Using Choice Experiments to Estimate Consumer Valuation: the Role of Experimental Design and Attribute Information Loads

Zhifeng Gao[a*], Lisa O. House[a] and Xiaohua Yu[b]

[a] Food and Resource Economics Department, P.O. Box 110240, University of Florida, Gainesville, FL 32611, USA

[b] Courant Research Center "Poverty, Equity and Growth." Platz der Göttinger Sieben 3, Goettingen, 37073, Germany

Abstract

With fixed dimensionality of choice experiments, previous simulation results show that D-optimal design with correct a priori information generates more accurate valuation. In the absence of a priori information, random designs and designs incorporate attribute interactions result in more precise valuation estimates. In this paper, Monte Carlo simulations demonstrate that the performances of different design strategies are affected by attribute information loads in choice experiments. Consumer valuation estimates in simulation settings vary with the number of attributes.

JEL classification: C9, D63

Keywords: Choice Experiments, Information Loads, Valuation

Forthcoming in ***Agricultural Economics***.

Acknowledgement: The authors gratefully acknowledge helpful comments from Ted C. Schroeder, two anonymous reviewers, and the editors.

Using Choice Experiments to Estimate Consumer Valuation: the Role of Experimental Design and Attribute Information Loads

Introduction

In the last 10 years, choice experiments (CEs) have been widely used to elicit consumer valuation of nonmarket goods and marketable goods with novel attributes. A simple search in Econlit with the key words “choice experiments” and “valuation” results in 211 studies in total, and in 2008 alone, there were 40 studies, about 4 times of that in 2000. The increasing popularity of CEs is partially in response to recognized problems with contingent valuation by a NOAA panel in 1990 (Hausman 1993), as well as the ability to easily identify the trade-offs among different product attributes relative to other approaches. In addition, CEs are consistent with Lancaster’s theory of utility maximization (Lancaster 1972). Further, continuous development in discrete choice modeling, such as multinomial logit models, generalized extreme value models, and mixed logit models, among others, has led to convenient availability in commercial software (e.g., Limdep and Stata). This development provides researchers with a set of powerful tools to study consumer choices corresponding to different assumptions of consumer preferences.

However, despite the rapid growth in CE research, design of CEs remains one of the most challenging issues. For instance, in a simple CE that has two alternatives, with each alternative having four two-level attributes, the total combination of attribute levels or choice sets is 256 (2^4x2^4). In most cases it is not feasible to ask respondents to make 256 continuous choice decisions, and this is a very simplistic example. The number of choices that respondents face grows exponentially as the design of the CE becomes more complicated, and designs often become more complicated to reflect real-life choices. One solution may be randomly choosing 8, 16 or 20 choice sets from the 256 sets for respondents to choose from, but this might deteriorate the statistical efficiency of data analysis. Therefore, the major challenge is to design statistically efficient experiments to provide enough information for accurately eliciting consumer preferences, and at the same time, make the length of choice experiments reasonable, such that cognitive burdens on survey participants are minimized.

Previous research has contributed to the development of design strategy to more precisely elicit consumer valuation. In the contingent valuation framework, Kanninen (1993a) derives the optimal experimental designs for double-bounded dichotomous choice models based on D-optimal, C-optimal and Fiducial Method criteria[1]. Kanninen (1993b) further develops the sequential C-optimal approach to collect a priori information on true parameters required to generate optimal designs. D-optimal designs for multinomial CEs are derived with the assumption that all attributes are quantitative variables. This allows the use of the numerical optimization to search over all possible choice designs without restrictions. Results show this approach generates better designs for main-effects, multinomial CEs than traditional designs such as main-effects design and shift design (cyclical design) (Kanninen 2002). Most recently, Louviere et al. (2008a) developed an approach to collect additional information using a rank-order explosion method to estimate models at individual levels. Other researchers (e.g. Scarpa and Rose 2008) developed CE designs that minimize the C-error in CEs as a function of willingness-to-pay (WTP), in comparison with the traditional designs that minimize D-error and A-error which are both functions of parameters in the consumer utility function.

[1] D-optimal design is based on the parameters; C-optimal design and Fiducial Method design are based on the WTP.

However, despite of the recent development in CE designs, no general agreement has been reached on what the "best design" of CEs is. Chrzan and Orme (2000) argue that each design approach has its advantages in capturing certain types of effects and there is no superior design for all purposes. For instance, the D-optimal and C-optimal (by minimizing C-error) designs proposed by Kanninen (2002) and Scarpa and Rose (2008) require the knowledge of true parameters of consumer preferences. Although previous study can be used as a source for the determination of true consumer preferences, this information is not always readily available. Using Monte Carlo simulation, results from studies demonstrate D-optimal design with poor quality a priori information is less promising than shift design without any a priori information (Ferrini and Scarpa 2007). With correct a priori information, D-optimal design generates a more accurate valuation of products or services (Carlsson and Martinsson 2003; Ferrini and Scarpa 2007). At present, the designs that don't require a priori information proposed by Louviere, Hensher and Swait (2000) are still the most popular in empirical studies because of their simplicity and availability in commercial software. In a comparison of six designs that assume no a priori information, Lusk and Norwood (2005) demonstrate that random designs perform best in WTP estimates.

Existing literature using simulations to study CE design performance compares different design strategies with fixed design dimensionality (e.g., number of alternatives, number of attributes and number of attribute levels). For instance, Carlsson and Martinsson (2003) construct pair-wise CEs with 4 attributes; Ferrini and Scarpa (2007) conduct a simulation based on pair-wise CEs with 4 attributes, each having 3 levels; and Scarpa and Rose (2008) demonstrate their study based on CEs containing three alternatives with four attributes. With recent field studies showing that design dimensionality, especially the number of product attributes in CEs, affects consumer preference and valuation (Hensher 2006, Islam, Louviere, and Burke 2007, Rose et al. 2008; Louviere et al. 2008b; Gao and Schroeder 2009), it is legitimate and also important to seek answers on the impacts of attribute information on WTP estimates in simulation scenarios and investigate the performance of different CE designs under various attribute information loads. This is because the lack of research on performance of different CE designs under various attribute information hinders our ability to infer efficiency of different CE design strategies. In addition, determining the influence of attribute information loads on WTP estimates will help to determine whether the impacts of additional information found in field studies are a result of the changes in respondent cognitive ability or just the statistical properties of the CEs when the number of attributes varies. The identification of the sources of the impacts of information loads on consumer is useful for better explanations of respondent changing behaviors in CEs.

The purpose of this article is to determine the performance of different CE designs on welfare estimates such as WTP under different attribute information loads. A secondary objective is to investigate the effects of information loads on consumer valuation in simulation scenarios. We will compare the most widely used CE designs in empirical works, with the assumption that no a priori information on consumer preference is available. The reason for this is the unavailability of high quality information in most cases and the fact that low quality a priori information will produce less promising results. We believe our research will provide valuable information to researchers focusing in the application of CEs by helping them to choose more appropriate design strategies based on the attribute information determined before implementation of CEs. In addition, identifying the impacts of dimensionality on the statistical properties of CEs may help identify the true source of impacts of information loads on consumer preferences and enhance our understanding of the effects of dimensionality on respondent valuation.

Choice Experiments and Experimental Design

In a CE, predetermined attributes that are believed to have the largest impacts on consumer choice decisions comprise a series of alternatives (profiles or choices). Two or more alternatives are used to form a choice set, and a sequence of choice sets composes a CE. Respondents are asked to choose one alternative from each choice set in a CE. Based on random utility theory, consumers will choose an alternative from each choice set to maximize her/his utility. Consumer preferences for products and product attributes can be elicited from their sequential choice decisions. However, in most cases, enumerating all combinations of product attributes is not feasible because the number of combinations of product attributes is grows exponentially, resulting in very large number of choice sets. Too many choice sets may hinder the consumer's ability to make rational choice decisions in a short time. Therefore, one of the major challenges of conducting a CE is to design experiments that are simple enough for respondents to make efficient choice decisions, while at the same time providing enough information for researchers to accurately elicit consumer preferences.

Various CE designs have been discussed and used by researchers, including orthogonal main effects design, D-optimal design, C-optimal design, fractional factorial design and shifted (or cycled) design, etc. There is a general agreement that CEs generated from the design should result in the minimum variance of coefficient estimates. In a linear model, the variance of coefficient estimates does not depend on the true parameters, such that $Var(\beta) = \sigma^2 (X'X)^{-1}$, where X is the design matrix and σ^2 is the variance of the random error in the linear model. Therefore, the D-optimal design which maximizes the D-efficiency $100 \cdot \frac{1}{N \mid (X'X)^{-1} \mid^{1/A}}$, should be the best design in linear models, where N is the number of observations in the design, and A is the number of attributes. Balanced orthogonal designs automatically result in designs with 100% D-efficiency because $(X'X)^{-1} = \frac{1}{N} \cdot I$ where I is an identity matrix. A design that deviates from orthogonal design will have a D-efficiency less than 100%, with the minimum efficiency being zero (Kuhfeld, Tobias and Garrat 1994). The procedures that can generate balanced orthogonal design, D-optimal design, etc. for linear models are readily available in SAS and other software which make the design of experiments of linear models simple. However, good designs based on the standards in a linear model may not hold for CEs. This is because the models used in CEs are usually nonlinear, and the variance of parameter estimates not only depends on the design matrix, but also on the true parameters in the models.

The variance matrix of the parameter estimates from CEs depends on the assumption of the random component in consumer random utility function $U_j = \beta' x_j + \varepsilon_j$, where U_j is the consumer utility of consuming product j, β is a vector of parameters, x_j is a vector of attributes for alternative j and ε_j is the random component with a certain type of probability distribution. Consumers will choose an alternative from a choice set to maximize her/his utility and the probability of choosing alternative j is $P_j = \text{Prob}(\beta' x_j + \varepsilon_j > \beta' x_i + \varepsilon_i,\ \forall i \neq j)$. In a multinomial logit model (MNL), a prevailing model used to estimate consumer preferences in CEs, ε_j has identical independent Gumbel distribution with constant variances of $\pi^2 / 6$. The probability that a consumer chooses alternative j from a choice set with n choices is $P_j = \frac{e^{\beta' x_j}}{\sum_{i=1}^{n} e^{\beta' x_i}}$. The maximum likelihood estimator is consistent and asymptotically normal distributed with mean β , and the asymptotic covariance matrix $\Omega = (Z'PZ)^{-1} = [\sum_{n=1}^{M}\sum_{j=1}^{n} z_{jn}^{'} P_{in} z_{jn}]^{-1}$, where $z_{jn} = x_{jn} - \sum_{i=1}^{n} x_{in} P_{in}$, and M is the number of choice sets for all respondents (McFadden 1974). The formula of the variance of the parameters in MNL implies that it is not possible to choose a design strategy that minimizes the variance of parameter estimates without knowing the true parameters in the consumer utility function. In this case, simulation studies have been used to examine the performance of different designs in parameter and WTP estimates (Carlsson and Martinsson 2003; Ferrini and Scarpa 2007; Lusk and Norwood 2005).

Kanninen (1993b, 2002) has demonstrated a sequential procedure to gain information on the parameter values in the consumer utility function, which can be used to design D-optimal or C-optimal CEs. However, in empirical studies, the D-optimal and C-optimal designs with a priori information have not been used often. This is because obtaining high-quality prior information of the true parameters of utility function may be difficult in most cases and designs maximizing D-efficiency in MNL with low quality a priori information usually lead to inferior estimations compared to those designs without a priori information (Ferrini and Scarpa 2007). Therefore, in this article we design CEs assuming no a priori information on consumer preference is available, which is the most common case in empirical studies. Particularly, we extend the work of Lusk and Norwood (2005) by investigating the efficiency of different CE designs regarding consumer valuation estimates under various information loads.

Comparison of Different Experimental Designs under Various Information Loads

Monte Carlo Simulation

To investigate the performance of different CE designs under various attribute information loads, we conducted several Monte Carlo simulations where the true consumer utility functions are known, such as $U_j = \beta' x_j = V_j$. Pair-wise CEs are designed without a priori information on consumer utility functions. In the pair-wise CEs, respondents are assumed to choose one alternative from two choices in a choice set. For each alternative in the CEs, the utility V_i is calculated using the presumed true parameters β and the attribute levels x_j. The random variable, ε_i, following the Gumbel distribution, is independently drawn with the number of observations equal the total number of alternatives in a simulation (2 × the number of choice sets in a choice experiment × the sample size). As a result, the assumptions of multinomial logit models are strictly satisfied. The random variable ε_i, which simulates unobservable consumer preferences, is added to V_j. As such, consumer random utilities $U_i = V_i + \varepsilon_i$ are compared across alternatives in a choice set, and the alternatives with the highest random utility level are assigned a number of one to simulate consumer choices. MNL models are estimated using simulated consumer choices and the attribute levels in CEs. The WTP estimate for each attribute is calculated as the negative value of the ratio of attribute and price coefficients, $WTP_k = -\frac{\beta_k}{\beta_0}$, where *k* indicates the *kth* attribute of alternatives in CEs and *0* indicates price. The procedure is repeated 500 times to compose a Monte Carlo simulation, resulting in 500 WTP for each attribute, for each simulation scenario. The performance of different CE designs with respect to the evaluation of attributes is compared using mean squared error: $MSE_k = \frac{1}{N}\sum_{i=1}^{N}(WTP_{ki} - WTP_{kt})^2$, and mean absolute relative error: $RSE_k = \frac{1}{N}\sum_{i=1}^{N}\left|(WTP_{ki} - WTP_{kt}) / WTP_{kt}\right|$, where WTP_{ki} is the simulated WTP of *kth* attribute, WTP_{kt} is the corresponding true WTP, and N is the number of simulations. Because consumer true WTP for different attributes varies in the simulation, the RSE_k provides a relative measure such that the impact of different CE designs can be compared across attributes. The CE design that results in the minimum MSE_k and RSE_k is considered the best design for estimating WTP for *kth* attribute. In addition, because the best design for different attributes may differ, the average mean square error and average mean absolute relative error are calculated as $MSE = \frac{1}{M}\sum_{k=1}^{M} MSE_k$ and $RSE = \frac{1}{M}\sum_{k=1}^{M} RSE_k$, respectively, where k indicates the *kth* attribute and M indicates the number of attributes in a CE. The *MSE* and *RSE* help determine the overall best design when there are multiple attributes in a CE. In addition, the measure of D-error is calculated as a function of the determinant of variance of parameter estimates $DEF = \frac{1}{N}\sum_{i=1}^{N} Det(\Omega(\beta^i))^{1/A}$, where β^i is the parameter estimates in the *ith* simulation. D-error can be used to evaluate the efficiency of different designs in parameter estimates[2].

Comparison Scenarios

We employ four types of CE designs and evaluate their performances under four attribute information loads, three levels of sample size and two types of utility functions (table 1). Therefore the total number of simulation scenarios is 4×4×3×2. The CE designs include randomly drawn choice sets from $3^N \times 3^N$ full factorial design (RD), where N indicates the number of attributes in a CE, 3 is the level of each attribute; main effects only design drawn from $3^N \times 3^N$ full factorial design (ME); main effects design maximizing D-efficiency with a minimum

[2] Other criteria to evaluate the performance of different CE designs have been extensively discussed in Scarpa and Rose (2008).

number of choice sets drawn from $3^N \times 3^N$ full factorial design, without a priori information (MD); and the design pairing alternatives generated from 3^N full factorial design using the bin method (RP). We excluded designs which incorporate interactions between attributes because consumer utility is assumed to be linear in attributes in most empirical studies. In addition, adding two-way interactions significantly increases the choice sets in a CE, especially when the attribute information loads is high, thus may result in heavy cognitive burden on survey respondents. Smaller main effects only designs may be preferable even if larger designs with interactions have statistical advantages (Lusk and Norwood 2005). The number of attributes in CEs, each with three levels, is 2, 3, 4 and 5, implying attribute information loads from low to high. The minimum number of attributes is two, because *price* and another product attribute must be included to calculate the WTP estimates. The maximum number of attributes is five, because larger number of attributes typically leads to CEs with more choice sets, which are difficult to administer and may quickly result in respondents' fatigue or information overload. Similar to Lusk and Norwood (2005) both continuous and discrete functions are assumed to be true consumer utility functions in simulations. The sample size measured by the total number of choice sets for all respondents are selected to be 250, 500 and 1000, representing the surveys with small, middle and large sample sizes, respectively. The description of different simulation scenarios is provided in Table 1.

The true consumer utility functions employed in our study are:

(1) $V_{ij}^{c} = \alpha_j + \beta_0 \cdot P + \sum_{k=1}^{n} \beta_k \cdot x_{ijk}$ for continuous utility functions and

(2) $V_{ij}^{d} = \alpha_j + \beta_0 \cdot P + \sum_{k=1}^{n} \beta_{ka} \cdot x_{ijka} + \sum_{k=1}^{n} \beta_{kb} \cdot x_{ijkb}$ for discrete utility functions. Where α_j is the alternative specified constant, β s are parameters to be estimated, P is product price and n equals 1, 2, 3 or 4 denoting the number of attributes in a CE. In equation (1), attributes such as x_{ijk} enter the utility function are specified as continuous variables, whereas in equation (2) attributes enter the utility function as discreet variables, denoted using dummy variables such as x_{ijka} and x_{ijkb}[3]. The true parameters in continuous utility functions are $\alpha_j = 1$, $\beta_0 = -1$, $\beta_1 = 2$, $\beta_2 = 3$, $\beta_3 = 4$ and $\beta_4 = 0.1$. The true parameter of dummy variables in discrete utility functions are $\beta_{1a} = 2$, $\beta_{1b} = 1$, $\beta_{2a} = 3$, $\beta_{2b} = 2$, $\beta_{3a} = 4$, $\beta_{3b} = 3$, $\beta_{4a} = 5$, $\beta_{4b} = 4$.

With the results from different simulation scenarios, the *MSE* and *RSE* can be compared 1) across sample size to study the effect of sample size, 2) across different designs to investigate the performance of different designs with small, mid and large sample sizes, and 3) across the number of attributes to investigate the effect of information loads on WTP estimates.

Identification of Factors Affecting WTP Estimate

Pair-wise comparisons across the sample size, design strategy and number of attributes can help identify the impact of a single factor on WTP estimates. However, the effects of different factors may be compounded so that the pair-wise comparisons are not sufficient. Three simple models are estimated as:

(3) $WTP_i = \alpha + \beta_1 \cdot Des_i + \delta_1 \cdot Sam_i + \lambda_1 \cdot Att_i + \kappa_1 \cdot Att_i^2 + \varepsilon_i$

(4) $SE_i = \alpha + \beta_2 \cdot Des_i + \delta_2 \cdot Sam_i + \lambda_2 \cdot Att_i + \kappa_2 \cdot Att_i^2 + \varepsilon_i$ and

(5) $RE_i = \alpha + \beta_3 \cdot Des_i + \delta_3 \cdot Sam_i + \lambda_3 \cdot Att_i + \kappa_3 \cdot Att_i^2 + \varepsilon_i$

[3] Because each attribute has three levels, two dummy variables are created for each attribute.

Where WTP_i is the willingness-to-pay estimate, $SE_i = (WTP_i - WTP_t)^2$ is the squared difference between the estimated WTP and the true WTP, $RE_i = |(WTP_i - WTP_t)/WTP_t|$ is the relative difference between the estimated WTP and the true WTP, $Des_i = [ME_i \; MD_i \; RP_i]'$ is a vector of dummies denoting the design strategy, $Sam_i = [M_i \; L_i]'$ is a vector of dummies denoting the sample size, Att_i is the number of attributes in CEs. αs, βs, δs, λs and κs are coefficients to be estimated, and ε_i represents the stochastic errors. A quadratic term of Att_i is added in the models because Swait and Adamowicz (2001) demonstrated that the consumer preference variance has a quadratic relationship with the complexity of the decision environment. Gao and Schroeder (2009) have shown that consumer WTP has a quadratic form with the number of attributes in CEs. However, it is not clear whether the relationship is from changing consumer preference as a result of the complexity of decision making, or a statistical property of CEs with the increasing number of attributes in designs. Models (3), (4) and (5) are estimated for both continuous and discrete utility functions.

Results

WTP Estimates and MSE in Various Scenarios with Continuous Utility Function

Table 2 reports the mean WTP_k and MSE_k[4] of each attribute and *MSE* of different designs in different simulation scenarios when the true utility functions are continuous. Results show that most WTP estimates are significantly different from the true values at the 5% level when the sample sizes are small. However, in some cases, larger sample size also results in WTP

[4] RSE_k was also calculated. However, our study shows that this measure provides similar information as MSE_k, they are not reported in this article.

estimates that are significantly different from the true values. For instance, with the RD designs, when the sample size is large, the WTP estimates of all attributes become significantly different from the true values in the CE with four attributes. The fact that those significant WTP estimates are accompanied with smaller MSEs indicates that the insignificance of WTP estimates with smaller sample sizes may be the result of larger variance of WTP estimates. The significance of WTP estimates is also related to the number of attributes in CEs. For instance, when the number of attributes in the CEs is four, the WTP estimates are more likely significantly different from the true WTP than when the number of attributes is lower. In addition, the WTP estimates may have nonlinear relationships with the number of attributes in the CEs. This phenomena implies that the changing WTP with different attribute information loads in field studies demonstrated by previous research may be the result of the statistical property of CE design rather that consumer response to new information.

For all CE designs, the *MSEi* of individual attribute i, and average *MSE* of all attributes in the CEs decrease with the increase in sample size, implying WTP estimates are improved from using larger sample sizes. The impacts of attribute information on the performance of CEs are not clear. In some cases, with a larger number of attributes in the CEs, the *MSEi* was larger (e.g. attribute X1 with RD design and small sample size), however, in other cases with a larger number of attributes, the resulting *MSEi* was smaller (e.g. attribute X2 with MD design and larger sample size). These results imply the best strategy to improve WTP estimation may not be keeping the number of attributes at the minimum level, at least from a statistical perspective. The relationship between the *MSEi*, *MSE* and CE designs is not clear. It may depend on the attribute information and sample size in the CEs. For instance, when the number of attributes is two, ME design results in the smallest *MSE* of attribute X1 at all sample size levels. However, when the number of attributes increases to three, RD design has the smallest *MSE* for attribute X1 at all sample size levels. These results indicate that comparing the performance of different designs with a fixed number of attributes, which is a common practice in previous studies, may not provide enough information to evaluate the efficiency of different design strategies.

WTP Estimates and MSE in Various Scenarios with Discrete Utility Function

The mean WTP_k and MSE_k[5] of each attribute and *MSE* of different designs in different simulation scenarios when the true utility functions are discrete are reported in Table 3. Results show that most WTP estimates are not significantly different from the true values at the 5% level. Most of the significant WTP estimates are in CEs with small sizes and where the number of attributes in the CEs is at its highest level, five. With middle and larger sample sizes, all except one WTP estimate (WTP estimate of X1a with MD design and 4 attributes), are not significant from the true WTP. Similar to that with continuous utility function, the WTP estimates may have nonlinear relationship with the number of attributes in the CEs.

Increased sample sizes result in smaller *MSEi* and *MSE*, implying WTP estimates are improved with larger sample sizes in the case of the discrete utility function. The number of attributes in CEs may have nonlinear relationships with the *MSEs*. For instance, the *MSE* for attribute X1 in the RP design is the smallest when the number of attributes in the CEs is three. However, in other designs, such as the RD design, the *MSE* for attribute X1 increases as the number of attributes in the CEs increases, implying worse performance of the CE designs with a

---

[5] With a discrete utility function, there are two WTP estimates for each attribute, corresponding to that in the continuous utility function. We only report the one WTP estimate for each attribute, because the second WTP estimate delivers similar information regarding the impact of attribute information, sample size and design strategy. RSE_k was also calculated, as that in the continuous utility function; but they are not reported in this article.

larger number of attributes. These results indicate that there are no dominant designs with regards to information loads in CEs, thus the right design should be chosen based on the number attribute determined by researchers.

Best Designs in Various Scenarios

The best CE designs under different information loads and sample sizes are shown in Table 4. For each individual attribute, the best designs are selected based on the minimum *MSEi* and *RSEi*. The best overall designs are also selected based on the minimum average *MSE* and *RSE* rules. In general, both rules result in the same best design. However, in some cases, the best designs differ between *MSE* and *RSE* rules. This may be a result of the fact that the *MSE* punishes extreme values in the sample more severely than the *RSE*, because it is based on the squared difference between WTP estimates and the true WTP compared to RSE which uses the absolute value of the ratio of deviation of WTP estimates from the true WTP and the true WTP. In addition, the best designs based on minimum *DEF* are also reported in Table 4, helping us to compare different designs based on WTP estimates and parameter estimates.

Results in table 4 show that except for two cases (when the number of attributes in CEs is 5, with a large sample size and continuous utility function; and with a small sample size and discrete utility function) the best designs based on minimum *MSE and RSE* are the same. The best designs depend on the number of attributes in CEs. For instance, when the number of attributes is two, ME design is the best design with the continuous utility function, while RD design is the best for attributes X1 when the number of attributes is three. The best designs also depend on the sample size. For example, with the continuous utility function, RP design is the best design for attributes X1 when the sample size is small and large, while RD design is the best when the sample size is middle. The best designs are significantly different between continuous and discrete utility functions. For instance, when the number of attributes in the CEs is two, the best design is ME for continuous utility functions, but MD for discrete utility functions. However, when the number of attributes in the CEs is five, the best design for both continuous and discrete utility functions tends to be the same – RD design is the best design. There is no clear relationship between the best designs based on *MSE* and *DEF* measures. In some cases, *MSE* and *DEF* result in the same best designs (e.g. when the number of attributes is two), and in other cases, they result in different best designs (e.g. when the number of attributes is three). Because the *DEF* is based on the parameter estimates, and *MSE* is based on WTP estimates, a function of parameters of attributes and price, it is reasonable to have the inconsistence between the best designs resulted from *MSE* and *DEF* measures. This inconsistency also indicates that the best design based on maximum D-efficiency may not be the best design focusing on WTP estimates. In this case, the C-efficiency developed by Scarpa and Rose (2008) should be used to search the best CEs if high quality a priori information on consumer preferences is available.

Factors Affecting WTP Estimates

Results of model (3) for continuous utility functions are reported in table 5. Most of the coefficients are significantly different from zero at the 5% level. The positive coefficients of design strategy indicated that RD, ME and RP designs tend to result in higher WTP estimates for attributes X1, X2 and X3, compared with the MD design. The statistical tests demonstrate that for attributes X1, X2 and X3, ME design results in statistically significant larger WTP estimates than RD and RP design at the 5% level. However, for the WTP for X4, those 4 design strategies are equivalent with regard to the WTP estimate. One potential reason for this phenomenon may be that the true WTP for attribute X4 has a much smaller scale compared with the WTP for other attributes (0.1 vs. 2, 3 and 4). Another reason may be that X4 is only presented in the CEs with five product attributes, and performance of design strategies are significantly affected by the number of attributes in the CEs. Additional models for attributes X1, X2, and X3 with the WTP estimates from CEs with 5 attributes also show that all the three design strategies (RD, ME, and RP) are equivalent.

For attributes X1 and X4, the significant negative coefficients of the sample size imply that middle and large sample sizes result in smaller WTP estimates. The insignificant coefficients of Middle for attributes X2 and X3 indicate that a larger sample size does not have significant impacts on WTP estimates. The impacts of attribute information on WTP estimates depend on the particular attribute of interest. For instance, the insignificance of coefficients of # of Attributes and # of Attributes Squared for attribute X1, indicate that attribute information does not have significant impacts on the WTP estimates. However, those coefficients are statistically significant from zero at the 10% level for attribute X2, indicating a quadratic relationship between the number of attributes and the WTP estimates for X2.

Results of model (4) and (5) for continuous utility functions (reported in table 5) show that RD, ME and RP designs result in significant lower squared difference between true and estimated WTP (*SE*) and relative difference between true and estimated WTP (*RE*) than MD design, indicating the overall better performance of RD, ME and RP design in WTP estimates. The statistical tests of the equivalence of different designs show that RD and RP result in significantly smaller *SE* and *RE*, indicating the overall better performance of RD and RP designs. RD and RP designs are not statistically significantly different based on *SE* measures, however, based on the *RE* measure, the RD design is better with regards to the WTP estimates of X1. The significant positive coefficient of # of Attributes for attribute X1 indicates that increasing the number of attributes in the CE significantly reduces the performance of CEs in WTP estimates of attributes X1 and X2. However, the significant negative coefficient of # of Attributes Squared (in the equation of *SE* for Xi) for attribute X2 implies the quadratic relationship between the performance of CEs and the number of attributes in CEs. The *SE* increases with an increasing number of attributes and then decreases after the number of attributes reaches a certain level. The results of model (5) indicates that for both attributes X1 and X2, there is a quadratic relationship between the performance of CEs and the number of attributes in CEs, if the divergence between the true and estimated WTP is measured in relative measures. The significant negative coefficients of Middle and Large indicate that increasing sample size for CEs significantly improves the performance of CEs with regards to WTP estimates.

Results of model (3) (table 6) for a discrete consumer utility function provide similar information compared with the results for the continuous utility function. First, RD, ME and RP result in statistically significant (5% level) and larger WTP estimates for all attributes than the MD design. Middle and large sample sizes do not result in statistically significant WTP estimates compared to small sample sizes. The number of attributes in CEs does not have a significant impact on WTP estimates for attribute X1a, X3a and X4a, however, it has a quadratic relationship with the WTP estimates for attribute X2a at the10% level.

Results of models (4) and (5) for discrete utility functions (table 6) imply that RD, ME and RP design significantly improve the WTP estimates with smaller *SE* and *RE* than MD design. Based on the scale of the absolute value of coefficient estimates of design strategy, RD design has the best overall performance, followed by the RP design, which is the same result found in the case of continuous utility functions. For attributes X1a and X2a, the *SE* and *RE* measures have statistically significant and convex relationships with the number of attributes in CEs. Both

SE and *RE* first decrease, then increase as the number of attributes in the CEs increases. This result implies that keeping the number of attributes at either the minimum (2 in our case) or the maximum (5 in our case) may not be the best strategy for WTP estimates.

The positive coefficients for the # of Attributes for attribute X3a indicate that increasing the number of attributes in the CEs significantly increases both *SE* and *RE* measures (5% level). This is different from the case of continuous utility functions, in which increasing the number of attributes resulted in decreased *SE* and *RE* for attribute X3. In addition, middle and large sample sizes of the CEs result in significantly smaller *SE* and *RE*, however, there is no significant difference in *SE* when the sample sizes of CEs change from middle to large, indicating that large sample sizes may not always result in better WTP estimates with the discrete consumer utility function.

Conclusion

The increasing application of CEs in consumer valuation reflects the comparative advantage of this technique in valuation of multiple product attributes and in estimation of tradeoffs between different attributes. On the other hand, the popularity of CEs demands more research in this area to study the impact of design parameters on consumer valuation estimates. Although used by many researchers, the effort to address the second issue is not adequate. Carlsson and Martinsson (2003), Lusk and Norwood (2005), and Ferrini and Scarpa (2007), among others are the few to study the impacts of design strategies on consumer valuation estimates using Monte Carlo simulation. However, in those studies, the impacts of CE design are evaluated with fixed attribute information loads. With the current field studies finding that attribute information loads affect consumer WTP estimates (e.g. Rose et al. 2009; Islam, Louviere, and Burke. 2007), it is worthwhile to investigate the performance of different design strategies under different information loads using simulation techniques. This helps to identify the sources that affect consumer valuation in field studies—are the impacts of information loads due to the change in consumer cognitive ability to make choice decisions, the substitute and complement effects between product attributes, or the pure statistical properties of CEs? In this paper we extend previous research by investigating the change in consumer WTP estimates under different attribute information loads, design strategies, sample sizes and utility function types to address those questions.

Results in this article delivered similar information with previous studies that larger sample size can significantly improve the WTP estimates with a continuous consumer utility function. If all other design parameters such as information loads and sample size are controlled, RD design is the overall best design; however, this design is not significantly better than RP design in some cases. The performance of design strategies may also depend on the information loads in the CEs — when the number of attributes is five, RD design is the best with both continuous and discrete cases; when the number of attributes is three, RP and MD are the best with continuous and discrete utility functions, respectively.

The fact that the WTP estimates of some attributes have a quadratic relationship with the number of attributes in CEs calls for further research to investigate the impacts of attribute information on consumer valuation—how much of the changes in consumer valuation demonstrated by the previous research is due to the changes in consumer cognitive ability to processing information, or because of the statistical property of the CEs with different number of attributes? For some attributes, the quadratic relationships between the performance and the number of attributes in the CEs indicates that keeping the attribute information as small as possible may not be the best design strategy to improve consumer valuation estimates. In our case, the optimal attribute number in the CEs is three in the discrete case. This conclusion depends on the particular attributes studied.

Choosing the best design strategy is only a small step toward conduction a successful CE. Many factors are correlated so that tradeoffs must be made between statistical performance, feasibility to administer the CEs, the potential problem of omitted variables in consumer utility functions, and budget constraints of the researchers. For example, the number of attributes in a CE determines the minimum choice set in ME and MD designs. ME design may be an overall better design than MD design, but the choice sets in ME design is always equal to or larger than that in an MD design. The increased number of choice sets in an ME design may increase the cognitive burden on respondents, which may result in less accurate estimation. Restricting information loads in a CE may reduce cognitive burden on respondents and improve the statistical property of choice experiments. However, if the omitted information includes important variables determining consumer preference in real-world purchases, less attribute information may result in biased estimation of consumer preference. All those factors should be carefully evaluated before conducting a CE. Further research may add more design parameters such as the number of choices in choice sets and the number of levels of attributes in CEs to study the performance of various design strategies. Further, it is worthwhile expressing the deviation of WTP estimates from the true WTP as a function of major factors that may affect consumer valuation subject to research budget, and seek the optimal strategies regarding CE design, attribute information loads, sample size, and other design parameters.

References:

Carlsson, F., and P. Martinsson. 2003. “Design Techniques for Stated Preference Methods in Health Economics.” *Health Economics* 12:281-94.

Chrzan,K., and B. Orme. “An Overview and Comparison of Design Strategies for Choice-Based Conjoint Analysis.” Sawtooth Software Research Paper Series, Sawtooth Software, Sequim, WA, 2000.

Ferrini, S. and R. Scarpa. 2007. “Designs with a priori Information for Nonmarket Valuation with Choice Experiments: A Monte Carlo Study.” *Journal of Environmental Economics and Management* 53:342-63.

Gao, Z. and T.C. Schroder. 2009. “Effects of Label Information on Consumer Willingness-To-Pay for Food Attributes.” *American Journal of Agricultural Economics forthcoming*.

Hausman, J.A., ed. 1993. *Contingent Valuation: A Critical Assessment*. New York: North Holland: p503.

Hensher, D.A. 2006. “Revealing Difference in Willingness to Pay due to the Dimensionality of Stated Choice Designs: An Initial Assessment.” *Environmental and Resource Economics* 34:7–44.

Islam, T., J.J. Louviere, and P.F. Burke. 2007. “Modeling the Effects of Including/Excluding Attributes in Choice Experiments on Systematic and Random Components.” *International Journal of Research in Marketing* 24:289–300.

Kanninen, B.J. 1993a. “Optimal Experimental Design for Double Bounded Dichotomous Choice Contingent Valuation.” *Land Economics* 69:138-146.

Kanninen, B.J. 1993b. “Design of Sequential Experiments for CV Studies.” *Journal of Environmental Economics and Management* 25: 1-11.

Kanninen, B.J. 2002. “Optimal Designs for Multinomial Choice Experiments.” *Journal of Marketing Research* 39:214-227.

Kuhfeld, W. F., R. D. Tobias, and M. Garratt. 1994. "Efficient Experimental Design with Marketing Research Applications." *Journal of Marketing Research* 31: 545-557.

Lancaster, Kelvin. 1972. *Consumer Demand: A New Approach* New York: Columbia University Press.

Louviere, J.J., D.A. Hensher, and J.D. Swait. 2000. *Stated Choice Methods: Analysis and Application* Cambridge, UK: Cambridge University Press.

Louviere J.J, D. Street, L. Burgess, N. Wasi, T. Islam and A. A. J. Marley. 2008a. “Modeling the Choices of Individual Decision-Makers by Combining Efficient Choice Experiment Design with Extra Preference Information.” *Journal of Choice Modeling* 1:128-163.

Louviere J.J, T. Islam, N. Wasi, D. Street and L. Burgess. 2008b. “Designing Discrete Choice Experiments: Do Optimal Designs Come at a Price?” *Journal of Consumer Research* 35:360-375.

Lusk, J.L., and F.B. Norwood. 2005. “Effect of Experimental Design on Choice-Based Conjoint Valuation Estimates.” *American Journal of Agricultural Economics* 87:771-85.

McFadden D. 1974. In *Frontiers in Econometrics*, Zarembka P (ed.) Academic Press: New York, 105-142.

Rose, J. M., D. A. Hensher, S. Caussade, J. d. D. Ortuzar, and R.-C. Jou. 2008 "Identifying Differences in Willingness to Pay Due To Dimensionality in Stated Choice Experiments: A Cross Country Analysis." *Journal of Transport Geography* 17: 21-29.

Scarpa, R. and J. M. Rose 2008. "Design Efficiency for Non-Market Valuation with Choice Modelling: How to Measure it, What to Report and Why." *The Australian Journal of Agricultural and Resource Economics* 52:253-282.

Swait, J. and W. Adamowicz. 2001. "Choice Environment, Market Complexity, and Consumer Behavior: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice." *Organizational Behavior and Human Decision Processes* 86:141-167.

Table 1 Description of Simulation Scenarios

CE Design	Description	# of Attribute (N)	# of Choice Sets in Experiment	Utility Function	Sample Size
RD	Random drew choice	2	N/A	Continuous	243/486/972
	set from $3^N x 3^N$ full	3	N/A	/Discrete	250/500/1000
	factorial design	4	N/A		250/500/1000
		5	N/A		250/500/1000
ME	Main effects only	2	9	Continuous	252/504/1008
	design drawn from	3	27	/Discrete	270/513/1026
	$3^N x 3^N$ full factorial	4	27		270/513/1026
	design	5	27		270/513/1026
MD	Main effects minimum	2	9	Continuous	252/504/1008
	design with maximized	3	13	/Discrete	260/507/1001
	D-efficiency	4	17		255/510/1003
		5	21		252/504/1008
RP	Random pair	2	9	Continuous	243/495/999
	alternatives generated	3	27	/Discrete	243/486/999
	from full factorial	4	81		243/486/972
	design	5	243		243/486/972

Table 2 WTP Estimates and MSE is Different Simulation Scenarios (Continuous Utility Function)

# of Attribute	Attribute	RD			ME			MD			RP		
		s=S	s=M	s=H	s=S	s=M	s=H	s=S	s=M	s=H	s=S	s=M	s=H
2	X1	2.01[a]	1.96*	1.98	1.95*	1.97*	1.98	2.01	1.98	1.97*	2.18*	2.05	1.99
		(0.31)[b]	(0.13)	(0.08)	(0.15)	(0.08)	(0.05)	(0.31)	(0.16)	(0.07)	(0.89)	(0.30)	(0.12)
3	X1	2.08*	1.98	1.97	2.16*	2.03	2.01	2.09*	2.02	2.00	2.02	1.97	1.97*
		(0.57)	(0.21)	(0.10)	(1.02)	(0.39)	(0.19)	(0.72)	(0.33)	(0.19)	(0.49)	(0.22)	(0.10)
	X2	3.14*	2.98	2.97	3.25*	3.06	3.02	3.15*	3.02	3.01	3.03	2.97	2.96
		(1.33)	(0.47)	(0.23)	(2.07)	(0.79)	(0.38)	(2.02)	(0.96)	(0.53)	(0.93)	(0.42)	(0.20)
	MSE	{0.95}[c]	{0.34}	{0.17}	{1.55}	{0.59}	{0.28}	{1.37}	{0.64}	{0.36}	{0.71}	{0.32}	{0.15}
4	X1	2.02	2.00	1.94*	2.93*	2.13*	2.04*	1.37*	1.78*	2.25*	2.13*	2.03	1.97
		(0.58)	(0.32)	(0.14)	(4.55)	(0.52)	(0.14)	(4.91)	(2.79)	(1.30)	(0.72)	(0.36)	(0.17)
	X2	3.03	2.99	2.91*	3.98*	3.17*	3.04	1.70*	2.59*	3.47*	3.22*	3.05	2.98
		(1.25)	(0.68)	(0.31)	(5.37)	(0.89)	(0.24)	(13.45)	(8.66)	(4.52)	(1.60)	(0.72)	(0.33)
	X3	4.04	3.98	3.87*	5.87*	4.26*	4.07*	2.25*	3.42*	4.64*	4.30*	4.09	3.95
		(2.25)	(1.23)	(0.57)	(18.26)	(1.97)	(0.49)	(27.97)	(16.35)	(8.46)	(2.85)	(1.31)	(0.62)
	MSE	{1.36}	{0.74}	{0.34}	{9.40}	{1.12}	{0.29}	{15.44}	{9.27}	{4.76}	{1.72}	{0.80}	{0.37}
5	X1	2.22*	2.10*	2.01	2.27*	2.08*	2.04	1.46*	2.13	2.06	2.28*	2.07	2.01
		(1.59)	(0.54)	(0.24)	(2.23)	(0.53)	(0.26)	(6.94)	(4.98)	(0.65)	(1.85)	(0.55)	(0.25)
	X2	3.29*	3.13*	3.02	3.46*	3.14*	3.06	2.05*	3.17	3.11	3.39*	3.08	3.02
		(2.91)	(1.04)	(0.44)	(5.56)	(1.31)	(0.65)	(19.54)	(13.71)	(1.87)	(3.86)	(1.24)	(0.57)
	X3	4.42*	4.21*	4.03	4.60*	4.19*	4.09	3.05*	4.24	4.12	4.55*	4.12	4.03
		(5.55)	(1.94)	(0.83)	(9.70)	(2.23)	(1.10)	(22.34)	(16.61)	(2.19)	(7.31)	(2.16)	(1.00)
	X4	0.12	0.10	0.10	0.11	0.09	0.10	0.15*	0.08	0.10	0.11	0.10	0.10
		(0.06)	(0.02)	(0.01)	(0.08)	(0.02)	(0.01)	(0.13)	(0.08)	(0.01)	(0.07)	(0.02)	(0.01)
	MSE	{2.53}	{0.88}	{0.38}	{4.39}	{1.02}	{0.50}	{12.24}	{8.84}	{1.18}	{3.27}	{0.99}	{0.46}

Notes: We removed 5% of the lowest and highest extreme value in WTP estimates, because some of those values are not reasonable.

One (*) asterisk represents the 0.05 level of statistical significance.

[a] Mean WTP from 450 simulated WTP of attribute Xi (i=1,2,3,4).

[b] Mean Squared Errors of attribute Xi.

[c] Average Mean Squared Errors of all attributes in a CE. For experiments with two attribute, MSE of X1=MSE, because only WTP for X1 exists.

Table 3 WTP Estimates and MSE in Different Simulation Scenarios (Discrete Utility Function)

# of Attribute	Attribute	RD			ME			MD			RP		
		s=S	s=M	s=H	s=S	s=M	s=H	s=S	s=M	s=H	s=S	s=M	s=H
2	X1a	1.99	1.99	1.99	1.94*	1.97	1.99	1.99	1.99	1.98	1.96	1.97	1.98
		(0.30)	(0.14)	(0.07)	(0.25)	(0.13)	(0.07)	(0.23)	(0.10)	(0.05)	(0.37)	(0.18)	(0.08)
	MSE	{0.21}	{0.09}	{0.05}	{0.16}	{0.09}	{0.05}	{0.16}	{0.07}	{0.04}	{0.22}	{0.11}	{0.05}
3	X1a	1.98	1.99	1.99	2.03	1.98	1.97	1.99	1.95*	1.97	2.00	1.98	1.98
		(0.35)	(0.19)	(0.09)	(0.45)	(0.20)	(0.11)	(0.32)	(0.15)	(0.08)	(0.34)	(0.17)	(0.07)
	X2a	2.98	2.99	2.99	3.05	2.99	2.96	2.99	2.96	2.96*	3.07	2.99	2.96*
		(0.63)	(0.35)	(0.18)	(1.01)	(0.47)	(0.24)	(0.65)	(0.32)	(0.16)	(0.86)	(0.43)	(0.18)
	MSE	{0.37}	{0.20}	{0.10}	{0.53}	{0.24}	{0.12}	{0.34}	{0.16}	{0.08}	{0.44}	{0.21}	{0.09}
4	X1a	2.00	1.98	1.98	1.97	1.98	1.98	2.26*	2.07*	2.02	2.03	1.98	1.98
		(0.90)	(0.21)	(0.09)	(0.85)	(0.21)	(0.11)	(2.67)	(0.60)	(0.27)	(1.09)	(0.22)	(0.12)
	X2a	3.03	2.99	2.97	2.91*	2.94	2.98	3.32*	3.07	3.03	3.07	2.96	2.99
		(0.90)	(0.43)	(0.21)	(0.85)	(0.49)	(0.25)	(2.67)	(0.70)	(0.34)	(1.09)	(0.40)	(0.23)
	X3a	4.02	3.98	3.95	3.90*	3.92*	3.96	4.41*	4.08	4.03	4.09	3.95	3.99
		(1.39)	(0.73)	(0.35)	(1.22)	(0.63)	(0.34)	(4.37)	(1.16)	(0.53)	(2.27)	(0.85)	(0.47)
	MSE	{0.70}	{0.35}	{0.17}	{0.66}	{0.35}	{0.18}	{2.39}	{0.64}	{0.30}	{1.00}	{0.38}	{0.21}
5	X1a	2.12*	1.99	1.98	2.21*	2.04	1.99	1.44*	2.10	2.03	2.14*	2.03	1.99
		(0.78)	(0.22)	(0.11)	(1.71)	(0.49)	(0.23)	(5.46)	(2.29)	(0.57)	(1.02)	(0.40)	(0.20)
	X2a	3.17*	3.00	2.98	3.28*	3.03	2.98	2.23*	2.97	2.97	3.20*	3.05	3.00
		(1.91)	(0.59)	(0.32)	(3.67)	(0.86)	(0.42)	(13.16)	(5.00)	(1.45)	(1.65)	(0.68)	(0.34)
	X3a	4.21*	4.02	3.99	4.19	4.00	3.95	3.16*	4.25	4.10	4.28*	4.07	4.01
		(2.85)	(0.95)	(0.53)	(5.79)	(1.63)	(0.73)	(26.63)	(9.25)	(1.98)	(2.99)	(1.18)	(0.62)
	X4a	5.31*	5.03	4.97	5.43*	5.07	4.97	3.85*	5.14	5.06	5.37*	5.11	4.99
		(4.78)	(1.53)	(0.80)	(9.31)	(2.36)	(0.97)	(26.45)	(10.14)	(2.40)	(5.15)	(2.13)	(1.02)
	MSE	{2.07}	{0.66}	{0.35}	{4.10}	{1.07}	{0.47}	{14.19}	{5.32}	{1.30}	{2.13}	{0.86}	{0.43}

Notes: We removed 5% of the lowest and highest extreme value in WTP estimates, because some of those values are not reasonable.

One (*) asterisk represents the 0.05 level of statistical significance.

[a] Mean WTP from 450 simulated WTP of attribute Xi (i=1,2,3,4). Corresponding to each attribute in continuous utility function, there are two attributes in discrete utility function.

[b] Mean Squared Errors of attribute Xi.

[c] Average Mean Squared Errors of all attributes in a CE.

Table 4 Best CE Designs in Different Scenarios based on MSEi, MSE, RSE and DEF

# of Attributes	Utility Function							
	Continuous				Discrete			
		Sample Size				Sample Size		
		Small	Middle	Large		Small	Middle	Large
2	MSE1[a]	ME	ME	ME	MSE1a	MD	MD	MD
	DEF	ME	ME	ME	DEF	MD	MD	MD
3	MSE1	RP	RD	RP	MSE1a	MD	MD	RP
	MSE2	RP	RP	RP	MSE2a	RD	MD	MD
	MSE/RSE	RP	RP	RP	MSE/RSE	MD	MD	MD
	DEF	RD	RD	RD	DEF	ME	ME	ME
4	MSE1	RD	RD	RD	MSE1a	ME	ME	RD
	MSE2	RD	RD	ME	MSE2a	ME	RP	RD
	MSE3	RD	RD	ME	MSE3a	ME	ME	ME
	MSE/RSE	RD	RD	ME	MSE/RSE	ME	RD	RD
	DEF	RP	RP	RP	DEF	RD	RD	RD
5	MSE1	RD	ME	RD	MSE1a	RD	RD	RD
	MSE2	RD	RD	RD	MSE2a	RP	RD	RD
	MSE3	RD	RD	RD	MSE3a	RD	RD	RD
	MSE4	RD	RD	RP	MSE4a	RD	RD	RD
	MSE/RSE	RD	RD	RD/RP	MSE/RSE	RD/RP	RD	RD
	DEF	RD	RD	RD	DEF	RD	RD	RD

Notes: [a] When the number of attributes in CEs are one, MSE=MSE1.

Table 5 Models of WTP Estimates, SE, RE on Design Strategy, Samples Size and Number of Attributes (Continuous Utility Function)

	WTP for				SE for				RE for			
	X1	X2	X3	X4	X1	X2	X3	X4	X1	X2	X3	X4
Design Strategy												
RD	0.10	0.24	0.47	0.00	-1.55	-6.29	-13.59	-0.04	-0.17	-0.31	-0.37	-0.34
	(0.00)[a]	(0.00)	(0.00)	(0.86)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ME	0.20	0.43	0.89	-0.01	-1.10	-5.33	-10.03	-0.04	-0.13	-0.25	-0.28	-0.28
	(0.00)	(0.00)	(0.00)	(0.15)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RP	0.13	0.27	0.55	-0.01	-1.45	-6.16	-13.11	-0.04	-0.15	-0.30	-0.35	-0.34
	(0.00)	(0.00)	(0.00)	(0.19)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample Size												
Middle	-0.06	-0.03	-0.07	-0.03	-0.96	-2.42	-6.55	-0.05	-0.14	-0.16	-0.20	-0.77
	(0.00)	(0.39)	(0.23)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Large	-0.06	-0.01	-0.03	-0.02	-1.49	-4.13	-10.12	-0.07	-0.23	-0.26	-0.32	-1.21
	(0.00)	(0.76)	(0.56)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Attributes	0.04	-0.28	0.08	N/A[d]	0.32	5.01	-0.78	N/A	0.11	0.38	-0.02	N/A
	(0.28)	(0.09)	(0.12)	N/A	(0.04)	(0.00)	(0.09)	N/A	(0.00)	(0.00)	(0.05)	N/A
# of Attributes Squared	0.00	0.05	N/A[c]	N/A	0.05	-0.54	N/A	N/A	-0.01	-0.05	N/A	N/A
	(0.62)	(0.07)	N/A	N/A	(0.13)	(0.01)	N/A	N/A	(0.00)	(0.00)	N/A	N/A
Constant	1.90	3.19	3.39	0.13	1.62	-0.37	23.95	0.11	0.29	0.03	0.86	2.32
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.84)	(0.00)	(0.00)	(0.00)	(0.63)	(0.00)	(0.00)
Statistical Test												
RD=ME	(0.00)	(0.00)	(0.00)	(0.21)	(0.00)	(0.00)	(0.00)	(0.42)	(0.00)	(0.00)	(0.00)	(0.27)
RD=RP	(0.07)	(0.49)	(0.25)	(0.26)	(0.24)	(0.64)	(0.46)	(0.95)	(0.00)	(0.34)	(0.14)	(0.99)
ME=RP	(0.00)	(0.00)	(0.00)	(0.90)	(0.00)	(0.00)	(0.00)	(0.46)	(0.01)	(0.00)	(0.00)	(0.27)
Middle=Large	(0.82)	(0.58)	(0.54)	(0.42)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R-Squared	0.007	0.009	0.016	0.004	0.054	0.065	0.077	0.054	0.137	0.16	0.151	0.113
# Observations	21600[b]	16200	10800	5400	21600	16200	10800	5400	21600	16200	10800	5400

[a] Numbers in parentheses are *p* values.
[b] Number of observations are derived from 450 simulations from each simulation scenario.
[c] On attribute X3, the # of attributes squared variable is nearly perfectly collinear with the number of attributes variable.
[d] On attribute X4, there is no variation in the # of attribute, only choice experiments with five attributes include X4.

Table 6 Models of WTP Estimates, SE, RE on Design Strategy, Samples Size and Number of Attributes (Discrete Utility Function)

	WTP for				SE for				RE for			
	X1a	X2a	X3a	X4a	X1a	X2a	X3a	X4a	X1a	X2a	X3a	X4a
Design Strategy												
RD	0.01	0.07	0.03	0.42	-0.78	-2.10	-6.19	-10.63	-0.11	-0.13	-0.19	-0.24
	(0.00)[a]	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ME	0.02	0.07	-0.02	0.47	-0.67	-1.80	-5.60	-8.79	-0.08	-0.10	-0.16	-0.19
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RP	0.02	0.09	0.06	0.47	-0.72	-2.07	-5.93	-10.23	-0.09	-0.12	-0.17	-0.22
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample Size												
Middle	-0.01	-0.03	0.00	0.10	-0.58	-1.53	-3.89	-7.38	-0.11	-0.12	-0.14	-0.18
	(0.64)	(0.16)	(0.96)	(0.21)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Large	-0.02	-0.05	-0.03	0.01	-0.80	-2.06	-5.25	-10.12	-0.18	-0.19	-0.22	-0.27
	(0.17)	(0.03)	(0.38)	(0.95)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Attributes	0.04	0.19	0.00	N/A[d]	-0.47	-3.57	3.40	N/A	-0.04	-0.20	0.12	N/A
	(0.13)	(0.09)	(0.95)	N/A	(0.00)	(0.00)	(0.00)	N/A	(0.00)	(0.00)	(0.00)	N/A
# of Attributes Squared	0.00	-0.03	N/A[c]	N/A	0.16	0.76	N/A	N/A	0.02	0.05	N/A	N/A
	(0.31)	(0.09)	N/A	N/A	(0.00)	(0.00)	N/A	N/A	(0.00)	(0.00)	N/A	N/A
Constant	1.94	2.71	4.02	4.65	1.49	7.22	-1.54	18.83	0.34	0.58	0.09	0.61
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Statistical Test												
RD=ME	(0.64)	(0.91)	(0.35)	(0.56)	(0.01)	(0.03)	(0.12)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
RD=RP	(0.73)	(0.42)	(0.44)	(0.53)	(0.21)	(0.80)	(0.49)	(0.59)	(0.00)	(0.40)	(0.03)	(0.30)
ME=RP	(0.90)	(0.49)	(0.09)	(0.97)	(0.23)	(0.06)	(0.39)	(0.05)	(0.11)	(0.00)	(0.30)	(0.01)
Middle=Large	(0.36)	(0.46)	(0.36)	(0.24)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R-Squared	0.001	0.001	0.000	0.007	0.061	0.056	0.069	0.09	0.169	0.163	0.158	0.165
# Observations	21600[b]	16200	10800	5400	21600	16200	10800	5400	21600	16200	10800	5400

Notes: The models for attributes Xib (i=1,2,3,4) are not presented in the table because the results are similar with those from models for attributes Xia.
[a] Numbers in parentheses are *p* values.
[b] Number of observations are derived from 450 simulations from each simulation scenario.
[c] On attribute X3a, the # of attributes squared variable is nearly perfectly collinear with the number of attributes variable.

[d] On attribute X4a, there is no variation in the # of attribute, only choice experiments with five attributes include X4a.