Constructing an Optimal Orthogonal Choice Design with Alternative-specific Attributes for Stated Choice Experiments

Li Tang*, Ph.D. Candidate
School of Transportation and Logistics,
Southwest Jiaotong University
No. 111 Second Ring Road North Section 1, Chengdu, Sichuan 610031, China
Phone: 86-13730670653
E-mail: tanglitraffic@gmail.com

Xia Luo, Ph.D., Professor
School of Transportation and Logistics,
Southwest Jiaotong University
No. 111 Second Ring Road North Section 1, Chengdu, Sichuan 610031, China
Phone: 86-028-86466625
E-mail: xia.luo@263.net

Yang Cheng, Ph.D., Assistant Researcher,
Department of Civil and Environmental Engineering,
University of Wisconsin-Madison, Madison, WI 53706, USA
Phone: 1-608-692-5499
E-mail: cheng8@wisc.edu

Fei Yang, Ph.D., Associate Professor
School of Transportation and Logistics,
Southwest Jiaotong University
No. 111 Second Ring Road North Section 1, Chengdu, Sichuan 610031, China
Phone: 86-13550025963
E-mail: yangfei_traffic@163.com

Bin Ran, Ph.D., Professor
Department of Civil & Environmental Engineering,
University of Wisconsin-Madison,
1415 Engineering Drive, Madison, WI 53706, USA
Phone: 1-608-262-0052 Fax: 1-608-262-5199
Email: bran@wisc.edu

Words: 4690 words + 6 tables =6190 words

* *Corresponding Author*

ABSTRACT

The stated choice (SC) experiment is generally regarded as an effective way to obtain data for discrete choice analysis. The SC experimental design method, which determines the rule to allocate different levels for each attribute in choice situations, will have a great impact on parameter estimation. The optimal orthogonal choice (OOC) design is one of the most efficient SC designs, by which more reliable parameter estimates can be achieved with an equal or lower sample size. However, OOC design can only be applied to utility models with generic attributes; using it to assign alternative-specific attribute levels is not fully discussed in literature. This paper provides a method to extend the use of OOC design to alternative-specific attributes. Column vectors for alternative-specific attributes are introduced and the value of each vector is forced to be orthogonal with other generic attributes in a same alternative. In this way, orthogonality of OOC design is kept within individual alternative but not necessarily across alternatives. The proposed method was compared with traditional orthogonal design and D-efficient design (another state-of-the-art efficient design method). Three experiments using field data on mode choice were conducted. The result shows that both proposed method and D-efficient design have a higher efficiency than the orthogonal design. In addition, under the complex experiment setting environment in real world, the proposed method outperforms D-efficient design in the sense that almost the same efficiency can be obtained while avoiding multiple iterations for optimal solution.

KEY WORDS: EFFICIENT DESIGN, OPTIMAL ORTHOGONAL, EFFICIENCY, LOGIT MODEL

1. INTRODUCTION

Revealed preference (RP) and stated preference (SP) data have long been regarded as the two main data paradigms for discrete choice analysis. The purpose of conducting RP or SP experiments is to collect data that can be used to estimate the independent influence of attributes on observed choices. Increasingly, the emphasis has shifted from 'stated preference' data to 'stated choice' (SC) data which resembles revealed preference data as closely as possible and can be analyzed with similar techniques, such as discrete choice logit or probit estimation (*1*,*2*). The popularity of SC experiments lies in two aspects: First, it can obtain choices on alternatives which do not exist in the current market. Thus analysts are able to predict, for example, the market share of a proposed transportation mode. Second, it can provide variability in attributes in a relatively small sample size, compared with RP experiment, so that better estimation of influence of each attribute on choice can be achieved. Usually, SC experiments present sampled respondents with a 'selected' number of different hypothetical choice situations, each consisting of a universal but finite set of alternatives defined on a number of attribute dimensions. Obviously, making choices among all the possible combinations of attribute levels is too many to accomplish for a single respondent. How analysts distribute attribute levels in an experiment design plays a big role. It may impact the performance of independent assessment for attributes contribution to observed choices. It also implies the ability of the experiment to detect statistical relationships that may exist within the data.

Historically, researchers have relied on orthogonal experimental designs, in which the attributes of the experiment are statistically independent by forcing them to be orthogonal (*3*). In this way, orthogonal designs theoretically allow for an independent determination of each attribute's influence upon the observed choices. There are several approaches to generate full or fractional factorial orthogonal design (*4*,*5*,*6*,*7*). While orthogonal design has long been used in practice, researchers in recent years argue the importance of orthogonality in SC data. They doubted the effectiveness when orthogonal design is used to estimate discrete choice model, not to mention whether orthogonality can be kept in reality (*8*,*9*). Orthogonality is important in linear models since it avoids multicollinearity problem and also minimize variance-covariance matrix of the estimated model. Unfortunately, discrete choice model is nonlinear. The derivation of its variance-covariance matrix is very different from the way in linear models. Therefore, keeping orthogonality of the parameters has little to do with minimizing their standard errors.

Acknowledgement of this fact has led researchers to develop the so-called efficient experimental designs. These designs are capable of producing more efficient data in the sense that more reliable parameter estimates can be achieved with an equal or lower sample size. To date, two kinds of efficient designs arouse more attention. One is called optimal orthogonal choice (OOC) design, which aims at maximizing attribute level differences as well as keeping orthogonality within alternatives. Another is the D-efficient design, the core of which is minimizing D-error, a statistic corresponding to the asymptotic variance-covariance (AVC) matrix of the discrete choice model, to get the smallest asymptotic standard error (i.e., square roots of the variances). The fundamental idea of OOC design can be derived from the work of Bunch and Louviere in 1994. They proposed a new strategy of 'shifting codes' in a two-level main effects design, called "Shifted Paris/ Fold-over" design (*5*). Later, Street and Burgess in 2004 construed optimal and near-optimal sets of pairs for estimation of main effects and two factor interactions (*10*). All the attributes are forced to have two levels at that time. Continuously, they developed the optimal design with asymmetric attributes in 2005 and 2006 (*11*,*12*). However, the current OOC method can only generate designs for generic attributes across alternatives (*8*). How alternative-specific attributes distribute is rarely discussed in literature. On the other side, D-efficient design does not suffer from these kinds of attribute setting and fitting problem. But to conduct a D-efficient design requires very complex computation work since the AVC matrix needs to be build and optimized. If OOC could be extended to a more general case, it can probably help analysts obtain high efficient designs in a more convenient way.

This paper targets on presenting the following contributions to the literature: first, a method which can extend the use of OOC design into alternative-specific attributes is provided. Real experiment environment is settled in case study to exam the feasibility of our proposed method. Second, a systematic and comprehensive evaluation is presented by comparing OOC design with traditional orthogonal design and the-state-of-art D-efficient design.

The remainder of this paper is organized as follows. In section 2, a detailed method of improving OOC design is introduced. In section 3, SC experiment designs using orthogonal, OOC, D-efficient design method separately are generated. Efficiency of each design in terms of D-error is calculated and carefully analyzed in section 4. Section 5 provides conclusions and points at avenues for further research.

2. IMPROVING OPTIMAL ORTHOGONAL CHOICE DESIGN

Usually, an experimental design may consist of several generic attributes and alternative-specific attributes (at least one of them). In this section, the proposed method is introduced in two stages: 1) generating OOC design for generic attributes, 2) adding column vectors for alternative-specific attributes.

2.1 Generating Design for Generic Attributes

The essential idea of OOC design is to maximize the differences of attribute levels across alternatives by forcing generic attributes never taking the same level over the experiment. In this way, parameters can be estimated in the largest extend of variety of attribute levels independently. For a general experimental design, we can assign levels for generic attributes following the basic process:

Step 1: generate a fractional factorial orthogonal design for alternative 1. *N* represents the number of choice situations of the design.

Step 2: choose some systematic changes to get the allocation of attribute levels in alternative 2 from alternative 1. Systematic changes are certain rules to decide how the attribute levels change from alternative 1 and will be discussed in later context.

Step 3: choose another systematic changes to get the allocation of attribute levels in alternative 3 from alternative 1.

Step 4: keep doing this until all the alternatives are determined.

It will be much easier to understand OOC design by starting with a binary attribute level design. Here, assume that L_{k^*} is the number of levels assigned to attribute k^* for alternative *j*, represented by 0, 1,…, L_{k^*}-1 and all the attributes are generic across alternatives. x^*_{kn} is the level for attribute k^* in choice situation n. In a design for 2 alternatives and 3 attributes each with 2 levels, an orthogonal design in 4 choice situations for alternative 1 can be firstly generated. Then 0's and 1's in alternative 1 are interchanged in alternative 2. Thus the levels of each attribute are forced to be different across alternatives. The result is shown in Table 1.

TABLE 1 Optimal Orthogonal Choice Design for 2 Alternatives with 3 Binary Attributes

	alternative 1			alternative 2		
choice situation	X_1^*	X_2^*	X_3^*	X_1^*	X_2^*	X_3^*
1	0	0	0	1	1	1
2	0	1	1	1	0	0
3	1	0	1	0	1	0
4	1	1	0	0	0	1

To generate OOC design for more alternatives, it is necessary to introduce S_{k^*} to represent the largest number of different pairs appeared between alternatives for a specific attribute. The equation of S_{k^*} is shown as follows. Where *J* stands for the number of alternatives in choice set.

$$S_{k^*}=\begin{cases}\left(J^2-1\right)/4 & l_{k^*}=2, J \ \ odd,\\ J^2/4 & l_{k^*}=2, J \ \ even,\\ \left(J^2-\left(l_{k^*}x^2+2xy+y\right)\right)/2 & 2<l_{k^*}\le J,\\ J\left(J-1\right)/2 & l_{k^*}\ge J\end{cases} \tag{1}$$

In an example of OOC design for 3 alternatives and 3 attributes each with 2 levels shown in **TABLE 2**, $S_{k^*}=\left(J^2-1\right)/4=2$ can be calculated for all the attributes. For instance, in the first choice situation (000, 110, 001), the attribute levels differ twice for each attribute (i.e. for attribute X_1^*, the levels are 010, creating 3 pairs (01, 10, 00) in which two of them (01, 10) are different).

Seeing from **TABLE 2**, the distribution of attribute levels in alternative 2 is obtained by interchanging 0's and 1's in X_1^* and X_2^* in alternative 1, and in alternative 3 it is obtained by interchanging 0's and 1's in X_3^* in alternative 1. These systematic changes can be also described as adding a generator in alternative 1 to get alternative 2 and adding another generator to get alternative 3. The addition is performed in modulo arithmetic according to the number of levels for a specific attribute. Here in the example, $L_{k^*}=2$ for all attributes, thus alternative 2 is obtained when a generator 110 is added to the choice situations in alternative 1 in modulo 2 arithmetic like this: $000+110\equiv 110$, $011+110\equiv 101$, and so on. Alternative 3 is generated in the same way by adding a generator 001 to alternative 1. Notice that the generators added to alternatives must have a value of $S_{k^*}=2$ (i.e. generators (000,110,001) used in **TABLE 2** can meet the requirement while another generators (000, 100, 010) can't not).

TABLE 2 Optimal Orthogonal Choice Design for 3 Alternatives with 3 Binary Attributes

	alternative 1			alternative 2			alternative 3		
choice situation	X_1^*	X_2^*	X_3^*	X_1^*	X_2^*	X_3^*	X_1^*	X_2^*	X_3^*
1	0	0	0	1	1	0	0	0	1
2	0	1	1	1	0	1	0	1	0
3	1	0	1	0	1	1	1	0	0
4	1	1	0	0	0	0	1	1	1

2.2 Adding Alternative-specific Attributes in OOC Design

Although OOC designs for any choice set size with any number of attributes each having any number of levels can be generated in similar way, this method will not work where alternative-specific attributes are introduced according to its basic idea. However, in transportation area, choice analysis always involves alternatives with their own specific attributes (i.e. parking fare for car or waiting time for public transit). To extend the use of OOC, column vectors are added for alternative-specific attributes on the base of OOC design. The total number of column vectors X_k equals to the total number of alternative-specific attribute k in the discrete choice model. Again, we assume L_k is the number of levels assigned to attribute k for alternative *j*, represented by 0, $1,\ldots,L_k$-1. The dimension of vector X_k equals to the total number of choice situations *N* in the former design generated for generic attributes. Here, it is noteworthy that the degree of freedom, which directly related to the minimum number of choice situations *N* of the design, should be kept after introducing alternative-specific attributes in the model. That is to say, the size of an OOC design only with generic attributes might be small. But when alternative-specific attributes are added, a larger size of design is required to reach at the same degree of freedom so that enough observations can be obtained for parameter estimation. Thus, when generating OOC design for generic attributes in the first step, the number of choice situation *N* should be determined by considering the total number of all the attributes including alternative-specific attributes.

To maintain the principle of OOC design, two constrains should be satisfied when calculating the value of X_k:

- Every level l_k in X_k appears at same times.
- The matrix that consists of X_k and X_{k^*} within the same alternative *j* is orthogonal (i.e. the covariance matrix of it is a diagonal matrix).

For alternatives with more than one specific attributes, the logic is calculating X_k s one by one. For example, X_1 , X_2 are two specific attributes and X_1^* , X_2^* are two generic attributes for alternative *j*. First, X_1 is determined by following the above rules. Then the feasible solution of X_2 is the one that being orthogonal with the new matrix constructed by X_1^*, X_2^* and X_1. It is very likely that the solution of the column vector is not unique, especially in a design with high dimension.

3. CASE STUDY

In this section, three experimental designs are generated using different methods (orthogonal, OOC and D-efficient design) for obtaining trip mode choice data on a corridor connecting downtown and suburb area in Chengdu. In section 3.1, a typical multinomial logit (MNL) model is formulated. Prior parameter values and attribute levels are settled. How to measure the statistical efficiency of SC experiments is discussed in section 3.2. Experimental design results are presented in section 3.3.

3.1 Model Formulation and Stimuli Refinement

The field data for the case study is collected in Chengdu, a large city located in southwest China. Four alternatives are involved: car, taxi, bus and subway. By far, car, taxi and bus are the existing transport modes. The subway, No. 2 West Extended Line, is still under construction and will be operating on June 1st, 2013. It starts from Xipu Town and ends at the Yingbin Avenue, then connected with the existing No. 2 Line. The opening of the subway is expected to change people's travel mode choice behavior and help release the heavy traffic congestion on the roads in the northwest area of the city. The new campus of Southwest Jiaotong University which is located in northwest suburb area and Shuhan Road which is located in the west part of downtown area are chosen as the origin and destination. Transport mode choice data will be collected based on the experimental design generated in this section.

According to a previous study (*13*) on comparing MNL and nested logit (NL), there is no statistical significant difference in terms of parameter calibration between MNL and NL models. Thus a typical MNL model in transportation is formulated and serves as the basis of most of the analyses in the subsequent section. The observed part of utility of every alternative is expressed as follows:

$$V^{car} = \beta_0^{car} + \beta_1 TT^{car} + \beta_2 TC^{car} + \beta_3 PF^{car}, \tag{2}$$

$$V^{taxi} = \beta_0^{taxi} + \beta_1 TT^{taxi} + \beta_2 TC^{taxi}, \tag{3}$$

$$V^{bus} = \beta_0^{bus} + \beta_1 TT^{bus} + \beta_2 TC^{bus} + \beta_4 WT^{bus}, \tag{4}$$

$$V^{sub} = \beta_1 TT^{sub} + \beta_2 TC^{sub} + \beta_4 WT^{sub}, \tag{5}$$

Where *TT* represents travel time, *TC* represents travel cost, *PF* represents parking fare and *WT* represents waiting time. For car users, *TC* equals to the fuel cost. For taxi users, *TC* equals to the money paid for the trip. And for bus and subway users, *TC* equals to the ticket price. *WT* is the time period between people arriving at the station/stop and before they getting on board. Seeing from Equations (2)-(5), parameters for *TT* and *TC* are generic across four alternatives and parameters for *WT* are generic across bus and subway. *PF* is the alternative specific parameter for alternative car. Thus, seven parameters are going to be estimated in total (three of them are alternative-specific constant, which has nothing to do with any attribute). The attribute levels and prior information about parameters are given in **TABLE 3** based on previous study results as well as to preserve realistic estimates for the private and public transport alternatives.

TABLE 3 Prior Parameter Values and Attribute Levels for Case Study

prior parameter values						
β_0^{car}	β_0^{taxi}	β_0^{bus}	β_1	β_2	β_3	β_4
1.2	-0.7	-1.2	-0.04	-0.25	-0.12	-0.02

attribute level										
car			taxi		bus			subway		
TT^{car}	TC^{car}	PF^{car}	TT^{taxi}	TC^{taxi}	TT^{bus}	TC^{bus}	WT^{bus}	TT^{sub}	TC^{sub}	WT^{sub}
24	7	8	24	27	45	1	5	20	2	3
30	9	10	30	33	55	2	8	24	3	4
35	11	15	35	40	65	3	10	27	4	5

In order to obtain better estimation of parameters, three levels are set for each attribute to maximize variations of the attribute as much as possible. The values of TT^{car}, TT^{taxi} and TT^{bus} are measured in free, normal and congested traffic flow. Since the No.2 West Extended Line has not operated now, we have to estimate the values of TT^{sub} by assuming the travel speed are 27 km/h, 30 km/h (current situation) and 35 km/h. The values of PF^{car} are given as the current, 25% and 50% increased price. The values of TC^{car} are calculated as the kilometers between the OD multiplied by the consumed oil price under 7.3 RMB/km, 7.7 RMB/km and 8 RMB/km. The values of TC^{taxi} are measured in free, normal and congested traffic flow. The values of TC^{bus} and TC^{sub} are based on the current price and plus/minus 1 RMB. The values of WT^{bus} and WT^{sub} are determined based on the departing time interval.

The number of choice situations (i.e. 36) is selected such that both attribute level balance and orthogonality can be achieved. Obviously, this number is too large for a single respondent. Thus, a block variable is introduced to divide the design into smaller parts (i.e. here we block the design into six parts so that six choice situations are provided to a single respondent). Each block is not orthogonal by itself, but in combination with other blocks. Attribute level balance is maintained as much as possible in each block.

3.2 Measure of Efficiency

To compare the statistical efficiency of SC experimental designs, a number of measurements have been proposed in the literature (5,*14*,*15*,*16*). The preferred measure among them is D-error, a statistic corresponding to the AVC matrix of the discrete choice model. To interpret the process of calculating it, here we briefly introduce the most well known multinomial logit model. Other models, like nested logit, can also be used to determine the value of D-error (*9*).

Assume an individual faced with alternative $j=1,2,...,J$ in choice situation $n=1,2,...,N$. The utility of an individual for alternative *j* in choice situation *n* can be expressed as:

$$U_{jn} = V_{jn} + \varepsilon_{jn}, \tag{6}$$

V_{jn} represents observed part of utility for each alternative *j* in choice situation *n*. It is assumed to be a linear additive function of several attributes with corresponding weights. The generic parameters and alternative-specific parameters can be denoted by β_k^*, $k=1,...,K^*$ and β_{jk}, $k=1,...,K_j$, respectively, with their associated attribute levels x_{jkn}^* and x_{jkn} for each choice situation *n*. Thus, the total number of parameters to be estimated is equal to $\overline{K} = K^* + \sum_{j=1}^{J} K_j$. V_{jn}, expressed as:

$$V_{jn} = \sum_{k=1}^{K^*} \beta_k^* x_{jkn}^* + \sum_{k=1}^{K_j} \beta_{jk} x_{jkn}, \quad \forall j=1,...,J, \forall n=1,...,N, \tag{7}$$

ε_{jn} is the unobserved component, independently and identically extreme value type Ⅰ distributed. The probability P_{jn} that an individual choose alternative *j* in choice situation *n* becomes:

$$P_{jn} = \frac{\exp\left(V_{jn}\right)}{\sum_{j=1}^{J} \exp(V_{jn})}, \quad \forall j = 1,...,J, \forall n = 1,...,N, \tag{8}$$

Considering the most popular way to estimate parameters is maximum likelihood estimation, the log-likelihood function of parameters for a single respondent can be expressed as:

$$L\left(\beta^*, \beta\right) = \sum_{n=1}^{N}\sum_{j=1}^{J} y_{jn} \log P_{jn}$$

$$= \sum_{n=1}^{N}\left[\sum_{j=1}^{J} y_{jn}\left(\sum_{k=1}^{K^*}\beta_k^* x_{jkn}^* + \sum_{k=1}^{K_j}\beta_{jk} x_{jkn}\right) - \log\left(\sum_{i=1}^{J}\exp\left(\sum_{k=1}^{K^*}\beta_k^* x_{jkn}^* + \sum_{k=1}^{K_j}\beta_{jk} x_{jkn}\right)\right)\right], \tag{9}$$

Where y represents the binary outcome of all choice situations. While alternative j is chosen in choice situation n, y_{jn} equals one, otherwise it is zero. Then the AVC matrix can be expressed as the second derivative of the log-likelihood function as follows:

$$\frac{\partial^2 L\left(\beta^*, \beta\right)}{\partial \beta_{k_1}^* \partial \beta_{k_2}^*} = \sum_{n=1}^{N}\sum_{j=1}^{J} x_{jk_1n}^* P_{jn}\left(x_{jk_1n}^* - \sum_{i=1}^{J} x_{ik_2n}^* P_{in}\right), \quad \forall k_1, k_2 = 1,...,K^*, \tag{10}$$

$$\frac{\partial^2 L\left(\beta^*, \beta\right)}{\partial \beta_{j_1k_1} \partial \beta_{k_2}^*} = \sum_{n=1}^{N} x_{j_1k_1n} P_{j_1n}\left(x_{j_1k_2n}^* - \sum_{i=1}^{J} x_{ik_2n}^* P_{in}\right),$$

$$\forall j_1 = 1,...,J, k_1 = 1,...,K_{j_1}, k_2 = 1,...,K^*, \tag{11}$$

$$\frac{\partial^2 L\left(\beta^*, \beta\right)}{\partial \beta_{j_1k_1} \partial \beta_{j_2k_2}} = \begin{cases} -\sum_{n=1}^{N} x_{j_1k_1n} x_{j_2k_2n} P_{j_1n} P_{j_2n}, & if\ j_1 \neq j_2; \\ \sum_{n=1}^{N} x_{j_1k_1n} x_{j_2k_2n} P_{j_1n}\left(1 - P_{j_2n}\right), & if\ j_1 = j_2. \end{cases} \quad \forall j_1 = 1,...,J, k_i = 1,...,K_{j_i}, \tag{12}$$

Equations (6)-(8) represent functions that allow generic and alternative-specific parameters. In the case where only generic parameters exist, only Eq. (10) remains, and when there are only alternative-specific parameters, Eq. (12) remains. In addition, if there are M identical respondents, these second derivatives are multiplied by M. But it is common to assume a single respondent (i.e. M=1) representative of all respondents, which is consistent with the MNL model form.

The AVC matrix can be obtained by taking the negative inverse of the expected second derivatives of the log-likelihood function of the model (*17*). Let $\left(\bar{\beta}^*, \bar{\beta}\right)$ denote the true values of the parameters. The Fisher information matrix I is defined as the expected values of the second derivative of the log-likelihood function:

$$I\left(\bar{\beta}^*, \bar{\beta}\right) = M \cdot \frac{\partial^2 L\left(\bar{\beta}^*, \bar{\beta}\right)}{\partial \beta \partial \beta^{'}}, \tag{13}$$

The AVC matrix can be expressed as a $\bar{K} \times \bar{K}$ matrix that equals to the negative inverse of the Fisher information matrix:

$$\Omega = -\left[I\left(\bar{\beta}^*, \bar{\beta}\right)\right]^{-1} = -\frac{1}{M}\left[\frac{\partial^2 L\left(\bar{\beta}^*, \bar{\beta}\right)}{\partial \beta \partial \beta^{'}}\right]^{-1}, \tag{14}$$

D-error is calculated by taking the determinant of the AVC matrix and scaling this value by the number of parameters $\bar{K}$. Since the calculation of D-error involves with the values of parameters, approaches to determine D-error have been improved in recent years. Early work assumed all parameters were zero, which means the analyst has no information of the true parameters values at all. This assumption results in the term D_z*-error* and is shown as:

$$D_z\text{-}error = -\left[\det \frac{\partial^2 L\left(0,0\right)}{\partial \beta \partial \beta^{'}}\right]^{-1/\bar{K}}, \tag{15}$$

Later work assumed non-zero priors that were known with certainty termed as D_p*-error* :

$$D_p\text{-}error = -\left[\det \frac{\partial^2 L\left(\bar{\beta}^*, \bar{\beta}\right)}{\partial\beta\partial\beta^{'}}\right]^{-1/\bar{K}}, \tag{16}$$

More recently, researchers have begun to examine efficient designs where the true population parameters are not known with certainty but can be drawn from Bayesian parameter distributions (with parameter θ), which is termed as D_b*-error* :

$$D_b\text{-}error = \int_{\bar{\beta}} \left[\det \frac{\partial^2 L\left(\tilde{\beta}^*, \tilde{\beta}\right)}{\partial\beta\partial\beta^{'}}\right]^{-1/\bar{K}} \phi\left(\tilde{\beta}\middle|\theta\right) d\tilde{\beta}. \tag{17}$$

In our case study, D_p*-error* is chosen as the statistic to measure the efficiency of experimental designs.

3.3 Experimental Design

Three different (attribute level balanced) designs with 36 choice situations are generated in this section. During the process of generating design, orthogonal coding is used to replace the real value of each attribute level. For a three-level attribute, -1, 0, 1 are assigned to represent low, medium and high level respectively. After getting the results, real values are substituted again. The final design results are shown in **TABLE 4** as well as D-error value for each design.

TABLE 4 Experimental Designs for Case Study

s	car			taxi		bus			subway			Block
	TT^{car} (min)	TC^{car} (RMB)	PF^{car} (RMB)	TT^{taxi} (min)	TC^{taxi} (RMB)	TT^{bus} (min)	TC^{bus} (RMB)	WT^{bus} (min)	TT^{sub} (min)	TC^{sub} (RMB)	WT^{sub} (min)	
Orthogonal design for MNL model (D-error=0.12362)												
1	24	7	10	30	33	55	2	5	20	2	4	1
2	30	9	8	24	27	45	2	8	24	2	4	1
3	30	11	15	30	40	45	1	5	27	3	4	1
4	35	9	10	35	27	65	1	10	20	3	4	1
5	24	7	15	35	27	45	3	10	27	2	5	1
6	35	11	8	24	40	65	3	5	20	2	5	1
7	30	9	8	24	33	55	1	5	20	3	3	2
8	24	7	10	30	27	45	1	8	24	3	3	2
9	30	9	15	35	40	65	3	8	24	3	5	2
10	35	11	10	30	33	55	3	10	27	3	5	2
11	35	7	8	35	27	55	2	8	20	4	5	2
12	24	11	15	24	33	45	2	5	24	4	5	2
13	35	11	10	30	40	65	2	8	24	4	4	3
14	30	9	15	35	33	55	2	10	27	4	4	3
15	35	11	8	24	27	45	1	10	27	4	3	3
16	24	7	15	35	40	65	1	5	20	4	3	3
17	24	9	10	24	33	65	3	10	24	2	3	3
18	30	7	8	30	40	55	3	8	27	2	3	3
19	35	7	15	24	40	55	2	10	24	3	3	4
20	24	11	8	35	33	65	2	8	27	3	3	4
21	35	9	15	30	33	45	3	8	20	4	3	4
22	30	11	10	35	27	55	3	5	24	4	3	4
23	30	7	8	30	33	65	1	10	24	4	5	4
24	24	9	10	24	40	55	1	8	27	4	5	4
25	24	11	8	35	40	55	1	10	24	2	4	5
26	35	7	15	24	33	65	1	8	27	2	4	5
27	35	9	10	35	40	45	2	5	27	2	3	5
28	30	11	15	30	27	65	2	10	20	2	3	5
29	24	9	8	30	27	65	3	5	27	4	4	5
30	30	7	10	24	40	45	3	10	20	4	4	5
31	30	7	10	24	27	65	2	5	27	3	5	6

32	24	9	8	30	40	45	2	10	20	3	5	6
33	24	11	15	24	27	55	3	8	20	3	4	6
34	35	7	8	35	33	45	3	5	24	3	4	6
35	30	11	10	35	33	45	1	8	20	2	5	6
36	35	9	15	30	27	55	1	5	24	2	5	6

Orthogonal optimal design for MNL model (D-error = 0.08471)

1	35	11	8	24	40	55	1	5	24	3	5	1
2	35	9	8	24	33	55	3	5	24	2	5	1
3	30	7	8	35	27	45	2	5	20	4	5	1
4	30	7	8	35	27	45	2	5	20	4	5	1
5	30	11	8	35	40	45	1	5	20	3	5	1
6	24	9	8	30	33	65	3	5	27	2	5	1
7	30	7	8	35	27	45	2	5	20	4	5	2
8	35	11	8	24	40	55	1	5	24	3	5	2
9	35	9	8	24	33	55	3	8	24	2	4	2
10	35	7	8	24	27	55	2	8	24	4	4	2
11	24	9	8	30	33	65	3	5	27	2	5	2
12	30	9	8	35	33	45	3	8	20	2	4	2
13	24	11	10	30	40	65	1	8	27	3	4	3
14	24	7	10	30	27	65	2	8	27	4	4	3
15	30	9	10	35	33	45	3	8	20	2	4	3
16	24	11	10	30	40	65	1	8	27	3	4	3
17	24	11	10	30	40	65	1	8	27	3	4	3
18	24	7	10	30	27	65	2	8	27	4	4	3
19	35	9	15	24	33	55	3	8	24	2	4	4
20	24	11	10	30	40	65	1	8	27	3	4	4
21	35	7	10	24	27	55	2	8	24	4	4	4
22	30	11	10	35	40	45	1	10	20	3	3	4
23	24	9	10	30	33	65	3	10	27	2	3	4
24	35	11	10	24	40	55	1	10	24	3	3	4
25	24	7	10	30	27	65	2	10	27	4	3	5
26	30	7	15	35	27	45	2	10	20	4	3	5
27	24	9	15	30	33	65	3	10	27	2	3	5
28	35	7	15	24	27	55	2	10	24	4	3	5
29	30	11	15	35	40	45	1	10	20	3	3	5
30	35	7	15	24	27	55	2	10	24	4	3	5
31	30	9	15	35	33	45	3	10	20	2	3	6
32	35	9	15	24	33	55	3	10	24	2	3	6
33	30	11	15	35	40	45	1	10	20	3	3	6
34	30	9	15	35	33	45	3	5	20	2	5	6
35	35	11	15	24	40	55	1	5	24	3	5	6
36	24	7	15	30	27	65	2	5	27	4	5	6

D-efficient design for MNL model (D-error = 0.08846)

1	35	9	8	24	33	45	2	5	20	3	5	1
2	35	9	15	24	33	45	1	10	20	3	3	1
3	24	7	10	30	40	65	3	8	27	4	4	1
4	35	11	8	24	33	45	1	5	20	2	5	1
5	30	7	10	30	40	55	3	8	24	4	4	1
6	24	7	10	35	33	65	3	8	27	4	4	1
7	24	9	15	35	33	65	2	5	27	3	5	2
8	30	11	15	35	27	45	1	5	24	2	5	2
9	35	11	8	24	27	55	2	10	20	2	3	2
10	24	11	15	35	27	65	1	10	27	2	3	2
11	24	11	15	35	27	65	1	5	27	2	5	2
12	35	7	10	24	40	45	3	8	20	4	4	2
13	35	9	8	24	27	55	3	8	20	3	5	3
14	24	11	8	35	27	65	1	10	27	2	3	3
15	30	9	8	30	33	55	2	5	24	3	5	3
16	35	9	15	24	27	45	2	10	20	3	3	3
17	30	7	10	30	40	55	3	5	24	4	4	3
18	24	7	15	35	40	65	3	8	27	4	4	3
19	35	11	15	24	33	45	1	10	20	2	3	4
20	24	11	8	35	33	65	1	10	27	2	3	4
21	35	11	15	30	40	45	1	10	20	2	3	4
22	30	9	10	30	33	55	2	10	24	3	3	4
23	30	7	10	30	40	55	3	8	24	4	3	4
24	24	11	8	35	27	65	1	5	27	2	5	4
25	30	7	10	24	40	55	3	8	24	4	4	5
26	30	7	10	30	40	45	3	8	24	4	4	5
27	30	7	10	30	40	55	2	8	24	4	4	5
28	30	9	8	30	33	55	2	10	24	3	3	5
29	24	7	10	35	40	65	3	8	27	4	4	5
30	30	7	10	30	40	55	3	10	24	4	4	5
31	35	9	8	24	27	45	2	8	20	3	4	6
32	35	9	15	24	27	45	2	5	20	3	5	6
33	24	11	15	35	33	65	1	5	27	2	5	6
34	30	9	8	30	27	55	2	5	24	3	5	6
35	35	9	15	24	33	45	1	5	20	3	5	6
36	24	11	8	35	27	65	2	10	27	2	3	6

4. RESULT ANALYSIS

Results in **TABLE 4** indicate that, the two efficient designs produce lower D-error value (0.08471 and 0.08846 for the OOC design and D-efficient design, respectively), while orthogonal design produces higher D-error value (0.12362). The D-error of the orthogonal design is 1.46 times greater than the D-error value of the OOC design and 1.40 times greater than that of D-efficient design.

Meanwhile, we are able to generate the AVC matrix for each design, shown in **TABLE 5**. The square roots of the diagonals of the AVC matrix represent the asymptotic standard errors of the parameter estimates. This suggests that on average, the asymptotic standard errors of the parameter estimates using the orthogonal design will be 1.18 to 1.21 times larger than these efficient designs. Clearly, the OOC and D-efficient designs are proved to be capable of providing more reliable parameter estimates than orthogonal design.

TABLE 5 Asymptotic Variance-covariance (AVC) Matrix for Case Study

	β_0^{car}	β_1	β_2	β_3	β_0^{taxi}	β_0^{bus}	β_4
AVC matrix for orthogonal design							
β_0^{car}	0.1880	-0.0014	-0.0117	-0.0043	0.1411	0.1410	-0.0017
β_1	-0.0014	0.0809	-0.0027	0.0002	0.0000	0.0009	0.0010
β_2	-0.0117	-0.0027	0.0806	0.0036	0.0030	0.0039	0.0035
β_3	-0.0043	0.0002	0.0036	0.1851	0.0003	0.0007	-0.0023
β_0^{taxi}	0.1411	0.0000	0.0030	0.0003	0.4245	0.1417	-0.0012
β_0^{bus}	0.1410	0.0009	0.0039	0.0007	0.1417	0.6064	0.0025
β_4	-0.0017	0.0010	0.0035	-0.0023	-0.0012	0.0025	0.1942
AVC matrix for optimal orthogonal choice design							
β_0^{car}	0.1907	-0.0015	-0.0199	-0.0037	0.1470	0.1379	-0.0048
β_1	-0.0015	0.0555	-0.0002	-0.0029	0.0009	-0.0001	0.0000
β_2	-0.0199	-0.0002	0.0764	0.0004	-0.0199	0.0086	0.0003
β_3	-0.0037	-0.0029	0.0004	0.1932	0.0008	0.0000	-0.0407
β_0^{taxi}	0.1470	0.0009	-0.0199	0.0008	0.4357	0.1379	-0.0069
β_0^{bus}	0.1379	-0.0001	0.0086	0.0000	0.1379	0.5984	-0.0087
β_4	-0.0048	0.0000	0.0003	-0.0407	-0.0069	-0.0087	0.1804
AVC matrix for D-efficient design							
β_0^{car}	0.1902	-0.0021	-0.0108	-0.0022	0.1411	0.1396	-0.0023
β_1	-0.0021	0.0455	0.0013	-0.0014	-0.0007	0.0010	-0.0012
β_2	-0.0108	0.0013	0.0496	0.0026	-0.0024	0.0001	-0.0039
β_3	-0.0022	-0.0014	0.0026	0.1757	-0.0026	0.0043	-0.0002
β_0^{taxi}	0.1411	-0.0007	-0.0024	-0.0026	0.4237	0.1394	-0.0028
β_0^{bus}	0.1396	0.0010	0.0001	0.0043	0.1394	0.6022	-0.0011
β_4	-0.0023	-0.0012	-0.0039	-0.0002	-0.0028	-0.0011	0.1551

On the other hand, the efficiency of the designs can be examined in terms of theoretical minimum sample size. Here, $\tilde{\beta}_k$ represents the prior value of each parameter and $se_1(\tilde{\beta}_k)$ is the asymptotic standard error of parameter assuming only a single respondent. t_T stands for the asymptotic *t*-value under the sample size T and is assume be a certain significance level (i.e. in our case study, t_T=1.98). Thus, the theoretical minimum sample size T_k^* for parameter *k* in an experimental design can be expressed as:

$$T_k^* \geq \left(\frac{t_T se_1\left(\tilde{\beta}_k\right)}{\tilde{\beta}_k} \right)^2 \qquad (18)$$

The theoretical minimum sample size of every parameter for the three different designs is calculated as shown in **TABLE 6**. Looking at individual parameters, for all the designs, the most difficult parameter for estimation (having the highest theoretical minimum sample size to be statistically significant in estimation) is β_4, needing a minimum sample size of 1903.35 in orthogonal design, 1151.62 in OOC design and 1520.14 in D-efficient design. Again, the OOC and D-efficient designs are confirmed to be much more efficient than orthogonal design, this time in terms of sample size requirements for individual parameter estimates.

TABLE 6 Theoretical Minimum Sample Size for Case Study

	β_0^{car}	β_1	β_2	β_3	β_0^{taxi}	β_0^{bus}	β_4
Minimum sample size for orthogonal design							
T_k^*	0.51	198.23	5.06	50.39	3.40	1.65	1903.35
Minimum sample size for optimal orthogonal choice design							
T_k^*	0.52	109.53	2.90	78.82	3.49	1.63	1151.62
Minimum sample size for D-efficient design							
T_k^*	0.52	111.49	3.11	47.83	3.39	1.64	1520.14

Furthermore, for OOC and D-efficient design, both of them can achieve almost the same efficiency in terms of either D-error value or theoretical minimum sample size, seeing from **TABLE 4** and **TABLE 6**. However, the complex experimental settings make it more difficult for D-efficient design to obtain the optimal solution. As a matter of fact, the current D-efficient design in **TABLE 4** catches the asymptotic optimal solution, of which the D-error equals 0.08846, after 265313 iterations. On the contrast, the computation work of OOC design is much simpler and the result can be worked out within a second. Seeing from this perspective, we can draw the conclusion that OOC design is superior to D-efficient design.

5. CONCLUSION

The SC experiment has been generally regarded as an effective method for discrete choice analysis, especially for newly introduced alternatives. Though orthogonal design has been used as the major experimental design method in practice, orthogonality is not that important in the nonlinear discrete choice models.

In this paper, a feasible approach to construct an OOC design with alternative-specific attributes is provided. To hold the original principle of OOC design, two stages are required: (1) generating the OOC design with only generic attributes using the former method, (2) adding column vectors one by one for alternative-specific attributes and making them orthogonal with other vectors within an alternative. With the proposed method, the attribute levels of generic attributes are in the maximum difference across the alternatives, while the distribution of alternative-specific attribute levels would not affect the orthogonality for each alternative.

The efficiency of the proposed method is examined by contrasting it with the conventional orthogonal design and another popular efficient design, D-efficient design. D-error, a common statistic corresponding with AVC matrix of the choice model, is chosen as the major measurement of experimental design efficiency. Also, minimum sample size based on each attribute parameter is calculated as the auxiliary criteria for the comparison.

Applying the proposed method to design a field SC survey in China, the results indicate the advantage of using OOC design in two aspects: (1) it is proved to be more capable of producing statistically significant parameter estimates than conventional orthogonal design, while has almost the same efficiency with D-efficient design. (2) The solving process of OOC design is relatively easy. The feasible solution can be obtained by a simple loop statement. Then the AVC matrix and D-error value can be calculated. On the other hand, to work out a solution using D-efficient design, the D-error value needs to be generated and compared time after time. The searching for an optimal solution is time consuming. Thus, it is believed that OOC design outperforms D-efficient design in the sense of avoiding multiple iterations and complex computation work.

ACKNOWLEDGEMENT

This study is supported by Fundamental Research Funds for the Central Universities (No. A0920502051307-03), National Natural Science Foundation of China (No. 51178403) and Sichuan Province Science and Technology Support Project (No. 2011FZ0050). Thanks for useful comments from Lin Chen. We also appreciate Dr. Yong Liao for great help in Matlab programming.

1 **REFERENCE**

[1] McFadden, D., 1986. The choice theory approach to market research. Marketing Science. V.5. no.4.

[2]Henser, D.A., Rose., J.M., Greene, W.H., 2005. Applied Choice Analysis: A Primer. Cambridge University Press, Cambridge.

[3] Louviere, J.J. and Hensher, D.A. and Swait, J.D., 2000. Stated Choice Methods: Analysis and Application, Cambridge University Press, Cambridge.

[4] Galilea, P. and Ortúzar, J.deD., 2005. Valuing noise level reductions in a residential location context. Transportation Research D 10, 305-322.

[5] Bunch, D.S., Louviere, J.J., Anderson, D., 1994. A comparison of Experimental Design Strategies for Choice-based Conjoint Analysis with Generic-attribute Multinomial Logit Models. Working Paper, Graduate School of Management, University of California, Davis.

[6] Louviere, J.J., Woodworth, G., 1983. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. Journal of Marketing Research 20, 350-367.

[7] Huber, J., Zwerina, K., 1996. The importance of utility balance and efficient choice designs. Journal of Marketing Research 33(August), 307-317.

[8]Rose, J.M., and Bliemer, M.C.J, 2005. Constructing efficient stated choice experimental design, ITLS working paper: ITLS-WP-05-07, April.

[9] Bliemer, M.C.J, Rose, J.M., Henser, D.A, 2009. Efficient stated choice experiments for estimating nested logit models, Transportation Research B 43(1), 19-35.

[10] Street, D.J., Burgess L., 2004. Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments, Journal of Statistical Planning and Inference, 118, 185-199.

[11] Burgess, L., Street, D.J., 2005. Optimal designs for choice experiments with asymmetric attributes, Journal of Statistical Planning and Inference, 134, 288-301.

[12] Street, D.J., Burgess, L., Louviere, J.J., 2006, Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments, Research in Marketing (22), 459-470.

[13] Yang, F., Li, X., Cao, S., Chen, L., 2012. Application of orthogonal design method for the logit model construction of public transport behavior and policy evaluation, working paper.

[14] Kessels, R., Goos, P. and Vandebroek, M., 2006. A comparison of criteria to design efficient choice experiments, Journal of Marketing Research,43(3), 409-419.

[15] Ferrini, S. and Scarpa, R., 2007. Designs with a-prior information for nonmarket valuation with choice-experiments: a Monte Carlo study, Journal of Environmental Economics and Management, 53, 342-363.

[16] Rose, J.M., Bliemer, M.C.J, 2007. Stated preference experimental design strategies. In: Hensher, D.A., Button, K.(Eds.), Transport Modeling, second ed., In: Handbooks in Transport, vol. 1 Elsevier Science, Oxford (Chapter 8).

[17] McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior, In Zarembka, P.(ed.), Frontiers in Econometrics, Academic Press, New York, 105-142.