Student: Juyoung Choi

Other students: Ping-Han Huang

Course: STP 598 (Case Studies in Design & Analysis of Experiments)

Program: Statistics MS

Instructor: Ming-Hung Kao

Date: Spring 2020

Comparison of Definitive Screening Designs and Standard Screening Designs

by

Ping-Han Huang

Juyoung Choi

Individual Effort Report

In completing this project, we equally divided all the tasks of the presentation and the report. Each of us took half of the tasks and completed it based on our discussion in periodic meetings and email correspondence. Specifically, Juyoung Choi was in charge of composing the parts of introduction and theoretical background and writing the PowerPoint presentation; Ping-Han Huang was responsible for composing the parts of simulation and discussion/conclusion and writing the project report. By typing down our names below, we all agree that this is a fair teamwork and appreciate each other's contribution.

Ping-Han Huang, April 23, 2020   *Juyoung Choi, April 23, 2020*

Table of Contents

1. Introduction

Screening designs are a kind of experimental designs used to screen multiple factors and identify active effects in the early stage of an investigation. The most common standard screening designs are Plackett-Burman designs and Resolution III/IV fractional factorial designs. While they are widely adopted in a variety of fields, these designs sometimes contain some undesirable features. For example, Plackett-Burman and Resolution III fractional factorial designs may have issues of aliasing some main effects with two-way interactions. Resolution IV fractional factorial designs may confound two-way interactions with each other. Additionally, in most of the standard screening designs the maximum of factor levels allowed is two, which may cause problems in the case where multilevel factors are needed to better address the factor-response relationship in the experiments.

In view of these limitations of standard screening designs, many researchers and investigators today adopt another kind of designs called Definitive Screening Designs (DSDs). DSDs are considered a more efficient way that not only avoids model ambiguity but also allow further analyses on nonlinear effects. In this study, we aim to structure a critical review of the fundamentals of DSDs, mainly focusing on how DSDs offer advantages over standard screening designs; and conduct a computer simulation to verify our findings from the review and further provide some insights on the application of DSDs.

2. Review of DSD Fundamentals

2.1 Structure

Definitive Screening Designs (Jones and Nachtsheim, 2011) are a class of statistical experimental designs that can be used for both screening and optimization. DSDs are able to handle both of quantitative and categorical factors. In this study, we focus on the case with three-level continuous factors. To take a close look at the DSD structure, here we introduce a full quadratic model with m factors that is assumed to follow a linear combination of main effects, two-way interactions, and quadratic effects:

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{i,j} + \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \beta_{jk} x_{i,j} x_{i,k} + \sum_{j=1}^{m} \beta_{jj} x_{i,j}^2 + \varepsilon_i$$

$$i = 1, \dots, 2m + 1,$$

where y_i is the response variable in the ith run, $x_{i,j}$ is the jth factor in the same i^{th} run, m represents the number of factors, $\beta_0, \dots, \beta_{mm}$ are the model coefficients, and the ε_i is the random error following an iid $N(0, \sigma^2)$. In a matrix form, we have $y = X\beta + \varepsilon$, where X is a 2m+1 by (m+2)(m+1)/2 model matrix.

Given the model statement above, the general design structure for m factors follows the form in the table. For m factors, there are 2m+1 runs based on m fold-over pairs and the overall center point. Each run has exactly one factor level at its center point and the others at extremes. The values of ±1 entries in the odd-numbered runs are determined by an optimal design algorithm. The values of ∓1 entries in the even-numbered runs are generated by the fold-over operation. Following up this structure, DSDs use the algorithm presented in the next section to perform screening.

Design Structure for m Factors

Foldover pair	Run (i)	$x_{i,1}$	$x_{i,2}$	$x_{i,3}$	$\cdots$	$x_{i,m}$
1	1	0	±1	±1	$\cdots$	±1
	2	0	∓1	∓1	$\cdots$	∓1
2	3	±1	0	±1	$\cdots$	±1
	4	∓1	0	∓1	$\cdots$	∓1
3	5	±1	±1	0	$\cdots$	±1
	6	∓1	∓1	0	$\cdots$	∓1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
m	2m − 1	±1	±1	±1	$\cdots$	0
	2m	∓1	∓1	∓1	$\cdots$	0
Centerpoint	2m + 1	0	0	0	$\cdots$	0

Factor levels span the columns $x_{i,1}$ through $x_{i,m}$.

2.2 Algorithm

The DSD algorithm consists of two parts, creating an initial design and optimizing the design. The specific steps in each part are as follows:

(1) Creating an initial design

- Assigning zeros to the underlying diagonal entries (as illustrated in the table above).
- Assigning -1 or 1 randomly to the odd-numbered rows of the rest of entries.
- Filling the even-numbered rows with the values in odd-numbered rows multiplied by -1.

(2) Optimizing the design

- In the nonzero entry of the first row, changing the value to 1 or -1 and the other value in the same fold-over pair to -1 or 1 respectively.
- Computing the determinant of the information matrix and changing the values in the second nonzero entry if the determinant is not improved.
- Repeating this process until no replacement occurs and then shuffling the rows of the design to make sure it is randomized.

4

2.3 Features

DSDs have seven preferable features over standard screening designs:

(1) The number of runs required is 2m+1, which is only a bit larger than twice of the number of factors.

(2) Main effects are completely independent of two-way interactions.

(3) Each two-way interaction is not completely confounded with the other two-way interactions, even though they may be correlated.

(4) All quadratic effects are estimable for continuous factors.

(5) Quadratic effects are orthogonal to main effects and not completely confounded with interaction effects.

(6) With the number of factors ranging from 6 to 12, DSDs are able to estimate all possible full quadratic models involving three or fewer factors with high statistical efficiency.

(7) The designs are able to detect nonlinearity and identify the responsible factors without confounding any terms up to second order.

3. Simulation

3.1 Settings

After reviewing the fundamentals of DSDs in the previous section, now we are curious how exactly DSDs perform compared to standard screening designs. In this section we conduct a computer simulation using SAS JMP to compare the performance of one DSD and two Plackett-Burman designs. Specifically, we construct a 13-run DSD, a 12-run Plackett-Burman design (denoted as PB design 1), and a 12-run PB augmented with one center point (denoted as PB design 2) that aims to match the run size of the 13-run DSD, as illustrated in Figure 1 below.

Figure 1

Design Structure

13-run DSD

	X1	X2	X3	X4	X5	X6
1	-1	1	0	1	-1	-1
2	1	1	-1	0	-1	1
3	1	-1	0	-1	1	1
4	1	-1	1	1	-1	0
5	0	-1	-1	-1	-1	-1
6	0	1	1	1	1	1
7	-1	0	1	-1	-1	1
8	0	0	0	0	0	0
9	-1	1	-1	-1	1	0
10	1	0	-1	1	1	-1
11	-1	-1	-1	1	0	1
12	1	1	1	-1	0	-1
13	-1	-1	1	0	1	-1

12-run PB

	Pattern	X1	X2	X3	X4	X5	X6
1	---+--	-1	-1	-1	1	-1	-1
2	++---+	1	1	-1	-1	-1	1
3	+++---	1	1	1	-1	-1	-1
4	+---+-	1	-1	-1	-1	1	-1
5	+-+++-	1	-1	1	1	1	-1
6	-+--+-	-1	1	-1	-1	1	-1
7	-+++--	-1	1	1	1	-1	-1
8	--+-++	-1	-1	1	-1	1	1
9	+--+-+	1	-1	-1	1	-1	1
10	++++++	1	1	1	1	1	1
11	-+-+++	-1	1	-1	1	1	1
12	--+--+	-1	-1	1	-1	-1	1

12-run PB + 1 center point

	Pattern	X1	X2	X3	X4	X5	X6
1	---+--	-1	-1	-1	1	-1	-1
2	-+-+++	-1	1	-1	1	1	1
3	+---+-	1	-1	-1	-1	1	-1
4	--+--+	-1	-1	1	-1	-1	1
5	++---+	1	1	-1	-1	-1	1
6	+--+-+	1	-1	-1	1	-1	1
7	000000	0	0	0	0	0	0
8	+-+++-	1	-1	1	1	1	-1
9	++++++	1	1	1	1	1	1
10	--+-++	-1	-1	1	-1	1	1
11	-+--+-	-1	1	-1	-1	1	-1
12	-+++--	-1	1	1	1	-1	-1
13	+++---	1	1	1	-1	-1	-1

3.2 Criteria

In our design comparison, we apply three types of criteria to evaluate the relative performance among the DSD and Plackett-Burman designs. The followings give some brief introduction to each criterion.

3.2.1 Power. The power refers to the probability that a design will detect active effects when they truly exist. It is computed for the hypothesis test $H_0: \beta_i = 0$. Given n is the number of observations and p is the number of terms other than the intercept in the model, we first obtain the α-level critical value F_c by solving:

$$\alpha = 1 - FDist(F_c, 1, n - p - 1).$$

Then we calculate the power as:

$$power = 1 - FDist(F_c, 1, n - p - 1, \lambda),$$

where λ is the noncentrality parameter for a noncentral F distribution.

3.2.2 Relative Prediction Variance. The relative prediction variance in SAS JMP is computed as:

$$\sigma_R^2 = \frac{x_i' var(\widehat{Y}) x_i}{\sigma^2} = \frac{x_i' var(x_i' \widehat{\beta}) x_i}{\sigma^2} = x_i'(X'X)^{-1} x_i,$$

where X is the model matrix, σ^2 is the error variance, $\widehat{\beta}$ is a vector of least squares estimates of the parameters, x_i' is the ith row of X.

3.2.3 Relative D-Efficiency. Given two designs d_i $(i = 1, 2)$ with corresponding model matrix $X(d_i)$, the relative D-efficiency is computed as:

$$D_e(d_1, d_2) = \left(\frac{|X(d_1)'X(d_1)|}{|X(d_2)' X(d_2)|} \right)^{1/p},$$

where p is the number of terms in the model.

3.3 Results

Our simulation results are composed of two parts: in the first part, we look into the three designs consisting of main effects only and their powers, relative predicted variance, and relative D-efficiencies. In the second part, we add two-way interactions to the designs and then conduct comparative analyses based on the same criteria.

3.3.1   *Main Effects Only.* We use the Compare Designs platform in SAS JMP to compare the performance of the three designs consisting of main effects only. In Figure 2 below, the PB design 2 performs the best for having the highest powers in all terms, while the 13-run DSD performs the worst among all. In terms of the relative prediction variance (left side of Figure 3), the PB design 2 again performs the best for having the uniformly smallest relative prediction variance, whereas the PB design 1 performs slightly worse. Further looking into the relative D-efficiencies (right side of Figure 3), we observe that the PB design 1 has higher relative D-efficiency compared to the 13-run DSD. Moreover, its relative D-efficiency increases further when a center point is added to the PB design 1.

Figure 2

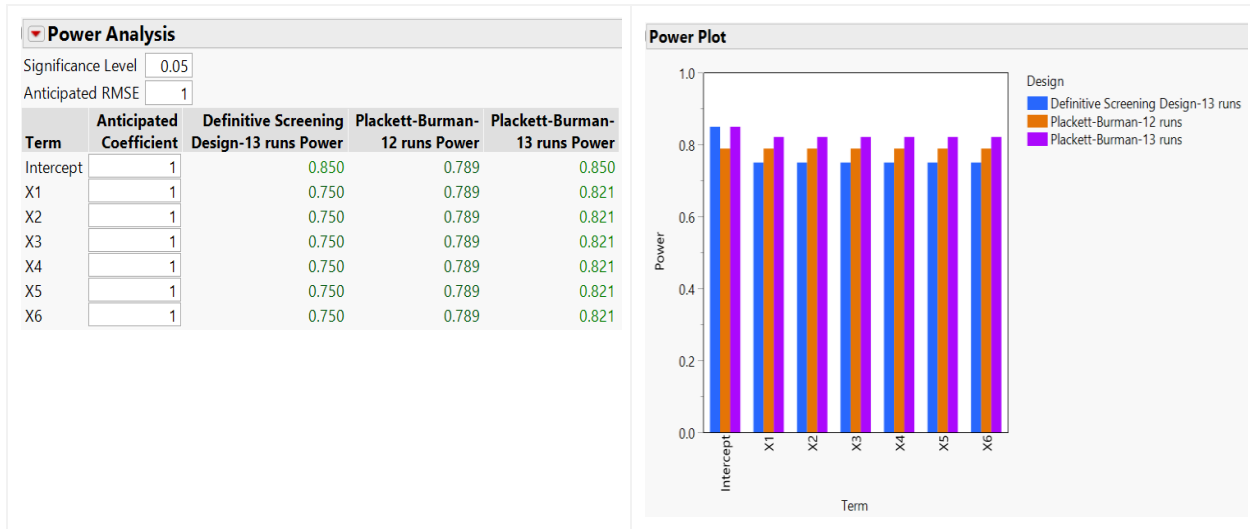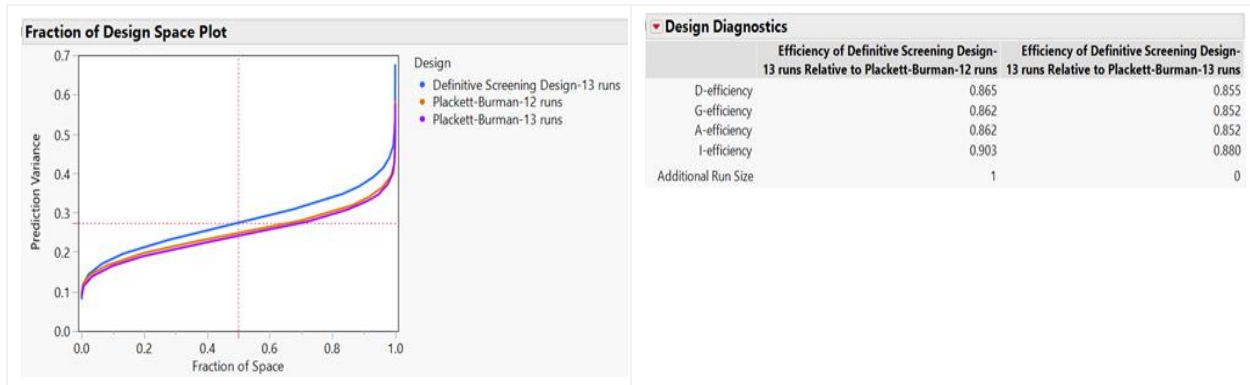*Power Analysis and Plot for Designs Consisting of Main Effects*

Figure 3

Relative Prediction Variance and Relative D-Efficiency for Designs Consisting of Main Effects

3.3.2   Main Effects + Two-Way Interactions. In this case, we add two-way interactions into the designs. Notice here we only include the two-way interactions that involve the first three factors, since the last three factors are confounding with other interaction terms and greatly reduce the power of the designs. In Figure 4 below, the 13-run DSD has the highest power in all factors especially in factors X4, X5, X6 and the two-way interaction terms. Additionally, the 13-run DSD has the uniformly smallest relative prediction variance among the three designs (left side of Figure 5). Nevertheless, the relative D-efficiency of the 13-run DSD is still the lowest compared to the other two PB designs (right side of Figure 5).

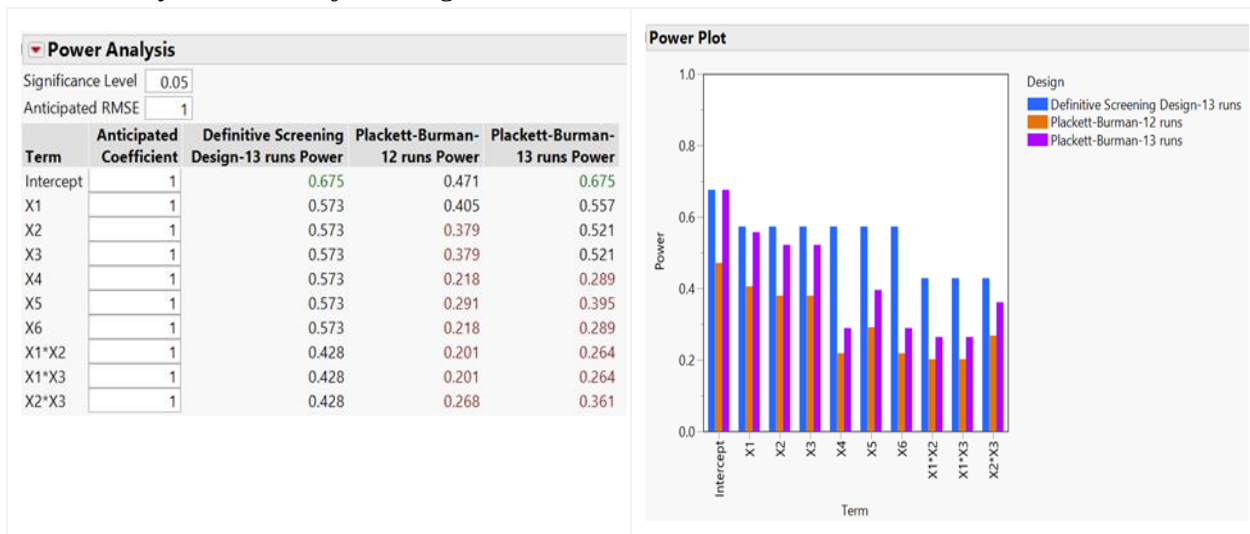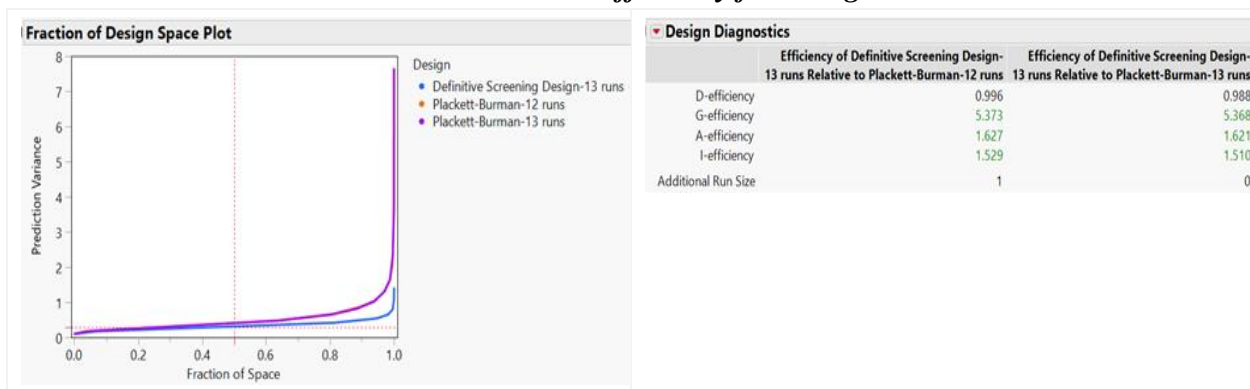**Figure 4**

Power Analysis and Plot for Designs with Interactions

Figure 5

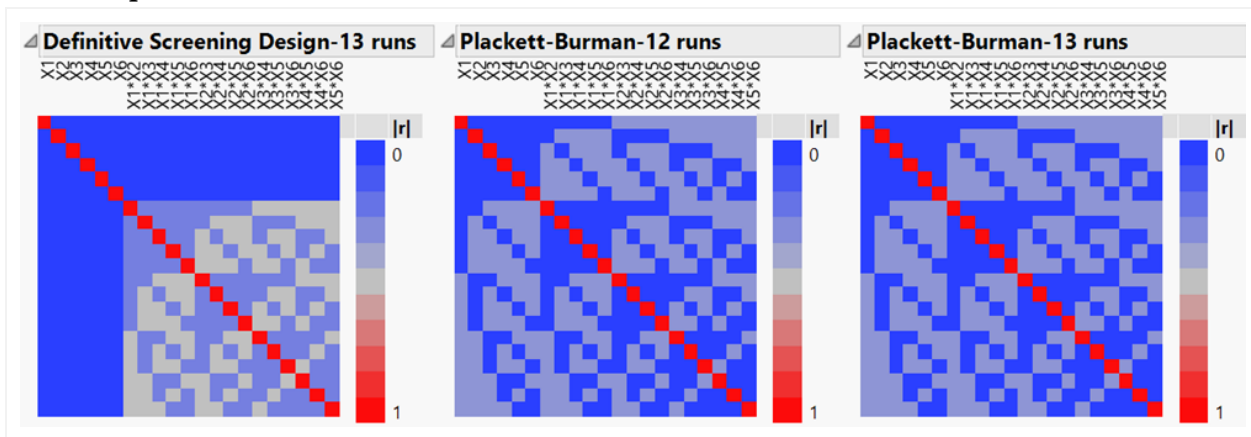*Relative Prediction Variance and Relative D-Efficiency for Designs with Interactions*

4. Discussion

There are two major findings in our simulation study. In the first part of our simulation, the 13-run DSD performs the worst in terms of power, relative prediction variance, and relative D-efficiency compared to the two Plackett-Burman designs. One possible reason why the relative prediction variance of the 13-run DSD is uniformly higher than of the other two designs is the different values of the diagonal elements of $X'X$ each design holds. The relative prediction variance $\sigma_R^2 = x_i'(X'X)^{-1}x_i$ (Section 3.2.2 on p.5) is inversely proportional to $X'X$. In our simulation, the value of the diagonal elements for the PB design 1 is 12, for the PB design 2 is 13−1=12, and for the DSD is 13−3=10. Therefore, the relative prediction variance of the 13-run DSD is the highest among the three designs.

In the second part of our simulation, the 13-run DSD performs better than the two Plackett-Burman designs in terms of power and relative prediction variance. Yet the relative D-efficiency given by the DSD is still the lowest. We are curious about the reasoning behind this result. First, we look into the color map (Figure 6 below). All of the main effects are orthogonal to the two-way interactions in the DSD, while some of the two-way interaction terms in the two Plackett-Burman designs are partially confounded with the main effects. We also observe that the correlations between the main effects and the two-way interaction terms in the two Plackett-Burman designs are not very strong. Thus, since our simulation designs here consist of less complicated aliasing structures and no quadratic terms, it is possible that the strengths of DSDs are not quite manifested in terms of the relative D-efficiency.

Figure 6

Color Map on Correlations

In further explaining this finding, Jones and Nachtsheim (2011) provide some insights from their study that gives similar results when comparing DSDs with standard orthogonal designs such as a Plackett-Burman design augmented with a center point. Their reasoning is that DSDs give up some efficiency for better estimation in which the main effect estimates are not biased by the two-way interaction terms. Unfortunately, our simulation structure may be too simple to clearly demonstrate this feature.

5. Conclusion

DSDs are a kind of novel three-level designs that requires a relatively small run size and avoids undesirable confounding of main effects and two-way interactions. They also allow the estimation of quadratic effects and the examination of factors that accounts for nonlinearity. Although in this study we focus on the comparisons between DSDs and Plackett-Burman designs with only main effects and two-way interaction terms, we observe that our 13-run DSD performs better and gives more satisfactory results when the model becomes more complicated with two-way interaction terms in the second part of the simulation. More advantageous features of DSDs may be found by constructing a more complex model including some quadratic effects.

Future study can also be conducted on how DSDs handle categorical factors. As noted by Jones and Nachtsheim (2013), the original DSDs have some constraint that all factors must be quantitative. Subsequently, the DSD-augment method and the ORTH-augment method are developed for DSDs to process two-level categorical factors. Nonetheless, some tradeoffs exist in the aliasing of main effects and two-way interactions when applying these two methods. Advanced research may be carried out to dig into these two methods and shed some light on the possible solutions to the confounding problem.

6. References

Comparison of a Definitive Screening Design with a Plackett-Burman Design. Retrieved February 16, 2020, from https://www.jmp.com/support/help/14-2/comparison-of-a-definitive-screening-design-with.shtml.

Jones, B., & Nachtsheim, C. J. (2011). A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects. Journal of Quality Technology, 43(1), 1–15. doi: 10.1080/00224065.2011.11917841.

Jones, B., & Nachtsheim, C. J. (2013). Definitive Screening Designs with Added Two-Level Categorical Factors. Journal of Quality Technology, 45(2), 121–129. doi: 10.1080/0022 4065.2013.11917921.