure 1), the use of covariates in the upper level model aims to increase the heterogeneity of the utility distribution by shrinking respondents of different subgroups to their means based on their own subgroups rather than the total mean (Figure 2, for a single 2-level covariate).

Figure 1: Single Sample Mean

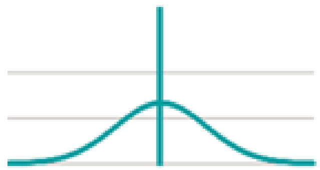
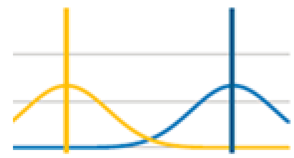


Figure 2: Multiple Sample Means



In theory this should do a lot to solve the potential problem of shrinkage of single respondents to the sample mean. However, in everyday work covariates have often not been found particularly effective in improving the overall model performance or in enhancing differentiation between subgroups in the simulations. If there is a sufficient number of choice tasks, covariates do not improve the model performance, because the lower level model dominates the solution. Even with a small number of choice tasks in most of our studies, covariates in general did not improve results; in some of our cases the covariates "washed out" and estimation converged to the same parameters as when no covariates were used.

We observed that in studies where each segment is represented by a sufficiently large sample size, HB without covariates converges towards same parameters as HB with a covariate model. But if the covariates are not really able to predict differences between sample segments, in the worst case they are just extra noise added to the model and can actually make it worse. Nevertheless the application of (the right) covariates in HB estimation will sometimes result in better distributed utilities and in an improvement of aggregate metrics, both at the total sample and subgroup level. The problem is just that we do not have or know the right covariates in all our projects.

Conclusion 1: Due to the lack of discriminating covariates, they are often not the solution to the issue of excessive shrinkage

Question 2: Why Don't We Simply Increase the Number of Choice Tasks in Order to Collect More Heterogeneity?

In our paper presented at the 2012 Sawtooth Software Conference we demonstrated there is a natural limit to how many choice tasks an individual respondent can answer. We called these limits Individual Choice Task Thresholds or simply ICTs. An ICT is the threshold past which an individual's further choices lead to poorer model fit rather than better, due to over-simplified responses or symptoms of exhaustion. In most of our studies we could see that in general, respondents had a diverse answering behavior and individually different choice task thresholds. For many respondents we got better or equally good hit rates and share predictions when we used only a smaller number of their choice tasks (the first ones, not the later ones) in order to avoid simplification.

Therefore we concluded that "Less is more," meaning that we should ask fewer choice tasks in order to improve results. Furthermore we learned that more choice tasks could even be dangerous, resulting into misleading results and interpretation. The analysis of the individual posterior distributions showed that a large number of respondents tend to simplify the answering in later choice tasks. In the first half of the choice tasks we saw higher number of attributes with significantly non-zero utilities than in later ones.

Conclusion 2: due to the ICT, we cannot solve the shrinkage issue by simply increasing the number of choice tasks.

Topic 3: Can Sample Size Compensate for the Limited Individual Information We Collect?

In order to answer this question Hein, Kurz & Steiner set up a research simulation experiment with 1,296 models:

Figure 3: Experimental Research with 1,296 Models (Hein et.al., 2013)

Factor	#Factor Levels	Factor Levels
attributes	4	6, 8, 10, 12
attribute levels	3	3, 4, 5
number of choice tasks (excl. 2 holdout tasks)	3	11, 13, 15
number of alternatives per choice task	3	3, 4, 5
number of respondents	3	500, 1000, 1500
sample	2	homogenous, heterogenous
error variance	2	standard (1,645), high (3,290)
experimental conditions in total: $2^2 * 3^4 * 4 = 1296$		

This simulation study clearly showed that for 6 and 8 attributes an increase of sample size could compensate for a decrease of the number of tasks (T) from 15 to 11 in terms of average RLHs. However, for a larger number of attributes (10 or 12, with 5 levels each), even a tripling of sample size from 500 to 1500 could not compensate for a relatively modest decrease of T from 15 to 13, in terms of average RLHs. We use T as the number of repeated measurements (number of tasks per respondent) in a choice model in our following explanations.

Furthermore the findings showed that a lack of individual information could not be compensated for by larger samples. HB does the best it can estimating part-worth utilities with a maximal amount of heterogeneity—but has no chance to provide good individual parameter accuracy if T is small. With small T one should consider using the upper level model for simulation.

Conclusion 3: Increased sample size is not a solution to cope with the limited number of choice tasks possible.

Issue 4: Parameter Settings in the HB Estimation

Each researcher has to define a hierarchical prior distribution of heterogeneity before estimating an HB model for discrete choice experiments. This distribution is usually intuitively chosen by the analyst. In practice most researchers currently use the multivariate normal distribution as the standard choice for their prior.

In a discrete choice data set, the observed choices y_{jt} are assumed to follow a multinomial logit distribution:

$$y_{jt} \sim MNL(X_{jt}, \beta_j) \; ; j = 1, \ldots, N; t = 1, \ldots, T;$$

where N is the number of respondents in the sample and T the number of choice sets each. The vector of part-worths β_j is different across respondents according to:

$$\beta_j = \Gamma z_j + u_j$$

where Γ is a matrix of coefficients relating the vector of part-worths β_j to a respondent's specific demographic variables z_j (i.e., the covariates). The product Γz_j accounts for the observed heterogeneity attributable to covariates, while u_j is a stochastic component representing the unobserved heterogeneity component. The distribution of u_j is of particular interest because it influences how β_j can vary across respondents independently of the covariates. In current practice, a standard multivariate normal distribution is almost always used as the default setup:

$$u_j \sim N(\mu, \Sigma)$$

is the "mid-level prior" governing how respondents differ, and its parameters come from the top-level prior that has "hyper-parameters" $\bar{\mu}$, υ and V set by the analyst (or by default, by the software):

$$\mu \sim N(\bar{\mu}, \Sigma \otimes a_{\mu}^{-1})$$
$$\Sigma \sim IW(\upsilon, V)$$

(The variance of the top-level prior is distributed as an Inverted Wishart variable [a multivariate generalization of the inverted chi-squared], scaled up by some identity-structured matrix. If the hyper-variance V is set very high, so the prior is very diffuse, the top-level model represented by the last two equations gives "free rein" to the data's determination of the parameters of the MVN for u_j in the middle-level model.)

If you believe in the above model of heterogeneity, the estimates result in consistent and efficient inferences about the unobserved population from the hierarchy, even for a small number of repeated measurements (T). But should you really believe in a MVN distribution of heterogeneity? For large T, the collection of in-sample posterior means are usually robust against misspecification of the upper level model. In other words, when T is large, lots of data overwhelms the priors, misspecified or not. However, in practical discrete choice experiments, T is almost always very small, so these assumptions matter more.

Consider a quick reminder of how hierarchical Bayesian methods work. Generally, three different levels are distinguished. The first level shows the hierarchical prior, with its parameters $\bar{\mu}$, υ and V, which are just chosen by the analyst (often just by accepting software defaults). The second level contains random effects or individual level coefficients β that represent consumer preferences for a given attribute. Finally, the third level is the data y. The index j represents the j^{th} respondent of a discrete choice data set and N denotes the total respondents participating in the survey. The index t stands for the t^{th} choice task and each respondent answers in total T choice tasks (Rossi, Allenby, McCulloch 2005).

According to the model, the β_j are determined by the hierarchical prior and then generate the data y (i.e., choices of the respondents). It implies that β_j for all unobserved consumers in the same market should be generated from the upper level model, assuming the upper level model is an acceptable representation of the population.

Using posterior means for this generalization can be risky from a theoretical perspective. There will be posterior uncertainty in the β_js for finite T that using posterior means completely ignores. On the other hand, the upper level model still provides the same insights about the population if N is large enough, even when T is small. This is a major theoretical reason to use a hierarchical model in the first place. A set of posterior means will only represent the population reasonably if both N and T become quite large. Large N overcomes any misspecification of the shape of the mid-level prior (i.e., MVN) and any problems in the hyper parameter values. Large T brings the posterior means into line with the data and reduces the influence of shrinkage. Large T also reduces the variance of respondents' posteriors, so that ignoring that uncertainty for each respondent is less problematic.

But in practice, most current choice simulators use posterior mean estimates of in-sample respondents to generalize beyond the sample of respondents interviewed and to predict preference shares for the population. This means that lower level model preferences are being used in the following form to generalize in simulations.

1. Use posterior draws $\{\beta^1{}_j, \ldots, \beta^R\}$ for $j \in \{1, \ldots, N\}$ calibration individuals with R draws saved for each
2. Calculate posterior means $\hat{\beta}_j = 1/R \sum^R{}_{r=1} \beta^r{}_j$ for each respondent
3. Calculate preference shares for alternative i in a simulation scenario defined by attributes X_j:

$$S_{ij} = 1/N \sum^N{}_{k=1} p(y_i | X_j, \hat{\beta}_k)$$

In other words, posterior means of in-sample respondents are used to generalize beyond the sample and to predict choices of consumers in the hypothetical market.

So far we introduced in our theoretical discussion how to generalize from an HB model using the lower level model by creating a distribution of part-worths that is—perfectly—a 1:1 representation of the degree of heterogeneity in the population. In the following, we will use a simulation study to show what happens with the answers of our respondent from the introductory example during the estimation process. This explains how the lower level model inferences will be influenced in terms of replicating the true amount of heterogeneity in the population. Therefore, we compare a distribution of in-sample posterior means to the distribution obtained from the posterior of the hierarchical prior. This comparison will be simulated for different numbers of choice sets per respondent T and sample sizes N. As we will see, the number of choice tasks T in a discrete choice model data set has a strong impact on the results.

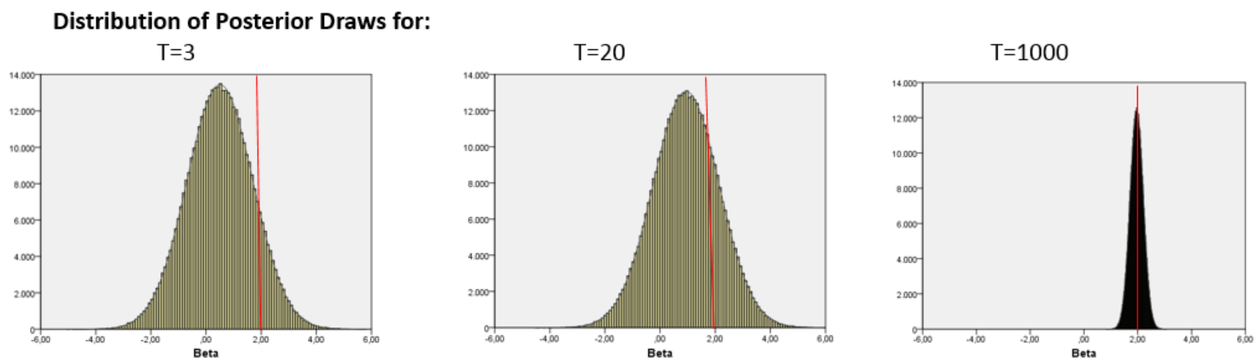
Our first simulation shows how inferences based on posterior means will be influenced when the number of repeated measurements T is small. For this simulation, we use a simple multinomial logit model setup (not hierarchical) that only includes one respondent (our woman with "rear wheel-drive" preference) without any heterogeneity.

Simulation Setup:

- MNL model (no hierarchy)
- Let $\beta = (2,2)$ represent data generating preferences (the "true" utilities)
- Use $\beta 0 = (0,0)$ and $\Sigma 0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ as informative priors
- $p = 3$ alternatives per choice task (as in our car example)
- $T1 = 3 < T2 = 20 < T3 = 1000$

Note that in these assumptions, the β_0 prior is not very close to the true β, but the prior variance Σ_0 is relatively tight. In other words, β is out in the tail of the MVN prior.

Figure 4



Red line represent the data generating preference

Figure 4 shows three graphs that correspond to the three different numbers of choice tasks T = 3, T = 20 and T = 1000 used in the simulation. The preferences, or betas, are plotted on the x-axis and their density on the y-axis (these graphs would be the same for either the first or second element of β). The vertical red line marks the true data generating preference (true β) that is equal to two. The solid black line traces the density function for the distribution of posterior draws. Figure 4 clearly shows that the posterior mean is little different from the prior when the number of repeated measurements T is small (T=3) and that, in this case, the posterior mean would not be very informative about the true location of the respondents preference. However, as the number of repeated measurements increases, the posterior mean becomes more and more accurate as to the true location of this respondent's preference.

The informative prior we used here is meant to mimic what we get from a population of respondents with relatively little heterogeneity. In our real life example, the large number of front-wheel-drive likers means relatively little heterogeneity. In this situation, posterior means of individual level coefficients will be shrunk toward the prior (the overall distribution of respondents) unless T is really large. This is the reason for the shrinkage from rear-wheel-drive preference to front-wheel-drive preference for our respondent.

In the next simulation we extended the results from Figure 4 and applied them in the context of hierarchical models. The purpose of the following simulation is to confirm the previous finding that posterior mean inferences will be shrunk strongly towards the prior if T is small, meaning that lower level model estimates are unable to discover the true heterogeneity distribution in the population under such conditions. On the other hand, we will show that inferences on the basis of the posterior of the hierarchical prior will be unbiased regardless of the number of repeated measurements T (for a more comprehensive explanation of this topic, see: Pachali, Kurz, Otter 2014).

Let $\beta_i \sim MVN(\bar{\beta}, V_\beta)$ with

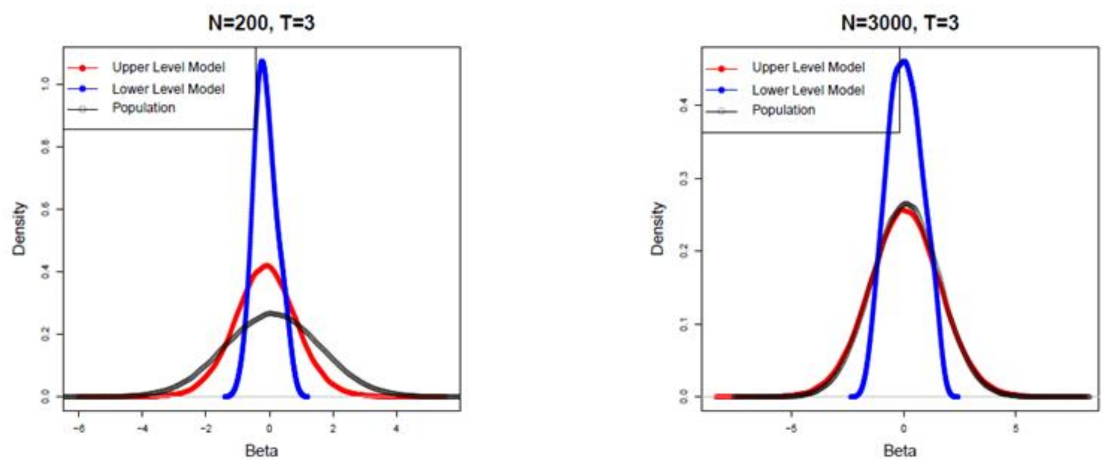
$$\bar{\beta} = (0,0.1,0.2,0.3,0.4) \text{ and } V_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1.5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2.5 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

We use this heterogeneity distribution to generate small (N = 200) and large (N = 3000) samples and then have each sample member participate in a short (T = 3) or a long (T = 60) discrete choice survey.

Consider combinations of small or large T and N:

N = 200, T = 3
N = 200, T = 60
N = 3000, T = 3
N = 3000, T = 60

Figure 5: Posterior Densities for Level "Rear-Wheel-Drive" with $\beta_2 \sim N(0.1, 1.5)$ and T = 3

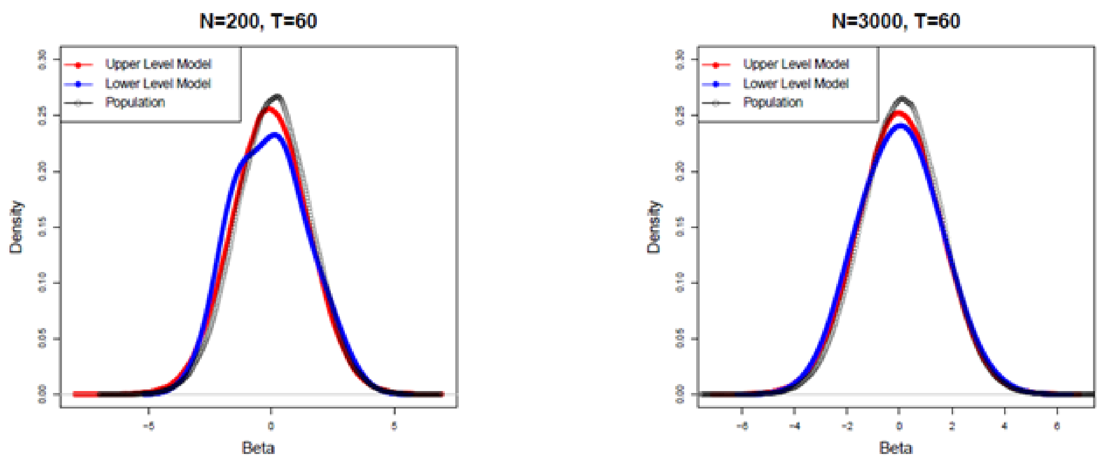


The posterior means severely underestimate the actual heterogeneity if T is small.

Figure 5 shows the distributions of posterior draws for the second part-worth. The figure on the left corresponds to the small sample N = 200 taking a short conjoint survey T = 3. The part-worth utility or beta is plotted on the x-axis while the y-axis shows the density. The black line (the lowest, flattest line) depicts the true distribution of heterogeneity for the part-worth "rear-wheel-drive." The blue line (the tallest, peakiest one) depicts the distribution of heterogeneity inferred from posterior means or the lower level model. Finally, the red line (the one in the middle) depicts the distribution of heterogeneity inferred from the posterior of the hierarchy, or the so-called upper level model. The figure on the right corresponds to the larger sample N = 3000 taking again a short conjoint survey T = 3. The color codes are the same (the black and red lines are nearly on top of each other).

Figure 5 shows how relying on posterior means of individual level coefficients severely underestimates the true amount of heterogeneity in both cases when the number of repeated measurements is small. This means even with a large sample size we are not able to capture enough heterogeneity that the "rear-drive" part-worth of our respondent isn't shrunk too much towards the mean.

Figure 6: Posterior Densities for "Rear-Wheel-Drive" with $\beta_2 \sim N$ (0.1, 1.5) and T = 60



When the number of repeated measurements T is high, the information in the individual level posteriors is essentially independent from the sample size N.

However, the differences in Figure 5 vanish once long conjoint surveys are considered where each consumer provides a lot of information. So we see that if we obtain enough individual information—i.e., a long interview for our respondent—our model captures enough heterogeneity to represent the correct "rear-drive" part-worth for our respondent, even if she has a preference very different from the population. Figure 6 shows that the form of generalization from the hierarchical model becomes less important if the number of repeated measurements increases and the information in the data get rich (in each half of this figure, the blue lower-level model line is the least peaked, the black population line is most peaked, and the red upper-level model line most peaked).

So far, it has been shown that lower level model inferences about the unobserved population are biased when the number of repeated measurements T is small. This bias is caused by systematically underestimating the true amount of heterogeneity in the population. On the other hand, upper level generalizations consistently estimate the true amount of heterogeneity in the population even if the number of repeated measurements is really small. However, this only holds true under the premise that the hierarchical prior distribution of heterogeneity is correctly specified. This is an important finding since T will almost always be small in practice due to time constraints associated with the questionnaire and due to the response quality problems we encounter if we exceed the ICTs.

Conclusion 4: The current practice of using posterior means to generalize from the HB model biases inference and decision when T is small. The bias is against heterogeneity and differentiation. In practice, T will always be relatively small, because clients are more demanding, models become bigger and bigger, and respondent time and attention is limited

(ICT). Generalizations from the upper level model are consistent, i.e., without bias and efficient, even for small T, so long as the hierarchical prior distribution of heterogeneity is not misspecified. Therefore we should invest more time in the specification of our hierarchical prior in order to derive better market insights.

Can We Capture Individual Level Behavior?

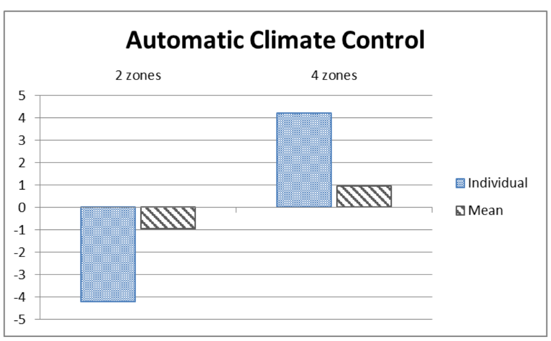
In day-to-day research practice researchers often talk about individual utility values estimated with hierarchical Bayes methods. However, pooling respondents together to get enough information and using multivariate normal distribution assumptions for estimating our models can result in much greater shrinkage than many practical users of HB are aware of.

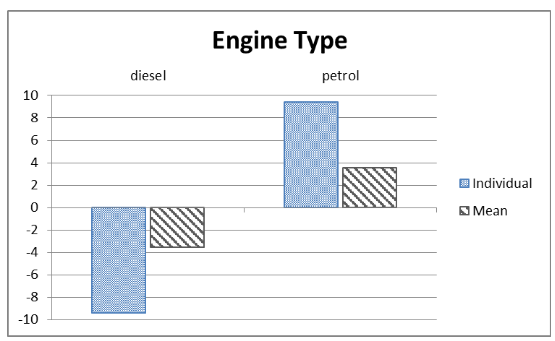
We analyzed a large-scale multinational study where we had a large database of recorded respondent observations in which they explained their preferences while taking the survey and compared these with their individual utilities. We aimed to get answers to the following questions:

Is it possible that individual preferences get washed out due to shrinkage?
Does it always happen?
Why does it happen?
What learnings can researchers take away?

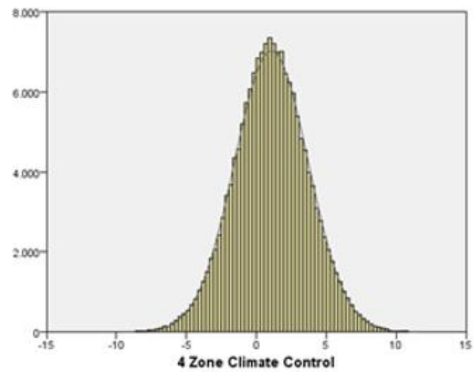
Let's start with cases where we found individual behavior perfectly captured in the individual utilities.

The following two examples are ones where the individual opinion is consistent with population mean. In these two examples of two level attributes the individual preference shows in the same direction as the sample mean:

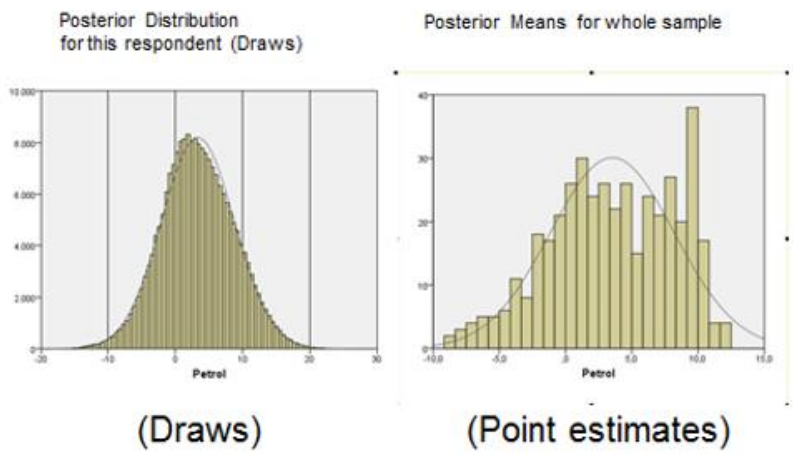




Furthermore the parameter captured a reasonable amount of variation as the following plot of estimation draws across all respondents shows:



Another interesting observation is that the mean over the last 1000 draws (posterior means) of the sample seems not be able to represent the true distribution (plotted line showing the nearly normal distributed real values):

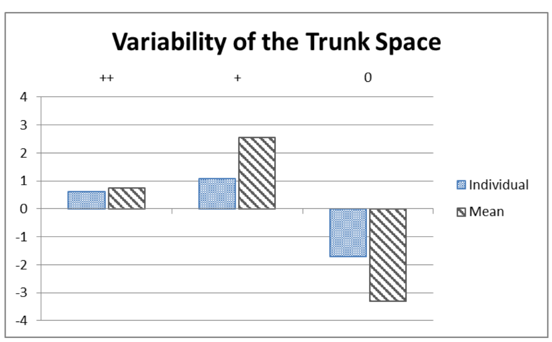


Let's now have a look at some cases where we did not capture individual level behavior:

In the first case the respondent's preference is not consistent with the population mean. This respondent gave a clear explanation of his preferences:

"I would say the first one because of variability of seats is good and more so the variability of the trunk space is better and the comfort of loading and unloading."

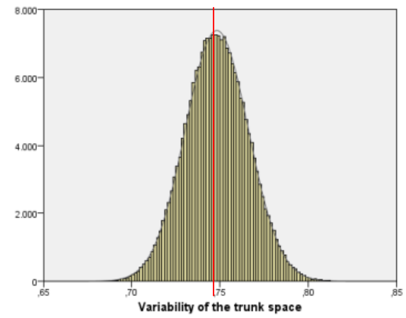
However, the individual utility values of this respondent look completely different:



There seems to be so much shrinkage that the order of preference between the first two levels of this respondent got reversed and there was not enough individual information to prevent that.

The posterior distribution for the sample of this case shows how narrowly distributed all draws are around the population mean:

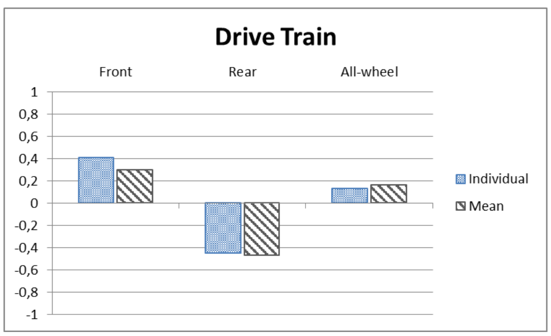
Posterior Distribution for ++ Level



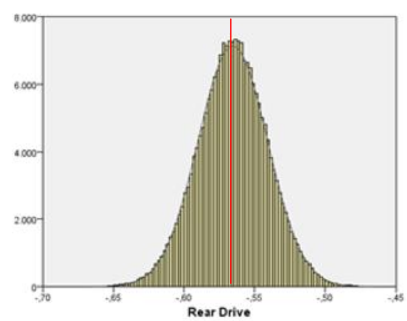
Another respondent among our examples has a clear preference:

“And plus it’s rear wheel drive.”—“And that’s important to you?”—“Yes.”

This respondent’s preference is also not in line with the population mean. As there is again not enough individual information to capture the difference, a strong shrinkage towards the population mean can be observed:



And again the parameter estimates did not capture much variance:



What Did We Learn from this Analysis?

Point estimates (posterior means) don’t always reflect the true variance (spread) in the population. However, we observed that posterior *draws* from the estimation process are usually doing better.

However, sometimes the draws too do not capture individual heterogeneity. This might be caused by several factors:

1. Too sparse individual information due to high model complexity (number of attributes and levels, alternative specific designs etc.) and limited number of choice tasks
2. Lack of covariates or (sometimes even worse) selection of the wrong covariates
3. Too small representation of a specific segment in the sample, caused by small share in the market (such as niche markets or exotic target groups)
4. Unusual individual answering behavior (i.e., outliers)

Bottom line: the failure to capture individuality is most severe for respondents who differ greatly from the population average. This is especially crucial when searching for segments or niches and when analyzing individual cases for any other purpose as well.

What If We Do Need to Search for Segments or Niches?

Before searching for segments or niches through individual results, researchers should apply more diagnostics to their studies and models. The following questions could provide some guidance:

- What does the tail-behavior of the distribution of the posterior draws look like? Are there any visual indications of shrinkage?
- Would we do better to apply mixtures of MVN's instead of using one single MVN?
- Are there effective covariates we could apply?
- Could we have collected more individual information, or reduced the number of parameters in the model?
- Can we simulate from draws or from the upper-level model, considering the distribution and structure of variance-covariance, instead of using point-estimates?

How to Avoid Mistakes in the Future?

From what we learned we advise practitioners to:

- Recognize that DCM—like almost every quantitative method—carries the danger of the mean fallacy error ("stuck in the middle").
- Always examine the individual distribution of your parameter estimates.
- Never forget that DCM models are generally excellent for the total market, but due to possible shrinkage effects they bear danger for small segments or niches
- Consider aggregated models such as Logit or distribution free LC, especially if there are attributes which are polarizing for a minority in the sample.
- Always try to understand the target group (consider pre-research in order to identify sub-segments or possible covariates).
- If one expects sparse data or specific sub segments the best practice depends on whether these segments are known up front:
    - If such segments are known in advance, apply covariates or sample-boost rare cells.
    - If such segments are unknown, apply diagnostics to the model (e.g., counts and parameter variations) and be aware of possible lack of individual level behavior in the model when running analysis and when drawing conclusions from the data.

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, S. 418–440, SAGE Publications Inc., Thousand Oaks.

Hein, M.; Kurz, P.; Steiner, W. (2013): Limits for Parameter Estimation in Choice-Based Conjoint Analysis: A Simulation Study, 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.

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

Pachali, M.; Kurz, P.; Otter, T. (2014): How to Generalize from a Hierarchical Model, 2014 ART Forum Presentation

Rossi, P; Allenby, G.; McCulloch, R. (2005): Bayesian Statistics and Marketing, Wiley, Hoboken NJ.

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