Web Appendix

Hierarchical Bayes Conjoint Choice Models – Model Framework, Bayesian Inference, Model Selection, and Interpretation of Estimation Results

By *Nils Goeken*, *Peter Kurz*, and *Winfried J. Steiner*

We further estimated the three choice models (HB-MNL, MoN-HB-MNL and DPM-HB-MNL models without z-variables) using part-worth utility functions as well for the brand-specific price attributes (leading to the estimation of a total of 94 part-worth parameters on an individual respondent level compared to 44 part-worth parameters in case the price attributes were coded linearly). We chose the same diffuse prior settings for model estimation as outlined in section 2.3 and set the fractional likelihood parameter to 1 as before. The MCMC sampler was again run for 210,000 iterations with a burn-in period of 110,000 iterations. To reduce possible correlation among the draws and to prevent internal storage problems, we once more used every 100th draw of the remaining 100,000 draws.

Tab. A1 shows the results for the measures of performance. Both the DPM-HB-MNL model and the MoN-HB-MNL model again returned one-component solutions for our data. Goodness-of-fit statistics (LL, PC, RLH, and IHR) largely improved for all three models, as was expected due to the much larger number of estimated parameters compared to the models with linear price attributes. However, the DPM-HB-MNL model now provided a much worse model fit than the HB-MNL and MoN-HB-MNL models, i.e. goodness-of-fit statistics here improved much less than for the other two models, compare *Tab. 2*. On the other hand, the DPM-HB-MNL model continued to yield the best predictive performance in terms of OHR (even if the credible intervals obtained for the three choice models overlap). Importantly, the predictive performance of all three choice models decreased somewhat compared to using linear price attributes (compare *Tab. 2*), indicating that the higher flexibility of using part-worth utility functions for the brand-specific price attributes did not pay off for prediction purposes but rather favoured overfitting of the data.

Model	# of mixture components	LL	PC	RLH	IHR	OHR
HB-MNL	M=1	[-18748.17;-17917.41]	[.760;.771]	[.717;.728]	[.864;.871]	[.503;.527]
MoN-HB-MNL	M=9[a]	[-19123.03;-18304.96]	[.755;.766]	[.712;.723]	[.861;.868]	[.503;.529]
DPM-HB-MNL	[M=1][b]	[-33558.75;-32611.34]	[.571;.583]	[.551;.561]	[.755;.763]	[.521;.546]

Notes: a: MoN-HB-MNL model estimated initially with nine components, allowing the components to be shut down in the posterior, and resulting in a one-component solution.
b: The number of components were obtained as a result a posteriori. The DPM-HB-MNL model returned one component for our data set (as indicated by [M=1]) as well.

Tab. A1: Goodness-of-fit and predictive accuracy statistics by model type. Shown are the 95% credible intervals of the posterior distributions

We further experimented with different values of the fractional likelihood parameter where smaller values enable a higher degree of consumer heterogeneity. For example, we obtained nine components for the MoN-HB-MNL model and two components for the DPM-HB-MNL model when setting the fractional likelihood parameter to 0.1 instead of 1.0 in the models with dummy-coded price attributes. As expected, more components allowed a still better model fit for the HB-MoN-MNL and HB-DPM-MNL models but decreased the predictive performance of the two models even further (compared to *Tab. A1*).