NUMERICAL ANALYSIS OF THE VALIDITY OF UNIFORM DESIGN IN STATED CHOICE MODELING

Donggen Wang
Department of Geography, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong

Pengfei Li
Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada. E-mail: p4li@uwaterloo.ca

Abstract:
This paper examines the statistical properties of uniform design for stated choice modeling. Estimation efficiency, prediction efficiency and test power (i.e., the ability to pick up significant effects or exclude insignificant effects) are selected as measures of statistical properties. Both uniform design and orthogonal design are used to generate profiles. The major experimental design strategies for multinomial logit models including shifted pairs, L^{2k}, 2^J block, all pairs and McFadden's sampling rule, are used to construct choice sets from the profiles generated by both orthogonal design and uniform design. Monte Carlo experiments are used to generate models, whose parameters vary in scale. The results show that the performance of uniform design in stated choice modeling is comparable to that of orthogonal design.

1. Introduction

Stated preference/choice methods have become established modeling tools for transportation studies. At the core of these methods is experimental design, which usually employs orthogonal design. The key of orthogonal design is to balance attribute levels and maintain orthogonalities between attributes. Consequently, the profiles generated by orthogonal design are usually sizable and rapidly increase with increasing numbers of attributes and/or levels. This is particularly significant for the case of attributes with uneven levels (e.g., for a case of seven attributes, two with 2 levels, three with 3 levels and two with 5 levels, at least 900 profiles will be needed by orthogonal design). Sometimes orthogonal design provides no solution for such cases. A large number of profiles imply huge costs of model development and substantial demand from respondents. Thus, orthogonal design imposes constraints on modelers regarding the numbers of levels and attributes to use. Recently, Wang and Li (2002) introduced a new experimental design method: uniform design, which has the potential to overcome the drawbacks of orthogonal design. This method selects from the s-dimensional space ('s' refers to the number of attributes) experimental points or profiles that are uniformly (or evenly) scattered in the space. The number of profiles produced by uniform design is substantially less than that by orthogonal design. In addition, uniform design can easily handle the cases of attributes with uneven levels and provide a reasonable number of profiles.

However, the advantages of uniform design over orthogonal design are not at no cost. Uniform design only requires that attribute levels are balanced in the profiles. Orthogonality, however, is not guaranteed. As both orthogonality and balance are usually considered important requirements for experimental design, if orthogonality is not maintained in uniform design, a number of questions regarding estimation and prediction may be raised. Firstly, are parameters estimated from uniform design close enough to the true parameters? Secondly, we often need to test the significance of effects in model development. In such a case, we may make two kinds of errors: inferring the significant effects to be insignificant or the insignificant effects to be significant. How would uniform design perform in this regard? Can uniform design limit the number of errors admissible? Thirdly, will the models developed from uniform design provide predictions at acceptable accuracy level? or specifically, are the market shares predicted from uniform design close enough to the true market share? Finally, how is uniform design comparable to orthogonal design in terms of these properties?

Wang and Li (2005) made an attempt to answer these questions. However, the study considered only fixed choice set design and additive models. Other choice set deign methods were not investigated. This paper extends the previous attempt to comprehensively examine the statistical properties of uniform design for the major experimental design strategies for multinomial logit models (Bunch, Louviere and Anderson, 1996), including shifted pairs, L^{2k}, 2^J block, all pairs and McFadden's sampling rule. Estimation efficiency, prediction efficiency and test power (i.e., the ability to pick up significant effects or exclude insignificant effects) are selected as measures of statistical properties. Both uniform design and orthogonal design are used to generate profiles. The aforementioned experimental design strategies for multinomial logit models are used to construct choice sets from the profiles generated by both orthogonal design and uniform design. Monte Carlo experiments are used to generate simulated models, whose parameters vary in the scales. The three measures of statistical properties are computed for and compared between different designs. We concentrate the comparison on the stated choice models and multinomial logit model is our main concern. For stated preference models, some existing results for studying the properties of uniform design are summarized. We firstly develop some notations and asymptotic properties about the maximum likelihood and predicted market shares for stated choice models as multinomial logit model is considered, and then introduce some measures for comparing uniform design and orthogonal design. Next we describe the experiment frameworks and the Monte Carlo experiments. Results on Monte Carlo experiments are followed by a summary of available results about uniform designs. Conclusions and Discussions are shown in the last section.

2. Principles of Uniform Design

Assume the research question is to establish the model between the response variable *Y* and *S* attributes, denoted by $X_1, \cdots, X_S$. The number of levels of these attributes is denoted by $q_1, \cdots, q_S$ respectively. The total number of all possible profiles is given by *N*= $q_1 \times q_2 \times \cdots \times q_S$. The collection of the *N* profiles is denoted by C^S. Selecting *n* profiles $Z_1, \cdots, Z_n$ from C^S can get *n* response value $Y_1 = f(Z_1), \cdots Y_n = f(Z_n)$. We can estimate the expectation of *Y*: *E*(*Y*) by the mean value of the *n* response values:

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} f(Z_i)$$

The research question here is how to select the *n* points, i.e. $Z_1,\cdots,Z_n$ from C^S, so that $\bar{Y}$ is the closest to *E*(*Y*), or has the highest accuracy. The upper bound of the gap between $\bar{Y}$ and *E*(*Y*) is given by the Koksma-Hlawka inequality (Niederriter 1992):

$$\Delta_n = | E(f(Z)) - \bar{Y} | \le V(f) D(Z_1, Z_2, ..., Z_n)$$

Where *V(f)* is the variation of the integrand *f* on C^S, which we cannot control; $D(Z_1,Z_2,\cdots,Z_n)$ is a discrepancy function, whose value is a function of the design (i.e., the *n* experimental profiles). According to the upper bound, the smaller the discrepancy, the higher the accuracy of the model is. On the other hand, the more uniform that the experimental points of a design are scattered, the smaller the discrepancy of the design is. Uniform design (UD) is such an experimental design method that selects the *n* experimental profiles $Z_1,\cdots,Z_n$, which are uniformly (or evenly) scattered in C^S. Uniformity can be measured by the so-called Star L*p*–discrepancy (Hua and Wang,1981; Niederreiter, 1992) or Centered L2–discrepancy (denoted by *CD*2) (Hickernell, 1998).

To compare uniform design with orthogonal design, assume there are *S* attributes each with $q_1,\cdots,q_S$ levels respectively. The number of profiles generated by the orthogonal design is determined by *t* × *m*, where *t* is the lowest common multiple of $q_1 \times q_2, q_1 \times q_2, \cdots, q_1 \times q_S, \cdots, q_{S-1} \times q_S$. The number of profiles generated by the uniform design is determined by *t*×*m*, where *t* is the lowest common multiple of $q_1,\cdots,q_S$. For example, suppose we have a case of six attributes; two with 2 levels, two with 3 levels and two with 7 levels. The lowest common multiple of 4,6,9,14,21 and 49 is 1764 (the least number needed by orthogonal design). The lowest common multiple of 2, 3 and 7 is 42 (the least number needed by uniform design). That is why the number of profiles generated by uniform design is substantially fewer than that by orthogonal design. Nevertheless, profiles generated orthogonal design are balanced and orthogonal, but that by uniform design are only balanced but not necessary orthogonal.

3. Notations and Definition of Measures

3.1 Notations

Consider a design consists of *N* choice sets, indexed by $n = 1,...,N$, and $J_n =$ the number of choice alternatives in choice set *n* (denoted by C_n). Let z_{jn} denote the vector of explanatory variable for the jth alternative in choice set *n*, β be the parameter vector with *p* parameters and *M* be the number of respondents for a given choice set. For multinomial logit model, the probability of choosing the *j*th alternative in C_n is given by:

$$P_{jn}(\beta) = \frac{\exp(z_{jn}'\beta)}{\sum_{i \in C_n} \exp(z_{in}'\beta)},$$

for $j = 1,...,J_n$. The maximum likelihood estimator of β (denoted by $\hat{\beta}^M$) has the following properties: $\hat{\beta}^M$ is consistent and asymptotically normal and $\sqrt{NM}(\hat{\beta}^M - \beta)$ is approximately multivariate normal distribution with zero mean and a covariance matrix $I(\beta)^{-1}$ as $NM \to \infty$, where

$$I(\beta) = \frac{1}{N}\sum_{n=1}^{N}\sum_{j\in C_n} P_{jn}(z_{jn} - \bar{z}_n)(z_{jn} - \bar{z}_n)'$$

and

$$\bar{z}_n = \sum_{j\in C_n} z_{jn} P_{jn}(\beta).$$

$I(\beta)$ is called as the "normalized information matrix" for estimated parameters (Bunch, *et al*, 1996).

Let C_T be the test choice set with J_T alternatives that we want to predict the market shares and x_j be the vector of explanatory variables for the jth alternative in C_T . $P_{jT}(\beta)$ and $\bar{x}$ are defined similarly as above. Denote $P(\beta, C_T)$ be a $(J_T - 1)\times 1$ vector with the *j*th element $P_{jT}(\beta)$ for $j = 1,..., J_T - 1$. Applying the delta method based on a truncated Taylor series expansion, we can get $P(\hat{\beta}, C_T)$, which is also consistent and asymptotically normal, and $\sqrt{NM}(P(\hat{\beta}^M, C_T) - P(\beta, C_T))$, which is approximately multivariate normal distribution with zero mean and a covariance matrix as $NM \to \infty$,

$$COV(\beta, C_T) = X' I(\beta)^{-1} X ,$$

where X is a matrix with the *j*th column $P_{jT}(\hat{\beta}, C_T)(x_j - \bar{x})$ for $j = 1,..., J_T - 1$. Similarly as above, $(X' I(\beta)^{-1} X)^{-1}$ is called as the "normalized information matrix" for predicted market shares. (Since $P_{J_T T}(\beta)$ can be determined by the first $J_T - 1$ predicted probabilities and the covariance matrix will be singular if we include $P_{J_T T}(\beta)$ in $P(\beta, C_T)$. Furthermore, if we can predict $P(\beta, C_T)$ accurately, the predictors of $P_{J_T T}(\beta)$ will also be accurate.)

3.2 Definition of measures

According to Lothar (1984) and Shao (1999), statistical properties include estimation efficiency, prediction efficiency and test power. For estimation efficiency, a good design is expected to estimate parameters as closely as possible to the true values. For prediction efficiency, a good design is expected to predict the market shares as accurately as possible. In many textbooks, prediction is considered as a special case of estimation. However, in conjoint analysis, the estimation of parameters and the prediction of market shares are treated equally important. Therefore, we need to analyze both estimation and prediction efficiency. Apart from these two properties, the ability to pick up significant attributes, i.e., test power, is also very important. A good design is expected to minimize the probability of making erroneous inference about the significance of the effects of attributes. Bunch and Batsell (1989) compared different estimators for a given design from the viewpoints of estimation and prediction efficiency. Bunch, et al (1996) compared different choice set strategies for profiles generated by orthogonal designs in terms of estimation efficiency. In this research, our objective is to compare the performance of uniform designs and orthogonal designs in terms of estimation efficiency, prediction efficiency and test power.

Estimation efficiency is often measured by the errors around the estimated parameters. The determinant of $I(\beta)^{-1}$ is widely used for assessing these errors (Bunch, Louviere and Anderson 1996, Kuhfeld, Tobias and Garratt 1994, Huber and Zwerina, 1996). Thus, the measure of estimation efficiency is defined as follows:

$ED = \det(I(\beta)^{-1})^{1/p}$.

The smaller the *ED* value, the shorter the lengths of confidence intervals are and the more efficient the estimators are.

In order to evaluate prediction efficiency, a test choice set consisting of four alternatives is generated, the attribute levels of which are randomly distributed between the upper bounds and lower bounds of the original attribute levels. Similar to estimation efficiency, the determinant of $COV(\beta, C_T)$ is used to measure the errors around the predicted market shares. That is,

$PD = \det(X' I(\beta)^{-1} X)^{1/(J_T - 1)}$.

The smaller the PD value, the more accurate the predicted market shares are.

There are two types of errors in significance test: type-I error which infers non-significant effects to be significant and type-II error which infers significant effects to be non-significant. The probability of making type-I error is often denoted by α, which is referred to significance level. Next we will consider how to calculate the probability of making type-II error. In significance test, if $\hat{\beta}^M{}_k$ lies in the interval $(-Z_{\alpha/2} \cdot \sigma_k\ , Z_{\alpha/2} \cdot \sigma_k)$, we will infer that the effect of k is not significant, where σ_k is the standard deviation of $\hat{\beta}^M{}_k$ and $Z_{\alpha/2}$ is upper $\alpha/2$ quintile of normal distribution. Thus the probability of making type-II error is $P(-Z_{\alpha/2} \cdot \sigma_k < \hat{\beta}^M{}_k < Z_{\alpha/2} \cdot \sigma_k)$ if $\beta_k \neq 0$. After simple calculation, we can get

$$PE_k = P(-Z_{\alpha/2} \cdot \sigma_k < \hat{\beta}^M{}_k < Z_{\alpha/2} \cdot \sigma_k) = \Phi(Z_{\alpha/2} - \frac{\beta_k}{\sigma_k}) - \Phi(-Z_{\alpha/2} - \frac{\beta_k}{\sigma_k}),$$

where $\Phi(x)$ is the cumulative distribution function of normal distribution with zero mean and one variance. Then the measure of test power is the average of probability of making type-II error, denoted by *APE*. The smaller the *APE*, the stronger the ability of a design is in picking up the significant effects.

While, *ED* and *PD* are considered as the large-sample or asymptotic properties for a given design, *APE* is a combination of small-sample and large-sample properties.

4. Study Design

Bunch and Bastell (1989) studied the performance of different estimators by considering seven different designs and then generalizing the result. We adopt this approach to compare the performance of orthogonal design and uniform design.

4.1 Choice set design

Let' us consider a design task, which involves three 2-level attributes and two 3-level attributes. By applying orthogonal design, a total of 36 profiles are needed. Using uniform design, one may generate 12 or 18 profiles. In order to apply the L^{2k} strategy for designing choice set, we select the 18 profiles design, i.e., $U(18, 2^3 3^2)$, which is listed in the Appendix. The profiles generated by the two design methods may be used to form choice sets by the following six commonly used strategies: all pairs, 2^J Block, BIBD, L^{lk} strategy, foldover and shifted designs. For details of these strategies, readers are referred to Bunch, Louviere and Anderson (1996). Since the foldover strategy may not guarantee the estimation of partworth utility model and the performance of BIBD is similar to that of 2^J block, we exclude these two strategies but consider only shifted pairs, L^{2k}, all pairs and 2^J Block. For profiles generated by orthogonal designs, L^{2k} strategy can guarantee the orthogonality between and within alternatives. Shifted pairs strategy, which includes foldover strategy as a special case for the two-level experiment, can only guarantee the orthogonality within alternatives. 2^J Block strategy makes the alternatives pairwise independent of each other across choice sets. Both 2^J block and all pairs strategies do not guarantee the orthogonality between and within alternatives. Table 1 compares the number and average size of choice sets by orthogonal design and uniform design for the four strategies of designing choice sets.

Table 1: Candidate designs for comparison study

Design	Number of choice sets	Average size of choice sets
By Orthogonal design		
Shifted pairs	36	2
2^J Block	40	18
L^{2k}	36	2
All pairs	630	2
By Uniform design		
Shifted pairs	18	2
2^J Block	20	9
L^{2k}	18	2
All pairs	153	2

Table 1 shows that the number of choice sets by orthogonal design is significantly larger than that by uniform design. In addition to the eight designs generated by the four strategies, we add one more choice set design, which is generated simply by sampling rules. Therefore, in total there are nine choice set designs in our experiments. Table 2 lists these designs.

4.2 Generating models for simulation

The true model affects the three measures of statistical properties defined earlier (i.e., *ED*, *PD* and *APE*). Different models will generate different results. In order to obtain reliable and general findings, we use Monte Carlo experiments to generate different models (similar approach is adopted by Bunch and Bastell, 1989 and Bunch, Louviere and Anderson, 1996).

Table 2: List of choice set designs

Designs	Profiles generated method	Choice set strategy
D_1	Orthogonal design	Shifted pairs
D_2	Orthogonal design	L^{2k}
D_3	Orthogonal design	All pairs
D_4	Orthogonal design	2^J Block
D_5	Uniform design	Shifted pairs
D_6	Uniform design	L^{2k}
D_7	Uniform design	All pairs
D_8	Uniform design	2^J Block
D_9	Orthogonal design	Sampling rule

Define the scale of parameters $Scale(\beta)=\sqrt{\sum_{k=1}^{p}\beta_k{}^2}$. By varying $Scale(\beta)$, we may generate different models. Chapman and Staelin (1982) discussed the effect of $Scale(\beta)$ and its relationship to the interpretation of the MNL model, which is referred to "Scale Theorem" in their article. The theorem tells that the larger the scale, the more extreme the probabilities are. We adopt their approach of standardizing the attributes to have zero mean and the variance of one. Since the true model is unknown to the researchers, it is useful to examine the three properties over a range of possibilities.

In order to differentiate two kinds of effects and guarantee the significant effects are significant different from errors, the models only include all linear main effects of the attributes. The true values of the parameters are generated by drawing 5 independent values from uniform distribution with a range of 0.2 to one and then rescaled so that $Scale(\beta)$ is satisfied. For the signs of the parameters, we assume that the partworth utilities monotonically increase along with the level increase for each attribute. That is to say, all signs for significant effects are positive. We select three scale values: 20% Exp Var, 45% Exp Var and 70% Exp Var. In these three situations, the $Scale(\beta)$'s are about 0.64, 1.16 and 1.96 respectively.

Apart from the model, the number of respondents M for each choice set is also influential on *APE* or the ability to pick up significant effects. To comprehensively compare *APE* values for different designs, we select 10 values for M: 10, 20, … , 100.

To account for the randomness, we generate 100 models for each scale value. For each model, we can calculate *ED* and *PD* for each of the nine designs. To compare the test power or the ability to pick up significant effects, we can also obtain 100 *APE* values for a given M and scale value.

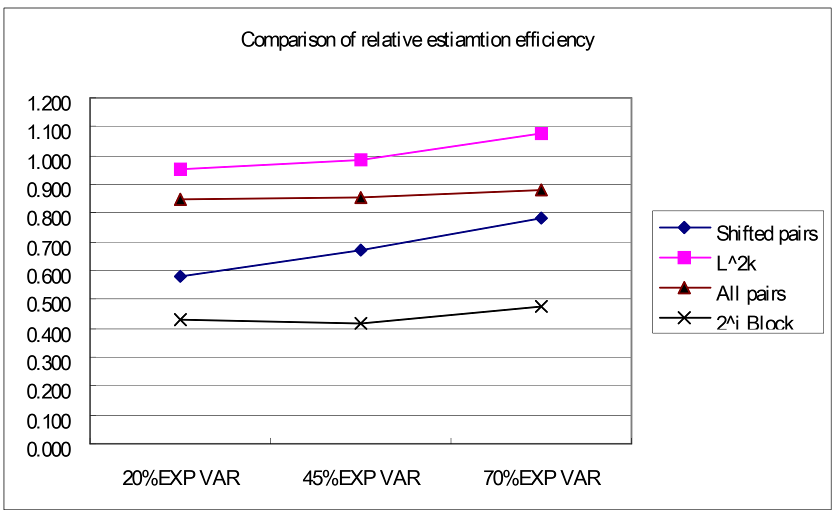
5. Discussion of Results

Based on the values of the three measures for the nine designs, we may now compare the statistical properties of uniform design with that of orthogonal design in terms of estimation efficiency, prediction efficiency and test power.

5.1 Estimation efficiency

As explained earlier, for each of the nine designs and the three scale values, a total of 100 *ED* values are produced. To evaluate relative estimation efficiency of uniform design to orthogonal design for the four strategies of designing choice sets and that of uniform design to sampling rule, we may analyze these data in two ways. Firstly, we may calculate the following eight ratios of *ED* values: $ED(D_1)/ED(D_5)$, $ED(D_2)/ED(D_6)$, $ED(D_3)/ED(D_7)$, $ED(D_4)/ED(D_8)$, $ED(D_5)/ED(D_9)$, $ED(D_6)/ED(D_9)$, $ED(D_7)/ED(D_9)$ and $ED(D_8)/ED(D_9)$. The 25% quartile, the median and the 75% quartile of the first four ratios for each scale value are listed in Table 3. The medians of the last four ratios are plotted in Figure 1. Secondly, we can use the idea of pairwise comparison to test whether *ED* values of orthogonal design, uniform design and sampling rule are the same for the given choice set strategy and scale value. The test statistic has *t*-distribution with 99 degree of freedoms. The results are listed in Table 4.

Figure 1: Comparing the relative estimation efficiency of sampling rule to uniform design



From Tables 3 and 4, we can see that for the design strategy of all pairs, uniform design has higher estimation efficiency than orthogonal design. All the test statistics are significant at the 0.05 significance level and the medians of ratios are larger than 1. For shifted pairs, the medians of the relative estimation efficiency for uniform design to orthogonal designs are larger than 0.98 and the test statistics are not significant for the last scale value, which suggests that the estimation efficiency of uniform design is not significantly different from that of orthogonal design. The estimation efficiency of orthogonal design is slightly higher than that of uniform design for L^{2k} and 2^J Block. For L^{2k}, the estimation efficiency of uniform design is on average 7% less than that of orthogonal design. In the case of 2^J Block, the advantage of the orthogonal design is more obvious especially for large scale values. The relative estimation efficiency of uniform design is 5%, 10% and 15% lower than that of orthogonal design for the three scale values, respectively. Why the relative estimation efficiencies decrease as scale values increase? Note that the average size of choice sets by orthogonal design is two times that by uniform design. Then utility by orthogonal design is more balanced than that by uniform design and the extreme probability in orthogonal design is much less than that of uniform design as the scale value increases. As a result, orthogonal design becomes more efficient than uniform design as scale value increases (Huber and Zwerina, 1996).

Table 3: Comparing the relative estimation efficiency of uniform design to orthogonal design

	$Scale(\beta) =$ 20% Exp Var	$Scale(\beta) =$ 45% Exp Var	$Scale(\beta) =$ 70% Exp Var
$ED(D_1)/ED(D_5)$			
25% Quantile	0.975	0.937	0.897
Median	0.983	0.988	1.024
75% Quantile	0.989	1.030	1.129
$ED(D_2)/ED(D_6)$			
25% Quantile	0.931	0.894	0.833
Median	0.937	0.929	0.925
75% Quantile	0.947	0.957	1.049
$ED(D_3)/ED(D_7)$			
25% Quantile	1.011	0.994	0.989
Median	1.015	1.017	1.037
75% Quantile	1.019	1.032	1.069
$ED(D_4)/ED(D_8)$			
25% Quantile	0.943	0.894	0.810
Median	0.952	0.907	0.855
75% Quantile	0.960	0.929	0.916

As for the comparison between uniform design and sampling rule, if the L^{2k} strategy is used, uniform design has slightly higher estimation efficiency than sampling rule for small scale value and conversely for large scale value. If the other three strategies are used, sampling rule is much worse than uniform design, especially when scale value is small.

Table 4: Testing the equivalence of *ED* values

Null Hypothesis	$Scale(\beta) =$ 20% Exp Var	$Scale(\beta) =$ 45% Exp Var	$Scale(\beta) =$ 70% Exp Var
$ED(D_1) = ED(D_5)$	-15.678	-3.942	0.464
$ED(D_2) = ED(D_6)$	-60.500	-13.816	-4.061
$ED(D_3) = ED(D_7)$	22.967	5.255	5.480
$ED(D_4) = ED(D_8)$	-43.663	-29.255	-14.900
$ED(D_5) = ED(D_9)$	-198.712	-60.095	-17.607
$ED(D_6) = ED(D_9)$	-24.951	-4.081	4.071
$ED(D_7) = ED(D_9)$	-88.978	-33.120	-14.914
$ED(D_8) = ED(D_9)$	-349.538	-132.742	-63.504

Note: $t_{0.975}(99) = 1.984$ and $t_{0.95}(99) = 1.660$

The estimation efficiencies of the four strategies decrease in the following sequence: 2^J Block>Shifted pairs>All pairs> L^{2k} if uniform design is used to construct profiles. Similar findings for orthogonal design were reported by Bunch, Louviere and Anderson (1996).

5.2 Prediction efficiency

In similar ways, we may compare the prediction efficiency by calculating eight ratios between the prediction efficiency of uniform designs and that of orthogonal design. The first four ratios are shown in Table 5. To account for randomness, the 25% quartile, the median and the 75% quartile are tabulated. The medians of the last four ratios are plotted in Figure 2. The values of test statistics for testing the null hypothesis that the prediction efficiency is equal are presented in Table 6.

Table 5 shows that uniform design has slightly stronger prediction ability than orthogonal design for the shifted pairs and all pairs strategies. The relative prediction efficiency of uniform design is on average 2% higher for shifted pairs and more than 6% higher for all pairs. Furthermore, the test statistic for shifted pairs and all pairs are significant for all scale values at the 0.05 significance level, which means the predicted market shares obtained from uniform design are more accurate than that from orthogonal design. For L^{2k} , uniform design has approximately the same prediction efficiency as orthogonal design. The relative prediction efficiency is only about 1% different at small scale value (20% Exp Var) and at large scale values (45% Exp Var or more). For 2^J Block, orthogonal design is more efficient than uniform design, especially for large scale values. The prediction efficiencies of uniform design are about 15%, 26% and 38% less for the three scale values, respectively. The larger size of choice sets by orthogonal design may also be the major reason for this large difference.

Table 5: Comparing the relative prediction efficiency of uniform design to orthogonal design

	$Scale(\beta)=$ 20% Exp Var	$Scale(\beta)=$ 45% Exp Var	$Scale(\beta)=$ 70% Exp Var
$PD(D_1)/PD(D_5)$			
25% Quantile	1.007	0.948	0.900
Median	1.022	1.020	1.030
75% Quantile	1.040	1.076	1.209
$PD(D_2)/PD(D_6)$			
25% Quantile	0.970	0.944	0.890
Median	0.989	1.012	1.012
75% Quantile	1.009	1.064	1.184
$PD(D_3)/PD(D_7)$			
25% Quantile	1.059	1.052	1.054
Median	1.065	1.083	1.118
75% Quantile	1.071	1.110	1.172
$PD(D_4)/PD(D_8)$			
25% Quantile	0.835	0.675	0.550
Median	0.848	0.727	0.622
75% Quantile	0.871	0.767	0.717

Figure 2: Comparing the relative prediction efficiency of sampling rule to uniform design

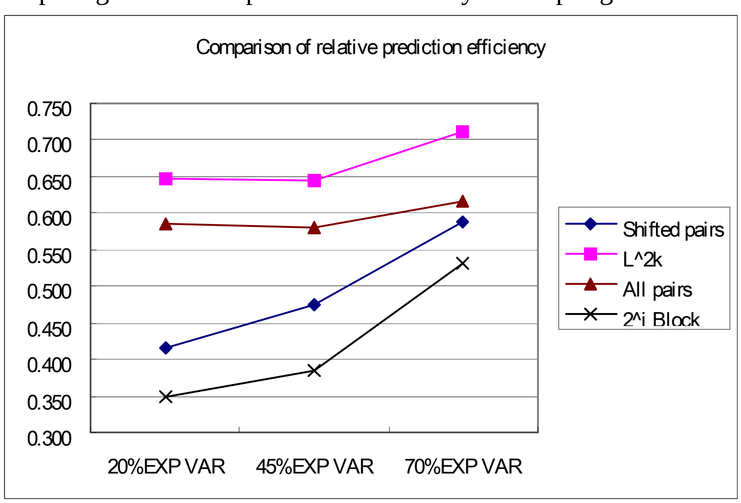


Table 6. Testing the equivalence of *PD* values

Null Hypothesis	$Scale(\beta)=$ 20% Exp Var	$Scale(\beta)=$ 45% Exp Var	$Scale(\beta)=$ 70% Exp Var
$PD(D_1)=PD(D_5)$	9.333	1.839	3.116
$PD(D_2)=PD(D_6)$	-4.439	0.750	0.434
$PD(D_3)=PD(D_7)$	71.597	23.232	15.617
$PD(D_4)=PD(D_8)$	-45.483	-41.176	-22.071
$PD(D_5)=PD(D_9)$	-153.309	-46.971	-21.979
$PD(D_6)=PD(D_9)$	-146.832	-50.446	-19.164
$PD(D_7)=PD(D_9)$	-158.432	-56.222	-27.103
$PD(D_8)=PD(D_9)$	-185.208	-65.213	-32.877

Note: $t_{0.975}(99)=1.984$ and $t_{0.95}(99)=1.660$

For sampling rule and uniform design, the prediction efficiency of uniform design is much higher than sampling rule, which can be easily seen in Figure 2. Similar to estimation efficiency, the prediction efficiencies of the four strategies by uniform design follow this order: 2^J Block>Shifted pairs>All pairs> L^{2k}.

5.3 Test power

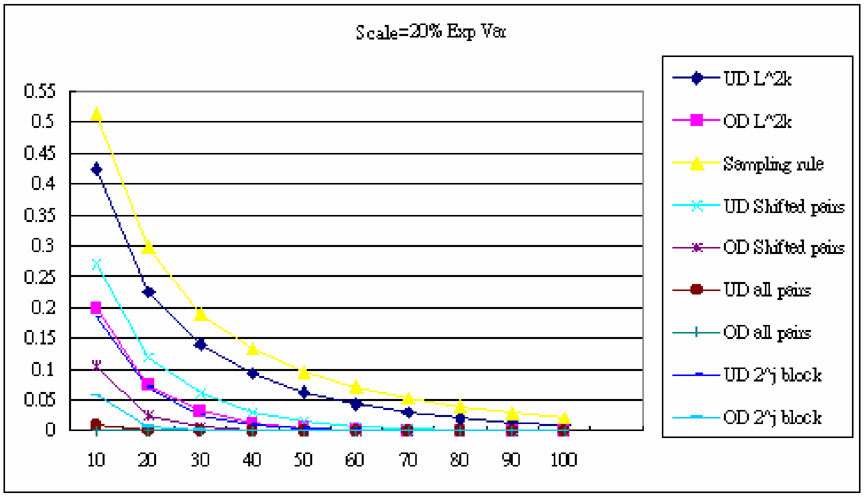
To compare the test power or the ability to pick up significant effects, we calculate a median APE value for each of the nine designs and for each *M* and scale value. The median APE values of the nine designs vs. *M* are plotted in Figure 3.

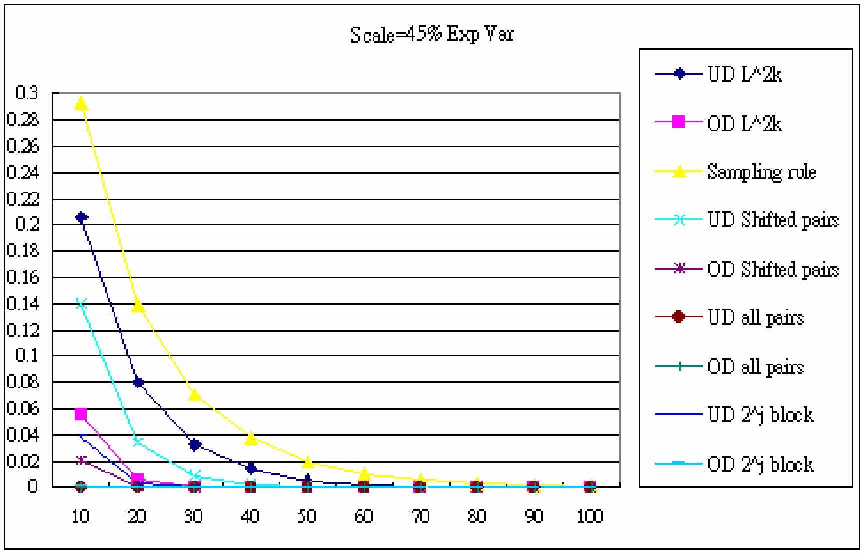
From the figure, the following findings may be derived: Firstly, when MN is the same, the *APE's* of uniform design and orthogonal design are approximately the same for all four choice set design strategies. Secondly, uniform design is superior to sampling rule in test power for all four strategies. Thirdly, 2^J Block has the strongest ability to pick up significant effects and is followed by shifted pairs, all pairs and L^{2k} for both orthogonal design and uniform design. Fourthly, when M becomes large, uniform design and orthogonal design has approximately the same *APE*, for example, *M*=60 in 20% Exp Var, *M*=30 in 45% Exp Var and 70% Exp Var. For shifted pairs the difference in *APE* values is less than 0.1%.

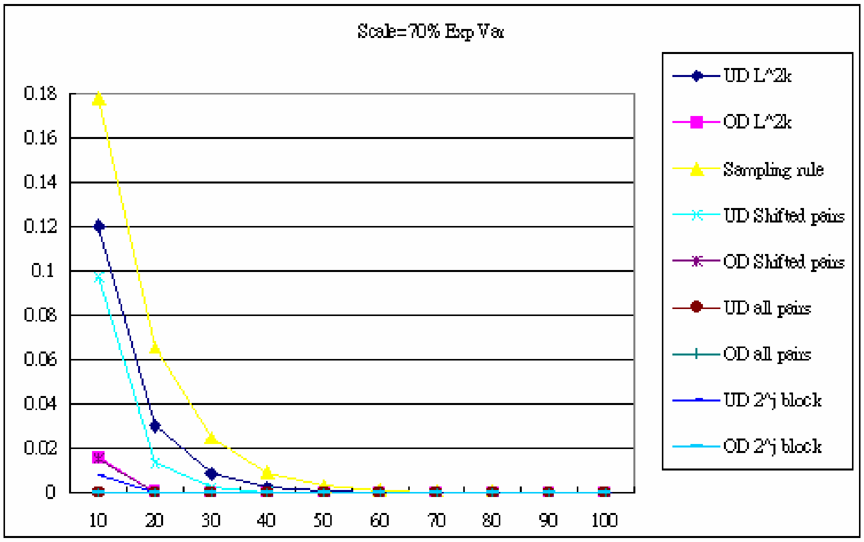
6. Conclusions

Orthogonal design and uniform design are two experimental designs to generate profiles. The advantage of uniform design over orthogonal design is that uniform design generates substantially fewer profiles particularly for attributes with uneven levels. In this research, we compared uniform design with orthogonal design and sampling rule on three statistical properties when shifted pairs, L^{2k} strategies, all pairs and 2^J block strategies are used to generate choice sets. From the comparison, the following conclusions are made.

Figure 3: Comparison of *APE* for three scale Values







In terms of estimation efficiency, when shifted pairs and L^{2k} strategies are used to construct choice sets, orthogonal design is slightly better than uniform design. The relative estimation efficiency will decrease no more than 2% and 7.5% respectively for shifted pairs and L^{2k}. The prediction efficiency of uniform design may is 2% higher than that of orthogonal design if shifted pairs strategy is used and the prediction efficiencies are approximately the same for orthogonal design and uniform design if L^{2k} strategy is used. For the strategy of all pairs, uniform design seems more efficient than orthogonal design in both estimation and prediction. Maybe due to the larger average size of choice set, orthogonal design performs better in both estimation and prediction efficiencies when the 2^J block strategy is used. However, when the scale value of the parameters is small, for example, less than 20% Exp Var, the result is also admissible (about 95% for estimation efficiency and 85% for prediction efficiency) for 2^J block strategy. When MN is the same, the *APE's* of uniform design and orthogonal design are approximately the same for four choice set strategies. Therefore for the first three choice set strategies, uniform design is a good choice if the number of profiles of orthogonal design is not applicable. And when the scale value is smaller than 20% Exp Var, uniform design is also acceptable for 2^J block strategy, since uniform design can not only reduce the number of choice set size but also reduce the number of average choice set size.

References:

Bunch D S, Bastell R R, 1989 "A Monte Carlo comparison of estimators for the multinomial logit model" *Journal of Marketing Research*, 26 (February) 56-68.

Bunch D S, Louviere J J, Anderson D A, 1996 "A comparison of experimental design strategies for choice-based conjoint analysis with generic-attribute multinomial logit model" Working paper.

Chapman R G, Richard S, 1982 "Exploiting rank-ordered choice data within the stochastic utility model" *Journal of Marketing Research*, 19 (August), 288-301

Chatterjee K, Fang K T, Qin H, 2001, "Uniformity in factorial designs with mixed levels". Technical Report MATH-322, Hong Kong Baptist University

Cheng, C S 1980, "Orthogonal arrays with variable numbers of symbols " *Annals of Statistics*, 8, 447-453

Dey A, 1985, "Orthogonal fractional factorial designs", John Wiley and Sons.

Eccleston J A, Hedayat A, 1974, "On the theory of connected designs: characterization and optimality". *The annals of statistics* **2**, 1238-1255

Fang K T, Lin D K J, Winker P, Zhang Y, "Uniform design: theory and application". *Technometrics* **42 (3)**, 237-248

Hickernell F J, 1999, "Goodness-of-fit statistics, discrepancies and robust design". *Statist.&Prob. Letters*. **44**, 73-78

Hickernell, F. J. (1998). "A Generalized Discrepancy and Quadrature Error Bound." *Math. Comp.* 17, 299-322.

Hickernell F J, Liu M Q, 2002, "Uniform designs limit aliasing". *Biometrika* **89 (4)**, 893-904

Hua, L. K. and Y. Wang (1981). *Application of Numbers Theory to Numerical Analysis*. Springer and Science Press, Berlin and Beijing.

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

Kuhfeld W F, Tobias R D, Garratt M, 1994 “Efficient experimental design with marketing research application” *Journal of Marketing Research*, 31 (November), 545-557

Lin D K J, “Another look at first-order saturated designs: the *p*-efficient designs”. *Technometrics* **35**, 284-292

Lothar S, 1984, “Applied statistics: a handbook of techniques ”. New York: Springer

McFadden D, 1974 “Conditional logit analysis of qualitative choice behavior” in *Frontiers in Econometrics*, Paul Zarembka, Ed. New York: Academic Press, 105-142

McFadden D, 1978 “Modelling the Choice of Residential Location” in A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (eds.), *Spatial Interaction Theory and Planning Models*, 75-96

Niederreiter, H. (1992). “Random Number Generation and Quasi-Monte Carlo Methods.” *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics,* Philadelphia.

Nguyen N K, “Cutting experimental designs into blocks”. “*Australian&New Zealand Journal of Statistics*”. **43**, 489-498

Shao J, 1999, **“**Mathematical statistics”. New York: Springer.

Xie M Y, Fang K T, 2000, “Admissibility and minimaxity of the uniform design in nonparametric regression model”. *J. Statist. Plan. Inference*, **83**, 101-111

Xu H, 2002, “An algorithm for constructing orthogonal and nearly-orthogonal arrays with mixed levels and small runs”. *Technometrics*, **44**, 356-368

Xu H, Wu C F J, 2001, “Generalized minimum aberration for asymmetrical fractional factorial designs”. *Annals of Statistics* **29**, 549-560

Wang D G, Li J K, 2002, "Handling large numbers of attributes and/or large Numbers of levels in conjoint experiments". *Geographical Analysis* **34 (4),** 350-362

Weins D P, 1991, “Designs for approximately linear regression: two optimality properties of uniform designs”. *Statist.&Prob. Letters*. **12** 2308-2042

Appendix: uniform designs $U(18, 2^3 3^2)$ and $U(18, 2^6 3^4)$

1	2	3	4	5	6	7	8	9	10
0	0	0	0	0	0	0	0	0	1
0	0	0	1	0	1	0	1	2	0
0	0	1	0	1	0	0	2	1	2
0	0	1	1	1	1	1	0	2	2
0	1	0	0	1	1	1	1	1	1
0	1	0	1	1	0	1	2	0	0
0	1	1	0	0	1	2	0	1	0
0	1	1	1	0	0	2	1	0	2
0	0	0	0	0	0	2	2	2	1
1	1	1	1	1	1	0	0	0	1
1	1	1	0	1	0	0	1	2	0
1	1	0	1	0	1	0	2	1	2
1	1	0	0	0	0	1	0	2	2
1	0	1	1	0	0	1	1	1	1
1	0	1	0	0	1	1	2	0	0
1	0	0	1	1	0	2	0	1	0
1	0	0	0	1	1	2	1	0	2
1	1	1	1	1	1	2	2	2	1

Note: 1. The above design is $U(18, 2^6 3^4)$ with the $e_{wd} = 0.9822$.

2. Choosing columns 1, 2, 3, 7 and 8 can generate $U(18, 2^3 3^2)$ with $e_{wd} = 0.9967$.

Table of Figure 1. Comparing the relative estimation efficiency of sampling rule to uniform design

	$Scale(\beta) =$ 20% Exp Var	$Scale(\beta) =$ 45% Exp Var	$Scale(\beta) =$ 70% Exp Var
$ED(D_5)/ED(D_9)$			
25% Quantile	0.575	0.622	0.727
Median	0.583	0.670	0.783
75% Quantile	0.592	0.698	0.872
$ED(D_6)/ED(D_9)$			
25% Quantile	0.934	0.925	0.973
Median	0.951	0.983	1.073
75% Quantile	0.964	1.027	1.167
$ED(D_7)/ED(D_9)$			
25% Quantile	0.838	0.820	0.828
Median	0.847	0.851	0.880
75% Quantile	0.856	0.886	0.928
$ED(D_8)/ED(D_9)$			
25% Quantile	0.428	0.405	0.448
Median	0.432	0.417	0.477
75% Quantile	0.440	0.429	0.502

Table of Figure 2. Comparing the relative prediction efficiency of of sampling rule to uniform design

	$Scale(\beta)=$ 20% Exp Var	$Scale(\beta)=$ 45% Exp Var	$Scale(\beta)=$ 70% Exp Var
$PD(D_5)/PD(D_9)$			
25% Quantile	0.406	0.435	0.523
Median	0.415	0.474	0.588
75% Quantile	0.428	0.527	0.688
$PD(D_6)/PD(D_9)$			
25% Quantile	0.633	0.607	0.635
Median	0.646	0.645	0.710
75% Quantile	0.659	0.683	0.811
$PD(D_7)/PD(D_9)$			
25% Quantile	0.582	0.558	0.561
Median	0.585	0.581	0.616
75% Quantile	0.593	0.598	0.651
$PD(D_8)/PD(D_9)$			
25% Quantile	0.338	0.368	0.480
Median	0.350	0.385	0.530
75% Quantile	0.359	0.404	0.572