TCA/HB Compared to CBC/HB for Predicting Choices Among Multi-Attributed Products

Daniel Baier, Marcin Pełka, Aneta Rybicka, and Stefanie Schreiber

Abstract For some years, choice-based conjoint analysis (CBC) has demonstrated its superiority over other preference measurement alternatives. So, e.g., in a recent study on German and Polish cola consumers, the superiority of CBC over traditional conjoint analysis (TCA) was striking. As one reason for this superiority, the usage of hierarchical Bayes for CBC parameter estimation was mentioned (CBC/HB). This paper clarifies whether this really makes the difference: Hierarchical Bayes is also used for TCA parameter estimation (TCA/HB). The application to the above mentioned data shows, that this improves the predictive validity compared to TCA but is still inferior to CBC/HB in "high data quality cases". However, in "low data quality cases" TCA/HB is superior to CBC/HB.

Daniel Baier and Stefanie Schreiber
Brandenburg University of Technology Cottbus-Senftenberg, Chair of Marketing and Innovation Management, Erich-Weinert-Straße 1, 03046 Cottbus, Germany,
✉ daniel.baier@tu-cottbus.de, stefanie.schreiber@tu-cottbus.de

Marcin Pełka and Aneta Rybicka
Wroclaw University of Economics, Department of Econometrics and Computer Science, Nowowiejska 3, 58-500 Jelenia Góra, Poland,
✉ marcin.pelka@ue.wroc.pl, aneta.rybicka@ue.wroc.pl

ARCHIVES OF DATA SCIENCE, SERIES A
KIT SCIENTIFIC PUBLISHING
Vol. 1, No. 1, S. 77–87, 2016
DOI 10.5445/KSP/1000058747/05
ISSN 2363-9881

1 Introduction

In marketing and market research, the application of preference measurement methods for modeling choices among multi-attributed products has a long history. Maybe the best known family of methods is conjoint analysis (CA). CA started its success in the 1960s and 1970s as an approach that allows to estimate part worths for attribute-levels from rankings or ratings of attribute-level-combinations using regression-like procedures (see, e.g., Green and Rao, 1971). Green et al (2001) used in their overview on CA methods the term traditional CA (TCA) for approaches that rely on ranking and/or rating data. It should be mentioned that in TCA the usage of MONANOVA or ANOVA for parameter estimation leads in many cases to rather similar results (see, e.g., Green and Srinivasan, 1978). This is mostly ascribed to the misfit between few observations and many parameters at the desired individual modeling level. Also, in TCA, despite many methodological improvements over the years (e.g. adaptive conjoint analysis), the basic five application steps remained the same (see, e.g., Green et al, 2001):

1. Determine the attributes and levels that influence the customer's choice decisions.
2. Design a set of fictional attribute-level-combinations (stimuli) for data collection.
3. Collect preferential evaluations of these stimuli from a sample of customers.
4. Derive part worths for the attribute-levels using regression-like procedures.
5. Predict the choices of each customer in an assumed market scenario and aggregate them to market shares or sales volumes.

However, today, not TCA but choice-based CA (CBC) (see, e.g., Louviere and Woodworth, 1983; Sawtooth Software, 2013a) is most frequently applied (see, e.g., Selka and Baier, 2014; Selka et al, 2014, for recent overviews on commercial applications). With CBC, in step 3, instead of rating or ranking attribute-level-combinations, the respondents are repeatedly confronted with sets of attribute-level-combinations (so-called choice sets) and asked to select the most preferred ones. Then, in step 4, a multinomial logit model is used for estimation. However, even more severe as TCA, CBC suffers from the misfit between few observations and many parameters at the individual modeling level. As a consequence only pooled models could be estimated (assuming that groups of customers have identical part worths). Here, hierarchical Bayes

(HB) methods for CBC part worth estimation (see, e.g., Allenby and Lenk, 1994; Sawtooth Software, 2009) provided the solution: Observations are shared across respondents during estimation. So, it is possible to estimate individual part worths from few observations per respondent.

Comparison studies (see, e.g., Elrod et al, 1992; Oliphant et al, 1992; Vriens et al, 1998; Moore et al, 1998; Moore, 2004; Karniouchina et al, 2009; Baier et al, 2015) have shown that CBC outperforms TCA in many cases, especially when CBC/HB is used for parameter estimation. However, since also HB methods exist to estimate TCA model parameters (see, e.g., Lenk et al, 1996; Baier and Polasek, 2003; Sawtooth Software, 2013b; Baier, 2014), the question remains unanswered whether CBC/HB is also superior to TCA/HB. This paper tries to close this gap. In Sect. 2 TCA/HB and in Sect. 3 CBC/HB are shortly described. Then, in Sect. 4, the data from Baier et al (2015) are used to compare TCA/HB and CBC/HB. Section 5 closes with a short conclusion and outlook.

2 Hierarchical Bayes Traditional Conjoint Analysis (TCA/HB)

Let $\mathbf{y}_1,\ldots,\mathbf{y}_n \in \mathbb{R}^m$ describe observed preferential evaluations from n respondents ($i = 1,...,n$) w.r.t. to m stimuli ($j = 1,...,m$). y_{ij} denotes the observed preference value of respondent i w.r.t. stimulus j. $\mathbf{X} \in \mathbb{R}^{m\times p}$ denotes the characterization of the m stimuli by p variables. In case of nominal or ordinal attribute-levels a dummy- or an effect-coding is used. The observed evaluations are assumed to come from the following (lower hierarchical) model:

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \quad \text{for } i = 1,\ldots,n \quad \text{with } \boldsymbol{\varepsilon}_i \sim N(\mathbf{0},\sigma^2\mathbf{I}), \tag{1}$$

with $\mathbf{I}$ as the identity matrix, σ^2 as an error variance parameter.

The individual part worths $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ are assumed (higher hierarchical model) to come from a multivariate normal distribution with mean part worth vector $\boldsymbol{\mu} \in \mathbb{R}^p$ and a (positive definite) covariance matrix $\mathbf{H} \in \mathbb{R}^{p\times p}$:

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu},\mathbf{H}) \quad i = 1,\ldots,n. \tag{2}$$

For estimating the model parameters $(\boldsymbol{\mu},\mathbf{H},\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n,\sigma^2)$, Bayesian procedures provide a mathematically tractable way that combines distributional information about the model parameters with the likelihood of the observed data. The result of this combination, the empirical posterior distribution of the model parameters, is generated by (Gibbs) sampling a sequence of draws from the conditional distributions of the model parameters (see, e.g., Lenk et al, 1996; Baier and Polasek, 2003; Sawtooth Software, 2013b; Baier, 2014, for details) using iteratively the following four steps (starting, e.g., with random values as estimates for the model parameters):

1. Use present estimates of $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ and $\mathbf{H}$ to generate a new estimate of $\boldsymbol{\mu}$. $\boldsymbol{\mu}$ is assumed to be distributed normally with mean equal to the average of the $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ and covariance matrix equal to $\mathbf{H}$ divided by the number of respondents. Randomly draw a new estimate of $\boldsymbol{\mu}$ from this distribution.
2. Use present estimates of $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ and $\boldsymbol{\mu}$ to draw a new estimate of $\mathbf{H}$ from an inverse Wishart distribution in the following way:
   - Calculate $\mathbf{G} = p\mathbf{I} + \sum_{i=1}^{n}(\boldsymbol{\mu} - \boldsymbol{\beta}_i)'(\boldsymbol{\mu} - \boldsymbol{\beta}_i)$.
   - Apply a Cholesky decomposition to $\mathbf{G}^{-1}$ s.t. $\mathbf{G}^{-1} = \mathbf{F}\mathbf{F}'$.
   - Draw $n+p$ vectors $\mathbf{u}_i$ from $N(\mathbf{0}, \mathbf{I})$, calculate $\mathbf{S} = \sum_{i=1}^{n+p}(\mathbf{F}\mathbf{u}_i)(\mathbf{F}\mathbf{u}_i)'$.
   - Set $\mathbf{H} = \mathbf{S}^{-1}$.
3. Use present estimates of $\boldsymbol{\mu}$, $\mathbf{H}$, and σ^2 to draw new estimates of $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ from the following conditional distributions ($i = 1, ..., n$):

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}_i, \mathbf{G}) \text{ with } \mathbf{G} = (\mathbf{H}^{-1} + \sigma^{-2}\mathbf{X}'\mathbf{X})^{-1}, \;\; \boldsymbol{\mu}_i = \mathbf{G}(\mathbf{H}^{-1}\boldsymbol{\mu} + \sigma^{-2}\mathbf{X}'\mathbf{y}_i).$$

4. Use present estimates of $\boldsymbol{\mu}$, $\mathbf{H}$, and $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ to generate a new estimate of σ^2 by a similar – but scalar – approach as in step 2.

The final estimates of the model parameters are obtained by averaging the repeated draws from the above four steps. Here often the draws from the first – so-called burn-in – iterations are omitted.

3 Hierarchical Bayes Choice-Based Conjoint Analysis (CBC/HB)

CBC differs from TCA insofar that respondents are repeatedly confronted with (choice) sets of attribute-level-combinations (stimuli) and asked to select the most preferred one. So, at the (lower hierarchical) level, the choice of stimulus j' out of J alternatives ($j = 1, ..., J$) has to be modeled. As usual in multinomial logit models

$$p_{ij'} = \frac{exp(\mathbf{x}_{j'}\boldsymbol{\beta}_i)}{\sum_{j=1} exp(\mathbf{x}_j\boldsymbol{\beta}_i)} \tag{3}$$

is the probability that the respondent i selects stimulus j' (assuming an independently, identically type I extreme distributed additional error in the utilities, see, e.g., Louviere and Woodworth, 1983; Sawtooth Software, 2013a). $\mathbf{x}_j$ denotes the characterization of the alternative j in this choice task ($j = 1, ..., J$). As with TCA/HB, the individual part worths $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ are assumed (higher hierarchical model) to come from a multivariate normal distribution with mean part worth vector $\boldsymbol{\mu} \in \mathbb{R}^p$ and a (positive definite) covariance matrix $\mathbf{H} \in \mathbb{R}^{p \times p}$:

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \mathbf{H}) \quad i = 1, \ldots, n. \tag{4}$$

Again, for estimating the model parameters $(\boldsymbol{\mu}, \mathbf{H}, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n)$, Bayesian procedures are used. The steps are similar as above, only the draws w.r.t. the $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n)$ differ. Here, a Metropolis Hastings algorithm has to be used. For details see Allenby and Lenk (1994) and Sawtooth Software (2009).

4 Empirical studies: Experiments and Results

For testing whether CBC/HB is superior to TCA/HB the data from Baier et al (2015) are used. The multi-attributed product under investigation was – as already mentioned – cola to be bought in the supermarket with the attributes brand (with levels Coca Cola, Pepsi Cola, other brand), flavor (Cola, Cola with orange, Cola with lemon, Cola with cherry), calorie content (normal, light, zero), caffeine content (caffeinated, caffeine-free), price (0.59 €/l, 0.69 €/l, 0.79 €/l, 0.89 €/l), and bottle size (0.5 l, 1 l, 1.5 l, 2 l). The usage of the unit price per volume (€/l) is somewhat problematic since the respondents are used to buy colas at absolute prices (€), but they were explicetely informed about this difference from the usual buying situation.

The data collection took place at two universities near the German-Polish border. The first experiment in each country was an offline-experiment with a TCA task. For TCA, 25 stimuli were generated using orthogonal plans as proposed by SPSS Conjoint to the above number of attributes and levels. In Germany 199 respondents participated in the TCA experiment, in Poland 194. The second experiment in each country was an online-experiment with a CBC task. The respondents were confronted with 18 choice sets, each consisting of four attribute-level-combinations plus a no-choice option. The number of stimuli and choice sets is somewhat high in both experiments, but the students received an incentive and accepted the (complicated) tasks. In Germany 169 respondents participated in the CBC experiment, in Poland 225. All experiments (also the online-experiments) were performed in a controlled laboratory situation: Interviewers informed the respondents about their tasks and observed the answering process. As an incentive for participating, in all experiments, the respondents received a voucher for a small bottle of cola in the cafeteria of their university. All experiments closed with the same holdout choice task (eight identical holdout choice sets were presented) to evaluate the predictive validity. All experiments were performed during four weeks in May and June 2013.

As discussed in Baier et al (2015), the data collection in all four experiments was possible without problems. However, in all experiments, data inspection showed that there were cheating respondents when filling out the questionnaires: Some obviously didn't sort the stimuli with great efforts (resulting in "similar" orderings compared to the stimuli numbers), some used some simplifying rules when selecting stimuli in the choice sets (e.g. always selecting the first stimulus in the set). The total number of such directly observable cheaters was small (e.g. about 10 % in the Polish samples, 5 % in the German samples), but this supported the impression that the Polish samples were data "of lower quality" than the German samples.

The TCA data were analyzed in Baier et al (2015) using MONANOVA as implemented in SPSS Conjoint whereas the CBC data were analyzed using Sawtooth Software's CBC/HB software (Sawtooth Software, 2009). The "directly observed" low data quality of the Polish samples is reflected in the model fit: So, e.g., 16 from the 194 respondents in the Polish TCA sample showed a Pearson correlation of 0.7 or lower when comparing the observed and the estimated preference values. In the German TCA sample only three respondents showed such a low model fit. On average, the Polish respondents showed a Pearson correlation of 0.858 whereas the German respondents showed a Pearson correlation of 0.982. Similar differences can be observed between the Polish and German CBC samples. So, we refer to the two German experiments in the following as a "high data quality case" whereas the two Polish experiments are referred to as the "low data quality case". Especially in the "high data quality case", the CBC experiment showed in Baier et al (2015) a clear superiority w.r.t. predictive validity.

Now, for the research question in this paper, we analyzed also the TCA data with hierarchical Bayes procedures. We used Sawtooth Software's Hierarchical Bayes Regression software for this purpose (Sawtooth Software, 2013b) with 50,000 draws as burn-ins and 10,000 draws for calculating the parameter estimates. We used constraints w.r.t. the price levels in order to prevent

Table 1 Averaged standardized part worths and attribute importances for the different samples in Germany and Poland (TCA/HB=Hierarchical Bayes Traditional Conjoint Analysis, CBC/HB=Hierarchical Bayes Choice-Based Conjoint Analysis); sample differences between Germany and Poland were t-tested; *: significant at α=.05, **: at α=.01, ***: at α=.001

		Averaged standardized part worths (std. dev.)							
		Germany				Poland			
Attribute	Level	TCA/HB (n=199)		CBC/HB (n=169)		TCA/HB (n=194)		CBC/HB (n=225)	
Brand	Coca Cola	.181	(.201)	.172	(.127)	.141	(.157)	.144	(.172)
	Pepsi Cola	.072	(.108)	.075	(.073)	.119	(.144)	.121	(.141)
	Other	.091***	(.141)	.015	(.035)	.120*	(.156)	.088	(.164)
Flavor	Cola	.236	(.188)	.241	(.140)	.111	(.137)	.164***	(.162)
	W. orange	.119	(.152)	.094	(.111)	.055	(.078)	.072	(.109)
	W. lemon	.153	(.146)	.132	(.117)	.125	(.130)	.105	(.111)
	W. cherry	.089*	(.124)	.060	(.100)	.150***	(.115)	.101	(.135)
Calorie	Normal	.167	(.156)	.193	(.167)	.086**	(.084)	.064	(.080)
	Light	.064	(.097)	.086*	(.081)	.086***	(.085)	.037	(.056)
	Zero	.042	(.076)	.046	(.096)	.037	(.067)	.048	(.071)
Caffeine	Caffein.	.128***	(.138)	.083	(.092)	.039	(.062)	.035	(.049)
	C.-free	.009	(.036)	.014	(.045)	.034	(.049)	.026	(.060)
Price	.59 €/l	.029	(.072)	.087***	(.095)	.097	(.071)	.130*	(.148)
	.69 €/l	.022	(.049)	.060***	(.065)	.031	(.037)	.100***	(.074)
	.79 €/l	.011	(.029)	.038***	(.044)	.002	(.001)	.086***	(.079)
	.89 €/l	.000	(.000)	.013***	(.025)	.000	(.000)	.044***	(.094)
Bottle size	0.5 l	.085***	(.093)	.027	(.049)	.049	(.059)	.056	(.090)
	1 l	.079*	(.085)	.061	(.052)	.106**	(.090)	.059	(.075)
	1.5 l	.054	(.064)	.049	(.050)	.112**	(.094)	.071	(.075)
	2 l	.033	(.051)	.048**	(.056)	.093	(.090)	.075	(.108)

	Averaged attribute importances (std. dev.)							
	Germany				Poland			
Attribute	TCA/HB		CBC/HB		TCA/HB		CBC/HB	
Brand	.231*	(.197)	.187	(.116)	.254	(.160)	.256	(.192)
Flavor	.292	(.180)	.279	(.133)	.256	(.125)	.255	(.170)
Calorie	.189	(.150)	.241***	(.141)	.136***	(.083)	.105	(.086)
Caffeine	.138***	(.134)	.096	(.091)	.073*	(.061)	.061	(.065)
Price	.029	(.072)	.103***	(.090)	.097	(.071)	.186***	(.140)
Bottle size	.122***	(.090)	.094	(.061)	.184***	(.080)	.137	(.121)

degeneration. The internal validity of TCA/HB was 0.767 (averaged R^2 across respondents) for the German sample and 0.418 for the Polish sample. Please note, as in Baier et al (2015) for TCA and CBC/HB, the “lower quality" of the Polish data. The internal validity of CBC/HB (see Baier et al, 2015) was 0.647 (averaged root likelihood value across respondents) for the German and 0.598 for the Polish sample. However, the respondents with a low model fit were not removed from the further analyses since we explicetely wanted to demonstrate the ability of the different procedures to deal with the “low data quality" problem.

Afterwards, for comparison reasons, the estimated part worths at the respondent level were standardized in the usual way so that – for each respondent in each experiment – the maximum possible value for a stimulus is 1 and the minimum 0. Also the individual attribute importances in each experiment were calculated via the difference of the highest and the lowest part worth for levels of the corresponding attribute. Table 1 gives the averaged part worths for all four experiments (TCA/HB, CBC/HB in Germany and Poland) and also averaged importances. From Table 1 one can easily see, that – more or less – the results across nationality (also: “data quality", see above) and methods are similar: “Flavor" is – on average – the most important attribute when selecting colas, followed by “brand" and “calorie". The importance of “price" differs between German and Polish consumers but also between TCA/HB and CBC/HB. So, e.g., with CBC/HB, the importance of price is much higher than with TCA/HB. This can be partly ascribed to the well-known fact that “simple" or “quantifiable" attributes are more looked at in the choice than in the ranking setting (see, e.g., Karniouchina et al, 2009; Baier et al, 2015), but also to the constraining of the price parameters during TCA/HB estimation.

The most important comparison deals with the predictive validity. Here the responses to the holdout choice tasks (identical in all experiments) have to be compared with model predictions (assuming that the respondent selects the stimulus with highest predicted sum of part worths in each holdout choice set). There are two possibilities to calculate them: One could control all holdout choices (including the selections of the no-choice option, in total n=9,880 selections) or one could control only the holdout choices where a holdout stimulus was selected (excluding the selection of the no-choice option, in total n=7,803 selections). The fair comparison would be the second one (see, e.g., Karniouchina et al, 2009; Baier et al, 2015), since TCA resp. TCA/HB do not collect data to predict the no-choice option, only CBC resp. CBC/HB are able to give such predictions.

Table 2 First Choice Hit Rates (FCHR). (TCA/HB=Hierarchical Bayes Traditional Conjoint Analysis, CBC/HB=Hierarchical Bayes Choice-Based Conjoint Analysis); "With NC" stands for all choices of respondents in the holdout tasks including the no-choice selections, for "Without NC" the no-choice selections are excluded; a binomial test was applied to compare the results: FCHR of one method was assumed and checked whether FCHR of the other is higher; *: significant at α=.05, **: at α=.01, ***: at α=.001

	First Choice Hit Rates			
	Germany		Poland	
Holdout choices considered	TCA/HB	CBC/HB	TCA/HB	CBC/HB
With NC	.456	.710***	.339	.335
Without NC	.648	.726*	.378	.323

So, for the second one, in Germany, CBC/HB outperforms TCA/HB with First Choice Hit Rate (FCHR, the percentage of correctly predicted selections in the holdout choice sets) values of 0.726 (for CBC/HB) and 0.648 (for TCA/HB). The TCA/HB value is better than the value 0.626 for TCA in Baier et al (2015), but still worse than CBC/HB. See Table 2. In the Polish experiments, TCA/HB (FCHR=0.378) outperforms CBC/HB (FCHR=0.323). See Table 2. However, since the FCHR values for the Polish experiments are very low, one should nevertheless conclude that CBC is superior with respect to prediction, especially in "high data quality cases".

5 Conclusions and outlook

The analyses have shown, that – especially in "high data quality cases" – for choice predictions, CBC/HB outperforms TCA and also – as a result of this paper – TCA/HB. However, especially in markets with "low data quality", the TCA/HB approach competes well, especially when the no-choice options are neglected. Of course, the comparisons between traditional and choice-based methods have not come to an end, one needs more such comparisons.

However, already from the comparison in this paper, one can draw ideas for methodological improvements: The main improvement with respect to predictive accuracy but also with respect to model fit comes from the motivation of the respondents to answer the questionnaires carefully. The "low data quality" problem in the Polish samples can only partly be compensated by the better HB estimation procedure. So, it seems that improvements that focus on better data collection and respondent motivation procedures are of higher value for marketing research practice than better parameter estimation procedures.

References

Allenby GM, Lenk PJ (1994) Modeling household purchase behavior with logistic regression. Journal of the American Statistical Association 89(428):1218–1231

Baier D (2014) Bayesian Methods for Conjoint Analysis-Based Predictions: Do We Still Need Latent Classes?, Springer International Publishing, Cham, pp 103–113. DOI 10.1007/978-3-319-01264-3_9

Baier D, Polasek W (2003) Market simulation using Bayesian procedures in conjoint analysis. In: Schwaiger M, Opitz O (eds) Exploratory Data Analysis in Empirical Research, Studies in Classification, Data Analysis, and Knowledge Organization, Springer, Berlin, pp 413–421, DOI 10.1007/978-3-642-55721-7_42

Baier D, Pełka M, Rybicka A, Schreiber S (2015) Ratings-/Rankings-Based Versus Choice-Based Conjoint Analysis for Predicting Choices, Springer, Berlin, pp 205–216. DOI 10.1007/978-3-662-44983-7_18

Elrod T, Louviere JJ, Davey K (1992) An empirical comparison of ratings-based and choice-based conjoint models. Journal of Marketing Research 29(3):368–377

Green PE, Rao VR (1971) Conjoint measurement for quantifying judgmental data. Journal of Marketing Research 8(3):355–363

Green PE, Srinivasan V (1978) Conjoint analysis in consumer research: Issues and outlook. Journal of Consumer Research 5(2):103–123

Green PE, Krieger AM, Wind Y (2001) Thirty years of conjoint analysis: Reflections and prospects. Interfaces 31(3b):56–73

Karniouchina E, Moore WL, Van der Rhee B, Verma R (2009) Issues in the use of ratings-based versus choice-based conjoint analysis in operations management research. European Journal of Operational Research 197(1):340–348

Lenk PJ, DeSarbo WS, Green PE, Young MR (1996) Hierarchical Bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. Marketing Science 15(2):173–191

Louviere JJ, Woodworth G (1983) Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data. Journal of Marketing Research 20(4):350–367

Moore WL (2004) A cross-validity comparison of ratings-based and choice-based conjoint analysis models. International Journal of Research in Marketing 21(3):299–312, DOI 10.1016/j.ijresmar.2004.01.002

Moore WL, Gray-Lee J, Louviere JJ (1998) A cross-validity comparison of conjoint analysis and choice models at different levels of aggregation. Marketing Letters 9(2):195–207, DOI 10.1023/A:1007913100332

Oliphant K, Eagle T, Louviere J, Anderson D (1992) Cross-task comparison of ratings-based and choice-based conjoint. In: Metegrano M (ed) Sawtooth Software Conference Proceedings, Sawtooth Software Inc., Ketchum, Idaho, pp 383–404

Sawtooth Software (2009) The CBC/HB System for Hierarchical Bayes Estimation Version 5.0 Technical Paper. Sawtooth Software Inc., Sequim, Washington, URL `http://www.sawtoothsoftware.com/support/technical-papers/hierarchical-bayes-estimation/cbc-hb-technical-paper-2009`

Sawtooth Software (2013a) The CBC System for Choice-Based Conjoint Analysis Version 8. Sawtooth Software Inc., Orem, Utah

Sawtooth Software (2013b) HB-Reg v4 for Hierarchical Bayes Regression. Sawtooth Software Inc., Orem, Utah

Selka S, Baier D (2014) Kommerzielle Anwendung auswahlbasierter Verfahren der Conjointanalyse: Eine empirische Untersuchung zur Validitätsentwicklung. Marketing ZFP 36(1):54–64

Selka S, Baier D, Kurz P (2014) The validity of conjoint analysis: An investigation of commercial studies over time. In: Spiliopoulou M, Schmidt-Thieme L, Janning R (eds) Data Analysis, Machine Learning and Knowledge Discovery, Springer International Publishing, Cham, pp 227–234, DOI 10.1007/978-3-319-01595-8_25

Vriens M, Oppewal H, Wedel M (1998) Rating-based versus choice-based latent class conjoint models: An empirical comparison. Journal of the Market Research Society 40(3):237–248