KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Business and Economics

# Efficient and robust willingness-to-pay designs for choice experiments: some evidence from simulations

Bart Vermeulen, Peter Goos, Riccardo Scarpa and Martina Vandebroek

DEPARTMENT OF DECISION SCIENCES AND INFORMATION MANAGEMENT (KBI)

KBI 0809

# Efficient and robust willingness-to-pay designs for choice experiments: some evidence from simulations

**Bart Vermeulen**[1]

Faculty of Business and Economics, Katholieke Universiteit Leuven

Naamsestraat 69, B-3000 Leuven, Belgium

Email: bart.vermeulen@econ.kuleuven.be

Tel: +32 (0)16 32 69 63

Fax: +32 (0)16 32 66 24

**Peter Goos**

Faculty of Applied Economics, Universiteit Antwerpen

Prinsstraat 13, B-2000 Antwerpen, Belgium

Email: Peter.Goos@ua.ac.be

Tel: +32 (0)3 220 40 59

Fax: +32 (0)3 220 48 17

**Riccardo Scarpa**

Economics Department, University of Waikato

Private Bag 3105, Hamilton 3240, New Zealand

Email: rscarpa@waikato.ac.nz

Tel: +64 (0) 7 838 4045

Fax: +64 (0) 7 838 4331

**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

[1]Corresponding author

# Efficient and robust willingness-to-pay designs for choice experiments: some evidence from simulations

**Abstract**

We apply a design efficiency criterion to construct conjoint choice experiments specifically focused on the accuracy of marginal $WTP$ estimates. In a simulation study and a numerical example, the resulting $WTP$-optimal designs are compared to alternative designs suggested in the literature. It turns out that $WTP$-optimal designs not only improve the estimation accuracy of the marginal $WTP$, as expected on the basis of the nature of the efficiency criterion, but they also considerably reduce the occurrence of extreme estimates, which also exhibit smaller deviations from the real values. The proposed criterion is therefore valuable for non-market valuation studies as it reduces the sample size required for a given degree of accuracy and it produces estimates with fewer outliers.

Keywords: willingness-to-pay, optimal design, choice experiments, conditional logit model, robust

# 1 Introduction

Since the early nineties the number of studies using conjoint choice experiments as a tool to estimate the value of attributes of complex goods has vastly increased. Whereas previous studies employing this stated preference method were mostly directed to predict choice behavior and market shares, the increasing emphasis on estimation of implied values of product or service attributes poses new challenges. One such challenge is the development and testing of specific design criteria for experiments aimed at this specific purpose and their comparative evaluation with more established criteria. This paper intends to contribute towards this effort.

The objective of conjoint choice experiments is to model respondents' choices as a function of the features of a good or a service. For that purpose, the respondents are presented with a series of choice tasks, in each of which they are asked to indicate their favorite alternative. Alternatives are described by means of attributes and their levels. Because the potential combinations of attributes, levels and their allocations in choice tasks are typically many more than can be handled in the course of the interview, experimental design techniques are used to select from the full factorial a suitable arrangement of choice tasks.

The observed choices are then typically analyzed invoking random utility theory by means of discrete choice models. In valuation studies the estimates of the utility coefficients are often used to calculate marginal rates of substitution ($MRS$) with respect to the cost coefficient and interpreted as consumers' marginal willingness-to-pay ($WTP$) for this attribute. A substantial number of stated preference studies have recently used choice experiments as a tool to derive value estimates. Examples of studies of this kind have been published not only in the conventional fields of application of stated choice, such as in marketing [22], transportation choice [9], environmental economics [3] and health care economics [21], but have also appeared in food [16], livestock [20] and crop choice [15], as well as in cultural [17] and energy economics [2]. In this articulated research programme the conditional logit model has been the dominant approach to data analysis.

In logit models of discrete choice the precision of estimates of utility coefficients, and consequently of the marginal $WTP$, is to a large extent determined by the quality of the data. This gives to the design of the conjoint choice experiment an extremely important role. An efficient design maximizes the information in the experiment and in this way guarantees accurate utility coefficient estimates at a manageable sample size. Creating an efficient conjoint choice design involves selecting the most appropriate alternatives and grouping them into choice sets according to an efficiency criterion. In this paper we propose and test the performance of an efficiency criterion to create choice experiments specifically leading to accurate marginal $WTP$ estimates. Based on a simulation study and a numerical example, the resulting $WTP$-optimal designs are compared with designs which focus on the precision of the estimated utility coefficients and with other commonly used conjoint choice design strategies in terms of the accuracy of the marginal $WTP$ estimates.

The plan of this paper is as follows. In the next section, we introduce the conditional logit model that is typically used to analyze the choices of the respondents and to derive the marginal $WTP$ estimates. In Section 3, we start by giving a short overview of the existing literature on design of conjoint choice experiments used for valuation issues. Next, in Section 3, we present an efficiency criterion focusing on precise marginal $WTP$ estimates and define the corresponding $WTP$-error.

In Section 4, we compare $WTP$-optimal designs with other commonly used designs in terms of the $WTP$-error and discuss the results of a simulation study in which designs obtained with different criteria are evaluated on the basis of their accuracy in terms of marginal $WTP$ estimates. In addition, we examine the designs with respect to the estimation accuracy of the utility coefficients, which also remains an important criterion. Finally, in Section 5, we illustrate the performance of the $WTP$-optimal designs in an application related to the marginal willingness to donate for environmental projects.

# 2 The conditional logit model and the marginal WTP

In this section, we briefly review the conditional logit model which is commonly used to analyze the data of a conjoint choice experiment and also provide a brief definition of the concept of marginal 'willingness-to-pay'.

## 2.1 The conditional logit model

Data from a conjoint choice experiment are usually analyzed by the widely-known conditional logit model. The utility of alternative $j$ in choice set $k$ for respondent $n$ is expressed as

$$U_{nkj} = \beta_1 x_{1kj} + \ldots + \beta_M x_{Mkj} + \epsilon_{nkj}. \tag{1}$$

The utility $U_{nkj}$ consists of two components: a deterministic component $\beta_1 x_{1kj} + \ldots + \beta_M x_{Mkj}$, or in vector notation $\mathbf{x}'_{kj}\boldsymbol{\beta}$, and a stochastic component $\epsilon_{nkj}$. In the deterministic component, the $M$-dimensional vector $\boldsymbol{\beta}$, which is assumed common for all respondents, contains the utility coefficients of the discrete choice model. These coefficients reflect the importance of the underlying $M$ attributes of the good or service. The $M$-dimensional vector $\mathbf{x}_{kj}$ describes the bundle of these $M$ attributes of alternative $j$ in choice set $k$. The stochastic error term $\epsilon_{nkj}$ captures the unobserved factors influencing the utility experienced by the respondent. The error terms are assumed to be independent and identically extreme value distributed. The probability that respondent $n$ chooses alternative $j$ of choice set $k$ is then

$$P_{nkj} = \frac{\exp(\mathbf{x}_{kj}^{'}\boldsymbol{\beta})}{\sum_{i=1}^{J}\exp(\mathbf{x}_{ki}^{'}\boldsymbol{\beta})}. \tag{2}$$

## 2.2 The marginal willingness-to-pay (WTP)

The marginal rate of substitution ($MRS$) is the rate which measures the willingness of individuals to give up one attribute of a good or service in exchange for another such that the utility of the good or service remains constant. So, it quantifies the trade-off between the two attributes and thus their relative importance. When the trade-off is made with respect to the price of a good or a service, the $MRS$ is called the marginal willingness-to-pay ($WTP$). In this way, the marginal $WTP$ for an attribute measures the change in price that compensates a change in the attribute $m$, while all other attributes are held constant. To estimate the marginal $WTP$, one of the attributes $x$ has to be the attribute price $p$ in the utility expression (1). Mathematically, the trade-off between the attribute $x_m$ and the price $p$ can be written as

$$dU = \beta_m dx_m + \beta_p dp = 0 \tag{3}$$

or as

$$WTP \equiv \frac{dp}{dx_m} = -\frac{\beta_m}{\beta_p}. \tag{4}$$

Expression (4) shows that the marginal $WTP$ is computed as minus the ratio of the coefficients for attribute $m$ and the price $p$ (see e.g. [10]).

# 3 Constructing optimal designs to estimate the WTP

In this section, we provide an overview of the literature on design of experiments to estimate the marginal $WTP$ and present an efficiency criterion that focuses on the accurate estimation of this substitution rate.

## 3.1 Designs for experiments to estimate the WTP

Despite an increasing number of applications of conjoint experiments for valuation issues, the literature on efficient designs for this purpose is scarce. In [11] optimal designs were developed for the double-bounded dichotomous contingent valuation experiment and focused specifically on the accuracy of estimation of marginal $WTP$. In a single-bounded dichotomous contingent valuation experiment, the marginal $WTP$ for a change in the attributes of a product or a service is estimated by asking the respondent whether he/she is prepared to pay a certain amount of money for this change. In the double-bounded experiment, this initial bid is followed by a second bid which is higher if the answer to the first bid was affirmative and lower otherwise. In [11] $D$-optimal designs, $c$-optimal and designs based on the so-called fiducial method were compared. While $D$-optimal designs minimize the determinant of the variance-covariance matrix of the estimated utility coefficients in the double-bounded logit model, $c$-optimal designs minimize the variance of a function of the estimated utility coefficients. The function under investigation was the marginal $WTP$, the variance of which was approximated using the delta-method [7]. The $c$-optimality criterion consisted of the sum of the approximate variances of the marginal $WTPs$. The third design strategy examined was based on minimizing the fiducial interval of the marginal $WTP$. The three design strategies were examined in terms of the variance of the estimated marginal $WTP$ and it turned out that the $c$-optimal and fiducial designs performed better than the $D$-optimal designs. However, the difference between the three design strategies was small for the double-bounded logit model.

There is almost no literature on the design of conjoint choice experiments to estimate the marginal $WTP$ precisely. In [5], results were reported on the accuracy of the marginal $WTP$ estimates obtained using a shifted, a locally $D$-optimal and a Bayesian $D$-optimal design. To obtain shifted designs, a starting design which involves a number of alternatives equal to the number of choice sets of the desired design is used as a base. By increasing all attribute levels of the first alternative of the initial design by one, the second alternative of the first choice set of the shifted design is found (if an attribute already is at its highest level, this level is changed to the lowest level admissible for that attribute). Another increase of the attribute levels of the first alternative of the initial design then leads to the third alternative of the first choice set of the shifted design. This procedure continues till the desired number of alternatives in a choice set in the final design is obtained. In a similar fashion, the other choice sets of the shifted design are found. $D$-optimal designs minimize the generalized variance of the utility coefficients of the discrete choice model by minimizing the determinant of their variance-covariance matrix. As this variance-covariance matrix is a function of the utility coefficients themselves, prior knowledge about these coefficients is required to develop $D$-optimal designs. Locally $D$-optimal designs address this problem by using a point estimate for the utility coefficients as input to the search procedure for $D$-optimal conjoint choice designs, while Bayesian $D$-optimal designs assume a prior distribution for the utility coefficients to formally account for the uncertainty about their values at this stage of the investigation. In [5] it was concluded that substantial improvements in marginal $WTP$ estimation accuracy were achieved when a Bayesian $D$-optimal design was used, provided an informative prior distribution had been specified. The gain in precision increased with the utility coefficients' absolute magnitudes. Using a Bayesian $D$-optimal design constructed with an uninformative prior distribution led to estimates for the $WTP$ that were less precise than those obtained using a shifted design.

In [19], locally $A$-, $D$- and $c$-optimal conjoint choice designs as well as random and orthogonal designs were compared based on their relative efficiency. The $c$-optimality criterion was comprised of the sum of the variances of the marginal $WTP$, which were approximated by the delta-method. The authors concluded that the $D$-optimal design performed surprisingly well in terms of the $c$-efficiency criterion, while other designs did substantially worse. Conversely, the $c$-optimal design scored relatively high in terms of the other efficiency criteria.

Numerous applications of choice experiments developed with the main goal of estimating the marginal $WTP$ used other design strategies. Most of them were based on fractional factorial (orthogonal) designs, [see e.g. 8, 18], and a few on full factorial designs [6]. Relatively few used $D$-efficiency criteria without or with sequential updating. For example, [15] generated a design of 40 profiles maximizing a local $D$-efficiency criterion. This design was then divided in 4 subdesigns of

10 profiles which were randomly presented to the respondents. An example of sequential updating based on Bayesian $D$-efficiency is reported in [25], where desirable efficiency gains are associated with this approach. To measure the marginal $WTP$ for eco-labels, [22] applied designs developed by Sawtooth Software providing minimal overlap, level balance and orthogonality. It should be clear that, despite the increasing number of applications of valuing attributes of a product or a service, the experimental design literature on conjoint choice experiments has not yet dedicated sufficient attention to developing design optimality criteria for the specific purpose of accurately estimating marginal $WTP$. Therefore, the goal of this paper is to provide a first contribution to close this gap by focusing on optimal designs for conjoint choice experiments to estimate the marginal $WTP$ accurately.

## 3.2 Bayesian WTP-optimal conjoint choice designs

In this paper, it is assumed that the goal of a conjoint choice experiment is to provide an accurate assessment of the marginal $WTP$ for the attributes of a product or service. This means that a $WTP$-optimal design minimizes the variance of minus the ratio of the coefficient for attribute $m$ and the price $p$, which makes that a $WTP$-optimal design can be classified in the broader class of $c$-optimal designs. The asymptotic variance of the marginal $WTP$ is approximated using the delta-method

$$\begin{aligned}\widehat{var}\left(\widehat{WTP}_m\right) &= \widehat{var}\left(-\frac{\hat{\beta}_m}{\hat{\beta}_p}\right) \\ &\approx \frac{1}{\hat{\beta}_p^2}\left(var(\hat{\beta}_m) - 2\left(\frac{\hat{\beta}_m}{\hat{\beta}_p}\right)cov(\hat{\beta}_m,\hat{\beta}_p) + \left(\frac{\hat{\beta}_m}{\hat{\beta}_p}\right)^2 var(\hat{\beta}_p)\right).\end{aligned} \tag{5}$$

Since there are $M-1$ attributes of the product besides the price, we are interested in accurate estimates of the $M-1$ marginal $WTP$ values. The proposed criterion $C_{WTP}$ is defined as the sum of these approximate variances. The efficiency of a design in terms of the proposed criterion is then expressed by

$$WTP\text{-error} = \sum_{m=1}^{M-1} \widehat{var}(\widehat{WTP}_m). \tag{6}$$

The choice design minimizing the $WTP$-error, is called the $WTP$-optimal design. Minimizing expression (6) implies a more accurate assessment of the marginal $WTP$ and, consequently, smaller confidence intervals for it.

As the conditional logit model is a non-linear model, the $WTP$-efficiency depends on the utility coefficients which are unknown at the moment of planning the experiment. A Bayesian approach ([23, 24, 14]) assuming a prior distribution $f(\boldsymbol{\beta})$ on $\boldsymbol{\beta}$ is then appropriate. The Bayesian version of the $WTP$-error is denoted by the $WTP_b$-error and is defined as the expected value of the $WTP$-error over the prior distribution

$$WTP_b\text{-error} = E_\beta \left[ \sum_{m=1}^{M-1} \widehat{var}\left(WTP_m\right) \right] = \int_{\Re^{M-1}} \widehat{var}\left(WTP\right) f(\boldsymbol{\beta}) d\boldsymbol{\beta}. \tag{7}$$

We approximate the integral in (7) by means of 100 quasi-random Halton draws ([28, 1, 29]) from the prior distribution and averaging the $WTP$-error over these draws. This systematic sample replaces the commonly used 1,000 pseudo-random draws from the prior distribution but yields the same designs in terms of efficiency. Evidently, the use of the systematic sample with good coverage properties saves a considerable amount of computation time and reduces simulation variance. The alternating-sample algorithm, described in [14], was applied to search for the design satisfying the Bayesian $WTP$-optimality criterion discussed in this paper, i.e. the design minimizing the $WTP_b$-error.

# 4 Evaluation of the Bayesian WTP-optimal designs

In this section we report the results from a simulation study devised to evaluate the proposed $WTP$-optimality criterion and we compare this criterion to several other popular design strategies based on the $WTP_b$-error. First, we briefly explain how we obtain $WTP$-optimal designs. Next, we present the benchmark designs used as a comparison. Then, we describe our findings across designs in terms of the $WTP_b$-error. Finally, we introduce some additional criteria to evaluate the different designs and report the results of a simulation study assuming in turn correct and incorrect prior information.

## 4.1 Computational aspects

The experiment used in this section involves two three-level attributes and one two-level attribute which are all effects-type coded. Besides these attributes, the price of a good is also included taking two levels that are linearly coded as 1 and 2. This implies that the number of parameters $M$, contained within $\boldsymbol{\beta}$, equals 6.

In our simulation study we assumed that 75 respondents are participating in the experiment and that each respondent indicates his favorite alternative out of the three available ones in each of 12 choice sets, thereby collecting 900 discrete choices. Each respondent evaluates the same choice sets.

As explained in Section 3.2, to create Bayesian $WTP$-optimal designs one needs to make assumptions about a prior distribution for the utility coefficients in the choice model. As a prior distribution, we used a $6$-dimensional normal distribution with mean [-0.5 0 -0.5 0 -0.5 -0.7]. The first $5$ elements of the mean vector correspond to the utility coefficients associated with the 3 attributes and the last element corresponds to the price coefficient. As can be noted in the mean vector, we assume that price and utility are negatively related, which is common in economic literature. The variance-covariance matrix of the prior distribution is taken equal to

$$\begin{pmatrix} 0.5\ \mathbf{I}_{M-1} & \mathbf{0}_{(M-1)\times 1} \\ \mathbf{0}_{1\times(M-1)} & 0.05 \end{pmatrix}, \tag{8}$$

where $\mathbf{I}_i$ is the $i$-dimensional identity matrix. This prior distribution follows the recommendations formulated in [13]. The variance of the price coefficient is smaller than the variance of the other utility coefficients, and sufficiently so to ensure that only negative price coefficients are drawn in the Bayesian optimal design approach.

## 4.2 Benchmark designs

To evaluate the performance of the Bayesian $WTP$-optimal design, we compare it with a Bayesian $D$-optimal design and three standard designs generated by Sawtooth Software. The development of Bayesian $D$-optimal designs is extensively described in [23, 12, 5, 13].

Sawtooth Software offers the user three design options. A first option, labeled 'complete enumeration', is a design which is constructed following three principles: level balance, minimal attribute level overlap within one choice set and orthogonality. Henceforth, we refer to this design as a (near-)orthogonal design. A second option, labeled 'random design', consists of constructing a design by means of randomly choosing the levels of the attributes within its possible values. Despite the randomness of this strategy, this option does not permit two identical alternatives in one choice set. The last option, labeled 'balanced overlap method', is the middle course between the random design and the (near-)orthogonal design. This option allows for a moderate attribute level overlap within one choice set.

### 4.3 Comparison in terms of the WTP$_b$-error

Table 1 displays the $WTP_b$-error approximated by the average $WTP$-error over the 100 Halton draws from the prior distribution for the Bayesian $WTP$-optimal design and the four benchmark designs. The table shows that the Bayesian $WTP$-optimal design is the most appropriate design to estimate the marginal $WTP$ accurately as it has the smallest $WTP_b$-error. It is followed by the Bayesian $D$-optimal design for which the error is almost 20 % higher. The errors of the other benchmark designs are at least twice as high as the Bayesian $WTP$-optimal design suggesting that these standard designs perform poorly when it comes to estimating the marginal $WTP$. The poor performance of the standard designs is in line with the results reported in [19].

## 4.4 Simulation study

After introducing the evaluation criteria, we elaborate on the findings of a simulation study. First, we focus on the case where the prior distribution used to generate the design is correct. Then, we turn our attention towards a situation where the prior distribution contains wrong information.

### 4.4.1 Criteria for evaluation

Based on simulated observations for all choice sets of the different designs discussed above, we estimated the utility coefficients $\boldsymbol{\beta}$ of the conditional logit model and used these coefficient estimates to calculate the marginal $WTP$ estimates. Comparing these estimates with their true values allowed us to calculate the expected mean squared error

$$EMSE_{WTP}(\hat{\boldsymbol{\beta}}) = \int_{\Re^{M-1}} (\widehat{\boldsymbol{W}}(\hat{\boldsymbol{\beta}}) - \boldsymbol{W}(\boldsymbol{\beta}))'(\widehat{\boldsymbol{W}}(\hat{\boldsymbol{\beta}}) - \boldsymbol{W}(\boldsymbol{\beta}))f(\hat{\boldsymbol{\beta}})d\hat{\boldsymbol{\beta}}, \quad (9)$$

where $f(\hat{\boldsymbol{\beta}})$ represents the distribution of the estimated utility coefficients and $\widehat{\boldsymbol{W}}(\hat{\boldsymbol{\beta}})$ and $\boldsymbol{W}(\boldsymbol{\beta})$ are vectors containing the $M-1$ marginal $WTP$ estimates and the real marginal $WTP$ values, respectively. Note that the $EMSE_{WTP}$ captures the bias and the variability in the marginal $WTP$ estimates. Evidently, a small $EMSE_{WTP}$ is preferred over a large one. In our simulation study, we approximated (9) for a given value of $\boldsymbol{\beta}$ by generating 1,000 data sets for 75 respondents. We computed $EMSE_{WTP}$ values for 75 parameter $\boldsymbol{\beta}$ values drawn from a $6$-dimensional normal distribution. Since the estimated price coefficient enters the marginal $WTP$ computation non-linearly, a poorly estimated price coefficient can result in unrealistic marginal $WTP$ estimates and consequently in unreasonably high values of $EMSE_{WTP}$. Therefore, to get a clear view on the results, we report the natural logarithm of the $EMSE_{WTP}$ values. The problem of unrealistic marginal $WTP$ estimates has already been described by several authors, (among others [27, 26] who also propose alternative solutions).

Additionally, we examine the accuracy of the estimates of the utility coefficients obtained by the different designs. In the same way as $EMSE_{WTP}$ we compute the expected mean squared error of the utility coefficients as

$$EMSE_{\beta}(\hat{\boldsymbol{\beta}}) = \int_{\Re^M} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) f(\hat{\boldsymbol{\beta}}) d\hat{\boldsymbol{\beta}}. \quad (10)$$

A small rather than a large value of $EMSE_{\beta}$ indicates more accurate estimates of the utility coefficients.

#### 4.4.2 Design performance under correct priors

First we computed the $EMSE_{WTP}$ value for each of 75 utility $\boldsymbol{\beta}$ coefficients drawn from the prior distribution used to develop the design which implies that the prior distribution contains correct information. Figure 1 visualizes the 75 $\log(EMSE_{WTP})$ values for each design in box-plots. It is clear that the Bayesian $WTP$-optimal design results in the most accurate marginal $WTP$ estimates. Because of the logarithmic scale used for the $EMSE_{WTP}$ values in Figure 1, the two box-plots of the Bayesian $WTP$- and $D$-optimal design seem not to differ considerably. However, the Bayesian $WTP$-optimal design leads to marginal $WTP$ estimates that are substantially more accurate than the Bayesian $D$-optimal design. The difference in accuracy is even larger between the Bayesian $WTP$-optimal design and the standard designs. Furthermore, the Bayesian $D$-optimal design and the standard designs produce larger and more $EMSE_{WTP}$ outlier values than the Bayesian $WTP$-optimal design.

Outliers were defined as values larger than $Q3 + 6 \cdot IQR$ with $IQR$ the interquartile range and $Q3$ the third quartile. Table 2 shows the average number of outliers among the marginal $WTP$ estimates per utility coefficients vector and three summary statistics related to $EMSE_{WTP}$ with (displayed in parentheses) and without marginal $WTP$ outliers. Even if we exclude the outliers for each design option, the Bayesian $WTP$-optimal design still results in more accurate marginal $WTP$ estimates. In this case, the average $EMSE_{WTP}$ value when using a $D$-optimal design is about 10% higher than the average $EMSE_{WTP}$ value when using a $WTP$-optimal design. The random design exhibits the worst performance yielding an average $EMSE_{WTP}$, which is three times larger than that produced by the $WTP$-optimal design.

Finally, Figure 2 visualizes the 75 $EMSE_{\beta}$ values resulting from estimating the 75 utility coefficients $\boldsymbol{\beta}$. The box-plots clearly indicate that the Bayesian $D$- and $WTP$-optimal designs produce substantially more accurate estimates for the utility coefficients than what is achieved by means of standard designs. We notice that the difference in estimation accuracy of the utility coefficients between the former two designs is only minor. This means that the Bayesian $WTP$-optimal design leads to estimated utility coefficients almost as precise as the Bayesian $D$-optimal design. This suggests the conclusion that the Bayesian $WTP$-optimal designs lead to the most accurate marginal $WTP$ estimates and are not much inferior to Bayesian $D$-optimal designs in terms of estimation accuracy for the utility coefficients.

### 4.4.3 Design performance under incorrect priors

In the previous section, we studied the performance of a Bayesian $WTP$-optimal design assuming that the prior distribution on $\boldsymbol{\beta}$ used to create the design contains correct information on the utility coefficients. In this section, however, we relax this assumption and examine the scenario of incorrectly specified prior parameters. Sensitivity to the use of correct prior assumption was found in previous studies based on Bayesian $D$-optimal designs [5] and it is an issue which warrants further investigation. In a first scenario, the real parameters $\boldsymbol{\beta}$ generating the responses of the 75 respondents come from a 6-dimensional normal distribution with mean [0 0 0 0 0 -0.7] and variance-covariance matrix

$$\begin{pmatrix} \mathbf{I}_{m-1} & \mathbf{0}_{(M-1)\times 1} \\ \mathbf{0}_{1\times(M-1)} & 0.05 \end{pmatrix}. \tag{11}$$

The responses are generated using the same designs as in the previous section. The distribution from which the real parameters are drawn covers–instead–a broad range of preference structures related to the attributes. This implies that the researcher posed incorrect assumptions on the preferences for the attributes, except for the price, at the moment of designing the experiment.

Figure 3 depicts 75 $\log(EMSE_{WTP})$ values for this scenario and shows that, even if we use incorrect prior information at the design stage, the Bayesian $WTP$-optimal design measures the marginal $WTP$ in a more accurate way than the competing designs. We notice that the standard designs have substantially larger extreme values for the marginal $WTP$ than the optimal designs.

Figure 4 visualizes the $EMSE_\beta$ values for the different designs. As previously noted for the case under correct prior information, this figure shows that the Bayesian $WTP$-optimal design estimates the model parameters more precisely than the standard designs and almost as precisely as the Bayesian $D$-optimal design.

A second scenario with incorrect prior was also explored. In this case the real utility coefficients came from a 6-dimensional normal distribution with mean $[-1\ \ 0\ \ -1\ \ 0\ \ -1\ \ -1]$ and variance

$$\begin{pmatrix} 0.25\ \mathbf{I}_{M-1} & \mathbf{0}_{(M-1)\times 1} \\ \mathbf{0}_{1\times(M-1)} & 0.05 \end{pmatrix}. \tag{12}$$

This scenario represents the case where the respondents have a stronger preference for the highest attribute level than was assumed in the prior distribution utilized to construct the optimal designs. The results for this scenario are not displayed here because they are very similar to those for the previous scenario.

In summary, the results obtained from this simulation study clearly suggest that the Bayesian $WTP$-optimal design produces more accurate marginal $WTP$ estimates than any of the other designs, including the Bayesian $D$-optimal one. This increased accuracy is to a large extent robust to the specification of the prior information used to construct the design. Moreover, and this is a novel result, the Bayesian $WTP$-optimal design yields considerably smaller and fewer extreme values for the marginal $WTP$ estimates than those produced by the benchmark designs. This is an important contribution in solving the problem of unrealistically large marginal $WTP$ estimates. The Bayesian $WTP$-optimal design offers the additional advantage that it results in parameter estimates almost as precise as the Bayesian $D$-optimal design, suggesting that $WTP$-efficiency does not come at a large loss in efficiency of utility coefficient estimates.

# 5 The willingness to donate for environmental projects

We illustrate the practical advantages of using $WTP$-optimal designs in an example described in [4] by comparing $WTP$-optimal designs with the design strategy used by the authors in terms of the marginal $WTP$ estimation accuracy.

In [4] a choice experiment was performed to value the willingness to donate for environmental projects. Three attributes were included in the study: the amount of money the respondents received, the donation they gave to an environmental project and the type of environmental project. In the choice experiment the respondents had to make a trade-off between the money they received and the donation they gave to an environmental project. The amount of money the respondent received took three levels: 35 kr, 50 kr and 65 kr ('kr' refers to Swedish Krona, the currency of Sweden where the experiment was performed). There were three values for the donation: 100 kr, 150 kr and 200 kr. The donations were given to one of the following three environmental projects: the rainforest, the Mediterranean Sea or the Baltic Sea which were dummy coded taking the Baltic Sea as a reference.

The experiment consisted of 14 choice sets of size two. There were 35 respondents who participated in the experiment which generated 490 observations in total. In [4] a locally $D$-optimal design was used, based on the information of a pilot study which estimated the marginal willingness to donate for environmental projects equal to five. As the pilot study did not allow the estimation of the utility coefficients of the environmental projects, these were set to zero to generate the design.

As an alternative, we propose a locally $WTP$-optimal design around the point estimate [0.2 1 0 0], which is in accordance with the information coming from the pilot study as reported by the authors. The elements of the vector correspond to the utility coefficients of the money the respondents received, the donation and the environmental projects, respectively. The results for this prior point estimate are representative for other prior point estimates which are omitted for brevity of the paper. Additionally, we developed a Bayesian $WTP$-optimal design. As a prior, we choose a normal distribution with mean [0.2 1 0 0], which reflects the information of the pilot study, and variance-covariance matrix $0.5 \cdot I_4$, where $I_4$ is the four dimensional identity matrix to reflect the uncertainty about the utility coefficients. Other prior distributions taking into account the information of the pilot study lead to similar results. The alternating-sample algorithm, described in [14], was used as the search procedure to find the Bayesian $WTP$-optimal design.

Based on simulated choices generated by the utility coefficients of the original study, shown in Table 3, and the three designs under study, the marginal willingness to donate was computed from the estimated utility coefficients of the conditional logit model. We estimated the marginal $WTP$ for 1,500 data sets. To evaluate the designs we display the 1,500 estimated marginal $WTP$ values resulting from the use of the different designs in Figure 5. Moreover, we compute three evaluation criteria which are the mean squared error ($MSE$), the bias and the relative absolute error ($RAE$), all defined in Table 4. The $MSE$ offers the advantage of capturing the bias and the variability of the estimated marginal $WTP$ values.

The box-plots in Figure 5 clearly show that the use of locally and Bayesian $WTP$-optimal designs results in fewer and smaller outlying estimates for the marginal $WTP$. Table 4 shows that there were even no outliers for the locally $WTP$-optimal design. The Bayesian $WTP$-optimal design reduced the number of outliers to one. This outlier is considerably smaller in size than the ones obtained with the locally $D$-optimal design. In Table 4, we notice that the $MSE$ for the $WTP$-optimal design is substantially smaller than the $MSE$ for the locally $D$-optimal design. We note that the bias of the marginal $WTP$ using the $D$-optimal design is slightly smaller than that for the $WTP$-optimal design. This suggests that the variance of the marginal $WTP$ is smaller for the latter than for the former design type.

There is a substantial difference between the marginal willingness to donate resulting from the pilot study and the final experiment. This implies that the prior information was not correct. We hence examine the scenario in which no information about the unknown utility coefficients is available. In this case, we compare a Bayesian $WTP$-optimal to a Bayesian and a locally $D$-optimal design. To develop the Bayesian designs, we specify the normal distribution with mean $[0\ \ 0\ \ 0\ \ 0]$ and variance $I_4$ as prior, which reflects lack of information about the sign and size of the unknown utility coefficients. The locally $D$-optimal design was developed with point estimate $[0\ \ 0\ \ 0\ \ 0]$ as this could have been a possible design strategy in an original study without any prior information. We display the distribution of the marginal $WTP$ values in Figure 6 and show the values for the three evaluation criteria in Table 5.

The box-plots in Figure 6 show that we avoid more and larger outlier values by using the Bayesian $WTP$-optimal design. However, we notice that the locally $D$-optimal design reduces the number of outliers as well. The reason for this is that the parameter vector $[0\ \ 0\ \ 0\ \ 0]$ used to construct the design coincidentally approximates the real parameter vector. This also explains why the local design performs better than the Bayesian design. However, if we look at Table 5, the $MSE$ values for both $D$-optimal designs are larger than for the $WTP$-optimal design. Since the bias does not differ considerably between the three designs, this suggests that the variance of the estimated marginal $WTP$ values is larger for both $D$-optimal designs than for the $WTP$-optimal one. For the locally $D$-optimal design this can also be seen in Figure 6 by observing that the size of the box is slightly larger and the whiskers of the box-plot are slightly longer.

# 6 Conclusion

Despite the expanding use of conjoint choice experiments to estimate the value of attributes of complex goods, the literature on the design of these experiments has not paid sufficient attention to developing design criteria addressing this specialized use of conjoint choice experiments. In this paper, following [11], we apply the $c$-optimality criterion to create optimal designs for conjoint choice experiments to estimate the marginal $WTP$ accurately and subject these designs to a series of comparisons with other more conventional designs. We use simulation and alternatively assume correct and incorrect prior information about the utility coefficients generating the true responses. The results show that the Bayesian $WTP$-optimal designs consistently produce marginal $WTP$ estimates that are substantially more accurate than those produced by other designs, including the Bayesian $D$-optimal designs, which under correct information were found to dominate more conventional designs in similar comparisons reported in [5]. Our results remain valid even if the prior information is not entirely correct. Importantly, for non-market valuation, the Bayesian $WTP$-optimal designs lead to smaller and fewer extreme values for the marginal $WTP$ estimates. Finally, we note that the advantages afforded by the Bayesian $WTP$-optimal design come at a negligible cost in terms of a loss of efficiency in the utility coefficient estimates when compared to results obtained from a Bayesian $D$-optimal design. The $c$-efficiency criterion would therefore appear to be a potentially valuable criterion in experimental design for conjoint choice experiments undertaken for the purpose of attribute valuation.

# References

[1] G. Baiocchi, *Monte carlo methods in environmental economics*, Applications of simulation methods in environmental and resource economics (R. Scarpa and A. Alberini, eds.), Springer Publisher, Dordrecht, The Netherlands, 2005, pp. 317–340.

[2] S. Banfi, M. Farsi, M. Filippini, and M. Jacob, *Willingness to pay for energy-saving measures in residential buildings*, Energy Economics **30** (2008), no. 2, 503–516.

[3] P. C. Boxall and W. L. Adamowicz, *Understanding heterogeneous preferences in random utility models: a latent class approach*, Environmental and Resource Economics **23** (2002), no. 4, 421–446.

[4] F. Carlsson and P. Martinsson, *Do hypothetical and actual marginal WTP differ in choice experiments?*, Journal of Environmental Economics and Management **41** (2001), 179–192.

[5] S. Ferrini and R. Scarpa, *Designs with a-priori information for nonmarket valuation with choice-experiments: a Monte Carlo study*, Journal of Environmental Economics and Management **53** (2007), no. 3, 342–363.

[6] G. D. Garrod, R. Scarpa, and K. G. Willis, *Estimating the benefits of traffic calming on through routes: A choice experiment approach*, Journal of Transport Economics and Policy **36** (2002), 211–231.

[7] A. S. Goldberger, *A course in econometrics*, Harvard University Press, London, U.K., 1991.

[8] R.R. Hearne and Z.M. Salinas, *The use of choice experiments in the analysis of tourist preferences for ecotourism development in Costa Rica*, Journal of Environmental Management **65** (2002), 153–163.

[9] D.A. Hensher and C. Sullivan, *Willingness to pay for road curviness and road type*, Transportation Research Part D **8** (2003), 139–155.

[10] A. R. Hole, *A comparison of approaches to estimating confidence intervals for willingness to pay measures*, Health Economics **16** (2007), 827–840.

[11] B. J. Kanninen, *Optimal experimental design for double-bounded dichotomous choice contingent valuation*, Land Economics **69** (1993), no. 2, 138–146.

[12] R. Kessels, P. Goos, and M. Vandebroek, *A comparison of criteria to design efficient choice experiments*, Journal of Marketing Research **43** (2006), no. 3, 409–419.

[13] R. Kessels, B. Jones, P. Goos, and M. Vandebroek, *Recommendations on the use of Bayesian optimal designs for choice experiments*, Research Report KBI_0617, Department of Decision Sciences and Information Management, Katholieke Universiteit Leuven, 18pp.

[14] ______, *An efficient algorithm for constructing Bayesian optimal choice designs*, Journal of Business and Economic Statistics (Forthcoming).

[15] S.C. Kimenju, U.B. Morawetz, and H. De Groote, *Comparing contingent valuation methods, choice experiments and experimental auctions in soliciting consumer preference for maize in Western Kenya: preliminary results*, Paper prepared for presentation at the African Econometric Society 10th Annual Conference on Econometric Modeling in Africa, Nairobi, Kenya, 2005.

[16] J. L. Lusk, J. Roosen, and J.A. Fox, *Demand for beef from cattle administered growth hormones or fed genetically modified corn: a comparison of consumers in France, Germany, the U.K. and the U.S.*, American Journal of Agricultural Economics **85** (2003), 16–29.

[17] E.R. Morey and K. Rossmann, *Using stated-preference questions to investigate variations in willingness to pay for preserving marble monuments: Classic heterogeneity, random parameters, and mixture models*, Journal of Cultural Economics **27** (2003), 215–229.

[18] N. Mtimet and L.M. Albisu, *Spanish wine consumer behavior: A choice experiment approach*, Agribusiness **22** (2006), 343–362.

[19] J. M. Rose and R. Scarpa, *Designs efficiency for nonmarket valuation with choice modelling: how to measure it, what to report and why*, Australian Journal of Agricultural and Resource Economics (Forthcoming).

[20] E. Ruto, G. Garrod, and R. Scarpa, *Valuing animal genetic resources: a choice modeling application to indigenous cattle in Kenya*, Agricultural Economics **38** (2008), 89–98.

[21] M. Ryan, *A comparison of stated preference methods for estimating monetary values*, Health Economics **13** (2004), 291–296.

[22] K. Sammer and R. Wüstenhagen, *The influence of eco-labelling on consumer behaviour - results of a discrete choice analysis for washing machines*, Business Strategy and Environment **15** (2006), 185–199.

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

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

[25] R. Scarpa, D. Campbell, and W. G. Hutchinson, *Benefit estimates for landscape improvements: sequential Bayesian design and respondents rationality in a choice experiment study*, Land Economics **83** (2007), no. 4, 617–634.

[26] R. Scarpa, M. Thiene, and K. Train, *Utility in WTP space: a tool to address confounding random scale effects in destination choice to the Alps*, American Journal of Agricultural Economics (Forthcoming).

[27] G. Sonnier, A. Ainslie, and T. Otter, *Heterogeneity distributions of willingness-to-pay in choice models*, Quantitative Marketing and Economics **5** (2007), no. 3, 313–331.

[28] K. Train, *Discrete choice methods with simulation*, Cambridge University Press, New York, 2003.

[29] J. Yu, P. Goos, and M. Vandebroek, *Efficient conjoint choice designs in the presence of respondent heterogeneity*, Marketing Science (Forthcoming).

## Tables and Figures

| *Design type* | *WTP$_b$-error* |
|---|---|
| $WTP$-optimal | 8.136 |
| $D$-optimal | 9.534 |
| Bal. overlap | 16.173 |
| (Near-)Orthogonal | 18.403 |
| Randomized | 27.261 |

**Table 1:** $WTP_b$-errors for the $WTP$-optimal design and the four benchmark designs

| *Simulation statistics* | $WTP$*-opt.* | $D$*-opt.* | *Bal.Overl.* | *(Near-)Orth.* | *Random* |
|---|---|---|---|---|---|
| Average n. of outliers per parameter set $\boldsymbol{\beta}$ | 9.4 | 14.4 | 18.3 | 24.1 | 38.3 |
| Average of $EMSE_{WTP}$ | 0.112<br>(0.139) | 0.125<br>(0.262) | 0.198<br>(0.579) | 0.223<br>(18.268) | 0.321<br>(46.583) |
| Minimum of $EMSE_{WTP}$ | 0.002<br>(0.002) | 0.002<br>(0.002) | 0.003<br>(0.003) | 0.003<br>(0.003) | 0.004<br>(0.004) |
| Maximum of $EMSE_{WTP}$ | 0.745<br>(4.347) | 0.895<br>(90.786) | 1.440<br>(111.700) | 1.736<br>(12893.140) | 2.576<br>(39747.270) |

**Table 2:** The average number of outliers per parameter set $\boldsymbol{\beta}$ and summary statistics of $EMSE_{WTP}$ values with and without outliers over 75 parameter sets $\boldsymbol{\beta}$. Values obtained with outliers are given in parentheses.

| *Variable* | *Coefficient* | *St. error* |
|---|---|---|
| Money | 0.033 | (0.010) |
| Donation | 0.021 | (0.003) |
| Mediterranean | –0.885 | (0.148) |
| Rainforest | –0.088 | (0.145) |
| Marginal $WTP$ Donation | 0.636 | |

**Table 3:** Estimated utility coefficients and standard error (in parentheses) of the attributes in the original study

| *Simulation statistics* | | *Locally* *D-optimal* | *Locally* *WTP-optimal* | *Bayesian* *WTP-optimal* |
|---|---|---|---|---|
| $MSE$: | $\frac{1}{R}\sum_{r=1}^{R}(\widehat{W}_r - W)^2$ | 0.044 | 0.033 | 0.029 |
| $Bias$: | $\frac{1}{R}\sum_{r=1}^{R} \mid \widehat{W}_r - W \mid$ | 0.114 | 0.127 | 0.120 |
| $RAE$: | $\frac{1}{R}\sum_{r=1}^{R} \frac{\lvert\widehat{W}_r - W\rvert}{W}$ | 20.009 | 17.945 | 18.805 |
| Number of outliers | | 8 | 0 | 1 |

**Table 4:** Evaluation criteria for a locally $D$-optimal design and locally and Bayesian $WTP$-optimal designs using the information of a pilot study based on $R$=1500 data sets

| *Simulation statistics* | | *Bayesian* *D-optimal* | *Locally* *D-optimal* | *Bayesian* *WTP-optimal* |
|---|---|---|---|---|
| $MSE$: | $\frac{1}{R}\sum_{r=1}^{R}(\widehat{W}_r - W)^2$ | 0.019 | 0.024 | 0.017 |
| $Bias$: | $\frac{1}{R}\sum_{r=1}^{R} \mid \widehat{W}_r - W \mid$ | 0.093 | 0.116 | 0.095 |
| $RAE$: | $\frac{1}{R}\sum_{r=1}^{R} \frac{\lvert\widehat{W}_r - W\rvert}{W}$ | 14.668 | 18.161 | 14.940 |
| Number of outliers | | 2 | 0 | 0 |

**Table 5:** Evaluation criteria for a Bayesian $WTP$-optimal and Bayesian and locally $D$-optimal designs in the absence of prior information from a pilot study

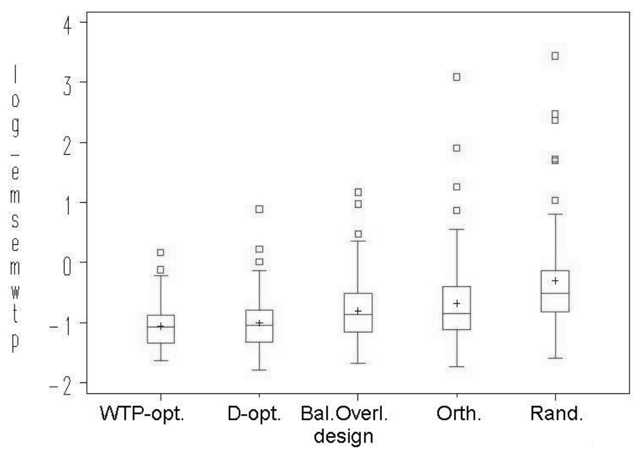


**Figure 1:** Log$(EMSE_{WTP})$ values for the different designs assuming a correct prior distribution

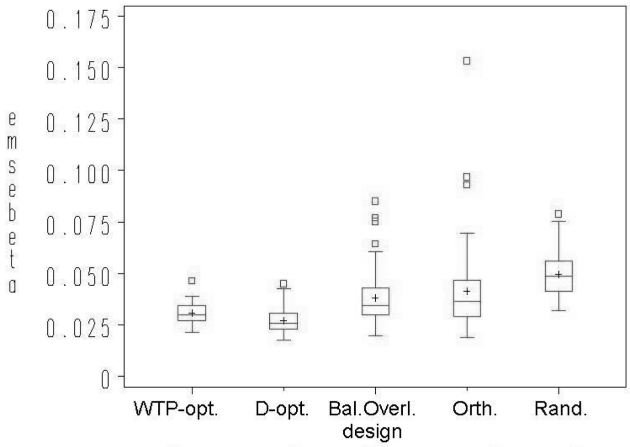


**Figure 2:** $EMSE_{\beta}$ values for the different designs assuming correct prior parameters

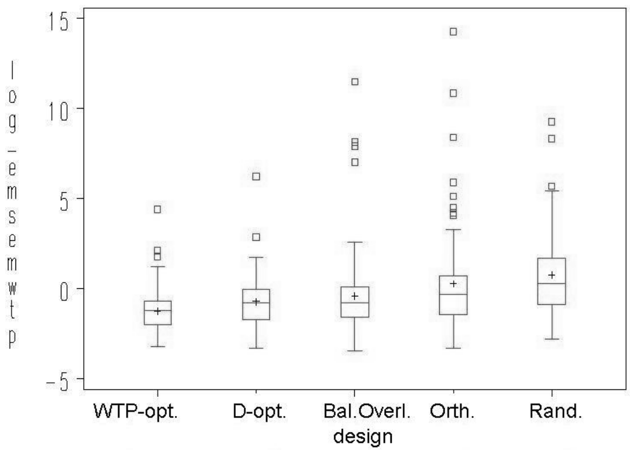


**Figure 3:** Log($EMSE_{WTP}$) values for the different designs posing incorrect assumptions on the preference for the attributes, except for the price

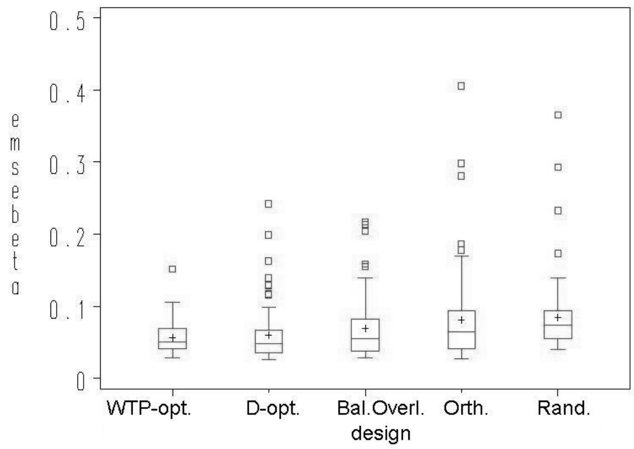


**Figure 4:** $EMSE_\beta$ values for the different designs assuming indifference of individuals for the attributes

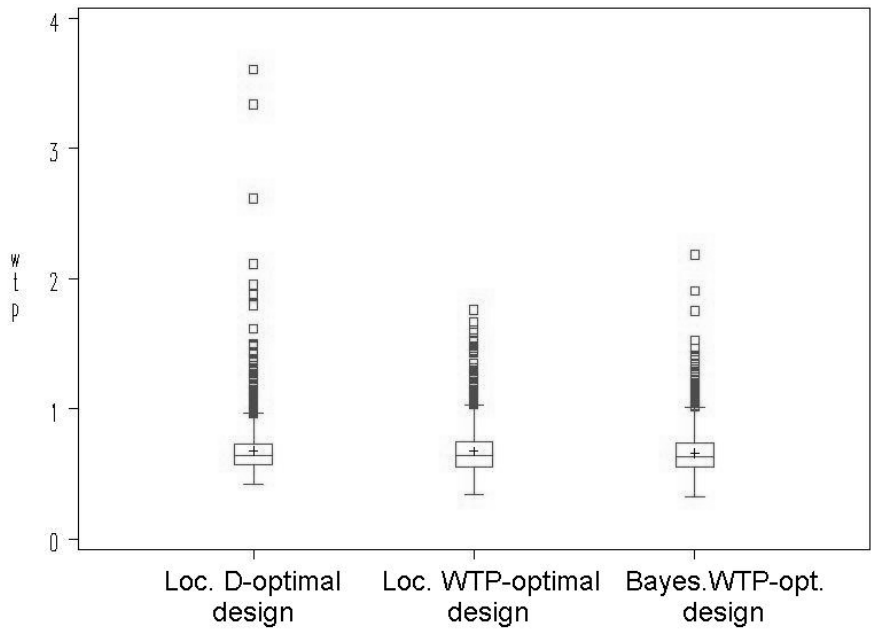


**Figure 5:** Marginal $WTP$ estimates from a locally $D$-optimal, locally $WTP$-optimal and a Bayesian $WTP$-optimal design using the prior information of the pilot study for 1500 simulated data sets

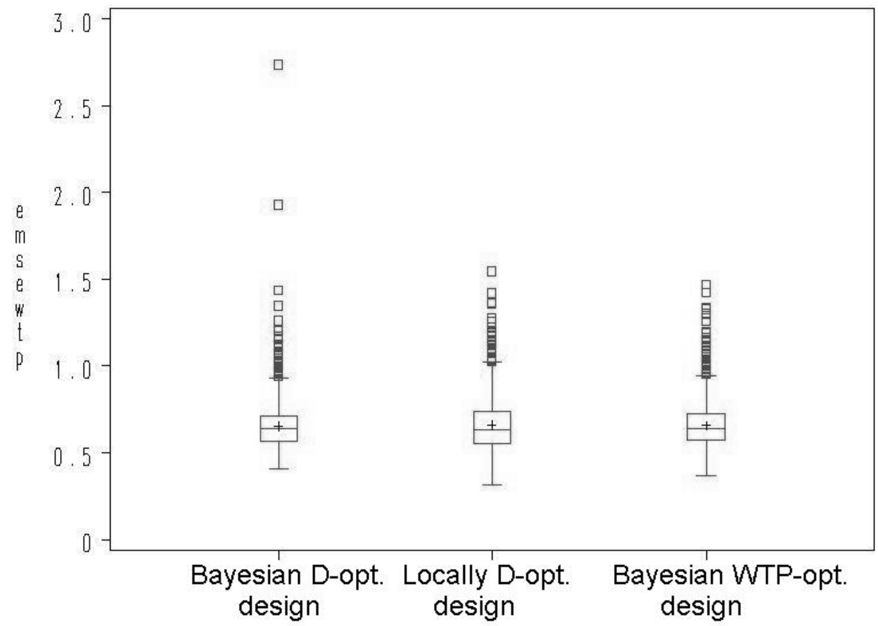


**Figure 6:** Marginal $WTP$ estimates from a Bayesian $WTP$-optimal, a Bayesian $D$-optimal and a locally $D$-optimal design in the absence of prior information from a pilot study for 1500 simulated data sets