Intern. J. of Research in Marketing 28 (2011) 378–388

Contents lists available at SciVerse ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity

Jie Yu [a], Peter Goos [b,c], Martina Vandebroek [a,d,*]

[a] Faculty of Business and Economics, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium
[b] Faculty of Applied Economics & StatUa Center for Statistics, University of Antwerp, Prinsstraat 13, B-2000 Antwerpen, Belgium
[c] Erasmus School of Economics, Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands
[d] Leuven Statistics Research Centre, KULeuven, Celestijnenlaan 200B, B-3001 Leuven, Belgium

ARTICLE INFO

Article history:
First received in 10 September 2010 and was under review for 4 $^{1}/_{2}$ month
Available online 12 September 2011

Area Editor: Joel C. Huber

Keywords:
Adaptive Bayesian design
Conditional logit model
Conjoint-choice experiments
Consumer heterogeneity
Panel mixed logit model
Response accuracy
Response error
Sequential design

ABSTRACT

We propose an efficient individually adapted sequential Bayesian approach for constructing conjoint-choice experiments, which uses Bayesian updating, a Bayesian analysis, and a Bayesian design criterion to generate a conjoint-choice design for each individual respondent based on the previous answers of that particular respondent. The proposed design approach is compared with three non-adaptive design approaches, two aggregate-customization approaches (based on the conditional logit model and on a mixed logit model), and the (nearly) orthogonal design approach, under various degrees of response accuracy and consumer heterogeneity. A simulation study shows that the individually adapted sequential Bayesian conjoint-choice designs perform better than the benchmark approaches in all scenarios we investigated in terms of the efficient estimation of individual-level part-worths and the prediction of individual choices. In the presence of high consumer heterogeneity, the improvements are impressive. The new method also performs well when the response accuracy is low, in contrast with the recently proposed adaptive polyhedral approach. Furthermore, the new methodology yields precise population-level parameter estimates, even though the design criterion focuses on the individual-level parameters.

1. Introduction

Choice-based conjoint analysis is widely used for product design, segmentation, and marketing strategy decisions. In choice-based conjoint studies, each respondent is requested to indicate his/her preferred product or service from each of a list of choice sets. Research on choice-based conjoint studies has shown that a careful design increases estimation and prediction accuracy and reduces experimental costs by requiring fewer respondents and/or fewer questions. The first papers on efficient designs for choice experiments aimed at the efficient estimation of the conditional logit model (also referred to as the multinomial logit model by some authors in the conjoint-choice design literature), which assumes that the respondents have the same preferences or part-worths, denoted by $\boldsymbol{\beta}$, for the attributes studied in the experiment. The conditional logit model can be estimated by maximum likelihood methods and the efficiency of a design in these papers was assessed using the Fisher information matrix. As the information matrix depends on the parameters $\boldsymbol{\beta}$ to be estimated, it has been assumed in the design stage that either the parameters are zero (yielding utility-neutral designs), that the parameters are known (yielding locally optimal designs), or that the likely values of the parameters can be specified by a prior distribution (yielding Bayesian optimal designs).

Currently, however, the mixed logit model that takes into account the heterogeneity in consumer preferences is commonly used. The mixed logit model assumes that the individual part-worths $\boldsymbol{\beta}_n$ follow a certain distribution across consumers, i.e., $\boldsymbol{\beta}_n \sim f(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})$. In the presence of consumer heterogeneity, one is interested not only in the average preference of an entire population, given by the population mean part-worth vector $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and the heterogeneity involved, which is expressed by the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$, but also in the preferences of individual respondents. Mixed logit models are typically estimated by a hierarchical Bayes estimation procedure, yielding posterior distributions for the individual- and population-level parameters (see, e.g., Allenby & Rossi, 1999; Arora & Huber, 2001; Arora, Allenby, & Ginter, 1998; Lenk, DeSarbo, Green, & Young, 1996, and Train, 2003). However, as was argued by Hensher and Greene (2003), the estimation of mixed logit models requires much better data quality than the conditional logit model because these models are intended to capture a much greater amount of true behavioral variability in choice making.

* Corresponding author at: Faculty of Business and Economics, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium. Tel.: +32 16 32 69 75; fax: +32 16 32 66 24.
E-mail addresses: jie.yu@econ.kuleuven.be (J. Yu), peter.goos@ua.ac.be (P. Goos), martina.vandebroek@econ.kuleuven.be (M. Vandebroek).

0167-8116/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2011.06.002

Although efficient designs for mixed logit models are much more difficult to obtain, several researchers have recently tackled this problem. The main attempts are listed in Table 1, together with Arora and Huber (2001) who optimized designs for conditional logit models and whose aggregate-customization approach is used later for comparison purposes. Sándor and Wedel (2002) were the first to construct efficient conjoint-choice designs that explicitly accounted for respondent heterogeneity. When constructing their designs, they assumed that the population mean preference $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and its covariance $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ were known. Sándor and Wedel's locally optimal design approach was improved by Yu, Goos, and Vandebroek (2009) who allowed for uncertainty about $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ and constructed Bayesian designs. Both studies, however, utilized the Fisher information matrix of the cross-sectional mixed logit model, which ignores the fact that repeated observations are made for each respondent. Nevertheless, the construction of the optimal designs was computationally very demanding because computing the Fisher information matrix requires the numerical evaluation of several high-dimensional integrals over the heterogeneity distribution of the model parameters $\boldsymbol{\beta}_n$.

Except for Toubia, Hauser, and Simester (2004) and Toubia, Hauser, and Garcia (2007), who use a polyhedral heuristic, only Bliemer and Rose (2010) have searched for optimal designs for the panel mixed logit model that correctly deals with the correlation pattern in the data. Bliemer and Rose (2010) do not use a Bayesian approach because this turned out to be computationally infeasible. In addition, the search for locally optimal designs is very time consuming, even for the small examples that they consider. This is because the information matrix for the panel mixed logit model can only be computed using extensive simulations. As a result, the approach in Bliemer and Rose (2010) yields locally optimal designs that are rather sensitive to misspecification of the prior parameter values.

This article presents a new approach to developing efficient designs for estimating the panel mixed logit model while taking into account the uncertainty of the model parameters. A key feature of this approach is that we use individually adapted conjoint-choice designs that capture the individual preferences maximally. We show that the approach leads to the efficient estimation of individual-level and population-level parameters. For web-based questionnaires, these individually optimized (referred to as individually adapted sequential Bayesian or IASB) designs show great promise.

The IASB approach is different from the aggregate-customization approach that was used in most papers on efficient designs for conjoint-choice experiments. The latter approach, which was used in Sándor and Wedel (2002), Yu et al. (2009), and Bliemer and Rose (2010), involves only one design that is optimal for the average respondent but not optimal for each individual respondent. Thus, in the aggregate approach, every respondent evaluates the same choice sets. A disadvantage of this type of design is that, when the respondent heterogeneity is high, the aggregate design might be inefficient for those respondents whose part-worths are located far from the population mean because the aggregate design is optimized for the average respondent. The advantage of using different designs for different respondents was demonstrated by Sándor and Wedel (2005). They constructed a small set of optimal designs that were randomly assigned to the respondents and concluded that this approach was especially beneficial in the presence of respondent heterogeneity. It is thus crucial to take into account respondent heterogeneity when constructing conjoint-choice designs and to use different designs for different respondents.

A novelty of our approach is that we use a Bayesian approach not only for the purpose of design construction but also to update the prior information for each individual respondent after each response. The Bayesian data analysis allows us to use a small initial design for each respondent and to update prior information on a consumer's preferences repeatedly during the survey. The small initial design and the continuous updating of prior information enable us to exploit any available information from each respondent to the largest possible extent for constructing the subsequent choice set(s). This approach ensures that each individual design is tailor-made for every respondent and does an excellent job capturing individual-level preferences. The total set of choices made by all respondents is finally used to estimate the panel mixed logit model by the hierarchical Bayes methodology. We show that, although the IASB approach procedure focuses on the precise estimation of the individual-level parameters, it also yields precise population-level parameter estimates.

The only other approach that uses a tailor-made design for each respondent is the adaptive method introduced in the literature by Toubia et al. (2004). In that approach, each choice set for a given respondent is constructed sequentially using polyhedral methods based on the respondent's previous responses so that the design is also tailored to each of the respondents. This adaptive approach tends to perform better than aggregate-customization approaches when the response error is low but it does not perform well when the response error is high. The poor performance in the presence of low response accuracy inspired Toubia et al. (2007) to adapt the original polyhedral method so that it takes into account response error. We compare our results briefly with those of Toubia et al. (2004, 2007) in Section 6.

In the next section, we present the panel mixed logit model for analyzing data from choice-based conjoint experiments in the presence of consumer heterogeneity. In Section 3, we describe the construction of IASB designs. In Section 4, we evaluate our approach for different levels of response accuracy and different levels of respondent heterogeneity by means of a simulation study. We describe a real-life application of the IASB approach in Section 5. Finally, in Section 6, we compare our results with those of the polyhedral method, discuss the usefulness of a stopping rule and summarize our main findings.

2. The panel mixed logit model

In a choice-based conjoint experiment, every respondent evaluates several choice sets with a certain number of alternatives so that repeated observations are made for each respondent. We denote the number of choice sets evaluated by each respondent by S and the number of alternatives in each of these choice sets by K. When the respondents are heterogeneous, every respondent has his/her own preference structure.

Table 1
Important references related to this paper.

Paper	Model	Design criterion	Type of design
Arora and Huber (2001)	Conditional logit model	Locally optimal ($\boldsymbol{\beta}$ assumed to be known)	Aggregate customization (customized for an average respondent)
Sándor and Wedel (2002)	Cross-sectional mixed logit model	Locally optimal ($\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ assumed to be known)	Aggregate customization
Sándor and Wedel (2005)	Cross-sectional mixed logit model	Locally optimal ($\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ assumed to be known)	Differentiated designs (randomly assigned)
Toubia et al. (2004, 2007)	Panel mixed logit model	Polyhedral methods (updated priors on $\boldsymbol{\beta}_n$)	Individually optimized designs
Yu et al. (2009)	Cross-sectional mixed logit model	Bayesian (use of prior on $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$)	Aggregate customization
Bliemer and Rose (2010)	Panel mixed logit model	Locally optimal ($\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ assumed to be known)	Aggregate customization
This paper	Panel mixed logit model	Bayesian (updated priors on $\boldsymbol{\beta}_n$)	Individually optimized designs

In the panel mixed logit model, the preference structure of a given respondent, n, is captured by the vector of part-worths β_n. That vector is a p-dimensional parameter vector containing the effects of different attribute levels on the utility specific to respondent n. Conditional on β_n, the probability that respondent n chooses alternative k ($k=1,\ldots,K$) in choice set s ($s=1,\ldots,S$) is

$$p_{ksn}(\beta_n) = \frac{\exp\left(x'_{ksn}\,\beta_n\right)}{\sum_{k=1}^{K}\exp(x'_{ksn}\,\beta_n)}, \tag{1}$$

where $\mathbf{x}_{ksn}$ is a p-dimensional vector characterizing the attributes of alternative k in choice set s for respondent n.

We denote the choices from respondent n for the S choice sets of K alternatives using the KS-dimensional vector $\mathbf{y}_n^S$, whose elements y_{ksn} equal one if respondent n chooses alternative k in choice set s and zero otherwise. The likelihood of a given sequence of choices $\mathbf{y}_n^S$ for respondent n can then be written as

$$L\left(y_n^S \middle| \mathbf{X}_n^S, \beta_n\right) = \prod_{s=1}^{S}\prod_{k=1}^{K}(p_{ksn}(\beta_n))^{y_{ksn}}, \tag{2}$$

where $\mathbf{X}_n^S$ is a matrix containing the characteristics of each of the alternatives in the S choice sets that have been assigned to respondent n.

We assume that, for each respondent, β_n is randomly drawn from a p-variate normal distribution $f(\beta_n|\mu_\beta, \Sigma_\beta)$, with mean μ_β and covariance matrix Σ_β. Under this assumption, the likelihood, unconditional on β_n, of respondent n's sequence of choices $\mathbf{y}_n^S$ is

$$\begin{aligned} L\left(\mathbf{y}_n^S \middle| \mathbf{X}_n^S, \mu_\beta, \Sigma_\beta\right) &= \int L\left(\mathbf{y}_n^S \middle| \mathbf{X}_n^S, \beta_n\right) f(\beta_n|\mu_\beta, \Sigma_\beta)\,d\beta_n, \\ &= \int\left(\prod_{s=1}^{S}\prod_{k=1}^{K}(p_{ksn}(\beta_n))^{y_{ksn}}\right) f(\beta_n|\mu_\beta, \Sigma_\beta)\,d\beta_n. \end{aligned} \tag{3}$$

This model is a panel mixed logit model because it explicitly takes into account the fact that repeated observations are collected from the same individuals. The use of a single parameter vector β_n across all choice sets evaluated by a given respondent ensures that the model captures the within-respondent correlation across repeated choices. Essentially, the panel mixed logit model assumes that every individual respondent behaves according to the conditional logit model, as expressed in Eq. (1), and it combines the individual logit models into a population-level model.

In this paper, we use the hierarchical Bayes approach to obtain the individual-level and population-level parameter estimates for the panel mixed logit model. The corresponding likelihood function of the observed responses for N respondents is

$$\begin{aligned} L\left(\mathbf{y}_{full} \middle| \mathbf{X}_{full}, \mu_\beta, \Sigma_\beta\right) &= \prod_{n=1}^{N} L\left(\mathbf{y}_n^S \middle| \mathbf{X}_n^S, \mu_\beta, \Sigma_\beta\right), \\ &= \prod_{n=1}^{N}\int\left(\prod_{s=1}^{S}\prod_{k=1}^{K}(p_{ksn}(\beta_n))^{y_{ksn}}\right) f(\beta_n|\mu_\beta, \Sigma_\beta)\,d\beta_n, \end{aligned} \tag{4}$$

where $\mathbf{y}_{full}$ contains the responses for the N respondents and $\mathbf{X}_{full}$ is the full design matrix, which concatenates the choice designs $\mathbf{X}_1^S$, $\mathbf{X}_2^S,\ldots, \mathbf{X}_N^S$ of each of the N individual respondents.

The key feature of hierarchical Bayes estimation is that the initial heterogeneity distribution of the individual preferences is used as a prior for the individual-level parameters for each respondent. In the Gibbs sampling procedure, the prior distribution is updated for each respondent separately using the choice data from that individual only, and the population-level parameters of the heterogeneity distribution are updated using the updated individual-level parameters of all respondents. Finally, point estimates are obtained from the averages of these posterior distributions. More details can be found in Appendix A and in Train (2003).

3. Construction of efficient conjoint-choice designs

3.1. Bayesian approach to design selection

As demonstrated by Bliemer and Rose (2010), the construction of optimal aggregate-customization designs for the panel mixed logit model is very complex because the only practical way to compute the Fisher information matrix for this model is through simulation. In this paper, we propose an individually adapted sequential Bayesian (IASB) approach as an alternative solution to construct choice designs that provide maximum information on the respondent heterogeneity.

The idea of the IASB approach is to generate each respondent's choice sets in a Bayesian framework based on the previous responses given by the respondent and based on the conditional logit model for whose information matrix a computationally convenient analytical expression is available. Note that a conditional logit model involves only one vector of part-worths. In general, these part-worths are assumed to be common for all respondents, and the model is estimated based on the choices of all respondents. In this paper, to construct the individual designs, we fit a different conditional logit model for each respondent separately to allow for the heterogeneity in the population. At the end of the survey in the IASB approach, each respondent ends up with a design that is tailored to his/her preference structure, given by his/her parameter vector β_n. To explain how the IASB approach works, we first introduce the Bayesian D-optimality criterion for selecting a conjoint-choice design.

The Bayesian design criterion that we use is based on the generalized Fisher information matrix of the conditional logit model rather than the ordinary Fisher information matrix, which so far has been used in the literature on the optimal design of choice-based conjoint experiments. The generalized Fisher information matrix is a more natural choice for a framework in which Bayesian estimation is used (see, e.g., Pilz, 1991), and it has been shown by Yu, Goos, and Vandebroek (2008) that this matrix gives a more reliable approximation of the posterior covariance matrix in the case of small data sets. The generalized Fisher information matrix is obtained by taking the negative expectation of the second derivative of the log-posterior density of the model parameters. For β_n, the part-worth vector of respondent n, it is

$$\mathbf{I}_{GFIM}(\beta_n|\mathbf{X}) = -E\left[\frac{\partial^2 \log q(\beta_n|\mathbf{Y},\mathbf{X})}{\partial\beta_n\partial\beta'_n}\right], \tag{5}$$

with $q(\beta_n|\mathbf{Y},\mathbf{X})$ representing the posterior distribution of β_n for a given design $\mathbf{X}$ and the corresponding responses $\mathbf{Y}$. When the prior distribution for β_n is a multivariate normal distribution with covariance matrix Σ_{β_n}, we obtain the following expression for the generalized Fisher information matrix (see Appendix B for a derivation of this result):

$$\mathbf{I}_{GFIM}(\beta_n, \mathbf{X}) = \sum_{s=1}^{S}\mathbf{X}'_s\left(\mathbf{P}_s^n - \mathbf{p}_s^n\mathbf{p}_s^{n'}\right)\mathbf{X}_s + \Sigma_{\beta_n}^{-1}, \tag{6}$$

with $\mathbf{X}_s$ the design matrix for choice set s, $\mathbf{P}_s^n = diag\,[p_{1sn}, p_{2sn}, \ldots, p_{Ksn}]$, and $\mathbf{p}_s^n = [p_{1sn}, p_{2sn}, \ldots, p_{Ksn}]$. A one-dimensional measure of the quality, or the efficiency, of a conjoint-choice design $\mathbf{X}$ for estimating β_n is the D-error

$$D = \det\left[\mathbf{I}_{GFIM}(\beta_n|\mathbf{X})\right]^{-1/p},$$

where p is the dimension of the part-worth vector β_n. The expectation of the D-error over the prior distribution of these part-worths, $\pi(\beta_n)$, is the D_B-error

$$D_B = \int \det\left[\mathbf{I}_{GFIM}(\beta_n|\mathbf{X})\right]^{-1/p}\pi(\beta_n)\,d\beta_n. \tag{7}$$

The design that minimizes this D_B-error is the Bayesian D-optimal design for estimating the parameters β_n in the conditional logit model.

A critical issue in constructing efficient designs is to determine a suitable and realistic prior distribution for the model parameters. Several useful approaches for constructing such prior distributions have been presented in the literature (e.g., Sándor & Wedel, 2001). Additionally, historical data or small pilot studies can be used as a source of inspiration for constructing sensible prior distributions. Obviously, each of these approaches leads to population-level prior information. However, in the presence of respondent heterogeneity, this is unsatisfactory, as it is individual-level information that is relevant for the customization of the design to each individual respondent. Our IASB approach solves this problem by updating the population-level prior information within each individual questionnaire with the information that becomes available each time a respondent has finished evaluating a choice set.

3.2. Individually adapted sequential Bayesian designs

Our proposed IASB design procedure consists of two stages, i.e., an initial static stage involving S_1 choice sets followed by a fully adaptive sequential stage involving S_2 choice sets. A key feature of the IASB approach is that it builds individual designs based on the conditional logit model. This is in line with the panel mixed logit model, which assumes that every individual's choice behavior is described by the conditional logit model. The conditional logit model can therefore safely be used as a basis for tailoring the choice design to each individual respondent. This approach circumvents the computational problems associated with the construction of optimal choice-based conjoint designs for the panel mixed logit model.

In our work, we have studied two slightly different versions of the IASB approach. The first version assumes that the respondents enter the conjoint study one after another. This allows the researcher to update the prior information after each choice and to use the updated prior information to generate the S_1 initial choice sets for the next respondent who enters the study. The second version of the IASB approach that we investigated assumes that the respondents simultaneously complete their choice tasks, in which case the prior information cannot be updated between respondents. In this paper, we focus on the second version of the IASB approach, as it is applicable in every practical situation. This is not true for the first version because, even in on-line conjoint choice experiments, it is unrealistic to assume that there will be no respondents who complete their choice tasks simultaneously. Moreover, the performance of the two versions of the IASB approach is very similar. In Sections 3.2.1 and 3.2.2, we describe the initial static stage and the adaptive sequential stage, respectively. We label the two versions of the IASB approach IASB-SIM (respondents evaluate choice sets simultaneously) and IASB-OAA (respondents enter the study one after another) and assume that the number of choice sets S for each respondent is fixed.

3.2.1. Initial static stage

3.2.1.1. IASB-SIM approach. In the initial static stage of the IASB-SIM approach, we use a common initial prior distribution $\pi(\boldsymbol{\beta})$ for all N respondents reflecting the available prior information on the population heterogeneity from historical data or pilot studies. This prior distribution is used (i) as a basis to generate an initial design with S_1 choice sets for each respondent and (ii) as an input to the Bayesian analysis for a given respondent as soon as the first S_1 responses from that respondent become available.

To construct an initial design of S_1 choice sets for each respondent, we create a large initial design $\mathbf{X}$ consisting of N subdesigns $\mathbf{X}_n^{S_1}$ using the greedy approach introduced by Sándor and Wedel (2005). This approach sequentially optimizes $\mathbf{X}_1^{S_1}$, $\left(\mathbf{X}_2^{S_1}\middle|\mathbf{X}_1^{S_1}\right),\ldots,\left(\mathbf{X}_N^{S_1}\middle|\mathbf{X}_1^{S_1},\ldots,\mathbf{X}_{N-1}^{S_1}\right)$, where $\mathbf{X}_n^{S_1}$ contains the initial S_1 choice sets for the nth respondent. In the greedy approach, the optimal subdesign $\mathbf{X}_1^{S_1}$ is obtained by minimizing the D_B-error in Eq. (7) based on the initial prior distribution $\pi(\boldsymbol{\beta})$. Next, the subdesign $\mathbf{X}_2^{S_1}$ is obtained by minimizing the joint D_B-error of $\mathbf{X}_1^{S_1}$ and $\mathbf{X}_2^{S_1}$ in terms of $\mathbf{X}_2^{S_1}$ using the previously determined $\mathbf{X}_1^{S_1}$. In a similar fashion, the subdesigns $\mathbf{X}_3^{S_1},\ldots,\mathbf{X}_N^{S_1}$ are obtained. In all these computations, the common initial prior distribution $\pi(\boldsymbol{\beta})$ is utilized, as that prior cannot be updated between respondents in the IASB-SIM approach.

For each individual respondent n, the data from the initial static stage are analyzed in a Bayesian fashion using the conditional logit model. A Bayesian approach is necessary because, at this stage, too few observations are available to use maximum likelihood-based estimation methods. The output of the Bayesian analysis of the data from the initial static stage will be a posterior distribution that is a mix of the initial prior distribution and the individual-level information that has become available. That posterior distribution is then used as an input to the adaptive sequential stage of the experiment. If we denote the responses of respondent n corresponding to the S_1 initial choice sets by $\mathbf{y}_n^{S_1}$ and the design for these choice sets by $\mathbf{X}_n^{S_1}$, then the posterior distribution $q(\boldsymbol{\beta}_n|\mathbf{y}_n^{S_1},\mathbf{X}_n^{S_1})$ can be computed as

$$q(\boldsymbol{\beta}_n|\mathbf{y}_n^{S_1},\mathbf{X}_n^{S_1})=\frac{L\left(\mathbf{y}_n^{S_1}\middle|\mathbf{X}_n^{S_1},\boldsymbol{\beta}_n\right)\pi(\boldsymbol{\beta}_n)}{\int L\left(\mathbf{y}_n^{S_1}|\mathbf{X}_n^{S_1},\boldsymbol{\beta}_n\right)\pi(\boldsymbol{\beta}_n)d\boldsymbol{\beta}_n},\qquad(8)$$

where $L\left(\mathbf{y}_n^{S_1}\middle|\mathbf{X}_n^{S_1},\boldsymbol{\beta}_n\right)$ is the likelihood for the conditional logit model defined in Eq. (2). This posterior distribution summarizes the available information for respondent n after the initial S_1 choice sets have been evaluated, and it forms the basis for the adaptive sequential design stage of the experiment.

3.2.1.2. IASB-OAA approach. In the initial static stage of the IASB-OAA approach, there is no common initial prior distribution. Instead, a different prior distribution is utilized for every respondent: the prior distribution for the nth respondent $\pi_n(\boldsymbol{\beta})$ is obtained by updating the initial prior distribution $\pi_1(\boldsymbol{\beta})$ (which was used for the first respondent) based on the choice set evaluations of the first $n-1$ respondents. The initial design for respondent n, $\mathbf{X}_n^{S_1}$, is then found by optimizing the Bayesian D_B-error in Eq. (7) evaluated over $\pi_n(\boldsymbol{\beta})$ and taking into account the choice sets generated for the $n-1$ previous respondents. Respondent n's data from the initial stage of the IASB-OAA approach are then processed in exactly the same fashion as in the IASB-SIM approach. Additionally, the adaptive sequential stage is identical in the two versions of the IASB approach.

The key difference between the IASB-OAA approach and the IASB-SIM approach is thus that the former starts from an updated prior distribution for each respondent. The updated prior distribution for a given respondent n summarizes all of the available data up to the $(n-1)$th respondent and the initial prior distribution, i.e., the prior distribution assumed before the first respondent enters the study.

3.2.2. Adaptive sequential stage

In the adaptive sequential stage, the prior information for the design construction is updated sequentially after each choice, and each choice set is constructed using the newly updated prior. The whole procedure for this adaptive sequential stage can be described as follows. The posterior distribution $q(\boldsymbol{\beta}_n|\mathbf{y}_n^{S_1},\mathbf{X}_n^{S_1})$ obtained after the initial static part of the experiment is used as the prior distribution $\pi(\boldsymbol{\beta}_n)$ in Eq. (7) for constructing the (S_1+1)th choice set, whose design matrix we denote by $\mathbf{x}_n^{S_1+1}$. The new choice set $\mathbf{x}_n^{S_1+1}$ is chosen such that the D_B-error of the combined design $\left(\mathbf{X}_n^{S_1},\mathbf{x}_n^{S_1+1}\right)$, which is evaluated over $q(\boldsymbol{\beta}_n|\mathbf{y}_n^{S_1},\mathbf{X}_n^{S_1})$, is minimized. This guarantees a maximum of extra information on $\boldsymbol{\beta}_n$ from the new choice set.

To compute the D_B-error for the combined design $\left(\mathbf{X}_n^{S_1},\mathbf{x}_n^{S_1+1}\right)$, a large number of draws is required from the posterior distribution obtained after the initial static stage, $q(\boldsymbol{\beta}_n|\mathbf{y}_n^{S_1},\mathbf{X}_n^{S_1})$. As there is no closed-form expression for the posterior distribution, we use importance sampling, which is a discrete approximation technique discussed by Bedrick, Christensen, and Johnson (1997), Monahan and Genz (1997), and Rossi, Allenby, and McCulloch (2005). Note that we use extensible shifted lattice points instead of pseudo-Monte Carlo sampling to take draws from the posterior distribution. The importance sampling approach is explained in detail in Appendix C.

The choice set $\mathbf{x}_n^{S_1+1}$ that minimizes the D_B-error of $\left(\mathbf{X}_n^{S_1}, \mathbf{x}_n^{S_1+1}\right)$ is then evaluated by respondent n, and the (S_1+1)th observation is obtained. We then update the prior information based on all S_1+1 observations from respondent n using

$$q(\beta_n|\mathbf{y}_n^{S_1+1}, \mathbf{X}_n^{S_1}, \mathbf{x}_n^{S_1+1}) = \frac{L\left(\mathbf{y}_n^{S_1+1}\middle|\mathbf{X}_n^{S_1}, \mathbf{x}_n^{S_1+1}, \beta_n\right)\pi\left(\beta_n\right)}{\int L\left(\mathbf{y}_n^{S_1+1}\middle|\mathbf{X}_n^{S_1}, \mathbf{x}_n^{S_1+1}, \beta_n\right)\pi\left(\beta_n\right)d\beta_n}, \tag{9}$$

where $L\left(\mathbf{y}_n^{S_1+1}\middle|\mathbf{X}_n^{S_1}, \mathbf{x}_n^{S_1+1}, \beta_n\right)$ is the likelihood for the conditional logit model defined in Eq. (2) for the first S_1+1 choice set evaluations made by respondent n. The resulting posterior distribution is then used as the updated prior distribution for determining the (S_1+2)th choice set for that respondent, again by minimizing the D_B-error. All further choice sets are generated similarly until a pre-specified number of choice sets S is reached. The algorithm we used to optimize the designs $\mathbf{x}_n^{S_1+1}, \ldots, \mathbf{x}_n^{S_1+S_2}$ of the choice sets $S_1+1, \ldots, S_1+S_2(=S)$ is the coordinate-exchange algorithm of Meyer and Nachtsheim (1995), which was introduced in the conjoint-choice design literature by Kessels, Jones, Goos, and Vandebroek (2009).

The IASB approach results in choice-based conjoint designs that are tailored to the preference structure β_n for each respondent. Note that the ACBC option in the Sawtooth software also provides an adaptive approach that generates respondent-specific choice sets (Sawtooth Software Inc., 2009). It uses the direct elicitation of preferred attribute levels and attribute level combinations and concentrates on a subspace of the profiles leaving out unacceptable levels. This approach is designed to keep the respondents' interest and engagement and does not focus on the statistical properties of the resulting design.

In the next section, we describe the setup of the simulation study we conducted to demonstrate that the IASB approach enables us to obtain more precise information on the individual-level and on the population-level parameters than existing non-adaptive procedures and thus to get a better picture of the consumer heterogeneity.

4. Comparative simulation study

In this section, we give a detailed report of the performance of the IASB approach in terms of the estimation accuracy and the predictive capability in four different scenarios. We investigate different levels of response accuracy because this strongly affects the results of conjoint-choice studies. In the presence of low response accuracy (or high response errors), the responses from an individual respondent provide unreliable information about the individual-level part-worth vector β_n, which jeopardizes the quality of the choice sets generated using that information. Therefore, the real test of our method is the scenario involving low response accuracy. It is in this scenario that the polyhedral method of Toubia et al. (2004, 2007) was not fully convincing compared to an aggregate-customization approach. In addition to varying the level of response accuracy, we also use different degrees of consumer heterogeneity in our study.

The results we report are for the IASB-SIM approach, which assumes that the respondents enter the conjoint study simultaneously. We do not show the results for the IASB-OAA approach, which assumes that the respondents enter the study one after another, as the IASB-OAA approach appears to have no added value over the IASB-SIM approach, even though it is computationally more burdensome. One reason for this is that the IASB-OAA and IASB-SIM approaches mainly differ in the construction of the initial static design, which has a limited impact on the quality of the information obtained from the experiment. Indeed, the strength of the IASB approaches is in the adaptive sequential stage, where the choice sets are tailored to the respondents and which involves substantially more choice sets than the initial static stage. Another reason for the lack of added value of the IASB-OAA approach over the IASB-SIM approach is that the assumption that the respondents are heterogeneous implies that population-level information, even when updated as in the IASB-OAA approach, is not enough to characterize an individual respondent and thus to create a good (initial) design for him/her.

4.1. Setup of the simulation study

To study whether the IASB approach for constructing choice-based designs is robust to response error, we adopted the simulation structure proposed by Arora and Huber (2001) for hierarchical Bayes estimation and utilized by Toubia et al. (2004). For that purpose, we varied both the response accuracy and the consumer heterogeneity by selecting a low and a high level. We considered designs with three attributes, each at three levels. Following Arora and Huber (2001) and Toubia et al. (2004), we used evenly spaced part-worths for the population mean μ_β. For each attribute, the part-worth values ranged from $-a$ to $+a$, with $a=0.5$ in the case of low response accuracy and $a=3$ in the case of high response accuracy. When using effects-type coding for the three three-level attributes, this means that $\mu_\beta=[-0.5,0,-0.5,0,-0.5,0]'$ in the case of low response accuracy, and $\mu_\beta=[-3,0,-3,0,-3,0]'$ in the case of high response accuracy.

The way in which we defined the degree of consumer heterogeneity is also the same as in Arora and Huber (2001) and Toubia et al. (2004). We assumed that the individual-level parameters vary according to a normal distribution with covariance matrix $\Sigma_\beta = \sigma_\beta^2\mathbf{I}_6$ around the population mean μ_β. We refer to σ_β^2 as the heterogeneity parameter and assume that it is constant across all attribute levels. In the case of high respondent heterogeneity, we use $3a$ as the value of σ_β^2, and in the case of low respondent heterogeneity, we use $0.5a$ as the value of σ_β^2. Note that we use the same values for a and σ_β^2 as Toubia et al. (2004), who empirically confirmed that these values are representative of those that might be obtained in practice.

Using our simulation study, we benchmarked our IASB approach against three alternatives. The first design is the Bayesian aggregate-customization design for the cross-sectional mixed logit model studied by Yu et al. (2009). This is a Bayesian benchmark design constructed assuming consumer heterogeneity but neglecting the panel structure of the data. The second benchmark design, which ignores consumer heterogeneity, is the aggregate-customization design for the conditional logit model studied by Arora and Huber (2001) as an option to improve individual-level estimates using hierarchical Bayes estimation. We refer to the former design as CUS-MIX and to the latter as CUS-COND. As explained above, a common design, customized for the average respondent, is used for each respondent in these approaches. Both Arora and Huber (2001) and Toubia et al. (2004) compared the performance of the CUS-COND to that of an orthogonal design in terms of the accuracy of the individual-level parameter estimates and predictive capabilities. Therefore, the third benchmark for our IASB approach in the simulation study is a nearly orthogonal design. The reason that we use a nearly orthogonal design instead of an orthogonal design is technical; in our simulation study, we assume that the total number of choice sets that a respondent has to evaluate is $S=16$, with three alternatives per choice set. This is one of many scenarios for which no orthogonal conjoint-choice design exists and only a nearly orthogonal design can be found. We constructed a nearly orthogonal design using the Sawtooth software. For the IASB approach in our study, the initial static stage consists of $S_1=5$ choice sets. The remaining $S_2=11$ choice sets are constructed sequentially using the Bayesian procedure outlined in Section 3.2.2.

We assume that the number of respondents participating in the experiment is 250. For each of the four design approaches in our study, the panel mixed logit model was estimated using a hierarchical Bayes approach starting from weak prior information. We refer to Section 5 and Appendix A for details regarding our implementation of the estimation approach.

4.2. Results for estimation accuracy

We considered two criteria to assess the precision of the parameter estimates for each design approach. Both criteria are based on the root mean squared error of the parameter estimates for an individual respondent, $RMSE_{\beta_n}$. The $RMSE_{\beta_n}$ value assesses how different the individual-level parameter estimates of respondent n are from the true parameters used to generate the data. First, we investigated for what percentage of respondents, the IASB method, the two aggregate-customization approaches (CUS-MIX and CUS-COND) and the nearly orthogonal design yield the best $RMSE_{\beta_n}$ value. The percentages we obtained are shown in Table 2. The results show that the IASB approach produces the best estimates for the individual respondents' parameters substantially more often than the nearly orthogonal design and the CUS-MIX and CUS-COND approaches whenever the heterogeneity and/or the accuracy is high. In Scenario 1, which involves low response accuracy and virtually no consumer heterogeneity, the IASB approach is only marginally better than the benchmark designs when it comes to recovering individual part-worths.

To quantify the extent to which the IASB approach performs better than the benchmark approaches in terms of the accuracy of the parameter estimates, we also computed the percentage decrease in the $RMSE_{\beta_n}$ value obtained with the IASB approach over each of the benchmark approaches for each respondent. We averaged these percentage values over all respondents in the study. These averaged improvements are displayed in Table 3. The row labeled 'Nearly Orthogonal' shows the averaged percentage decrease in $RMSE_{\beta_n}$ value obtained with the IASB approach instead of the nearly orthogonal design approach. The rows labeled 'CUS-MIX' and 'CUS-COND' show how much better the IASB approach is compared to the aggregate-customization approaches for the cross-sectional mixed logit model and the conditional logit model, respectively. The positive percentage improvements for the IASB approach indicate that the IASB approach outperforms the three benchmark approaches in each of the four scenarios in our study. In addition, Table 3 also shows that the CUS-MIX approach, the aggregate-customization approach that accounts for heterogeneity, outperforms the CUS-COND and the nearly orthogonal design approaches.

In Fig. 1, we show how the $RMSE_{\beta_n}$ values of the 250 respondents are distributed in the four different scenarios. The solid, dotted, dashed and dash-dotted curves correspond to the distribution of the $RMSE_{\beta_n}$ values obtained from the IASB, CUS-MIX, CUS-COND and nearly orthogonal design approaches. The advantage of the proposed IASB approach is clearly shown in the figure, as this approach has the leftmost distribution in each of the scenarios. Note that the scale of the horizontal axis is different for each of the scenarios in the figure.

The success of the IASB approach is due to the fact that, in most cases, the convergence of the Bayesian estimates of the individual-level part-worths $\boldsymbol{\beta}_n$ to their true values is faster than with the benchmark approaches. To show this, we computed a Bayesian estimate after the evaluation of each of the 16 choice sets for each respondent n. We show the convergence of the estimates to the true values by plotting the $RMSE_{\beta_n}$ values for these 16 estimates for a number of randomly selected respondents in Fig. 2a for the case of high response accuracy and in Fig. 2b for the case of low response accuracy. The solid curve in each plot shows the $RMSE_{\beta_n}$ values for the IASB approach, whereas the dotted curve in each plot presents the $RMSE_{\beta_n}$ values for the CUS-MIX approach, which is the closest competitor to the IASB approach.

Table 2
Percentages of respondents for which the IASB approach, the CUS-MIX and CUS-COND approaches and the nearly orthogonal design approach yield the best individual part-worth estimates.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Response accuracy	Low ($a=0.5$)	Low ($a=0.5$)	High ($a=3$)	High ($a=3$)
Heterogeneity	Low ($\sigma^2_\beta=0.25$)	High ($\sigma^2_\beta=1.5$)	Low ($\sigma^2_\beta=1.5$)	High ($\sigma^2_\beta=9$)
Nearly orthogonal	27.73%	19.92%	10.55%	7.42%
CUS-COND	23.44%	17.19%	11.72%	5.47%
CUS-MIX	20.31%	27.34%	20.70%	16.41%
IASB	28.52%	35.55%	57.03%	70.70%

Table 3
Improvement in terms of estimation accuracy (as measured by the average percentage decrease in $RMSE_{\beta_n}$ value) achieved with the IASB approach compared to the nearly orthogonal design and the CUS-COND and CUS-MIX aggregate-customization approaches.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Response accuracy	Low ($a=0.5$)	Low ($a=0.5$)	High ($a=3$)	High ($a=3$)
Heterogeneity	Low ($\sigma^2_\beta=0.25$)	High ($\sigma^2_\beta=1.5$)	Low ($\sigma^2_\beta=1.5$)	High ($\sigma^2_\beta=9$)
Nearly orthogonal	5.7%	16.1%	26.6%	32.6%
CUS-COND	7.2%	18.4%	22.7%	32.1%
CUS-MIX	5.3%	13.6%	17.1%	20.9%

It is clear from Fig. 2 that, in most cases, the $RMSE_{\beta_n}$ values drop faster for the IASB approach than for the CUS-MIX approach and that the convergence of the parameter estimates to the true values is thus faster for the IASB approach. The convergence of the Bayesian estimates to the true values for $\boldsymbol{\beta}_n$ implies that the choice sets in the sequential adaptive stage of the IASB approach are based on continuously improving prior information about $\boldsymbol{\beta}_n$.

The previous results show that the individual-level information can be very precisely retrieved by the IASB approach. In Table 4, we compare the precision of the population-level estimates obtained by the four different design approaches. From this table, it can be concluded that the IASB approach also outperforms the nearly orthogonal and CUS-COND designs in all scenarios for estimating the mean preference $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ and corresponding heterogeneity $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$. It also outperforms the CUS-MIX approach in all but one case (see number in italics), which shows that the IASB approach in general yields better population-level estimates than designs that were constructed specifically for estimating these population-level parameters but that neglected the panel structure of the data.

4.3. Results for prediction accuracy

We also investigated the performance of the IASB approach, the CUS-COND and CUS-MIX approaches and the nearly orthogonal design in terms of prediction accuracy. First, we computed how often the choices of the individual respondents are predicted correctly. To compute this hit rate, we simulated choices for each respondent for all 2925 possible choice sets that could be formed by three three-level attributes, as was also done by Kessels, Goos, and Vandebroek (2006). The average hit rates across all respondents for the IASB approach and the three benchmark approaches are shown in Table 5. The results show that the IASB approach performs better than the three other approaches in the four scenarios we studied, showing once more that the individual preferences are well captured by the individually optimized designs. The IASB approach performs well not only in the case of high response accuracy but also when the response accuracy is low. The results also indicate that the IASB approach is especially valuable in the presence of high consumer heterogeneity, which is in line with the results from Section 4.2.

We also computed the average percentage decrease in the root mean squared prediction error, $RMSE_p$, achieved with the IASB

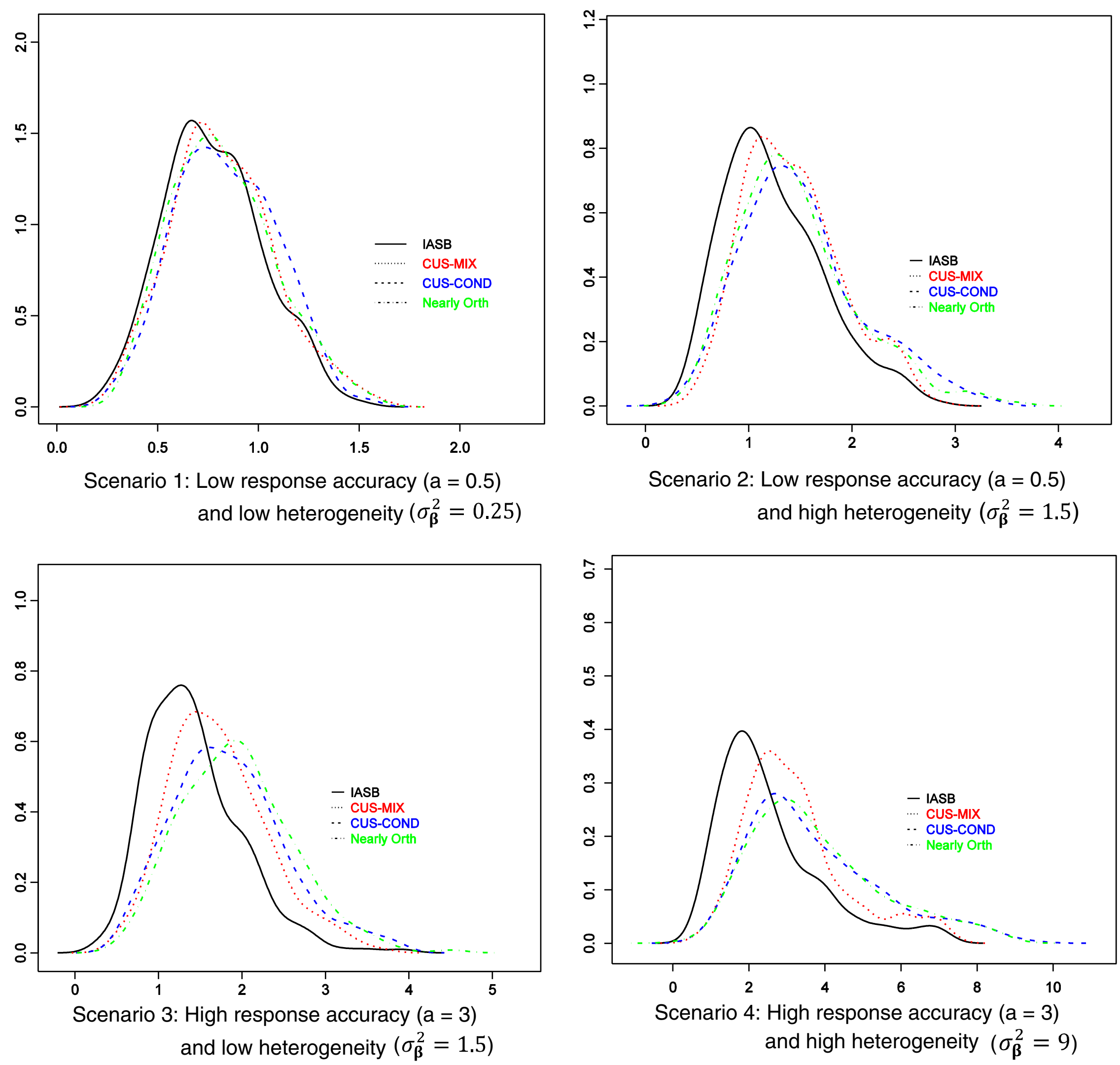

Fig. 1. Distribution of $RMSE_{\beta_n}$ values for the four different design approaches.

approach instead of the CUS-MIX, CUS-COND, and nearly orthogonal design approaches. This gave results that are very similar to those for the $RMSE_{\beta_n}$ values in Table 3. Therefore, we do not show them.

5. Practical application

We applied the IASB approach in a choice-based conjoint experiment to study the preferences of master's students at the Faculty of Business and Economics of the KULeuven in Belgium for the form in which their course notes are printed and distributed and to investigate whether the students are willing to pay extra for more environmentally friendly and/or more convenient ways to print and distribute course notes. Because Belgian students usually do not pay for their university education themselves and are virtually always heavily sponsored by their parents, our expectation was that we would observe a positive attitude toward environmentally friendly and convenient options. This IASB design application, involving four two-level attributes and one three-level attribute, demonstrates that it is perfectly feasible to use the computationally intensive IASB approach in practice. Each of the respondents in the study evaluated 16 choice sets with two alternatives. The attributes and their levels are displayed in Table 6. Several groups of students participated in the study, and the students in each of these groups completed the choice-based questionnaire simultaneously. For this reason, we utilized the IASB-SIM version of the IASB approach. In total, 120 students participated in the IASB experiment.

The mean and covariance of the prior distribution were obtained from a pilot study involving 50 respondents, other than those used in the actual experiment. To speed up the generation of the choice sets in the IASB approach, we applied a systematic sampling approach, based on a sample of 2048 extensible shifted lattice points (ESLP) rather than a large sample of pseudo-Monte Carlo draws. This is explained in detail in Appendix C. Without such a systematic sampling approach, the generation of the choice sets would have been prohibitively slow. However, in this case, the average computation time for generating an optimal choice set (on Dell PCs with a 3.16 GHz Intel processor and 4 GB RAM using the SAS procedure IML) was approximately 15 s.

After collecting all of the data, we estimated the panel mixed logit model using the hierarchical Bayesian approach, which combines Gibbs sampling with the Metropolis-Hastings (MH) algorithm, as explained in Appendix A. In our implementation of the Bayesian estimation approach, the starting values for the population mean μ_{β} and for all of the individual-level parameters were equal to zero. For the population covariance matrix, the starting value was $\mathbf{\Sigma}_{\beta} = p\mathbf{I}_p$, where $p = 6$ is the number of parameters in the study. We performed 100,000 iterations in total, 70,000 of which were used for burn-in. The remaining 30,000 draws were thinned by selecting every 10th

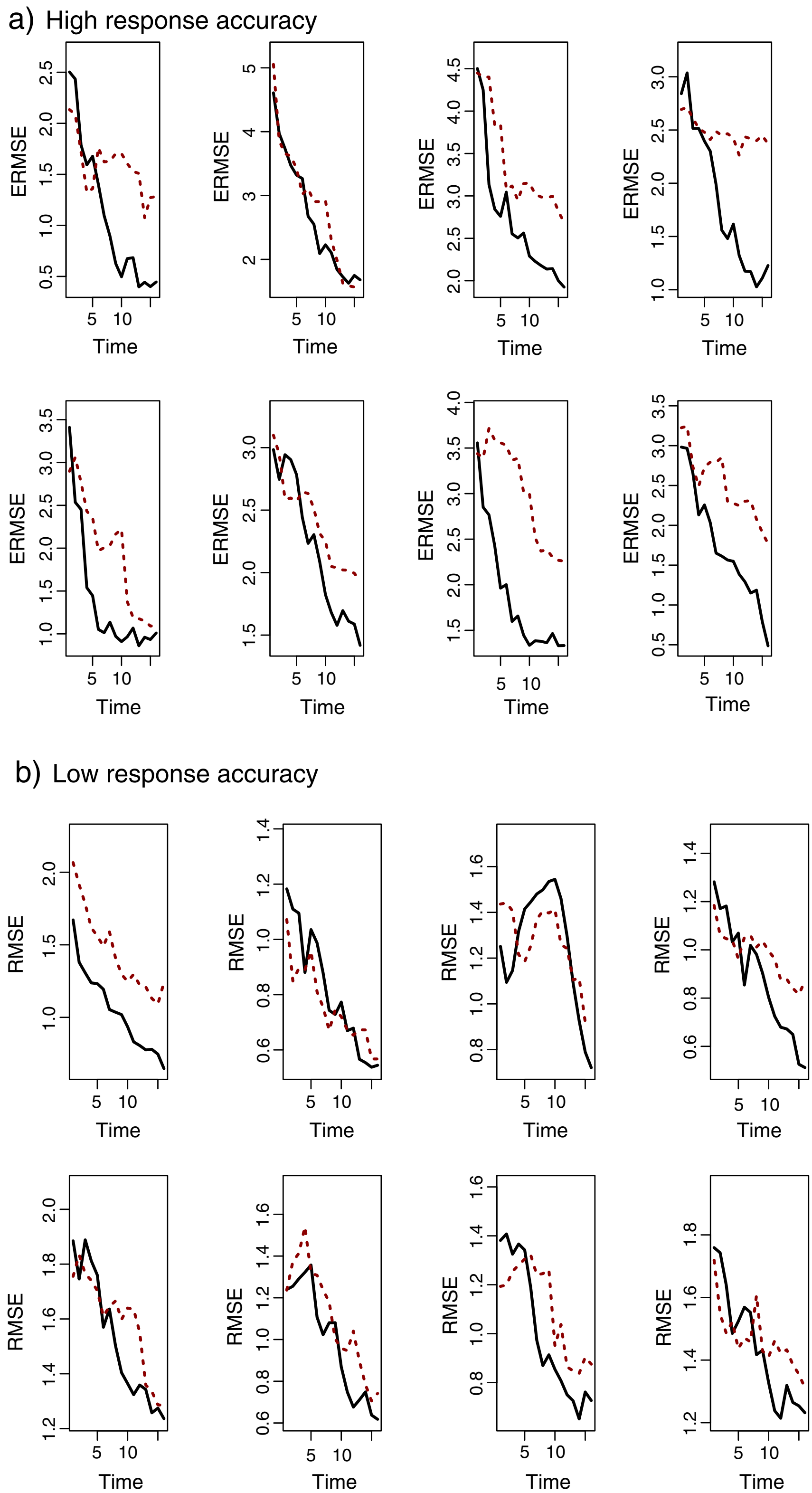

Fig. 2. $RMSE_{\beta_n}$ values for the IASB (solid curve) and CUS-MIX (dotted curve) approaches obtained for eight randomly selected respondents in the case of low respondent heterogeneity.

draw to remove the autocorrelations, yielding 3000 draws, which we used to determine the parameter estimates. The draws from the posterior distribution for each respondent's parameter vector β_n were obtained using the random-walk MH algorithm. One issue in this algorithm is to determine the size of a so-called jump, which is the distance between one draw for β_n and the next. We adjusted the value of the jump dynamically because this leads to a faster convergence than a fixed jump value. To do so, we utilized an acceptance rate of 0.30, which is in line with the optimal acceptance rates reported by Gelman, Carlin, Stern, and Rubin (1995).

Table 4
(*RMSE* values for μ_β | *RMSE* values for Σ_β) obtained with the nearly orthogonal design, the CUS-COND and CUS-MIX aggregate-customization approaches and the IASB approach.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Response accuracy	Low ($a=0.5$)	Low ($a=0.5$)	High ($a=3$)	High ($a=3$)
Heterogeneity	Low ($\sigma^2_\beta=0.25$)	High ($\sigma^2_\beta=1.5$)	Low ($\sigma^2_\beta=1.5$)	High ($\sigma^2_\beta=9$)
Nearly orthogonal	(0.15 | 0.37)	(0.24 | 0.99)	(0.39 | 1.33)	(0.86 | 7.34)
CUS-COND	(0.15 | 0.35)	(0.24 | 1.12)	(0.37 | 1.14)	(0.88 | 8.33)
CUS-MIX	(0.16 | 0.37)	(0.23 | 0.85)	(*0.29* | 1.02)	(0.72 | 5.56)
IASB	(0.13 | 0.31)	(0.20 | 0.78)	(0.35 | 0.95)	(0.71 | 5.32)

We used the means of the posterior distributions for the population-level parameters $\boldsymbol{\mu}_\beta$ and $\boldsymbol{\Sigma}_\beta$ as point estimates. These posterior means are displayed in Table 7. To interpret them, it is important to know that we used effects-type coding for the attribute levels. On average, the students prefer recycled paper and one-sided printing. The average student is very price sensitive (despite the sponsorship of the parents) and almost indifferent between course notes published in one or two volumes, and between the two different publishers. The relatively large diagonal elements of the covariance matrix $\boldsymbol{\Sigma}_\beta$, however, indicate that there is substantial heterogeneity among the respondents.

To visualize the heterogeneity in student preferences, we display the distribution of the individual-level estimates for the six part-worths in Fig. 3. It is clear from the figure that the students do not differ much in their preferences for the publisher, but there is substantial heterogeneity in the price sensitivity and in the strength of the preferences in favor of one-sided printing. Additionally, a substantial share of the student population prefers new paper over recycled paper (as opposed to the average student, who has a pretty strong preference in favor of recycled paper). The number of students in favor of course notes published in one volume is about equally large as the number of those in favor of course notes published in two volumes.

6. Conclusion and discussion

Motivated by a suggestion in Sándor and Wedel's 2005 paper on heterogeneous conjoint-choice designs, we propose an individually adapted sequential Bayesian design (IASB) approach for constructing choice-based conjoint experiments for the panel mixed logit model. Our approach is a Bayesian one that optimizes the next choice set to be shown to a respondent based on his/her previous choices and on population-level prior information. Because of recent computational advances, this individualized approach is practically feasible in online surveys.

We compare the proposed methodology with three non-adaptive benchmark design approaches (the CUS-MIX approach, the CUS-COND approach, and a nearly orthogonal design). The comparison involves a test under various degrees of response accuracy and respondent heterogeneity. The simulation study shows that, for retrieving individual-level information, the proposed IASB approach by far outperforms the non-adaptive benchmark designs unless the consumer heterogeneity and the response accuracy are both low. In such a scenario, the IASB approach is only marginally better than the benchmark designs. We pay special attention to the scenarios where the response accuracy is low as the quality of the IASB designs is influenced by errors in the responses. In general, every sequential method in which choice sets are constructed based on a respondent's choices will perform worse in the presence of low response accuracy than in the presence of high response accuracy, as the generation of the choice sets is based on unreliable information in the former case. The polyhedral method in Toubia et al. (2004), however, seems much more sensitive than the IASB approach to low response accuracy. Toubia et al. (2007) proposed an adapted version of the polyhedral method that is reported to be better than the original one when the response accuracy is low. Unlike the original method, the adapted polyhedral method takes response error explicitly into account when setting up individually adapted conjoint-choice designs. The results for the adapted polyhedral method in Toubia et al. (2007) appear to be better than the original method in Toubia et al. (2004), but they do not report any results for response accuracy levels as low as those found in Toubia et al. (2004), Arora and Huber (2001) and this article.

Several variations of the IASB methodology can be developed. One might, e.g., use a stopping rule, which ensures that the choice experiment for a given respondent is stopped as soon as a satisfactory amount of information on his/her individual-level part-worths has been obtained. We investigated a prediction-based stopping rule that works as follows. In the IASB approach, a Bayesian analysis is performed after the initial design and after each of the next choice sets has been evaluated. Based on the Bayesian parameter estimates obtained in each step, we predict the respondent's choice for the next choice set. As soon as the choices for three consecutive choice sets are predicted correctly, the experiment is stopped for that respondent. The main advantage of this stopping rule is that it leads to an increase in estimation accuracy for respondents with inconsistent answers, as these typically have to evaluate a larger number of choice sets. Respondents who make consistent choices do not need to evaluate many choice sets to produce accurate parameter estimates.

Table 5
Average individual hit rates with the nearly orthogonal design, the CUS-MIX and CUS-COND aggregate-customization approaches, and the IASB approach.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Response accuracy	Low ($a=0.5$)	Low ($a=0.5$)	High ($a=3$)	High ($a=3$)
Heterogeneity	Low ($\sigma^2_\beta=0.25$)	High ($\sigma^2_\beta=1.5$)	Low ($\sigma^2_\beta=1.5$)	High ($\sigma^2_\beta=9$)
Nearly orthogonal	77.16%	81.96%	89.13%	86.09%
CUS-COND	75.99%	80.44%	89.94%	85.95%
CUS-MIX	76.55%	81.64%	90.21%	87.76%
IASB	78.79%	85.18%	93.76%	93.47%

Table 6
Attributes and attribute levels in the real-life IASB experiment.

Attribute	Levels		
Type of paper	New		Recycled
Printing	One-sided		Two-sided
Number of volumes	One		Two
Distributor	Publisher 1		Publisher 2
Price	10 Euro	12.5 Euro	15 Euro

Table 7
Hierarchical Bayesian estimates of the population-level preference structure as given by the population mean μ_β and covariance Σ_β in the real-life IASB experiment.

Parameter		New paper	One-sided printing	One volume	Publ 1	10 Euro	12.5 Euro
$\boldsymbol{\mu}_\beta$		−0.489	1.479	0.218	0.143	3.018	0.090
$\boldsymbol{\Sigma}_\beta$	New paper	1.488	0.960	−0.044	0.077	−0.531	0.036
	One-sided printing	0.960	5.965	0.550	−0.614	−1.704	0.051
	One volume	−0.044	0.550	1.576	−0.028	−0.236	−0.085
	Publisher 1	0.077	−0.614	−0.028	0.626	0.102	0.045
	10 Euro	−0.531	−1.704	−0.236	0.102	7.157	0.176
	12.5 Euro	0.036	0.051	−0.085	0.045	0.176	0.512

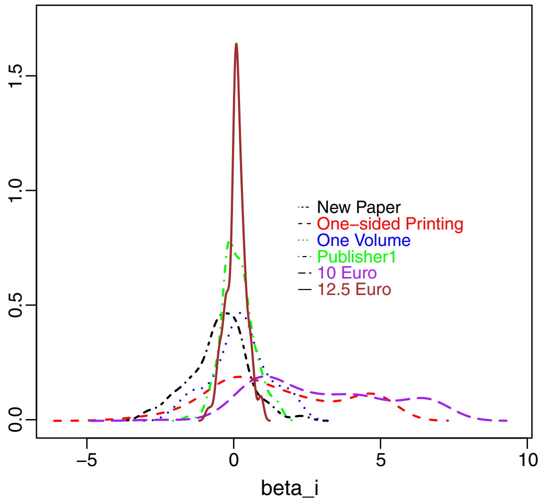

Fig. 3. Distributions of the individual-level part-worth estimates in the practical application.

Finally, using a real-life application, we also demonstrated that the construction of new choice sets in the IASB design approach is practically feasible within a reasonable amount of time. Therefore, there is little doubt that the IASB approach is a valuable tool for future conjoint-choice experimentation in marketing, even for problems that involve many more attributes than the real-life application we discussed here.

Acknowledgments

We are grateful to the referees for their constructive comments, which led to a substantial improvement of this paper. In addition, we acknowledge the financial support of the Research Council of the KULeuven (OT-project 04/07).

Appendix A. Hierarchical Bayes estimation of the mixed logit model

We use the hierarchical Bayes approach to estimate the mixed logit model. Here, we provide a summary of the different conditional posteriors that are involved. A detailed explanation of this estimation procedure can be found in Train (2003).

Assume the following uninformative prior distributions for $\boldsymbol{\mu}_\beta$ and $\boldsymbol{\Sigma}_\beta$: $k(\boldsymbol{\mu}_\beta)$ is a multivariate normal distribution with zero mean and large variances and $h(\boldsymbol{\Sigma}_\beta)$ is an Inverse Wishart distribution with parameters p and $\mathbf{I}_p$, $IW(p, \mathbf{I}_p)$, where p is the dimension of $\boldsymbol{\beta}_n$ and $\mathbf{I}_p$ is the p-dimensional identity matrix. Given that we have assumed that $\boldsymbol{\beta}_n \sim N(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$, the joint posterior distribution of $\boldsymbol{\mu}_\beta$, $\boldsymbol{\Sigma}_\beta$ and $\boldsymbol{\beta}_n$, for all N respondents, can be written as

$$K\left(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta, \boldsymbol{\beta}_{1, \dots,}\boldsymbol{\beta}_N \middle| \mathbf{y}_{full}\right) \propto \prod_{n=1}^{N} L\left(\mathbf{y}_n^S \middle| \mathbf{X}_n^S, \boldsymbol{\beta}_n\right) f\left(\boldsymbol{\beta}_n | \boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta\right) k\left(\boldsymbol{\mu}_\beta\right) h\left(\boldsymbol{\Sigma}_\beta\right).$$

Draws from this posterior can be obtained by iteratively taking draws from the conditional distributions of some of the parameters, given all other parameters:

- The conditional distribution of $\boldsymbol{\mu}_\beta$, given $\boldsymbol{\beta}_n$ and $\boldsymbol{\Sigma}_\beta$, is $N(\overline{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_\beta / N)$ with $\overline{\boldsymbol{\beta}}$ representing the sample mean of all $\boldsymbol{\beta}_n$.
- The conditional distribution of $\boldsymbol{\Sigma}_\beta$, given $\boldsymbol{\mu}_\beta$ and $\boldsymbol{\beta}_n$, is $IW\left(p + N, \frac{p\mathbf{I}_p + N\mathbf{S}}{p+N}\right)$ with $\mathbf{S}$ representing the sample covariance matrix of the $\boldsymbol{\beta}_n$.
- The distribution of $\boldsymbol{\beta}_n$ conditional on $\mathbf{y}_n^S$, $\boldsymbol{\mu}_\beta$ and $\boldsymbol{\Sigma}_\beta$ has no closed form, and we use a *Metropolis-Hastings random-walk* procedure to take draws from a normal proposal distribution.

With the initial values $\boldsymbol{\mu}_\beta^0$, $\boldsymbol{\Sigma}_\beta^0$ and $\boldsymbol{\beta}_1^{*0}, \dots, \boldsymbol{\beta}_N^{*0}$, the $(j+1)^{th}$ iteration of the *Gibbs sampler* works as follows:

- Draw $\boldsymbol{\mu}_\beta^{j+1}$ from $N\left(\overline{\boldsymbol{\beta}}^j, \boldsymbol{\Sigma}_\beta^j / N\right)$ where $\overline{\boldsymbol{\beta}}^j$ is the sample mean of the $\boldsymbol{\beta}_n^j$.
- Draw $\boldsymbol{\Sigma}_\beta^{j+1}$ from $IW\left(p + N, \frac{p\mathbf{I}_p + N\mathbf{S}^j}{p+N}\right)$ where $\mathbf{S}^j$ is the sample covariance matrix of the $\boldsymbol{\beta}_n^j$ around $\boldsymbol{\mu}_\beta^{j+1}$.
- Let $\boldsymbol{\eta}^{j+1}$ be a vector with p iid draws from $N(0,1)$. Compute $\tilde{\boldsymbol{\beta}}_n^{j+1} = \boldsymbol{\beta}_n^j + \rho \mathbf{C} \boldsymbol{\eta}^{j+1}$ where ρ is a scalar that is chosen such that the acceptance rate is close to 30% and $\mathbf{C}$ is the Choleski factor of $\boldsymbol{\Sigma}_\beta^{j+1}$. Calculate the ratio

$$F = \frac{L\left(\mathbf{y}_n^S | \tilde{\boldsymbol{\beta}}_n^{j+1}\right) f\left(\tilde{\boldsymbol{\beta}}_n^{j+1} | \boldsymbol{\mu}_\beta^{j+1}, \boldsymbol{\Sigma}_\beta^{j+1}\right)}{L\left(\mathbf{y}_n^S | \boldsymbol{\beta}_n^j\right) f\left(\boldsymbol{\beta}_n^j | \boldsymbol{\mu}_\beta^{j+1}, \boldsymbol{\Sigma}_\beta^{j+1}\right)}$$

and accept $\tilde{\boldsymbol{\beta}}_n^{j+1}$ as the new $\boldsymbol{\beta}_n^{j+1}$ with probability $min(F,1)$

Appendix B. Derivation of the generalized Fisher information matrix

As, up to a proportionality constant, the posterior distribution of $\boldsymbol{\beta}_n$ is the product of the likelihood $L(\mathbf{Y}|\boldsymbol{\beta}_n, \mathbf{X})$ and the prior distribution $\pi(\boldsymbol{\beta}_n)$, we have

$$\mathbf{I}_{GFIM}\left(\boldsymbol{\beta}_n | \mathbf{X}\right) = -E_{\mathbf{Y}}\left[\frac{\partial^2 \log q(\boldsymbol{\beta}_n | \mathbf{Y}, \mathbf{X})}{\partial \boldsymbol{\beta}_n \partial \boldsymbol{\beta}_n'}\right] = E_{\mathbf{Y}}\left[-\frac{\partial^2 \log \mathrm{L}(\mathbf{Y}|\boldsymbol{\beta}_n, \mathbf{X})}{\partial \boldsymbol{\beta}_n \partial \boldsymbol{\beta}_n'} - \frac{\partial^2 \pi(\boldsymbol{\beta}_n)}{\partial \boldsymbol{\beta}_n \partial \boldsymbol{\beta}_n'}\right]$$

The first term in this expression is the Fisher information matrix, which is equal to $\sum_{s=1}^{S} \mathbf{X}_s'(\mathbf{P}_s^n - \mathbf{p}_s^n \mathbf{p}_s^{n'})\mathbf{X}_s$ for a conditional logit model (see, e.g., Sándor & Wedel, 2001). For a normal prior distribution with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}_n}$,

$$\mathbf{I}_{GFIM}\left(\boldsymbol{\beta}_n | \mathbf{X}\right) = \sum_{s=1}^{S} \mathbf{X}_s'\left(\mathbf{P}_s^n - \mathbf{p}_s^n \mathbf{p}_s^{n'}\right)\mathbf{X}_s + \boldsymbol{\Sigma}_{\boldsymbol{\beta}_n}^{-1}.$$

Note that the generalized Fisher information matrix equals the usual Fisher information matrix in case $\boldsymbol{\Sigma}_{\boldsymbol{\beta}_n} \to \infty$; i.e., when an uninformative prior is used.

Appendix C. Importance sampling

To compute the D_B-errors required for the adaptive sequential stage of the choice experiments, a large number of draws is required from the posterior distribution $q(\boldsymbol{\beta}_n | \mathbf{y}_n^{\tilde{S}}, \mathbf{X}_n^{\tilde{S}})$, with $\tilde{S}$ representing the number of choice sets. As there is no closed-form expression for the posterior distribution, we use importance sampling. This method, which was discussed by Bedrick et al. (1997), Monahan and Genz (1997) and Rossi et al. (2005), works as follows:

1. Compute the mode $\boldsymbol{\beta}_n^*$ and the Hessian matrix $\mathbf{H}$ of the posterior distribution $q(\boldsymbol{\beta}_n | \mathbf{y}_n^{\tilde{S}}, \mathbf{X}_n^{\tilde{S}})$.
2. Take R draws, denoted by $\boldsymbol{\beta}_n^r$ $(r = 1, \dots, R)$, from the importance density, which is the multivariate student t distribution with mean $\boldsymbol{\beta}_n^*$, covariance matrix $-\mathbf{H}^{-1}$ and v degrees of freedom. In our computations, we used systematic draws, more specifically $R = 2048$ extensible shifted lattice points (transformed using the baker's transformation), rather than Monte Carlo draws, as this gives better approximations (see, e.g., Sándor & András, 2004 or Yu, Goos, & Vandebroek, 2010). For the degrees of freedom v, we used 8, but it should be pointed out that any value between 5 and 15 seems to give satisfactory results.

3. Compute the weight w^r for each draw β_n^r. Suppose that $q^*(\beta_n|\mathbf{y}_n^{\tilde{S}}, \mathbf{X}_n^{\tilde{S}})$ is the density kernel of $q(\beta_n|\mathbf{y}_n^{\tilde{S}}, \mathbf{X}_n^{\tilde{S}})$ and $\psi(\beta_n)$ is the kernel of the importance density; then, the weight w_r for the rth draw β_n^r is

$$w_r = \frac{q^*\left(\beta_n^r \middle| \mathbf{y}_n^{\tilde{S}}, \mathbf{X}_n^{\tilde{S}}\right)/\psi(\beta_n^r)}{\sum_{i=1}^R q^*\left(\beta_n^i \middle| \mathbf{y}_n^{\tilde{S}}, \mathbf{X}_n^{\tilde{S}}\right)/\psi(\beta_n^i)}.$$

4. The D_B-error for the combined first- and second-stage design given by $\mathbf{X}_n^{\tilde{S}}$ and $\mathbf{x}_n^{\tilde{S}+1}$ is approximated by

$$D_B\text{-error}\left(\mathbf{X}_n^{\tilde{S}}, \mathbf{x}_n^{\tilde{S}+1}\right) = \sum_{r=1}^R w_r\, D\text{-error}\left(\beta_n^r, \mathbf{X}_n^{\tilde{S}}, \ \mathbf{x}_n^{\tilde{S}+1}\right).$$

References

Allenby, G. M., & Rossi, P. E. (1999). Marketing models of consumer heterogeneity. *Journal of Econometrics*, *89*, 57–78.

Arora, N., Allenby, G. M., & Ginter, J. L. (1998). A hierarchical Bayes model of primary and secondary demand. *Marketing Science*, *17*, 29–44.

Arora, N., & Huber, J. (2001). Improving parameter estimates and model prediction by aggregate customization in choice experiments. *Journal of Consumer Research*, *28*, 273–283.

Bedrick, E. J., Christensen, R., & Johnson, W. (1997). Binomial regression: Predicting survival at a trauma center. *The American Statistician*, *51*, 211–218.

Bliemer, M. C. J., & Rose, J. M. (2010). Construction of experimental designs for mixed logit models allowing for correlation across choice observations. *Transportation Research Part B*, *44*, 720–734.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (1995). *Bayesian data analysis.* Suffolk: Chapman and Hall.

Hensher, D. A., & Greene, W. H. (2003). The mixed logit model: The state of practice. *Transportation*, *30*, 133–176.

Kessels, R., Goos, P., & Vandebroek, M. (2006). A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, *43*, 409–419.

Kessels, R., Jones, B., Goos, P., & Vandebroek, M. (2009). An efficient algorithm for constructing Bayesian optimal choice designs. *Journal of Business and Economic Statistics*, *27*, 279–291.

Lenk, P. J., DeSarbo, W. S., Green, P. E., & Young, M. R. (1996). Hierarchical Bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. *Marketing Science*, *15*, 173–191.

Meyer, R. K., & Nachtsheim, C. J. (1995). The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics*, *37*, 60–69.

Monahan, J., & Genz, A. (1997). Spherical-radial integration rules for Bayesian computation. *Journal of the American Statistical Association*, *92*, 664–674.

Pilz, J. (1991). *Bayesian estimation and experimental design for linear models.* Wiley.

Rossi, P. E., Allenby, G. M., & McCulloch, R. (2005). *Bayesian statistics and marketing.* Wiley series in probability and statistics: Wiley.

Sándor, Z., & András, P. (2004). Alternative sampling methods for estimating multivariate normal probabilities. *Journal of Econometrics*, *2*, 207–234.

Sándor, Z., & Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, *38*, 430–444.

Sándor, Z., & Wedel, M. (2002). Profile construction in experimental choice designs for mixed logit models. *Marketing Science*, *21*, 455–475.

Sándor, Z., & Wedel, M. (2005). Heterogeneous conjoint choice designs. *Journal of Marketing Research*, *42*, 210–218.

Sawtooth Software Inc. (2009). ACBC Technical paper. http://www.sawtoothsoftware.com/download/techpap/acbctech.pdf

Toubia, O., Hauser, J. R., & Simester, D. I. (2004). Polyhedral methods for adaptive choice-based conjoint analysis. *Journal of Marketing Research*, *41*, 116–131.

Toubia, O., Hauser, J., & Garcia, R. (2007). Probabilistic polyhedral methods for adaptive choice-based conjoint analysis: Theory and application. *Marketing Science*, *26*, 596–610.

Train, K. (2003). *Discrete choice methods with simulation.* : Cambridge University Press.

Yu, J., Goos, P., & Vandebroek, M. (2008). A comparison of different Bayesian design criteria for constructing efficient conjoint choice experiments. *2008, KULeuven, Faculty of Business and Economics, Research Report KBI 0817*.

Yu, J., Goos, P., & Vandebroek, M. (2009). Efficient conjoint choice designs in the presence of respondent heterogeneity. *Marketing Science*, *28*, 122–135.

Yu, J., Goos, P., & Vandebroek, M. (2010). Comparing different sampling schemes for approximating the integrals involved in the efficient design of stated choice experiments. *Transportation Research Part B*, *44*, 1268–1289.