



Sawtooth Software

RESEARCH PAPER SERIES

New Advances Shed Light on HB Anomalies

Bryan Orme,
Sawtooth Software, Inc.

New Advances Shed Light on HB Anomalies

Bryan Orme, Sawtooth Software
Copyright Sawtooth Software 2003

Hierarchical Bayes estimation for choice data represents one of the most successful new developments in our field. HB has proven robust for ratings-based conjoint, ACA, and full-profile CBC projects. Tests comparing HB to other methods of part worth estimation have generally favored HB. However, two anomalies specific to HB estimation have caused us some puzzlement and concern.

“Omitted” Level Estimation under Effects Coding

In “The Joys and Sorrows of Implementing HB Methods for Conjoint Analysis” presented at the 1999 HB Conference at Ohio State, our Chairman, Rich Johnson, showed that estimation of the last (omitted) level for each attribute under effects coding was problematic. The part worth for the omitted level (the reference level, constrained to be negative the sum of the other levels within the same attribute) had larger variance than other levels within the same attribute. The more levels for an attribute, the more pronounced the inflation of variance for the omitted level. Rich suggested the opposite was true for dummy coding (the omitted level has a lower variance relative to the other parameters). He concluded, “Thus, with HB, unlike OLS, it makes a difference how the data are coded. I don’t know what the practical consequences of this will turn out to be. But, like many other things about Bayesian analysis, it was a surprise to me.”

We have since learned that for most data sets the bias in variance for the omitted level under effects coding has been of little practical consequence. But, for particularly sparse data sets, especially when there are many levels within an attribute, the estimation of the omitted level can be severely compromised. Not only is the variance amplified, but for the most extreme cases, we’ve noted just recently that the point estimate itself can be significantly biased in the negative direction.

Partial-Profile CBC and HB Anomaly

In a paper presented at the 2001 Sawtooth Software Conference (“The Effects of Disaggregation with Partial Profile Choice Experiments”), Jon Pinnell and Lisa Fridley analyzed nine different partial-profile CBC studies that had been conducted by three different research agencies. Using our CBC/HB software and an aggregate logit solution, they compared the ability of the respective part worths to predict each individual’s choices to different random choice tasks that were “held out.” We would naturally expect that the HB solution should improve individual-level classification rates relative to an aggregate model where all respondents are pooled. Surprisingly, the authors found that HB did *worse* in four of the nine data sets, and offered no improvement for two others. We’ve puzzled over this finding, as it seems that the data borrowing mechanism in HB should appropriately leverage the more robust information available from the

population parameters relative to the relatively sparse data available at the individual level. Only recently do we have an explanation and a better solution.

Improving the Prior Covariance Matrix Specifications

In the hierarchical Bayes world, we begin with a prior assumption about individual and population parameters and update that information as new data are added. The population-level *priors* consist of a vector of means and a covariance matrix. The degrees of freedom are also specified for the prior covariance matrix, indicating how much weight should be given to the priors versus the data.

Academics have generally suggested using zero for the means, a covariance matrix proportional to an identity matrix (some positive constant across the diagonal, and zeros in the off-diagonal elements), and degrees of freedom for the covariance matrix equal to the number of parameters to be estimated (plus a small integer constant). When there is enough information available at the individual level relative to the number of parameters to be estimated, these prior assumptions have very little effect on the posterior part worth estimates. However, under extreme conditions such as having many levels of an attribute under effects coding, or estimating many parameters from sparse partial-profile CBC designs, proper specification of the prior covariance matrix matters.

With direction from Peter Lenk of the University of Michigan (a leading academic in HB) we have modified how our CBC/HB software sets the prior covariance matrix. If effects-coding is specified, we introduce appropriate negative covariances in the off-diagonal elements of the prior covariance matrix, to reflect the fact that levels *within* each attribute are necessarily negatively correlated. This step resolves the problems stated above for estimating the part worths for omitted levels. Next, we permit the user to set the prior variance along with the degrees of freedom for the prior covariance matrix, thereby tuning the assumed between-respondent variance and relative contribution of the priors versus the data. This tuning can be important for modeling sparse data sets, such as is the case with some partial-profile CBC designs.

We have examined two of the partial-profile CBC data sets that Pinnell and Fridley found problematic when they used our previous CBC/HB software. The assumptions regarding the priors in the old CBC/HB software seemed to lead to overfitting at the individual level (and reduction in hit rate accuracy to holdouts) for these data sets. Using the new CBC/HB software, we decreased the prior variance assumption (assuming greater homogeneity in the population) and increased the degrees of freedom for the prior covariance matrix (increasing the weight for the prior). The hit rates for the new HB solutions improved, and were now slightly higher than aggregate logit in both cases. We believe the small remaining difference in hit rate is due to a combination of low heterogeneity in the sample and little information available for each individual relative to the number of parameters to be estimated.

The important take-away for partial-profile CBC is that the previous failures that Pinnell and Fridley illustrated for HB were not inherent to HB or partial profile designs, but were

a result of the fixed priors we used in our software that were suboptimal for these particular data sets. With the newest CBC/HB software (v3), the researcher can tune the priors, avoiding overfitting in these unusual cases. We should emphasize, however, that the defaults in both the previous and current versions of CBC/HB software seem to work very well for most CBC data sets in practice—especially full-profile designs.

To learn more about the enhancements in the newest version of CBC/HB software, please refer to the CBC/HB v3 Technical Paper in our Technical Papers library at www.sawtoothsoftware.com.

More Flexible Priors in HB-Regression

The ability to tune the prior variance and degrees of freedom for the prior covariance matrix may be even more valuable in the context of our generalized HB program for regression-based problems, called HB-Reg. With our other conjoint-based HB systems, we could make reasonable assumptions about the relative scaling of the dependent variable and the conjoint part worth coefficients. In contrast, HB-Reg is a generalized system for analyzing many types of user-formatted data, and we cannot make general assumptions regarding the scaling of the variables and the related measures of variance. The previous version of HB-Reg assumed a fixed prior variance for betas. If there was relatively sparse information available at the respondent level (which is often the case) and if the data were scaled much differently from the prior variance assumption, the estimation could be sub-optimal or even incorrect.

The new version of HB-Reg (version 3) permits the user to modify the prior settings as described previously for CBC/HB. Additionally, for very advanced users we've provided a way to supply a user-specified prior covariance matrix. These improvements should increase the overall applicability and value of HB-Reg. The new Windows interface in version 3 should also make it easier to use, and the graphical display of the parameter estimates by iteration leads to easier determination of convergence.