KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Business and Economics

Sequential choice designs to estimate the distribution of willingness-to-pay

Vishva Danthurebandara, Jie Yu and Martina Vandebroek

DEPARTMENT OF DECISION SCIENCES AND INFORMATION MANAGEMENT (KBI)

KBI 1016

Sequential choice designs to estimate the distribution of willingness-to-pay

Vishva M. Danthurebandara
Faculty of Business and Economics, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium.
Email: Vishva.Danthurebandara@econ.kuleuven.be
Tel: +32 (0)16 32 69 63
Fax: +32 (0)16 32 66 24

Jie Yu
Faculty of Business and Economics, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium.
Email: Jie.Yu@econ.kuleuven.be
Tel: +32 (0)16 32 69 62
Fax: +32 (0)16 32 66 24

Martina Vandebroek
Faculty of Business and Economics & Leuven Statistics Research Centre, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium.
Email: Martina.Vandebroek@econ.kuleuven.be
Tel: +32 (0)16 32 69 75
Fax: +32 (0)16 32 66 24

Sequential choice designs to estimate the distribution of willingness-to-pay

Abstract

The concept of willingness-to-pay (WTP) has attracted the attention of marketeers because of its usefulness in many applications. Nowadays one aims at describing the market heterogeneity by estimating the distribution of WTP. However, this poses several problems that have been discussed repeatedly in the literature. Many authors report unrealistic, extreme or inaccurate individual-level WTP estimates.

We propose to use an adaptive sequential approach to construct conjoint choice designs for estimating the distribution of WTP. It uses Bayesian methods to generate individually optimized choice sets. These choice sets are computed sequentially based on the prior information of each individual which is updated after each choice. The choices made by all respondents are then used to estimate the mixed logit model which yields individual-level utility coefficients and corresponding individual-level WTP estimates from which the distribution of WTP can be derived.

This sequential approach is compared in a simulation study with two non-sequential designs: a semi-Bayesian D-optimal design for the conditional logit model and a nearly orthogonal design. The results shows that the sequential design performs much better than the benchmark designs. It yields more accurate individual-level WTP estimates and produces a more accurate picture of the heterogeneity.

1 Introduction

In market valuations, researchers often calculate marginal rates of substitution (MRS). By definition, MRS is the rate at which a consumer is ready to give up one product characteristic in exchange for another product characteristic while maintaining the same level of utility. MRS with respect to the price is called willingness-to-pay (WTP). The concept of WTP has become an essential component in cost-benefit analysis and has often been using for policy making, product planning and product management affairs. The following examples have been taken from the literature and illustrate the broad application of willingness-to-pay.

Sonnier et al. (2007) present two discrete choice experiments to illustrate the effect of the model parametrization on the accuracy of WTP estimates. The first study has been conducted on midsize sedans which are described by the attributes model, engine power, audio, safety level and price. One of the important findings is that most people have higher WTP-values for the Japanese models involved (*Toyota Camry* and for *Nissan Maxima*) than for the European *Volkswagen Passat*. On the other hand, most consumers are willing to pay more for *Volkswagen Passat* than for *Ford Taurus*. These results are used to determine the optimal price for a Ford Taurus such that the firm's expected profit is maximized and to predict the corresponding market shares. Sonnier et al. (2007) also analyze the data of a choice study that was conducted by the Eastman Kodak Company to assess the market for cameras. The results are used to price and predict market shares for a few hypothetical cameras. The heterogeneity in WTP has also been investigated by Scarpa et al. (2008) in a study about hikers' destination choices in the Alps conducted by the Italian Alpine Club. There were 18 different destination sites and five site attributes which describe the land-use of the site and the hiking conditions. They estimated a mixed logit model and obtained interesting results about the individuals' willingness-to-pay. For example, the results indicated that most of the hikers were willing to pay a significant amount to avoid an extra difficulty level, to obtain an extra safety and resting point and for a 1% increase in easily walkable trail length. The results of this WTP study were helpful in planning future tours and site maintenance in order to attract more hikers. Applications in transportation, the energy domain and the health sector can be found in Hensher et al., 2003; Banfi et al., 2008; Ryan et al., 2008.

In all the papers mentioned above, means and variances of the WTP-distribution are given and a cautious remark is made about the occurrence of a large number of outliers. Daly et al. (2009) however show that for the assumptions made about the distribution of the utility coefficients in most of these papers, the means, variances and higher order moments of the WTP distribution do not even exist. The accuracy of the estimated WTP-distribution can therefore only be assessed by comparing percentiles and individual-level estimates as we will illustrate in this paper.

We are not aware of studies that aim at constructing conjoint choice designs to estimate the distribution of the individual-level WTP-values efficiently. Vermeulen et al. (2008) compare different optimal design criteria to estimate WTP but they assume market homogeneity. With the approach in this paper we can obtain precise information on the preference of the individuals and hence capture market heterogeneity efficiently. We use a sequential procedure that was recently introduced in Yu et al. (2010). In this sequential approach, Bayesian methods are used for design selection as well as to update the prior information of each individual's utility coefficients after each response. The Bayesian method allows to use a small initial design for each respondent to start with and to update the prior information about the consumer preferences repeatedly during the study. In this way, the respondents are provided with individually optimized choice sets which capture the heterogeneity

most efficiently. In the final estimation stage, the Hierarchical Bayes method is used to estimate the panel mixed logit model. Yu et al. (2010) showed that this approach performs much better than the non-sequential choice designs for estimating the individual-level part-worths. We will use this sequential approach to estimate the individual WTP-values and investigate the efficiency of the procedure in this context.

This paper is organized as follows. In the next section we present the panel mixed logit model and the estimation methods involved, the concept of willingness-to-pay and the sequential design methodology we use to construct the choice sets. In section 3 we discuss our simulation study and the corresponding results and in section 4 we summarize our key findings.

2 Methodology

2.1 The panel mixed logit model

The panel mixed logit model is used in this paper to describe the population heterogeneity. The preferences of respondent n are represented by the individual-specific utility coefficient vector $\boldsymbol{\beta_n}$. The probability that respondent n chooses alternative k (k = 1,...,K) in choice set s (s = 1,...,S) is

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

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

Denote by $\mathbf{y}_n^S$ the $K \times S$-dimensional individual-specific vector containing the S choices from respondent n that correspond to the S choice sets of K alternatives. Conditional on $\boldsymbol{\beta_n}$, the likelihood function for a given $\mathbf{y}_n^S$ can be written as

$$L(\mathbf{y}_n^S|\mathbf{X}_n^S, \boldsymbol{\beta_n}) \;=\; \prod_{s=1}^{S}\prod_{k=1}^{K}(p_{ksn}(\boldsymbol{\beta_n}))^{y_{ksn}}, \tag{2}$$

where y_{ksn}, element of $\mathbf{y}_n^S$, is 1 if respondent n chooses alternative k in choice set s and 0 otherwise and $\mathbf{X}_n^S$ is a matrix containing the attribute values of each alternative in the S choice sets that has been assigned to respondent n.

We assume that the coefficient vector $\boldsymbol{\beta_n}$ is randomly drawn from a p-variate normal distribution with mean $\boldsymbol{\mu_\beta}$ and covariance matrix $\boldsymbol{\Sigma_\beta}$, that is $N(\boldsymbol{\beta_n} \mid \boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta})$. Then the likelihood of a given sequence of choices $\mathbf{y}_n^S$, unconditional on $\boldsymbol{\beta_n}$, for respondent n is

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

The probabilities of a single respondent in multiple choice situations will be correlated and the above formulation takes this dependency into account. That is, the coefficient vector $\boldsymbol{\beta_n}$ for a given respondent n appears in all choice sets and ensures that the model captures the correlation across repeated choices.

$L(\mathbf{y}_{full}|\mathbf{X}_{full}, \boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta})$ is then the likelihood function of the choices of all N respondents, unconditional on $\boldsymbol{\beta_n}$, and is defined as

$$L(\mathbf{y}_{full}|\mathbf{X}_{full}, \boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta}) \;=\; \prod_{n=1}^{N} L(\mathbf{y}_n^S|\mathbf{X}_n^S, \boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta}) \tag{4}$$

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 $\mathbf{X}_1^S, ..., \mathbf{X}_N^S$ of each individual respondent.

Maximizing this (simulated) likelihood function yields the maximum likelihood estimates for $\boldsymbol{\mu_\beta}$ and $\boldsymbol{\Sigma_\beta}$. Information on the individual parth-worths $\boldsymbol{\beta_n}$ can then be obtained through the procedure described in Train (2003). However, the panel mixed logit model can also be estimated through a Hierarchical Bayes approach which avoids the difficult optimization of the likelihood function. Bayesian estimation techniques have often been used in the recent discrete choice literature (Toubia et al., 2004; Arora and Huber, 2001; Rossi et al., 1996). Under the Bayesian approach, prior beliefs (prior distribution) about parameters are combined with sample information (likelihood) to create updated or posterior beliefs about the parameters.The joint posterior distribution for $\boldsymbol{\beta_n}$, $\boldsymbol{\mu_\beta}$ and $\boldsymbol{\Sigma_\beta}$ is

$$K(\boldsymbol{\beta_n}, \boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta} \mid \mathbf{y}_{full}) \;\propto\; L(\mathbf{y}_{full}|\mathbf{X}_{full}, \boldsymbol{\beta_n})\; N(\boldsymbol{\beta_n}|\boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta})\; k(\boldsymbol{\mu_\beta})\; h(\boldsymbol{\Sigma_\beta}) \tag{5}$$

where k and h are the hyper-prior distributions of $\boldsymbol{\mu_\beta}$ and $\boldsymbol{\Sigma_\beta}$ respectively.

Draws from the joint posterior distribution $K(\boldsymbol{\beta_n}, \boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta}|\ \mathbf{y}_{full})$ are obtained using Gibbs sampling, which can be used to take random draws from multi-parameter densities using full conditional draws (Casella and George, 1992). As such, we obtain posterior distributions for the population mean $\boldsymbol{\mu_\beta}$, for the population heterogeneity matrix $\boldsymbol{\Sigma_\beta}$ and for the individual-level coefficient vectors $\boldsymbol{\beta_n}$. A detailed discussion on HB estimation can be found in Train (2003).

2.2 The willingness-to-pay (WTP)

The marginal rate of substitution quantifies the trade-off between two attributes and thus their relative importance. MRS with respect to the price coefficient is called the willingness-to-pay (WTP). The WTP expresses the maximum amount a person is willing to pay, sacrifice or exchange to obtain a change in an attribute. In order to compute the WTP, one of the attributes of the experiment has to be the price. The mathematical formula of the WTP for attribute level m for a given respondent n is

$$WTP_{m_n} \;=\; -\frac{\beta_{m_n}}{\beta_{p_n}}, \tag{6}$$

where β_{m_n} and β_{p_n} are the utility coefficient of respondent n of attribute level m and the price coefficient, respectively (Vermeulen et al., 2008). We represent by the vector $\boldsymbol{WTP}_n$ all the WTP-values of respondent

n. The individual respondents' $\boldsymbol{WTP}_n$ estimates can then be used to obtain the heterogeneity distribution of WTP.

Models with convenient distributions for the utility coefficients however give inconvenient distributions for WTP that cannot be derived in closed form. For estimates of the price coefficient close to zero, the ratio will become extremely large and the resulting distribution gets long flat tails. As already mentioned, Daly et al. (2009) show that for many often used distributions for the utility coefficients such as the normal and triangular distributions, the distribution of WTP has infinite moments.

To estimate the posterior distribution of the individual-level $\boldsymbol{WTP}_n$ and the distribution of the WTP in the market, we will use the results of the HB estimation. For each respondent n and for each draw from the posterior distribution of $\boldsymbol{\beta_n}$ we compute the corresponding willingness-to-pay coefficients by taking the ratio of the utility coefficients and the price coefficient. The average of this posterior distribution of $\boldsymbol{WTP}_n$ is used as the estimate for $\boldsymbol{WTP}_n$. The distribution of these individual-level $\boldsymbol{WTP}_n$ estimates is used as the estimated heterogeneity distribution of willingness-to-pay.

2.3 Design construction

Traditionally in optimal design of experiments, one determines the choice sets by optimizing a function of the variance-covariance matrix of the estimates of interest but this becomes very complicated for the panel mixed logit model. By adopting an individual sequential approach, one ensures efficient estimation of the underlying individual conditional logit model which simplifies the design problem substantially. In this section, we briefly review the sequential design construction process of Yu et al. (2010). First we introduce the design criterion used and then explain the algorithm behind the sequential approach.

2.3.1 Bayesian D-optimality criterion

Assuredly, one of the most frequently used design criteria is the D-criterion. D-optimal designs minimize the determinant of the variance-covariance matrix of the parameters to be estimated. Many authors have discussed its advantages over other design criteria (Goos, 2002; Kessels et al., 2006a). As choice models are nonlinear in the parameters, the variance-covariance matrix depends on the values of the model parameters. Several solutions have been proposed for this. Anderson and Wiley (1992), for example, use zero prior parameter values to construct designs which are then called utility neutral designs. Huber and Zwerina (1996) use nonzero prior values for all model parameters, which leads to locally optimal designs, and show that if the prior values are close enough to the true values, the resulting design is more efficient than the utility neutral designs. Sándor and Wedel (2001) have introduced semi-Bayesian optimal designs in the marketing literature. They take the uncertainty about the prior values into account by integrating the design criterion out over a distribution of the prior values. They examine the situations where semi-Bayesian optimal designs are more efficient than the corresponding locally optimal designs. A detailed discussion on these design construction approaches can be found in Yu et al. (2008). All the design approaches mentioned are based on the asymptotic covariance matrix of the maximum likelihood estimates, so the prior knowledge on the model parameters is only taken into account in the design stage.

A true Bayesian optimal design is based on the expected covariance matrix of the posterior distribution of the parameters of interest where the expectation is taken over the marginal distribution of the data (Yu et al., 2008). An important advantage of this approach is that no asymptotic results are needed so that we can safely use these designs for our individual sequential approach. However, as no analytical results exist, one has to approximate the expected covariance matrix of the posterior distribution. Yu et al. (2008) compare different approximations and found that for small sample sizes the approximation based on the generalized Fisher information matrix (GFIM) is best. Therefore the Bayesian optimal design we derive to estimate $\boldsymbol{\beta}$ in the sequential procedure are based on the GFIM for the parameters $\boldsymbol{\beta}$ of the conditional logit model. The approximated covariance matrix is the negative expectation of the second derivative of the log posterior density, that is

$$I_{GFIM}(\boldsymbol{\beta}|\mathbf{X}) \;=\; -E[\frac{\partial^2 \log\, q(\boldsymbol{\beta}|\mathbf{Y},\mathbf{X})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}], \tag{7}$$

where $q(\boldsymbol{\beta}|\mathbf{Y},\mathbf{X})$ is the posterior distribution of $\boldsymbol{\beta}$ conditional on the design matrix $\mathbf{X}$ and the corresponding responses $\mathbf{Y}$. The determinant of this I_{GFIM} matrix will be used to assess the efficiency of the design. As this determinant depends on the parameter values, we consider the expected value of the determinant over a prior distribution of the parameter values, $\pi(\boldsymbol{\beta})$:

$$D_B - error \;=\; \int |I_{GFIM}(\boldsymbol{\beta}|\mathbf{X})|^{-\frac{1}{p}}\, \pi(\boldsymbol{\beta}) d\boldsymbol{\beta} \tag{8}$$

which is called the Bayesian D-error. The optimal design is the one that minimizes the $D_B - error$.

2.3.2 Sequential design approach

We use a sequential approach to select the choice sets. Unlike the aggregate-customization approach in which all respondents evaluate a common design optimized for the average respondent and which has been used frequently in the literature (Arora and Huber, 2001; Sándor and Wedel, 2001; Kessels et al., 2006b), the sequential approach allows to construct individual designs that can take the information on the respondent heterogeneity maximally into account. We will briefly revisit the two-step sequential approach (for more information see Yu et al., 2010).

Initial stage:

- Using a common prior distribution for all N respondents, $\pi(\boldsymbol{\beta})$, we generate N initial D-optimal designs with S_1 choice sets of size K using the greedy approach introduced by Sándor and Wedel (2005). We denote the corresponding design matrix by $\mathbf{X}_n^{S_1}$. For each individual n, the data from the initial stage $\mathbf{y}_n^{S_1}$ are analyzed in a Bayesian way using the conditional logit model. The output of this analysis is a posterior distribution for $\boldsymbol{\beta_n}$, denoted by $q(\boldsymbol{\beta_n}|\mathbf{y}_n^{S_1},\mathbf{X}_n^{S_1})$, that is a mix of the initial prior distribution $\pi(\boldsymbol{\beta})$ and the individual-level likelihood $L(\mathbf{y}_n^{S_1}|\mathbf{X}_n^{S_1},\boldsymbol{\beta_n})$. This posterior distribution is then used as input for the sequential stage of the experiment.

Sequential stage:

- The posterior distribution obtained from the initial stage, $q(\boldsymbol{\beta_n}|\mathbf{y}_n^{S_1},\mathbf{X}_n^{S_1})$, is used as the prior distribution for constructing the next choice set, $\mathbf{x}_n^{S_1+1}$ for respondent n. The new choice set is chosen by minimizing the $D_B - error$ of the combined design $(\mathbf{X}_n^{S_1},\mathbf{x}_n^{S_1+1})$.

- The new choice set $\mathbf{x}_n^{S_1+1}$ is assigned to respondent n and the prior information is updated with all S_1+1 observations. The resulting posterior distribution, $q(\boldsymbol{\beta_n}|\mathbf{y}_n^{S_1+1}, \mathbf{X}_n^{S_1}, \mathbf{x}_n^{S_1+1})$, is then used to obtain the next choice set, $\mathbf{x}_n^{S_1+2}$, by minimizing the $D_B - error$. This process is repeated until a pre-specified number of choice sets S is attained.

To approximate the $D_B - error$ in both the initial and sequential stages, it is necessary to have a large number of draws from the posterior distribution. Since there is no closed form expression available for the posterior distribution, we use importance sampling proposed by Bedrick et al. (1997) and also discussed in Yu et al. (2010).

3 Simulation Study

To investigate how efficient the individual WTP values can be estimated by this sequential approach and the Hierarchical Bayes estimation results, we perform a simulation study in which we compare the results obtained by different designs.

3.1 Simulation setup

We consider designs with four attributes, one with four levels and three with two levels including a price attribute at two levels. We also assume that the total number of choice sets the N=250 individuals have to evaluate is 16, with three alternatives per choice set (S=16 and K=3). We use $S_1 = 5$ choice sets in the initial stage and 11 choice sets are constructed sequentially.

We need a prior distribution $\pi(\boldsymbol{\beta})$ for the utility coefficients to optimize the design. In the discrete choice literature, the utility coefficients have been assumed to be uncorrelated by some authors (Sonnier et al., 2007; Bliemer and Rose, 2008) and correlated by others (Train and Weeks, 2005; Dotson et al., 2009). The WTP parameters are correlated in both cases, since the price coefficient enters the denominator of each WTP. We use prior values that were obtained in an empirical study with a similar setup by Dotson et al. (2009). They obtained individual-level parameters with mean $\boldsymbol{\mu_\beta} = [0.547, 0.405, 0.163, -0.398, -0.009, -0.540]$ and covariance matrix

$$\boldsymbol{\Sigma_\beta} = \begin{bmatrix} 1.097 & 0.684 & 0.484 & 0.144 & 0.151 & 0.028 \\ 0.684 & 1.535 & 0.525 & 0.3 & 0.375 & -0.066 \\ 0.484 & 0.525 & 0.867 & 0.066 & 0.392 & -0.007 \\ 0.144 & 0.3 & 0.066 & 1.323 & 0.264 & -0.157 \\ 0.151 & 0.375 & 0.392 & 0.264 & 2.056 & -0.184 \\ 0.028 & -0.066 & -0.007 & -0.157 & -0.184 & 0.354 \end{bmatrix}.$$

We will use a multivariate normal distribution with this estimated mean and covariance matrix as the prior distribution $\pi(\boldsymbol{\beta})$ to generate our designs.

The sequential design is compared with two non-sequential benchmark designs, a semi-Bayesian D-optimal design and a nearly orthogonal design. The semi-Bayesian D-optimal design is constructed for the conditional logit model and uses the same prior distribution $\pi(\boldsymbol{\beta})$ that we use for the sequential design. Further discussion of constructing semi-Bayesian D-optimal designs can be found in Sándor and Wedel (2001). The nearly orthogonal design, which is close to a utility neutral design, is constructed using SPSS software (SPSS Inc., 2007).

In the sequential approach we need to simulate the individual choices in each step. We will first use the prior distribution $\pi(\boldsymbol{\beta})$ as the heterogeneity distribution from which to generate the $\boldsymbol{\beta_n}$ that are considered to be the true utility coefficients of the respondents in the study. The individual-level $\boldsymbol{WTP}_n$ values calculated from these true utility coefficients are considered as the true WTP coefficients of these respondents. To evaluate the estimated heterogeneity distributions of WTP obtained with the different designs we need to know the true heterogeneity distribution of WTP. This distribution was obtained by simulating a huge number of utility coefficients from the prior distribution and calculating the corresponding WTP-values. To assess the effect of a misspecified prior distribution, we also simulate individual choices using $\boldsymbol{\beta_n}$ drawn from different heterogeneity distributions than the prior distribution used for constructing the design.

After generating the data using the different designs, we estimate for each of them the panel mixed logit model by the Hierarchical Bayes approach. We then estimate the individual-level $\mathbf{WTP}_n$ vectors for each design and compare them to the true values to assess their accuracy.

3.2 Simulation results

3.2.1 Estimation accuracy of the utility coefficients

Yu et al. (2010) show that the sequential approach improves the estimation accuracy of the utility parameters compared to an orthogonal design and a semi-Bayesian D-optimal design. As expected, we obtained similar results about the precision of the utility parameters. We summarize our results by the root mean squared error values for the individual utility coefficients $\boldsymbol{\beta}_n$, for the population mean $\boldsymbol{\mu}_\beta$ and for the population covariance matrix $\boldsymbol{\Sigma}_\beta$ obtained with different designs. For example, the RMSE for the population mean of the utility coefficient $\boldsymbol{\mu}_\beta$ can be defined as:

$$RMSE(\boldsymbol{\mu}_\beta) = [(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)'(\widehat{\boldsymbol{\mu}}_\beta - \boldsymbol{\mu}_\beta)]^{1/2} \tag{9}$$

where $\widehat{\boldsymbol{\mu}}_\beta$ is the mean of the posterior distribution of $\boldsymbol{\mu}_\beta$ obtained by the HB process explained in section 2.1 and $\boldsymbol{\mu}_\beta$ is the true population mean. A similar definition can be obtained for $\boldsymbol{\Sigma}_\beta$ when we put all the unique elements in a vector. For $\boldsymbol{\beta}_n$ we use the average of the RMSE-values of all respondents.

Table 1 presents the results. From the results in Table 1 we can see that with respect to the estimation accuracy

Table 1: RMSE values for the utility coefficients obtained under different design approaches

	Sequential	Near-Orthogonal	SB D-optimal
$\boldsymbol{\beta}_n$	1.0218	1.2503	1.2238
$\boldsymbol{\mu}_\beta$	0.0736	0.0894	0.0949
$\boldsymbol{\Sigma}_\beta$	0.5374	0.7380	0.6534

of $\boldsymbol{\beta}_n$, the sequential design is 17% better than the semi-Bayesian D-optimal design and 18% better than the orthogonal design. Similarly, the sequential design performs much better than the benchmark designs with respect to the estimation accuracy of $\boldsymbol{\mu}_\beta$ and $\boldsymbol{\Sigma}_\beta$.

3.2.2 Distribution of the estimated Individual-level WTP

We assumed that the utility coefficients come from a multivariate normal distribution. From Daly et al. (2009) we know that the implied distribution of WTP does not have finite moments in this case. Therefore, there is no point in reporting the population mean and the population covariance matrix of the willingness-to-pay values. To assess the performance of the different design approaches for estimating the distribution of WTP, we compute the percentiles of the estimated and the true individual-level WTP distributions and we look at the estimation errors of the individual-level estimates. We report only the results of the first WTP parameter as the results are similar for the other parameters.

Table 2 shows the percentiles of the heterogeneity WTP distributions obtained with the different designs. It is

Table 2: Percentiles of individual-level WTP distributions obtained with different designs

Percentile	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**
True	-8.326	-3.792	-0.952	0.582	2.001	5.003	9.590
Sequential	-7.209	-3.636	-0.670	0.555	2.028	4.568	8.112
Near-Orthogonal	-6.118	-2.736	-1.029	0.487	1.770	3.924	12.355
SB D-optimal	-15.537	-5.723	-1.439	0.393	2.469	6.006	11.146

clear from this table that the sequential design can estimate the true distribution of the individual-level WTP more precisely than all the benchmark designs.

To assess the accuracy, we also compute the estimation error of the $\boldsymbol{WTP}_n$ parameters for each respondent, which is the difference between the true individual-level WTP coefficient and its estimate. In Figure 2, the y-axis represents the individuals' estimation error of the first WTP parameter and the x-axis represents the individuals. From Figure 2, we can see that the Orthogonal design and the semi-Bayesian design lead to some extremely inaccurate estimates. The sequential design on the contrary avoids extremely inaccurate estimates. This is because the sequential design is tailored to each individual's preference structure and it allows us to catch more individual information. Therefore, the important merit of the sequential approach advocated here is that we can reduce the extremely inaccurate estimates considerably.

3.2.3 Robustness of the design on the misspecification of the prior distribution

In the previous section, we studied the performance of the sequential design approach assuming that the prior distribution on $\boldsymbol{\beta}$ used to construct the design contains correct information on the distribution of the utility coefficients. In this section we relax this assumption and examine the effect of misspecified prior distributions. We still use $\pi(\boldsymbol{\beta})$ defined in section 3.1 to generate the designs but the choices are generated with individual-level $\boldsymbol{\beta_n}$ that were drawn from other distributions. In each case the utility coefficients are drawn from a multivariate normal distribution with mean $\widetilde{\boldsymbol{\mu}}$ and covariance matrix $\boldsymbol{\Sigma_\beta}$, where

$$\widetilde{\boldsymbol{\mu}} = \boldsymbol{\mu_\beta} + \delta \mathbf{1}_p.$$

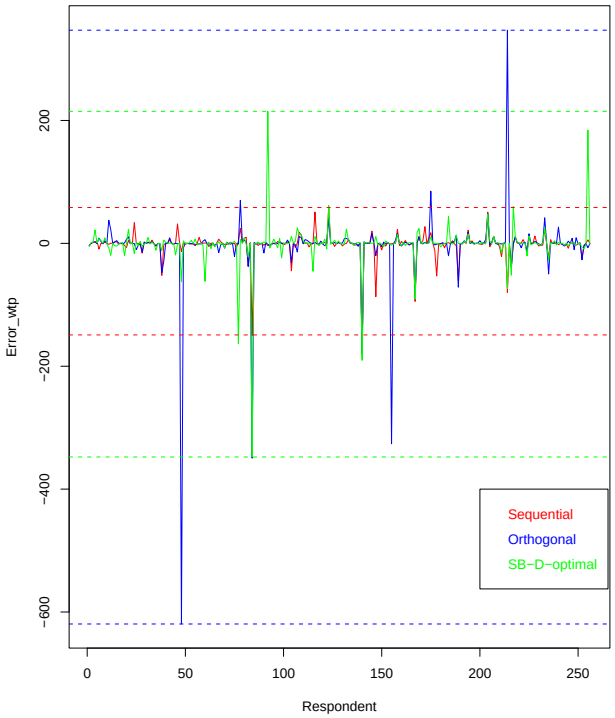

Figure 1: Estimation error of Individual-level estimates of the WTP

Note that we do not change the correlation structure so the prior information on the variances and covariances is still assumed to be correct. The parameter δ reflects how much the mean of the multivariate normal distribution that we use to generate the responses deviates from the mean value that we assumed while constructing the designs. We report the results for $\delta = -0.3$ to model a moderate deviation and $\delta = -0.6$ to denote serious misspecification. Table 3 presents the percentiles of the WTP distributions obtained under different designs and for the both δ-values.

Table 3: Percentiles of the heterogeneity WTP distribution obtained under different designs and for different mean deviation values

	Percentile	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**
δ = -0.3	True	-5.019	-2.608	-0.908	0.137	1.068	2.419	5.187
	Sequential	-5.765	-2.292	-0.751	0.198	1.074	2.415	5.092
	Near-Orthogonal	-8.812	-3.759	-1.070	0.068	1.113	2.475	3.822
	SB D-optimal	-7.336	-2.766	-1.041	-0.045	0.895	2.519	4.424
δ = -0.6	True	-3.391	-1.721	-0.901	-0.141	0.598	1.445	2.416
	Sequential	-2.686	-1.731	-0.890	-0.101	0.602	1.430	2.383
	Near-Orthogonal	-3.700	-2.106	-1.039	-0.212	0.679	1.749	2.934
	SB D-optimal	-3.505	-2.172	-0.912	-0.085	0.593	1.455	2.942

It is clear from this table that the sequential design can estimate the true heterogeneity WTP distribution more efficiently than the benchmark designs in both cases as the percentiles of the WTP distribution obtained by the sequential design are closest to those of the true WTP distribution in most of the cases. This indicates that the performance of the sequential approach is quite robust with respect to the misspecification of the prior distribution. This is due to the sequential nature of the approach in which we update the prior information

repeatedly over the process. Even if the starting information is inaccurate, the information will become better in each step of the sequential procedure.

4 Discussion and Conclusions

Several design and estimation techniques have been introduced in the marketing literature to estimate the distribution of willingness-to-pay (WTP) efficiently. In this paper, we use an adaptive sequential approach to construct choice designs to estimate the distribution of the WTP. The idea is motivated by findings of Yu et al. (2010) who developed the sequential approach to estimate the individual-level utility coefficients. In this approach Bayesian methods are used for the selection of choice sets and to update the prior information on each individual's preference after each choice. Finally the Hierarchical Bayes method is used to estimate the individual utility coefficients of the panel mixed logit model and the corresponding WTP values.

The sequential design was compared with two non-sequential designs: a semi-Bayesian D-optimal design for the conditional logit model and a nearly orthogonal design constructed for the same experimental setting. The simulation study shows that the sequential approach leads to designs that can estimate the true WTP distribution much more precisely than the benchmark designs by comparing the percentiles of the estimated WTP distributions based on the different designs with the percentiles of the true WTP distribution. Comparison of the estimation errors also shows that the sequential design avoids extremely inaccurate estimates of individual-level WTP. Moreover, the sequential approach also outperforms the other designs when the prior distribution is misspecified. This is to be expected as the prior information is updated repeatedly and is as such corrected in each step.

References

Anderson, D.A. and Wiley, J.B., 1992. Efficient choice designs for estimating availability cross-effect models. *Marketing Letters.* 3 357-370.

Arora, N., Huber, J., 2001. Improving parameter estimates and model prediction by aggregate customization in choice experiments. *International Journal of Forecasting.* 28 273-283.

Banfi, S., Farsi, M., Filippini, M., Jacob, M., 2008. Willingness to pay for energy-saving measures in residential buildings. *Energy Economics.* 30 503-516.

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., 2008. Construction of experimental designs for mixed logit models allowing for correlation across choice observations. Working paper ITLS-WP-08-13. Institute of Transport and Logistics Studies, University of Sydney.

Casella, G. and George, E., 1992. Explaining the Gibbs Sampler. *American Statistician.* 46 167-174.

Daly, A., Hess, S., Train, K., 2009. Assuring finite moments for willingness to pay in random coefficient models. *paper presented at the European Transport Conference, Leeuwenhorst, October 2009.*

Dotson, J., Lenk, P., Brazell, J.P., Otter, T., MacEachern, S.N., Allenby, G.M., 2009. A Probit Model with Structured Covariance for Similarity Effects and Source of Volume Calculations. Available at SSRN: http://ssrn.com/abstract=1396232

Goos, P., 2002. The optimal design of blocked and split-plot experiments. Springer, New York.

Hensher, D. and Sullivan, C., 2003. Willingness to pay for road curviness and road type. *Transportation Research Part D.* 8 139-155.

Huber, J., Zwerina, K., 1996. The importance of utility balance in efficient choice designs. *Journal of Marketing Research.* 33 307-317.

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

Kessels, R., Jones, B., Goos, P., Vandebroek, M., 2006b. An efficient algorithm for constructing Bayesian optimal choice designs. *Journal of Business and Economics Statistics.* In press.

Rossi, P.E., McCulloch, R.E., Allenby, G.M., 1996. The value of purchase history data in target marketing. *Marketing Science.* 15 321-340.

Ryan, M., Gerard, K., Amaya-Amaya, M. (eds), 2008. Using discrete choice experiments to value health and health care. 101-116.

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.

Scarpa, R., Thiene, M., Train, K., 2008. Utility in willingness to pay space: A tool to address confounding random scale effects in destination choice to the ASlps. *American Journal of Agriculture Economics.* 90(4) 994-1010.

Sonnier, G., Ainslie, A., Otter, T., 2007. Heterogeneity distributions of willingness to pay in choice models. *Quantitative Marketing and Economics.* 5 313-331.

SPSS Inc., 2007. SPSS Conjoint 16.0. Printed in the United States of America.

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

Train, K., 2003. Discrete choice methods with simulation. Cambridge University Press.

Train, K., Weeks, M., (2005). Discrete choice models in preference space and willingness to pay space. In: Scarpa, R., Alberini, A. (eds), Applications of simulation methods in environmental and resource economics. Springer Publisher, Dordrecht, The Netherlands.

Vermeulen, B., Goos, P., Scarpa, R., Vandebroek, M., (2008). Design criteria to develop choice experiments to measure the WTP accurately. Research Report KBI-0809, Department of Decision Sciences and Information Management, Katholieke Universiteit Leuven.

Yu, J., Goos, P., Vandebroek, M., 2008. A comparison of different Bayesian design criteria for constructing efficient conjoint choice experiments. Faculty of Business and Economics Research report. KBI-0817.

Yu, J., Goos, P., Vandebroek, M., 2010. Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity. Submitted for publication.