

## Research Article

# ***D*-Optimal Design for a Causal Structure for Completely Randomized and Random Blocked Experiments**

Zaher Kmail <sup>1</sup> and Kent Eskridge <sup>2</sup>

<sup>1</sup>*School of Interdisciplinary Arts and Sciences, University of Washington Tacoma, Tacoma, WA, USA*

<sup>2</sup>*Department of Statistics, University of Nebraska—Lincoln, Lincoln, NE, USA*

Correspondence should be addressed to Zaher Kmail; [zkmail@uw.edu](mailto:zkmail@uw.edu)

Received 26 April 2022; Revised 30 July 2022; Accepted 1 August 2022; Published 30 August 2022

Academic Editor: Muhammad Ahsan

Copyright © 2022 Zaher Kmail and Kent Eskridge. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most experimental design literature on causal inference focuses on establishing a causal relationship between variables, but there is no literature on how to identify a design that results in the optimal parameter estimates for a structural equation model (SEM). In this research, search algorithms are used to produce a *D*-optimal design for a SEM for three-stage least squares and full information maximum likelihood estimators. Then, a *D*-optimal design for the estimate of the model parameters of a mixed-effects SEM is obtained. The efficiency of each of the *D*-optimal designs for SEMs is compared with univariate optimal and uniform designs. In each case, the causal relationship changed the optimal designs dramatically and the new *D*-optimal designs were more efficient.

## 1. Introduction

Most experimental design literature on causal inference focuses on establishing a causal relationship between variables, but there is no literature on how to identify a design that results in the optimal parameter estimates for a structural equation model (SEM). This work was motivated by an agricultural research study by Milander and colleagues [1] where a SEM was proposed to determine the influence of seeding rate on the yield components of waxy maize and to better determine the interrelationship between yield and yield components by estimating the direct effect of the path model based on the biological understanding of the interrelationship among the endogenous and exogenous variables. The endogenous variables in the proposed system as described in Figure 1 were the yield components rows per ear, ear length, kernels per row, kernels per ear, grain yield, ear circumference, and kernel weight. Seeding rate was the only exogenous variable.

The goodness-of-fit parameters for the model were poor, so the data were reanalyzed to obtain a model that better fit the data, with the resulting model as described in Figure 2.

However, it did not give any insight into the interrelationship among many of the variables, which raised the question of whether or not the design was adequate for structural equation modeling.

The study was conducted over a period of three years and the original model was known before the data were collected [1]. Ideally, the researchers should have designed an experiment which would have allowed for the most precise estimates of the parameters, thus leading to more precise inferential statistics. Rather than uniformly applying the seeding rate levels across the experiment, the optimal design would have provided the researchers the number of replicates for each seeding rate level and which of the seeding rate levels should have been replicated more. However, they did not have the tools to obtain an optimal design for the proposed model nor were they able to take advantage of the information that was collected in the first year to adjust the design for subsequent years. There simply was no literature to support such an objective.

Historically, what drove the growth of optimal design theory was the need to have a design that is optimal for all parameters that need to be estimated [2, 3]. Technological

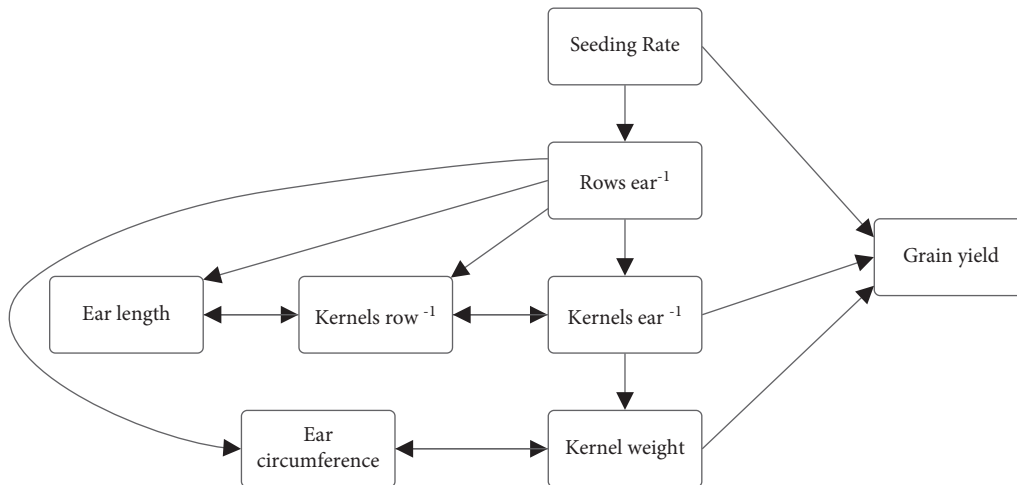


FIGURE 1: The proposed causal model for a maize yield component study published in Maydica by Milander et al. [1].

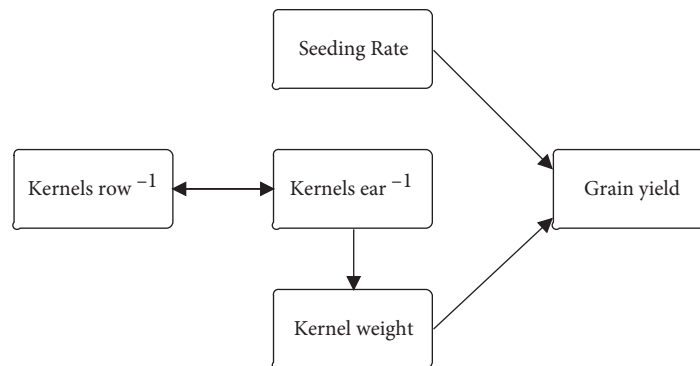


FIGURE 2: The data-supported causal model for the maize yield component study published in Maydica by Milander et al. [1].

advancements made it possible to develop the search algorithms that are needed to obtain optimal designs [4–8]. The new growth in the application of experimental design came about because of the lack of theoretical results in the optimal design of blocked and split-plot experiments, which made the computerized design algorithms popular and led to three proposals for point-exchange algorithms for constructing  $D$ -optimal blocked designs [9–15].

Traditional statistical analyses focus on univariate linear models, which have limitations because most random processes involve multiple dependent variables. This, in turn, led to the development of the multivariate linear model that does not address the causal relationships among the dependent (endogenous) variables, which is not the case with structural equation modeling [16, 17]. SEMs allow qualitative cause-effect information to be combined with statistical data and provide quantitative assessment of the cause-effect relationships among, and within, the endogenous and exogenous variables [17, 18].

Even though there have been many algorithms developed for univariate optimal designs and univariate mixed optimal designs, there is very limited literature about multivariate optimal design with examples like the work of Soumaya and fellow authors [19] and Schwabe [20]. There is literature in experimental design in structural equation

modeling but not in optimal design. The research in this area so far is to establish causal relationships (direction of the arrows) among the variables [21, 22]. This research extends to counterfactuals and experimental design where it is used to establish causal relationships [23]. However, no research exists to obtain an optimal design to estimate the parameters of a SEM.

The gap in literature demonstrates the need for this current research, which comprises three objectives. For a given causal structure, the first objective is to obtain a  $D$ -optimal design that allows the most precise estimates of the endogenous and exogenous parameters of that model. Then, for a given mixed causal structure with random blocks or split-plots, the second objective is to obtain a  $D$ -optimal design for estimating the parameters of the model. Finally, the efficiency of each of the  $D$ -optimal designs for causal structures will be compared with the optimal designs for the univariate case.

## 2. Methods

For the univariate case, the linear model is  $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$ , where  $\underline{y}$  is an  $n \times 1$  vector of observations,  $X$  is an  $n \times q$  design matrix of rank  $q$ ,  $\underline{\beta}$  is a  $q \times 1$  vector of unknown coefficients that are estimable, and  $\underline{\varepsilon}$  is an  $n \times 1$  identically

independently normally distributed vector with  $E(\underline{\varepsilon}) = 0$  and  $\text{Var}(\underline{\varepsilon}) = \sigma^2 I_{(n)}$  where  $I_{(n)}$  is the identity matrix with rank  $n$ . The least-squares estimate of  $\underline{\beta}$  is  $\underline{\hat{\beta}} = (X'X)^{-1}X'y$ . Our objective is to choose a design “ $X$ ” such that  $\max_X |\text{Var}(\underline{\hat{\beta}})| = \max_X |X'X|$  [7, 8]. Such a design “ $X$ ” is called a  $D$ -optimal design.

The same philosophy would be followed for a mixed model. The model for the data that comes from experiments with random blocks or split-plots is  $\underline{y} = X\underline{\beta} + Z\underline{u} + \underline{\varepsilon}$  where  $\underline{y}$ ,  $X$ ,  $\underline{\beta}$ , and  $\underline{\varepsilon}$  are defined above and  $\underline{u}$  is a random vector with  $E(\underline{u}) = 0$  and  $\text{Var}(\underline{u}) = \sigma_u^2 I_{(b)}$ .  $Z$  is an  $n \times b$  matrix of the form  $Z = \text{diag}[\underline{1}_{k_1}, \dots, \underline{1}_{k_b}]$  where  $\underline{1}_{k_i}$  is a  $k_i \times 1$  vector of ones. Then, it follows that  $V = \text{Var}(\underline{y}) = \sigma_\varepsilon^2 I_{(n)} + \sigma_u^2 ZZ'$ . Now it can be shown that  $\underline{\tilde{\beta}} = (X'V^{-1}X)^{-1}X'V^{-1}y$  where  $\text{Var}(\underline{\tilde{\beta}}) = (X'V^{-1}X)^{-1}$ . The objective becomes to choose the design matrix “ $X$ ” such that  $\max_X |X'V^{-1}X|$ . However,  $V$  is unknown. This presents a new challenge and shows that the more complicated the model, the more challenges can arise when choosing a  $D$ -optimal design.

The first solution to the above challenge is to assume that  $V = \text{Var}(\underline{y}) = \sigma_\varepsilon^2 I_{(n)} + \sigma_u^2 ZZ'$  is known by making assumptions about  $\sigma_\varepsilon^2$  and  $\sigma_u^2$ . The second solution is to rewrite  $V = \text{Var}(\underline{y}) = \sigma_\varepsilon^2 I_{(n)} + \sigma_u^2 ZZ' = \sigma_\varepsilon^2 (I_{(n)} + \eta ZZ') = \sigma_\varepsilon^2 (I_{(n)} + (\rho/1 - \rho)ZZ')$ , where  $\eta = (\sigma_u^2/\sigma_\varepsilon^2)$  and  $\rho = (\sigma_u^2/(\sigma_\varepsilon^2 + \sigma_u^2))$  [14, 15, 24]. This allows us to rewrite  $|X'V^{-1}X| = \sigma_\varepsilon^{-2q} |X'(I + \eta ZZ')^{-1}X|$  as the function to optimize a univariate linear model. It can be assumed without loss of generality  $\sigma_\varepsilon^2 = 1$ , thus  $|X'V^{-1}X|$  can be expressed in terms of variance ratio  $\eta$  or in terms of the correlation coefficient  $\rho$ . Therefore, the  $D$ -optimal design for a blocked or a split-plot experiment depends on the variance ratio. However, Goos and Vandebroek [14, 15] and Goos and Jones [13] argued that the dependence is minor and presented algorithms for constructing locally optimal designs for a given  $\eta$  or  $\rho$ . In some special cases, the optimal design is globally optimal and neither depends on the variance ratio nor the degree of correlation of which the most popular is the orthogonal design [14].

The previous method, or “traditional method” as it is referred to in literature, is based on optimization of  $|\text{Var}(\underline{\hat{\beta}})|$  or  $|X'V^{-1}X|$  and has two weaknesses [25]. The first weakness is that the  $D$ -optimal design depends on  $\eta$  or  $\rho$  and these values are not known prior to the experiment. The second weakness is that the optimal design focuses only on the fixed effect parameters and ignores the estimate of the random effect parameters. These weaknesses could have severe consequences on the estimate of the variance components, which severely affects the inferential statistics on the fixed effect parameters. Mylona et al. [25] provided scenarios for optimal designs that were constructed using the traditional  $D$ -optimality criterion and, in these designs, the random effects  $\sigma_u^2$  and  $\sigma_\varepsilon^2$  are not estimable.

Consequently, the two weaknesses led to a new composite  $D$ -optimality criterion. The new composite criterion was proposed by Mylona et al. [25]. The reason

for calling it composite is because this criterion takes into consideration both the fixed effect and the random effect parameters.

However,  $D$ -optimality criterion and orthogonal designs are both dependent on an assumed model. If the assumed model incorrectly represents the interrelationships between the endogenous and exogenous variables, it may lead to an increased bias error. Thus, in cases where there is no knowledge of the design specification, one of the most used designs is the uniform design in which the design points are scattered over the experimental region [26]. In comparison to  $D$ -optimal and orthogonal designs, uniform designs are less sensitive for model selection [27, 28]. Consequently, uniform design does not require pre-experiment model specification [29]. Uniform design could be used to explore the relationship between variables.

For the same reasons that led to the development of algorithms to account for all parameters of interest in the univariate case, our objective is to develop algorithms for SEMs that consider both the endogenous and exogenous parameters. In structural equation modeling, estimation methods are more complex and there are a variety of class estimators. This research focuses on three-stage least squares (3SLS) and full information maximum likelihood (FIML) methods because they are the only full-system estimation methods where all parameters are estimated simultaneously.

### 3. $D$ -Optimal Design for a Causal Structure Model

*3.1. The Information Matrix Approximation for 3SLS.* In most systems, researchers are interested in more than one endogenous variable and these variables directly and indirectly affect each other. To model such a study, we use SEMs. To demonstrate the methodology and introduce the new approach for  $D$ -optimal designs, the following examples will be introduced. A simple example is shown in Figure 3 with two exogenous variables and two endogenous variables.

Many phenomena can be modeled based on this SEM. For example, at the University of Washington Tacoma there are two pre-calculus pathways into Calculus I, a two-course sequence (TMATH 115 and TMATH 116) and a one-course sequence (TMATH 120). In this model,  $x_1$  represents the grade in TMATH 115,  $y_1$  represents the grade in TMATH 116,  $x_2$  represents the grade in TMATH 120, and  $y_2$  represents the Calculus I grade. If we are designing a study to estimate the parameters in Figure 3, our first objective would be to determine how many students should be enrolled in each pathway to obtain the best estimates for the three parameters  $\gamma_{11}$ ,  $\gamma_{22}$ , and  $b_{12}$ . There are two parameters that need to be estimated through the first pathway but only one through the second pathway. In this example, assigning half of the sample size to each pathway may not be optimal because there are more parameters on the first pathway than the second. So, the fundamental question here is, how many students should be assigned to each pathway to obtain the optimal estimates for all parameters?

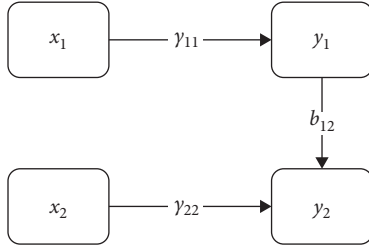


FIGURE 3: Identified SEM where each exogenous variable directly affects only one endogenous variable.

We will focus on the model in Figure 3. In this centered model, the  $k^{\text{th}}$  observation can be expressed mathematically as  $y_{k1} = \gamma_{11}x_{k1} + \varepsilon_{k1}$  and  $y_{k2} = b_{12}y_{k1} + \gamma_{22}x_{k2} + \varepsilon_{k2}$  where  $\begin{bmatrix} \varepsilon_{k1} \\ \varepsilon_{k2} \end{bmatrix} \sim N(0, \Sigma)$  for  $k = 1 \dots n$ . Some definitions are necessary to develop the model.  $Y_{(j)}$  is an  $n \times (p_j - 1)$  matrix that consists of the endogenous independent variables that are in the  $j^{\text{th}}$  equation.  $\underline{y}_{(j)}$  is an  $n \times 1$  vector that consists of the responses of the endogenous dependent variable for the  $j^{\text{th}}$  equation.  $X_{(j)}$  is an  $n \times q_j$  matrix that consists of the exogenous variables that are in the  $j^{\text{th}}$  equation,  $\underline{b}_{(j)}$  are the endogenous parameters of the  $j^{\text{th}}$  equation, and  $\underline{\gamma}_{(j)}$  are the exogenous parameters of the  $j^{\text{th}}$  equation.  $\underline{\varepsilon}_{(j)}$  is an  $n \times 1$  vector of random error for the  $j^{\text{th}}$  endogenous variable. Each equation in the system can be rewritten using vector notations. So,  $\underline{y}_{(j)} = Y_{(j)}\underline{b}_{(j)} + X_{(j)}\underline{\gamma}_{(j)} + \underline{\varepsilon}_{(j)}$  and  $\underline{\varepsilon}_{(j)} \sim N(\underline{0}, \sigma_{\varepsilon_{(j)}}^2 I_{(n)})$ . Now rewrite the  $j^{\text{th}}$  equation as  $\underline{y}_{(j)} = \begin{bmatrix} Y_{(j)} & X_{(j)} \end{bmatrix} \begin{bmatrix} \underline{b}'_{(j)} & \underline{\gamma}'_{(j)} \end{bmatrix}' + \underline{\varepsilon}_{(j)}$  where  $W_{(j)} = \begin{bmatrix} Y_{(j)} & X_{(j)} \end{bmatrix}$  and  $\underline{\delta}_{(j)} = \begin{bmatrix} \underline{b}'_{(j)} & \underline{\gamma}'_{(j)} \end{bmatrix}'$ . Therefore,  $\underline{y}_{(j)} = W_{(j)}\underline{\delta}_{(j)} + \underline{\varepsilon}_{(j)}$ . Multiply both sides by  $X'$ , use matrix notation, and use the Kronecker product to obtain  $(I \otimes X')\underline{y}^* = (I \otimes X')W^*\underline{\delta}^* + (I \otimes X')\underline{\varepsilon}^*$  where  $\underline{y}^* = [\underline{y}'_{(1)}, \dots, \underline{y}'_{(p)}]'$ ,  $W^* = \text{diag}(W_{(1)}, \dots, W_{(p)})$ ,  $\underline{\delta}^* = [\underline{\delta}'_{(1)}, \dots, \underline{\delta}'_{(p)}]'$ , and  $\underline{\varepsilon}^* = [\underline{\varepsilon}'_{(1)}, \dots, \underline{\varepsilon}'_{(p)}]'$  [30].

Then, to estimate  $\underline{\delta}^*$ , use the Generalized Least Square (GLS) to obtain  $\hat{\underline{\delta}}_{3SLS}^* = [W^{*'}(\Sigma^{-1} \otimes X[X'X]^{-1}X')W^*]^{-1}W^{*'}[\Sigma^{-1} \otimes X[X'X]^{-1}X']^{-1}\underline{y}^*$ . So, the covariance matrix of the 3SLS estimators is  $\text{Var}(\hat{\underline{\delta}}_{3SLS}^*) = [\hat{W}^{*'}(\Sigma^{-1} \otimes X[X'X]^{-1}X')\hat{W}^*]^{-1} + O(1/n)$  where  $\hat{W}^* = E(W)$  [31, 32], which is equivalent to the asymptotic information matrix for FIML [33]. Thus, the estimate of the covariance matrix would be

$$\hat{\text{Var}}\left(\hat{\underline{\delta}}^*\right) = \left[\hat{W}^{*'}\left(\hat{\Sigma}^{-1} \otimes X[X'X]^{-1}X'\right)\hat{W}^*\right]^{-1}. \quad (1)$$

The reason that we are interested in  $\hat{\text{Var}}(\hat{\underline{\delta}}^*)$  is because our objective is to obtain the design  $X$  that “minimizes”

$\hat{\text{Var}}(\hat{\underline{\delta}}^*)$ . Since normality is assumed throughout the research, then  $\hat{\text{Var}}(\hat{\underline{\delta}}^*)^{-1}$  is called the information matrix estimate and is denoted by  $\hat{M} = [\hat{\text{Var}}(\hat{\underline{\delta}}^*)]^{-1}$ . Thus, our objective would be to obtain the  $D$ -optimal design “ $X$ ” such that

$$\max_X \hat{M} = \min_X \left| \text{Var}\left(\hat{\underline{\delta}}^*\right) \right|. \quad (2)$$

In this paper, we assume that there are no constraints on the treatments or on the exogenous variables. However, if there are constraints on the treatments, such as if one of the treatments is limited, then the optimality criteria can be adjusted as is done in the univariate case, which is known as the optimal design of mixture experiments problem [13, 34–37].

*Example 1.* Based on equation (2), the  $D$ -optimal design for the causal structure below (Figure 4) assumes that the treatments are qualitative treatments for twenty observations and assumes that the true values for  $\gamma_{11}$ ,  $\gamma_{22}$ , and  $b_{12}$  are 8, 2, and 5, respectively Figure 4:

In vector notation, the model can be expressed as  $\underline{y}_{(1)} = \gamma_{11}\underline{x}_{(1)} + \underline{\varepsilon}_{(1)}$  and  $\underline{y}_{(2)} = b_{12}\underline{y}_{(1)} + \gamma_{22}\underline{x}_{(2)} + \underline{\varepsilon}_{(2)}$  where  $\begin{bmatrix} \underline{\varepsilon}_{(1)} \\ \underline{\varepsilon}_{(2)} \end{bmatrix} \sim N(\underline{0}, \Sigma \otimes I_{(20)})$ . In this example,  $\underline{y}_{(1)}$  is a  $20 \times 1$  vector that includes the responses of the first endogenous variable,  $\underline{y}_{(2)}$  is a  $20 \times 1$  vector that includes the responses of the second endogenous variable,  $\underline{x}_{(1)}$  is a  $20 \times 1$  vector that indicates whether or not the first treatment is applied on the  $k^{\text{th}}$  experimental unit as denoted by  $(x_{k1} = 1)$  or  $(x_{k1} = 0)$ ,  $\underline{x}_{(2)}$  is a  $20 \times 1$  vector that indicates whether or not the second treatment is applied on the  $k^{\text{th}}$  experimental unit as denoted by  $(x_{k2} = 1)$  or  $(x_{k2} = 0)$ ,  $\underline{\varepsilon}_{(1)}$  is a  $20 \times 1$  vector of random residuals for the first endogenous variable,  $\underline{\varepsilon}_{(2)}$  is a  $20 \times 1$  vector of random residuals for the second endogenous variable,  $\underline{0}$  is a  $40 \times 1$  vector of zeros,  $\Sigma = I_{(2)}$ , and  $I_{(20)}$  is a  $20 \times 20$  identity matrix. To obtain a  $D$ -optimal design, the algorithm starts with multiple random designs in order to avoid local optimality. For this specific example, 50 initial random designs were used, each design with size 20. The initial designs were constructed randomly where each  $x_{kj}$  value had a 50–50 chance of being a 1 or 0. To improve the design, the algorithm compares each point in the design with the candidate points  $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$  and makes a simultaneous exchange with the candidate point that improves the optimality criterion from equation (2) the most. The exchanges will continue until no further improvement is achieved or the improvement is sufficiently small. This procedure will be repeated for each of the 50 initial designs. The final designs will be compared. The design  $X$  that has the smallest  $\hat{\text{Var}}(\hat{\underline{\delta}}^*)$  from among the 50 final designs will be selected as the  $D$ -optimal design.

Using the previous algorithm, a  $D$ -optimal design for 20 observations for the model in Example 1 is given in Table 1.

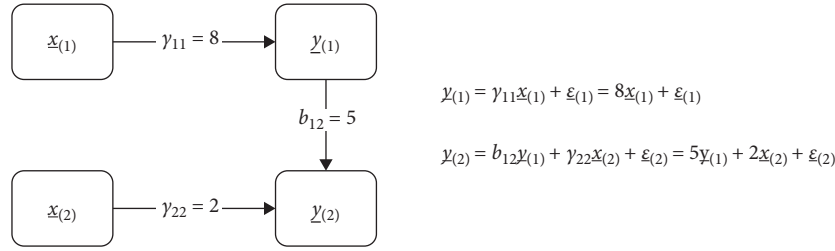


FIGURE 4: The path model for Example 1.

TABLE 1: 3SLS and FIML  $D$ -optimal design for Example 1.

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$x_1$	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1
$x_2$	0	0	1	1	0	1	1	0	1	1	0	1	0	1	1	1	0	1	0	0

In the  $D$ -optimal design for the causal structure for the model in Figure 4, the points (1, 1) and (1, 0) were each replicated nine times, and the point (0, 1) was replicated twice. The point (0, 0) was not replicated, which is expected because the model was assumed to have no intercept.

Table 1 was based on the true value of  $W^*$  where the true parameters are used in  $\hat{W}^*$ . However, the true values are unknown. There are two approaches that address this issue. First, when prior data are available, estimates of the parameters can be used to estimate  $\hat{\text{Var}}(\hat{\delta}_{3\text{SLS}}^*)$  instead of the true parameters. One question that arises with the use of this approach is the robustness of the design. To address the sensitivity of the true values of the parameters on the design, simulation is used to obtain the optimal design based on the estimates of the parameters.

When there is no prior data available, a second approach is a Bayesian approach which assumes a prior distribution for the endogenous variables, which is similar to the approach that is used by Mylona et al. [25] and Goos and Mylona [38] to address the variance components optimal design for blocked or split-plot experiments. This approach allows for the uncertainty of the endogenous parameters. Using the Bayesian approach to obtain a 3SLS  $D$ -optimal design for a causal structure will be the subject of future work.

Optimal design theory is currently limited to univariate or multivariate applications where one of the biggest weaknesses of these approaches is that they ignore any possible causal structure among the endogenous variables. Because of the lack of theoretical results and the algorithms to produce optimal designs for a causal structure, one approach is to use univariate optimal designs for causal models, which leads to a loss of efficiency. To obtain a univariate  $D$ -optimal design for Example 1, it is assumed that the treatment will affect both endogenous variables to avoid the situation where there is an optimal design for each equation. In this case, we obtain a  $D$ -optimal design for both treatment combinations. An optimal design in the univariate case for two qualitative treatments (*i.e.*  $y = \gamma_{11}x_{(1)} +$

$\gamma_{22}x_{(2)} + \varepsilon$ ) can be found in Table 2 where the four possible candidate points are (0, 0), (0, 1), (1, 0), and (1, 1). The three candidate points (0, 1), (1, 0), and (1, 1) are replicated seven times, seven times, and six times, respectively. For the same reasons as in the  $D$ -optimal design for a causal structure, the point (0, 0) was not replicated since the intercept was not included in the model.

The ratio of the determinants of the information matrices of the univariate versus the 3SLS designs and the  $D$ -efficiency of the two designs are used to compare the optimality of the designs [9, 10]. The results from Table 3 compare the efficiency of the univariate optimal design from Table 2 to the efficiency of the new  $D$ -optimal design from Tables 2 and 3.

When comparing the univariate optimal design to the 3SLS  $D$ -optimal design for a causal structure, there was an approximately 18% increase in the determinant of the asymptotic information matrix estimates. The univariate optimal design is about 94% as  $D$ -efficient as the 3SLS  $D$ -optimal design for a causal structure in all four cases.

One criticism over the comparison of the  $D$ -optimal designs obtained above is that they are based on the asymptotic information matrices, but the  $D$ -optimal designs were obtained for small samples sizes (20 observations). Thus, it is important to ensure that the results are consistent for small samples sizes, as well. To verify the comparison, we simulated data based on the causal structure  $D$ -optimal design and data based on the univariate optimal design 100,000 times each. We then estimated the parameters for both designs using the GLS estimate  $\hat{\delta}_{3\text{SLS}}^*$  and calculated the covariance matrix of the parameter estimates for both designs, whose elements are  $\text{cov}(\hat{\delta}_i, \hat{\delta}_j) = \sum_{j=1}^{10^5} \sum_{i=1}^{10^5} (\delta_i - \hat{\delta}_i)(\delta_j - \hat{\delta}_j) / (10^5 - 1)$  where  $\hat{\delta}_i$  is the  $i^{\text{th}}$  element of  $\hat{\delta}_{3\text{SLS}}^*$ . The determinants of the covariance matrices of the parameters estimates were computed. This process was repeated three times. The determinant of the covariance matrix of the parameters estimates for the 3SLS causal structure  $D$ -optimal design was consistently smaller than the determinant

TABLE 2: Classical optimal design for univariate model  $y = \gamma_{11}x_{(1)} + \gamma_{22}x_{(2)} + \varepsilon$ .

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$x_1$	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1
$x_2$	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0

TABLE 3: Parameter values, estimates, determinants of information matrices, their ratios, and  $D$ -efficiency based on a 3SLS  $D$ -optimal design and a univariate optimal design for three simulations.

Parameters and information matrices	Parameter values	Simulation 1 estimates	Simulation 2 estimates	Simulation 3 estimates
$\gamma_{11}$	8	7.91	8.25	8.34
$\gamma_{22}$	2	2.31	0.19	1.30
$b_{12}$	5	5.01	5.10	5.03
$ M  =  \text{Var}(\hat{\delta}_{\text{Univariate}}^*) ^{-1}$	110,656	108,062.4	117,621.26	120,228.29
$ M_{3\text{SLS}}  =  \text{Var}(\hat{\delta}_{3\text{SLS}}^*) ^{-1}$	134,784	131,624.88	143,268	146,443.48
$ M / M_{3\text{SLS}} $	82%	82%	82%	82%
$D^\ddagger$ - efficiency = $( M / M_{3\text{SLS}} )^{1/3}$	94%	94%	94%	94%

<sup>‡</sup> See [10].

TABLE 4: Determinants of the covariance matrices for the 3SLS causal structure  $D$ -optimal design and the 3SLS univariate optimal design based on three simulations each with 100,000 data sets.

	1 <sup>st</sup> 100,000 data sets	2 <sup>nd</sup> 100,000 data sets	3 <sup>rd</sup> 100,000 data sets
Determinant of the covariance matrix for the causal structure $D$ -optimal design	7,613,179.5	7,718,001.3	7,290,743.0
Determinant of the covariance matrix for the univariate optimal design	9,008,371.8	8,893,254.3	9,035,191.3
$ M / M_{3\text{SLS}} $	85%	87%	81%
$D$ - efficiency = $( M / M_{3\text{SLS}} )^{1/3}$	95%	95%	93%

of the covariance matrix of the parameters estimates for the univariate optimal design, as shown in Table 4.

The results from Table 4 were consistent with the results for the comparison of the asymptotic information matrices. The 3SLS causal structure  $D$ -optimal design consistently produced a smaller determinant for the covariance matrix of the parameters estimates than the 3SLS univariate optimal design. Specifically, the determinant of the covariance matrix of the parameter estimates for the 3SLS causal structure  $D$ -optimal design was about 15%–19% smaller than the determinant of the covariance matrix of the parameter estimates for the 3SLS univariate optimal design. Based on those results, the 3SLS causal structure  $D$ -optimal design was 5%–7% more  $D$ -efficient than the univariate optimal design. It is important to note that whether the true parameters or their estimates are used, the 3SLS  $D$ -optimal design for a causal structure did not change.

### 3.2. The Information Matrix Approximation for FIML.

The full information maximum likelihood (FIML) is another estimation method for SEMs. In this section, we will discuss optimal designs based on the FIML estimators and will investigate the relationship between the optimal designs.

Durbin [33] proposed a transformation for the maximum likelihood equations that simplified the computations and made it easier to study the properties of FIML estimators and their advantages over 3SLS estimators. However, even with the special notations, the FIML information matrix is

sophisticated and takes much more work to obtain than the 3SLS information matrix. Based on Durbin's [33] results, our objective is to obtain a design matrix " $X$ " such that  $\max_X |M_{\text{FIML}}| = \max_X |\overline{Q}G\overline{Q}|$ . The derivation and the special notation are noted in Appendix A. The reason that both  $D$ -optimal designs are considered is because the 3SLS information matrix is easy to obtain, but FIML estimators have advantages over 3SLS estimators as discussed by Durbin [33].

As with the 3SLS  $D$ -optimal design, the objective would be to obtain the design matrix  $X$  that will minimize the estimate of the determinant of the covariance of the FIML estimates or maximize the determinant of the FIML information matrix estimate. The performance of the  $D$ -optimal design for a causal structure will then be compared to the univariate optimal design by computing  $|\text{Var}(\hat{\delta}_{\text{FIML}})|$  through simulation. The design which consistently has the largest determinants of the information matrix would be the better design. In the following example, the objective will be to use the estimate of the information matrix to obtain the FIML optimal design.

For the causal structure in Example 1, the 3SLS  $D$ -optimal design is also the  $D$ -optimal design for the FIML estimates. Similar to 3SLS, the true parameters are unknown, so the FIML estimates are used to replace the true parameters for the three simulations. In practice, the information matrix is estimated based on the parameters estimates. This is similar to the approach of Goos and Vandebroek [14, 15] and Goos and Jones [13] in the case of optimal design for

random blocks or split-plots where they suggest using the estimates of the variance parameters or a reasonable guess, arguing that the design minimally depends on those values. The results obtained for a  $D$ -optimal design for a causal structure in this research are consistent with those of Goos and Vandebroek [14, 15] and Goos and Jones [13] where the optimal design did not change when replacing the true parameters with their estimates.

The results shown in Table 5 show the efficiency of the  $D$ -optimal design from Table 1 as compared to the efficiency of the univariate optimal design from Table 2.

The results for the FIML causal structure  $D$ -optimal design are similar to the results for the 3SLS causal structure  $D$ -optimal design. The optimal design for the univariate case was about 6% less  $D$ -efficient than the FIML design for the causal structure. The loss of efficiency demonstrates the importance of taking the endogenous parameters into account to obtain an optimal design to study the causal structure.

For both the 3SLS and FIML  $D$ -optimal designs, the goal of the three simulations is to show that the  $D$ -optimal design did not change even though the estimate of the parameters deviated from the true value. However, it may be problematic to base a conclusion on three simulations only; therefore, 300 data sets were simulated and then the parameters were estimated. These parameters were used to obtain a  $D$ -optimal design based on the estimate of the parameters. In each of the 300 simulations, the  $D$ -optimal design did not change and was equivalent to the  $D$ -optimal design in Table 1.

Another way to demonstrate the robustness of the design is to incrementally change the value of the parameter(s) from the true value, then obtain a  $D$ -optimal design based on the value of the parameters, and finally compare the  $D$ -optimal design to that which was obtained based on the true value. For the model in Figure 4, the  $D$ -optimal design depends only on  $\gamma_{11} = 8$  since  $W$  includes only  $\underline{y}_{(1)} = \gamma_{11}\underline{x}_{(1)} + \underline{\varepsilon}_{(1)}$ . Table 1 is the  $D$ -optimal design for  $\gamma_{11} = 8$ . Now, we will change the value of  $\gamma_{11}$  incrementally based on the interval  $8 \pm 6 = [14, 2]$ . We first start with 8 and then decrease  $\gamma_{11}$  incrementally by 0.1 (i.e. 8, 7.9, 7.8, ..., 2). Then, we obtain the  $D$ -optimal design based on these values and check whether it is equivalent to the  $D$ -optimal design in Table 1. Similarly, we start with 8 and then increase  $\gamma_{11}$  incrementally by 0.1 (i.e. 8, 8.1, 8.2, ..., 14). Then, again, we obtain the  $D$ -optimal design based on these values and check whether it is equivalent to the  $D$ -optimal design in Table 1. The  $D$ -optimal designs did not change even though the value of the parameters changed up to 75% from the true value. These results are consistent with Goos and Vandebroek [14, 15] and Goos and Jones [13] where they argued that the optimal design depended minimally on the specific parameter values.

#### 4. $D$ -Optimal Design for a Causal Structure Model with Random Blocks

In applied research, blocked designs are likely the most commonly used experimental designs since they can be used to account for variation attributable to sources other than

treatments and can considerably improve the precision of the experiment. The first objective is to establish the blocked causal structure and then to obtain the 3SLS and FIML estimators for the endogenous and exogenous parameters, which will be used to estimate their information matrices. The second objective of this section is to use the algorithm described in Section 1.1 in Example 1 to obtain a  $D$ -optimal design for both the endogenous and exogenous parameters. The last objective is to compare the efficiency of the  $D$ -optimal design to the classical univariate mixed model optimal design.

*4.1. 3SLS  $D$ -Optimal Design for a Causal Structure with Random Blocks.* The model with random blocks is similar to the model for a completely randomized design except that the block effect needs to be added to every endogenous variable. Therefore, for the  $i^{\text{th}}$  endogenous variable as the dependent variable, the model can be written as.

$$\underline{y}_{(i)} = Y_{(i)}\underline{b}_{(i)} + X_{(i)}\underline{\gamma}_{(i)} + Z\underline{u}_{(i)} + \underline{\varepsilon}_{(i)} \quad \text{where} \quad \begin{bmatrix} \underline{u}_{(i)} \\ \underline{u}_{(j)} \end{bmatrix} \sim N\left(\underline{0}, \begin{bmatrix} \sigma_{u_{(i)}}^2 & \sigma_{u_{(i)j}} \\ \sigma_{u_{(i)j}} & \sigma_{u_{(j)}}^2 \end{bmatrix} \otimes I_{(b)}\right)$$

and all of the other terms have been previously defined. If the blocks have the same size, then  $Z = \text{diag}\{1_k, 1_k, \dots, 1_k\} = I_{b \times b} \otimes 1_k$ , the elements of  $\underline{u}_{(i)}$  and  $\underline{\varepsilon}_{(i)}$  are assumed to be mutually independent and normally distributed with zero mean and variances  $\sigma_{u_{(i)}}^2$  and  $\sigma_{\varepsilon_{(i)}}^2$ , respectively.

Let  $W_{(i)} = [Y_{(i)} \ X_{(i)}]$  and  $\underline{\delta}_{(i)} = [\underline{b}_{(i)}' \ \underline{\gamma}_{(i)}']'$ . Then, the model with random blocks can be rewritten as  $\underline{y}_{(i)} = W_{(i)}\underline{\delta}_{(i)} + Z\underline{u}_{(i)} + \underline{\varepsilon}_{(i)}$ . Use matrix notation and then the Kronecker product to rewrite the previous model as  $(I \otimes X')\underline{y}^* = (I \otimes X')W^*\underline{\delta}^* + (I \otimes X'Z)\underline{u}^* + (I \otimes X')\underline{\varepsilon}^*$ .

Therefore,  $\hat{\delta}_{3SLS}^* = \{[(I \otimes X')W^*]^{-1} V^{-1} [(I \otimes X')W^*]\}^{-1} [(I \otimes X')W^*]^{-1} V^{-1} [(I \otimes X')\underline{y}^*]$  where  $V = \text{Var}[(I \otimes X'Z)\underline{u}^* + (I \otimes X')\underline{\varepsilon}^*]$ . Let  $\text{Var}(\underline{u}^*) = \Sigma_u \otimes I_{(b)}$  and  $\text{Var}(\underline{\varepsilon}^*) = \Sigma \otimes I_{(n)}$  where  $\Sigma_u$  is a  $p \times p$  nonnegative definite matrix. Then,  $\text{Var}[(I \otimes X'Z)\underline{u}^* + (I \otimes X')\underline{\varepsilon}^*] = (I \otimes X'Z)\text{Var}[\underline{u}^*](I \otimes X'Z)' + (I \otimes X')\text{Var}[\underline{\varepsilon}^*](I \otimes X')' = (I \otimes X'Z)\{\Sigma_u \otimes I_b\}(I \otimes X'Z)' + (I \otimes X')\{\Sigma \otimes I_n\}(I \otimes X')' = Z^*G^*Z^{*'} + R^*$  where  $z^* = (I \otimes X'Z)$  and  $G^* = \Sigma_u \otimes I_b$ .

Using the same approach as described in Section 1.1, the asymptotic information matrix would be  $M_{3SLS} = [(I \otimes X')\hat{W}^*]^{-1} V^{-1} [(I \otimes X')\hat{W}^*]$  where  $V = Z^*G^*Z^{*'} + R^*$ .

Based on the information matrix for the 3SLS estimates for the causal structure  $D$ -optimal design with random blocks, our objective would be to use the algorithm described in Section 1.1 in Example 1 to obtain the design  $X$  such

that  $\max_X |M_{3SLS}| = \max_X |[(I \otimes X')\hat{W}^*]^{-1} V^{-1} [(I \otimes X')\hat{W}^*]|$ . Since the information matrix depends on the unknown

parameters through  $\hat{\Sigma}$  and  $\hat{W}^*$ , we face the same predicament as in the case of the completely randomized  $D$ -optimal design for a causal structure where the parameters are

TABLE 5: Parameter values, estimates, determinants of information matrices, their ratios, and  $D$ -efficiency based on a FIML  $D$ -optimal design and a univariate optimal design for three simulations.

Parameters and information matrices	Parameter values	Simulation 1 estimates	Simulation 2 estimates	Simulation 3 estimates
$\gamma_{11}$	8	7.98	8.20	8.36
$\gamma_{22}$	2	2.05	1.17	2.04
$b_{12}$	5	5.03	4.98	4.98
$ M  =  \text{Var}(\hat{\underline{\delta}}_{\text{Univariate}}^*) ^{-1}$	110,656	110,169.65	116,153.07	120,709.06
$ M_{\text{FIML}}  =  \text{Var}(\hat{\underline{\delta}}_{\text{FIML}}^*) ^{-1}$	134,784	133,044.67	141,479.68	147,029.08
$ M / M_{\text{FIML}} $	82%	83%	82%	82%
$D^* - \text{efficiency} = ( M / M_{\text{FIML}} )^{1/3}$	94%	94%	94%	94%

\* See [10].

unknown. The solutions that were proposed for the completely randomized  $D$ -optimal design for a causal structure in Section 1.1 are appropriate for the causal structure  $D$ -optimal design with random blocks. Other approaches are also possible. For example, in an experiment with random blocks that is conducted in multiple sequential stages, the data of one block (year, lab, day, etc.) could be utilized to obtain initial estimates that are then used to obtain the  $D$ -optimal design for the rest of the blocks (other years, other labs, other days, etc.). Requiring preliminary information is a standard part of univariate design. If there is no prior data, then a reasonable guess can be used similar to the approach of Goos and Vandebroek [14, 15] and Goos and Jones [13]. In many experiments, multiple factors and limitations of the experimental material require the use of split-plot designs and/or incomplete block designs. Specifically, the way the factors are applied and the nature of the experimental material give rise to different sizes of experimental units and limits on block size. For this reason, an incomplete block design with three factors is used as an example in this section. In the next example, the objective is to use the asymptotic information matrix estimates to obtain a  $D$ -optimal design for a causal structure with four random blocks where the size of each block is four.

*Example 2.* Assume that there are three binary treatment factors (exogenous variables) with two endogenous variables for the causal structure given in Figure 5. Here, the  $\underline{x}_{(i)}$ ,  $i = 1, 2, 3$  each with a  $x_{kj}$  value, had a 50–50 chance of being a 1 or  $-1$ . Also, assume that the true values for  $\gamma_{11}$ ,  $\gamma_{22}$ ,  $\gamma_{31}$ , and  $b_{12}$  are 8, 2, 3, and 5, respectively Figure 5:

A  $D$ -optimal design for a causal structure with four random blocks with block size equal to four under the given assumptions can be found in Table 6.

The expected determinant of the information matrix is 5,222,400 and the previous design is not orthogonal. The expected determinant of the information matrix of the orthogonal design in Table 7 is 4,784,128. The orthogonal design is the optimal design for the univariate mixed model with three binary treatment factors and four blocks each block with size four.

One concern for both  $D$ -optimal and orthogonal designs is that they both require an assumed model structure. One popular alternative is the uniform design, which is “a space filling design” ([39], p. 131) that uniformly distributes experimental points across the design domain without model

pre-specification [39]. Therefore, we will compare the information of the uniform design to the orthogonal design and the new  $D$ -optimal design. Using a uniform design table [39, 40], a uniform design for 16 runs with 3 factors and an experimental cube of  $[-1, 1]^3$  is shown in Tables 7 and 8.

The expected determinant of the information matrix for the uniform design is 28,612, which is significantly smaller than the determinants of the  $D$ -optimal and the orthogonal design information matrices.

The results show that the new  $D$ -optimal design was approximately 2.2% more  $D$ -efficient than the orthogonal design, whereas the  $D$ -optimal design was 72.8% more  $D$ -efficient than the uniform design. The  $D$ -optimal design performed significantly better than the uniform design but is still comparable to the orthogonal design. Although orthogonal designs are desirable because they are universally optimal for first-order linear models and are also robust for nonlinear models, they also come with concerns in that they sometimes do not exist and require a large sample size [25]. It is also important to keep in mind that in Example 2, there is only one endogenous parameter. The significance in the increase in efficiency of the determinant may not be easily identified through our simple example. In a more complex design, there will be more endogenous variables and if those variables are interconnected in a more sophisticated way, then the increase of the determinant of the information matrix would be amplified.

It is worth noting the differences between the 3SLS causal structure  $D$ -optimal design and the orthogonal design. First, the orthogonal design completely confounds the three-way interaction ( $x_{i1}x_{j2}x_{k3}$ ) with the blocks where  $x_{i1}$  is the  $i^{\text{th}}$  element of  $\underline{x}_{(1)}$ ,  $x_{j2}$  is the  $j^{\text{th}}$  element of  $\underline{x}_{(2)}$ , and  $x_{k3}$  is the  $k^{\text{th}}$  element of  $\underline{x}_{(3)}$ . The number of replicates for each candidate point is 2. However, this is not the case in the 3SLS causal structure  $D$ -optimal design where the two-way interaction of ( $x_{j2}x_{k3}$ ) is confounded with Blocks 1 and 3, the three-way interaction of ( $x_{i1}x_{j2}x_{k3}$ ) is confounded with Block 4, and the two-way interaction of ( $x_{i1}x_{k3}$ ) is confounded with Block 2. In the 3SLS causal structure  $D$ -optimal design, the number of candidate points is not replicated equally. For example, the candidate point  $(-1, -1, -1)$  was replicated three times, but the candidate point  $(1, -1, -1)$  was replicated only once.

To understand the differences between the designs, it is important to contrast the asymptotic information matrix of the 3SLS  $D$ -optimal design for a causal structure with that of



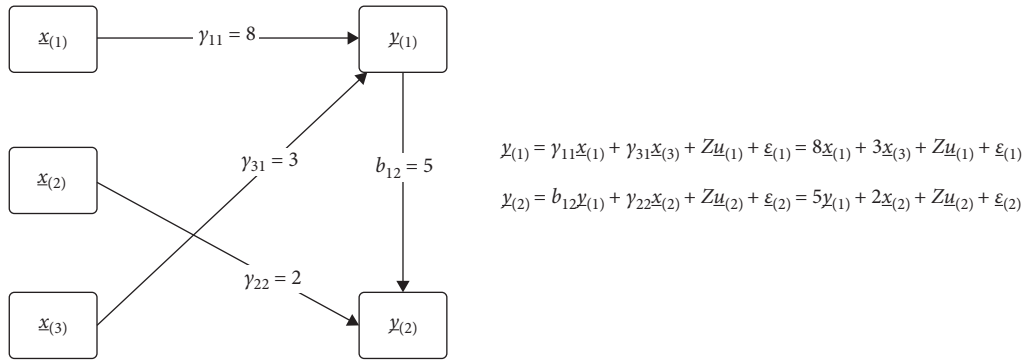


FIGURE 5: The identified path model for Example 2.

TABLE 6: A 3SLS causal structure  $D$ -optimal design with three binary factors and random blocks where each block size is equal to four.

Block 1			Block 2			Block 3			Block 4		
$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$
1	1	1	1	1	1	-1	1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	1	-1	1	1
1	-1	-1	-1	1	-1	1	-1	1	1	-1	1
-1	1	1	1	-1	1	1	1	-1	1	1	-1

TABLE 7: An orthogonal design for three binary treatment factors and four blocks of size four.

Block 1			Block 2			Block 3			Block 4		
$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$
-1	1	1	1	-1	-1	-1	1	-1	-1	-1	-1
1	-1	1	-1	-1	1	1	-1	-1	-1	1	1
1	1	-1	1	1	1	-1	-1	1	1	-1	1
-1	-1	-1	-1	1	-1	1	1	1	1	1	-1

the orthogonal design. In this way, we can see where information is gained and parameter estimates are improved for the former design and where information is sacrificed and parameter estimates worsen for the latter design. Tables 9 and 10 show the asymptotic information matrices for the model in Figure 5 for the orthogonal design and the 3SLS  $D$ -optimal design, respectively Table 9.

When the information of the estimates is compared, we can see that the exogenous parameters in both models have the same information except for the elements.

$\gamma_{11}, \gamma_{31}$ . However, the endogenous parameters in the 3SLS  $D$ -optimal design gained more information with an increase of about 16% over the orthogonal design. These results are expected because univariate optimal designs do not take the precision of the endogenous parameters into account, whereas  $D$ -optimal designs for causal structures do. Therefore, our new approach has the advantage of giving the optimal combination of treatments for the entire model and taking into account the precision of both the exogenous and endogenous parameters.

4.2. FIML  $D$ -Optimal Design for a Causal Structure with Random Blocks. As in the completely randomized  $D$ -optimal design for a causal structure, another important class of estimators is the maximum likelihood estimators. We will obtain the FIML estimators for a causal structure with

random blocks and then obtain the covariance matrix estimates for the estimators, or the information matrix estimates. The information matrix estimates will then be used to obtain a  $D$ -optimal design for the FIML estimators.

Starting with the notation from Section 2.1,  $\underline{y}^* - W^* \underline{\delta}^* = Z^* \underline{u}^* + \underline{\varepsilon}^*$  where  $\begin{bmatrix} \underline{u}^* \\ \underline{\varepsilon}^* \end{bmatrix} \sim N\left(\begin{bmatrix} \underline{0} \\ \underline{0} \end{bmatrix}, \begin{bmatrix} \Sigma_u \otimes I_{(b)} & 0 \\ 0 & \Sigma \otimes I_{(n)} \end{bmatrix}\right)$ . Then,  $\mathcal{L}(V, \underline{\delta}^* | \underline{y}^*, W) = (-np/2) \log(2\pi) - (1/2) \log |V| - (1/2) (\underline{y}^* - W^* \underline{\delta}^*)' V^{-1} \underline{y}^* - W^* \underline{\delta}^*$ . To maximize the likelihood,  $(\partial \mathcal{L} / \partial \underline{\delta}^*) = W^{*'} V^{-1} \underline{y}^* - W^{*'} V^{-1} W^* \hat{\underline{\delta}}^*$  and set the partial derivative equal to 0, which gives  $W^{*'} V^{-1} \underline{y}^* - W^{*'} V^{-1} W^* \hat{\underline{\delta}}^* = 0$ . Solving for  $\hat{\underline{\delta}}^*$ , makes  $\hat{\underline{\delta}}^* = (W^{*'} V^{-1} W^*)^{-1} W^{*'} V^{-1} \underline{y}^*$ . Then, find the second partial derivative and take its expected value to obtain the information matrix,  $(\partial^2 \mathcal{L} / \partial (\underline{\delta}^*) \partial (\underline{\delta}^*)) = -W^{*'} V^{-1} W^*$ .

Next,  $\text{Var}^{-1}(\hat{\underline{\delta}}^*) = E(W^{*'} V^{-1} W^*)$ . However, in general, we cannot find the expected value explicitly until the model is known since  $W^*$  includes fixed and random variables. But, for Example 2 above the information matrix is simplified to [30].

TABLE 8: A uniform design for 16 runs with 3 factors and an experimental cube of  $[-1, 1]^3$ .

#	1	2	3	$x_1$	$x_2$	$x_3$
1	1	12	7	-1	7/15	-3/15
2	5	9	1	-7/15	1/15	-1
3	3	4	4	-11/15	-9/15	-9/15
4	11	2	2	5/15	-13/15	-13/15
5	4	14	15	-9/15	11/15	13/15
6	13	7	6	9/15	-3/15	-5/15
7	2	6	12	-13/15	-5/15	7/15
8	14	13	3	11/15	9/15	-11/15
9	15	3	14	13/15	-11/15	11/15
10	10	11	13	3/15	5/15	9/15
11	16	10	9	1	3/15	1/15
12	8	8	16	-1/15	-1/15	1
13	9	5	8	1/15	-7/15	-1/15
14	7	16	5	-3/15	1	-7/15
15	6	1	10	-5/15	-1	3/15
16	12	15	11	7/15	13/15	5/15

TABLE 9: Asymptotic information matrix for Example 2 based on the orthogonal design.

	$\gamma_{11}$	$\gamma_{22}$	$\gamma_{31}$	$b_{12}$
$\gamma_{11}$	16	0	0	0
$\gamma_{22}$	0	16	0	0
$\gamma_{31}$	0	0	16	0
$b_{12}$	0	0	0	1,168

TABLE 10: Asymptotic information matrix for Example 2 based on the 3SLS  $D$ -optimal design.

	$\gamma_{11}$	$\gamma_{22}$	$\gamma_{31}$	$b_{12}$
$\gamma_{11}$	16	0	4	0
$\gamma_{22}$	0	16	0	0
$\gamma_{31}$	4	0	16	0
$b_{12}$	0	0	0	1,360

Based on the information matrix above, a FIML  $D$ -optimal design for Example 2 is equivalent to the design in Table 6. The determinant of the information matrix is 5,280,295.4. Neither the orthogonal design nor the uniform design was as efficient as the new  $D$ -optimal design. The determinants of the information matrices were 4,845,883.1 and 36,622 for the orthogonal design and uniform design, respectively. Thus, the  $D$ -optimal design was approximately 2.1% more  $D$ -efficient than the orthogonal design and 71.1% more  $D$ -efficient than the uniform design. These results are similar to the  $D$ -optimal design for Section 1.2 where both the 3SLS and the FIML estimates had the same  $D$ -optimal design.

## 5. Discussion

Optimal design theory and its applications are generally limited to univariate designs. However, most natural phenomena include multiple endogenous variables which are related in complicated ways. There are two weaknesses concerning univariate designs. First, univariate designs are

incapable of addressing more than one endogenous variable that could be affected by different exogenous variables. The second weakness is that univariate designs ignore the endogenous parameters. Our new approach overcomes both weaknesses. The  $D$ -optimality criteria were developed for both the 3SLS and FIML estimates for a causal structure and for a causal structure with random blocks. In addition, our approach is broadly applicable since it generalizes to any linear causal model with both fixed and random effects, and with any type of nesting and factorial structure.

Our results show that the  $D$ -optimal designs were the same for both FIML and 3SLS estimates. The reader may wonder why they should care about 3SLS  $D$ -optimal design since FIML estimators have traditionally been used in structural equation modeling since 3SLS estimators are not based on optimization. However, 3SLS information matrices are mathematically easier to obtain than FIML information matrices. The interest here is in optimal design, not estimation. Since there was no difference between the FIML and 3SLS  $D$ -optimal designs for both examples and, in general, since the FIML and 3SLS information matrices are asymptotically equivalent [33], we

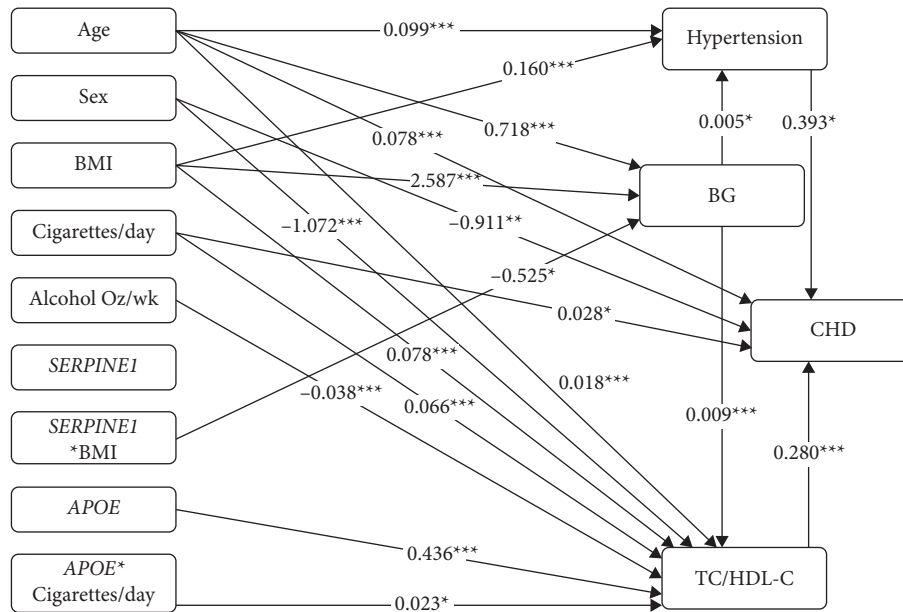


FIGURE 6: Path estimates of SEM of gene-by-environment interaction in the development of CHD from annals of human genetics by Mi et al. [42].

use the 3SLS information matrix both for its simplicity in obtaining a  $D$ -optimal design and to ease hesitancy for scientists to utilize the more complicated FIML information matrix.

In the example of the completely randomized causal structure, the new  $D$ -optimal design increased the information matrix determinant by at least 17% over the univariate optimal design. Because of this increase, the univariate optimal design was only 94% as  $D$ -efficient as the new  $D$ -optimal design. The information matrix of a causal structure depends on parameter values that are unknown. This raises the concern of whether or not the  $D$ -optimal design is robust. Based on the simulation results from Example 1, the  $D$ -optimal design is robust because it did not change when the parameter values changed.

The fact that the information matrix of a causal structure depends on unknown parameter values is not unique to causal structure  $D$ -optimal designs. The same issue arises for  $D$ -optimal designs for random effects, or the covariance parameters, in a split-plot or a blocked design as well where the values of the parameters are needed to obtain the  $D$ -optimal design. The results in our research are consistent with those of Goos and Vandebroek [14, 15] and Goos and Jones [13] where the optimal design did not change when replacing the true parameters for an estimate.

Similarly, for a causal structure model with random blocks, the endogenous parameters had a significant impact on the  $D$ -optimal design. In the univariate mixed model, orthogonal designs are universally optimal [14, 41]. However, it is not optimal for a causal structure model with random blocks. As shown in Example 2, the new  $D$ -optimal design increased the determinant of the information matrix by at least 9%, meaning that the orthogonal design was less efficient than the new  $D$ -optimal design. The orthogonal

design was consistently about 97.8% and 97.9% as  $D$ -efficient as the new 3SLS and FIML  $D$ -optimal designs, respectively.

While it may be a disadvantage that  $D$ -optimal designs require an assumed model structure, the results show that the uniform design, which focuses on distributing the design points uniformly over the experimental domain, is not optimal. The uniform design was about 27.2% as  $D$ -efficient as the new 3SLS  $D$ -optimal design and 28.9% as  $D$ -efficient as the new FIML  $D$ -optimal design. In both cases, the uniform designs consistently lost more than 70% of the information in contrast to the new  $D$ -optimal design. Such a large loss of information will result in inflating the variance and has serious consequences on inferential statistics and the power of the statistical tests.

The designs that were produced using the new criteria show that endogenous parameters have an important effect on  $D$ -optimal designs. The results further show the importance of taking advantage of prior knowledge of the interrelationships of endogenous and exogenous parameters to develop a  $D$ -optimal design using the algorithms and methodology developed here. An application for the use of our algorithms could be in research similar to the heart study by Mi and colleagues [42] who examined the complex interrelationships among physiological, genetic, and environmental factors and their interaction effects on coronary heart disease (CHD), as shown in Figure 6. The data in that study were collected and analysis was performed based on the data alone; there was no pre-planned design. To obtain the most precise estimates to describe the system, causal structure  $D$ -optimal design can be used to determine replication of the levels of each factor such as how many men and women to include in the study, how many of the subjects should be smokers, how many should be obese, or how to

determine an optimal sample of volunteers based on budgetary constraints.

Another practical use of our algorithms and methodology comes from the field of agriculture through research conducted by Campbell and colleagues [43], and Dhungana and colleagues [44] where the objective was to understand how gene-environment interactions (GEI) influence inter-related and complex traits such as grain yield. This subject of future work would help to answer questions like which genotypes should be chosen for the study and what is the combination of those genotypes to obtain optimal estimates. As a multi-stage experiment which took place over a period of three years, our methods could have been implemented to use the preliminary results from the first year of the study to answer those questions if the model was pre-determined.

**5.1. Limitations.** The fact that the algorithm requires the model to be predetermined is a challenge because in most natural phenomena the interrelationships in a system could be complex. In univariate designs, the challenge is reduced because there is only one endogenous variable that is directly affected by an exogenous variable. However, in structural equation modeling, not all exogenous variables affect endogenous variables directly and the interrelationship among the endogenous variables needs to be predetermined, as well. An additional challenge to obtain a  $D$ -optimal design through our algorithm is that some of the estimates of the endogenous parameters are required before an experiment is conducted, which is unknown prior to the experiment. Addressing these limitations is also the subject of future work by using a prior distribution and a Bayesian approach to allow for the uncertainty of the unknown endogenous and exogenous parameters. Finally, how our technique works with other designs such as  $A$ -optimal,  $C$ -optimal, or  $E$ -optimal designs is yet to be explored.

## Appendix

### A. Derivation of the FIML Information Matrix

For the structural equation model in the form  $YB + X\Gamma + E = 0$  where the rows of  $E$  are independently and normally distributed with vector mean zero and a  $p \times p$  variance matrix  $\Sigma$ , the log likelihood function can be given by  $\log L = \text{Constant} + n \log|B| + (n/2)\log|\Sigma^{-1}| - (1/2)\text{tr}[(YB + X\Gamma)'(YB + X\Gamma)\Sigma^{-1}]$ . First, to obtain the FIML estimators, we find the partial derivatives and set them to zero as follows:

$$\frac{\partial \log L}{\partial B} \quad (A.1)$$

$$\longrightarrow n \left( \hat{B}' \right)^{-1} - Y' (Y\hat{B} + X\hat{\Gamma})\hat{\Sigma}^{-1} = 0,$$

$$\frac{\partial \log L}{\partial \Gamma} \quad (A.2)$$

$$\longrightarrow -X' (Y\hat{B} + X\hat{\Gamma})\hat{\Sigma}^{-1} = 0.$$

Because of the invariance properties of the maximum likelihood (ML) estimators [45], it would simplify the derivative if we derive the likelihood with respect to  $\Sigma^{-1}$  instead of  $\Sigma$ . So,

$$\frac{\partial \log L}{\partial \Sigma^{-1}} \quad (A.3)$$

$$\longrightarrow \frac{n}{2} \hat{\Sigma} - \frac{1}{2} (Y\hat{B} + X\hat{\Gamma})' (Y\hat{B} + X\hat{\Gamma}) = 0.$$

Per Kmail [30], algebraic manipulation of equations (A.1) and (A.3) gives

$$\left( \hat{B}' \right)^{-1} \hat{\Gamma}' X' (Y\hat{B} + X\hat{\Gamma})\hat{\Sigma}^{-1} = 0. \quad (A.4)$$

Let  $W = [Y \ X]$ ,  $\hat{Y} = -X\hat{\Gamma}\hat{B}^{-1}$ ,  $\hat{W} = \begin{bmatrix} -X\hat{\Gamma}\hat{B}^{-1} & X \end{bmatrix}$ , and let  $C = \begin{bmatrix} \hat{B} \\ \hat{\Gamma} \end{bmatrix}$  and  $\hat{C} = \begin{bmatrix} \hat{B} \\ \hat{\Gamma} \end{bmatrix}$ . Now rewrite (equations (A.2) and (A.4)) in terms of these new notations. These equations can be written as

$$\begin{bmatrix} \left( \hat{B}' \right)^{-1} \hat{\Gamma}' X' (Y\hat{B} + X\hat{\Gamma})\hat{\Sigma}^{-1} \\ -X' (Y\hat{B} + X\hat{\Gamma})\hat{\Sigma}^{-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (A.5)$$

Then, by rewriting the previous equation,

$$\begin{bmatrix} \left( \hat{B}' \right)^{-1} \hat{\Gamma}' X' \\ -X' \end{bmatrix} (Y\hat{B} + X\hat{\Gamma})\hat{\Sigma}^{-1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (A.6)$$

Next,

$$\underbrace{\begin{bmatrix} -X\hat{\Gamma}\hat{B}^{-1} & X \end{bmatrix}'}_{\hat{W}} \underbrace{\begin{bmatrix} Y & X \end{bmatrix}}_{\hat{W}} \underbrace{\begin{bmatrix} \hat{B} \\ \hat{\Gamma} \end{bmatrix}}_{\hat{C}} \hat{\Sigma}^{-1} = 0. \quad (A.7)$$

Finally,

$$\hat{W}' \hat{W} \hat{C} \hat{\Sigma}^{-1} = 0. \quad (A.8)$$

These are under the same restrictions as (equation (A.2)) and (equation (A.4)), which implies that the only unknown elements of  $C$  in the left-hand side are equated to zero.

Let  $\hat{c}_{(j)}$  be the  $j^{\text{th}}$  column of  $C$  where  $j = 1, \dots, P$ . This leads us to conclude that the  $j^{\text{th}}$  element of  $\hat{c}_{(j)}$  is one since the diagonal of  $B$  consists of 1's. Also, other elements of  $\hat{c}_{(j)}$  are zero by the original model restrictions on  $B$  and  $\Gamma$ . Let  $m_j$  be the number of unknown elements in  $\hat{c}_{(j)}$  and let  $-\hat{\delta}_{(j)}$  be the  $m_j \times 1$  vectors of these unknowns. In other words,  $-\hat{\delta}_{(j)}$  represents the unknown coefficients of the  $j^{\text{th}}$  equation in the original model. Let the columns of  $W = [Y \ X]$  corresponding to unknown elements of  $\hat{c}_{(j)}$  be rearranged as  $n \times m_j$  matrix called  $Q_{(j)}$ . Also, let  $y_{(j)}$  represent the  $j^{\text{th}}$

column of  $Y_{n \times p}$ . The new notations will allow us to rewrite (equation (A.8)) without restrictions. Note that  $W\hat{\underline{c}}_{(j)} = \underline{y}_{(j)} - Q_{(j)}\hat{\underline{\delta}}_{(j)}$  since the coefficient of  $\underline{y}_{(j)}$  is one,  $Q_{(j)}$  has coefficients of  $-\hat{\underline{\delta}}_{(j)}$ , and the rest of columns of  $W$  have coefficients of zero. Thus,

$$W\hat{C} = [\underline{y}_{(1)} - Q_{(1)}\hat{\underline{\delta}}_{(1)}, \underline{y}_{(2)} - Q_{(2)}\hat{\underline{\delta}}_{(2)}, \dots, \underline{y}_{(p)} - Q_{(p)}\hat{\underline{\delta}}_{(p)}]. \tag{A.9}$$

By substituting this quantity in to (equation (A.8)), it becomes

$$\hat{W}' \sum_{i=k}^p \left( \underline{y}_{(k)} - Q_{(k)}\hat{\underline{\delta}}_{(k)} \right) [\sigma^{k1}, \sigma^{k2}, \dots, \sigma^{kp}] = 0, \tag{A.10}$$

where  $\sigma^{kj}$  is the  $kj^{\text{th}}$  element of  $\Sigma^{-1}$ . Consequently, the  $j^{\text{th}}$  column of the left-hand side of (equation (A.10)) is equal to  $\hat{W}' \sum_{k=1}^p (\underline{y}_{(k)} - Q_{(k)}\hat{\underline{\delta}}_{(k)})\sigma^{kj} = 0$ . Use similar notation as before and let  $\hat{Q}_{(j)}$  be the columns of  $\hat{W}'$  that correspond to unknown elements of the vector  $\hat{\underline{c}}_{(j)}$ . For that reason,  $\hat{Q}_{(j)} \sum_{k=1}^p (\underline{y}_{(k)} - Q_{(k)}\hat{\underline{\delta}}_{(k)})\sigma^{kj} = 0$ . is equivalent to the previous equation. The full system is obtained by stacking all of the equations and using a similar technique as the 3SLS. Let,

$$\begin{aligned} Q &= \begin{bmatrix} \hat{Q}_{(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \hat{Q}_{(p)} \end{bmatrix}, \\ \underline{y}^* &= \begin{bmatrix} \underline{y}_{(1)} \\ \vdots \\ \underline{y}_{(p)} \end{bmatrix}, \\ \hat{\underline{\delta}}^* &= \begin{bmatrix} \hat{\underline{\delta}}_{(1)} \\ \vdots \\ \hat{\underline{\delta}}_{(p)} \end{bmatrix}, \\ G &= \begin{bmatrix} \sigma^{11}I & \sigma^{12}I & \dots & \dots & \sigma^{1p}I \\ \sigma^{21}I & \sigma^{22}I & \sigma^{2p}I & & \\ \vdots & \ddots & \vdots & & \\ \vdots & \ddots & \vdots & & \\ \sigma^{p1}I & \sigma^{p2}I & \dots & \dots & \sigma^{pp}I \end{bmatrix}. \end{aligned} \tag{A.11}$$

Then, equation (A.10) simplifies to

$$\hat{Q}'G \left( \underline{y}^* - Q\hat{\underline{\delta}}^* \right) = 0, \tag{A.12}$$

which implies that,

$$\hat{Q}'GQ\hat{\underline{\delta}}^* = \hat{Q}'G\underline{y}^*. \tag{A.13}$$

The advantage of this notation is that the elements of  $\hat{\underline{\delta}}$  are unknown, and it is an unrestricted set of equations. The restrictions on  $B$  and  $\Gamma$  have been resolved by the notation. However, there is a problem in solving for  $\hat{\underline{\delta}}$  since  $G$  is unknown. In addition,  $\hat{Q}$  consists of columns of  $\hat{W} = \begin{bmatrix} -X\hat{\Gamma}\hat{B}^{-1} & X \end{bmatrix}$ , which is unknown because of  $\hat{\Gamma}$  and  $\hat{B}$ . Therefore, replace these values by initial values and replace  $\hat{\Sigma} \otimes I = (1/n)(YB + X\hat{\Gamma})'(YB + X\hat{\Gamma}) \otimes I$ . Thus, (equation (A.13)) becomes

$$\hat{Q}'\hat{G}Q\hat{\underline{\delta}}^* = \hat{Q}'\hat{G}\underline{y}^*. \tag{A.14}$$

Let the maximum likelihood estimate of  $\hat{W} = \begin{bmatrix} -XBA & X \end{bmatrix}$  be denoted by  $\hat{W}$  not by  $\hat{\bar{W}}$ . The estimate of  $\hat{Q}$  which is the arrangement of  $\hat{W}$  is denoted by  $\hat{Q}$ . Finally, the maximum likelihood estimate of  $\hat{\underline{\delta}}^*$  is  $\hat{\underline{\delta}}^*$ .

Using the results from (A.14),  $\hat{Q}'\hat{G}Q\hat{\underline{\delta}}^* = \hat{Q}'\hat{G}\underline{y}^*$ . However,  $A$ ,  $B$ , and  $\hat{G}$  are unknown. Therefore, an iterative approach needs to be used. Start by the 2SLS estimates which are of  $O(n^{-(1/2)})$  in probability. To obtain the second estimates, let  $\hat{\underline{\delta}}_{(2)}^* = \hat{\underline{\delta}}_{(1)}^* + d\hat{\underline{\delta}}^*$ . Using the first-order Taylor expansions, let  $\bar{Q}_{(1)}$  and  $\bar{G}_{(1)}$  be the values of  $\bar{Q}$  and  $\bar{G}$  evaluated at  $\hat{\underline{\delta}}_{(1)}^*$ . Then,  $\bar{Q}_{(1)}\bar{G}_{(1)}(Q\hat{\underline{\delta}}_{(1)}^* - \underline{y}^*) + Q_{(1)}\bar{G}_{(1)}Qd\hat{\underline{\delta}}^* + (dQ'\bar{G}_{(1)} + \bar{Q}_{(1)}d\bar{G})(Q\hat{\underline{\delta}}_{(1)}^* - \underline{y}^*) = 0$ . But since  $dQ'\bar{G}_{(1)}$  and  $\bar{Q}_{(1)}d\bar{G}$  are small in comparison to  $\bar{Q}_{(1)}\bar{G}_{(1)}$ , we ignore the third term and conclude that  $d\hat{\underline{\delta}}^* = (\bar{Q}_{(1)}\bar{G}_{(1)}Q)^{-1}\bar{Q}_{(1)}\bar{G}_{(1)}(\underline{y}^* - Q\hat{\underline{\delta}}_{(1)}^*)$ . Therefore,  $\hat{\underline{\delta}}_{(2)}^* = \hat{\underline{\delta}}_{(1)}^* + d\hat{\underline{\delta}}^* = \hat{\underline{\delta}}_{(1)}^* + (\bar{Q}_{(1)}\bar{G}_{(1)}Q)^{-1}\bar{Q}_{(1)}\bar{G}_{(1)}\underline{y}^* - \hat{\underline{\delta}}_{(1)}^* = (\bar{Q}_{(1)}\bar{G}_{(1)}Q)^{-1}\bar{Q}_{(1)}\bar{G}_{(1)}\underline{y}^*$ . Repeating the previous procedure, we can generalize the results for the  $(r + 1)$  estimates where  $\hat{\underline{\delta}}_{(r+1)}^* = (\bar{Q}_r\bar{G}_rQ)^{-1}\bar{Q}_r\bar{G}_r\underline{y}^*$ . Kmail [30] shows the asymptotic variance matrix of the estimates is the  $\text{Var}(\hat{\delta}_{\text{FIML}}) = (\bar{Q}'\bar{G}\bar{Q})^{-1} + O(n^{-1})$ .

### Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during the current study.

## Disclosure

This research is presented in part as the first authors dissertation research at <https://digitalcommons.unl.edu/statisticsdiss/21/>. Any mistakes, omissions, etc., are that of the authors alone.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors thank Dr. Erin Blankenship, Dr. Stephen Kachman, and Dr. Allan Peterson for their critical reviews of the first author's Ph.D. dissertation upon which this paper is based.

## References

- [1] J. J. Milander, S. C. Mason, G. Kruger, T. D. Galusha, and Z. Kmail, "Waxy maize yield and components as influenced by environment, water regime, and hybrid," *Maydica*, vol. 60, 2015.
- [2] S. Altan and J. Singh, *Recent Advances in Experimental Designs and Related Topics*, Nova Science Publishers, Inc, Hauppauge, NY, USA, 2001.
- [3] J. Kiefer and J. Wolfowitz, "Optimum designs in regression problems," *The Annals of Mathematical Statistics*, vol. 30, no. 2, pp. 271–294, 1959.
- [4] R. D. Cook and C. J. Nachtrheim, "A comparison of algorithms for constructing exact D-optimal designs," *Technometrics*, vol. 22, no. 3, pp. 315–324, 1980.
- [5] O. Dykstra, "The augmentation of experimental data to maximize  $|X'X|$ ," *Technometrics*, vol. 13, no. 3, pp. 682–688, 1971.
- [6] V. V. Fedorov, *Theory of Optimal Experiments*, Academic Press, New York, NY, USA, 1972.
- [7] T. J. Mitchell, "An algorithm for the construction of "D-optimal" experimental designs," *Technometrics*, vol. 16, no. 2, pp. 203–210, 1974.
- [8] T. J. Mitchell, "Computer construction of "D-optimal" first-order designs," *Technometrics*, vol. 16, no. 2, pp. 211–220, 1974.
- [9] J. E. Atkins and C. S. Cheng, "Optimal regression designs in the presence of random block effects," *Journal of Statistical Planning and Inference*, vol. 77, no. 2, pp. 321–335, 1999.
- [10] A. C. Atkinson and A. N. Donev, "The construction of exact D-optimum experimental designs with application to blocking response surface designs," *Biometrika*, vol. 76, no. 3, pp. 515–526, 1989.
- [11] C. S. Cheng, "Optimal regression designs under random block-effects models," *Statistica Sinica*, vol. 5, pp. 485–497, 1995.
- [12] R. D. Cook and C. J. Nachtsheim, "Computer-aided blocking of factorial and response-surface designs," *Technometrics*, vol. 31, no. 3, pp. 339–346, 1989.
- [13] P. Goos and B. Jones, *Optimal Design of Experiments: A Case Study Approach*, John Wiley & Sons, Hoboken, NJ, USA, 2011.
- [14] P. Goos and M. Vandebroek, "optimal response surface designs in the presence of random block effects," *Computational Statistics & Data Analysis*, vol. 37, no. 4, pp. 433–453, 2001.
- [15] P. Goos and M. Vanderbroek, "D-optimal split-plot designs with given numbers and sizes of whole plots," *Technometrics*, vol. 45, no. 3, pp. 235–245, 2003.
- [16] S. Gupta, "SEM for experimental designs: an information systems example," *Electronic Journal of Business Research Methods*, vol. 12, no. 1, pp. 27–40, 2014.
- [17] K. G. Jöreskog, "A general method for estimating a linear structural equation system," *ETS Research Bulletin Series*, vol. 1970, no. 2, 1970.
- [18] K. Bollen, *Structural Equations with Latent Variables*, John Wiley & Sons, Hoboken, NJ, USA, 1989.
- [19] M. Soumaya, N. Gaffke, and R. Schwabe, "Optimal design for multivariate observations in seemingly unrelated linear models," *Journal of Multivariate Analysis*, vol. 142, pp. 48–56, 2015.
- [20] R. Schwabe, "Optimal design of experiments for multivariate response in two-factor linear models," *Multidimensional Statistical Analysis and Theory of Random Matrices*, pp. 235–242, 1996.
- [21] T. R. Brick and D. H. Bailey, "Rock the MIC: the matrix of implied causation, a tool for experimental design and model checking," *Advances in Methods and Practices in Psychological Science*, vol. 3, no. 3, pp. 286–299, 2020.
- [22] S. Sussex, C. Uhler, and A. Krause, "Near-optimal multi-perturbation experimental design for causal structure learning," *Advances in Neural Information Processing Systems*, vol. 34, 2021.
- [23] P. J. Ferraro, "Counterfactual thinking and impact evaluation in environmental policy," *New Directions for Evaluation*, vol. 122, pp. 75–84, 2009.
- [24] B. Jones and P. Goos, "A candidate-set-free algorithm for generating D-optimal split-plot designs," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 56, no. 3, pp. 347–364, 2007.
- [25] K. Mylona, P. Goos, and B. Jones, "Optimal design of blocked and split-plot experiments for fixed effects and variance component estimation," *Technometrics*, vol. 56, no. 2, pp. 132–144, 2014.
- [26] A. M. Elsawah, "A novel non-heuristic search technique for constructing uniform designs with a mixture of two- and four-level factors: a simple industrial applicable approach," *Journal of the Korean Statistical Society*, 2022.
- [27] K. T. Fang, "The uniform designs: application of number-theoretic methods in experimental design," *Acta Mathematicae Applicatae Sinica*, vol. 3, pp. 363–372, 1980.
- [28] Y. Wang and K. T. Fang, "A note on uniform distribution and experimental design," *Chinese Science Bulletin*, vol. 26, no. 764, pp. 485–489, 1981.
- [29] K. T. Fang and R. Li, "Uniform design for computer experiments and its optimal properties," *International Journal of Materials and Product Technology*, vol. 25, no. 1/2/3, pp. 198–210, 2006.
- [30] Z. Kmail, "Optimal design for a causal structure," Doctoral dissertation, University of Nebraska—Lincoln, Lincoln, NE, USA, 2019.
- [31] J. D. Sargan, "Three-stage least-squares and full maximum likelihood estimates," *Econometrica*, vol. 32, no. 1/2, pp. 77–81, 1964.
- [32] A. Zellner and H. Theil, "Three-stage least squares: simultaneous estimation of simultaneous equations," *Econometrica*, vol. 30, no. 1, pp. 54–78, 1962.
- [33] J. Durbin, "Maximum likelihood estimation of the parameters of a system of simultaneous regression equations," *Econometric Theory*, vol. 4, no. 1, pp. 159–170, 1988.

- [34] B. Jones and P. Goos, "An algorithm for finding D-efficient equivalent-estimation second-order split-plot designs," *Journal of Quality Technology*, vol. 44, no. 4, pp. 363–374, 2012.
- [35] P. Goos, B. Jones, and U. Syafitri, "I-optimal design of mixture experiments," *Journal of the American Statistical Association*, vol. 111, no. 514, pp. 899–911, 2016.
- [36] P. Goos and U. Syafitri, "V-optimal mixture designs for the qth degree model," *Chemometrics and Intelligent Laboratory Systems*, vol. 136, pp. 173–178, 2014.
- [37] P. Goos, U. Syafitri, B. Sartono, and A. R. Vazquez, "A nonlinear multidimensional knapsack problem in the optimal design of mixture experiments," *European Journal of Operational Research*, vol. 281, no. 1, pp. 201–221, 2020.
- [38] P. Goos and K. Mylona, "Quadrature methods for bayesian optimal design of experiments with nonnormal prior distributions," *Journal of Computational & Graphical Statistics*, vol. 27, no. 1, pp. 179–194, 2018.
- [39] K. T. Fang and D. K. Lin, "Uniform experimental designs and their applications in industry," in *Handbook of Statistics 22 (131–170)*, R. Khattree and C. R. Rao, Eds., Elsevier, Amsterdam, Netherlands, 2003.
- [40] K. T. Fang, C. X. Ma, D. Maringer, Y. Tang, and P. Winker, "Uniform design tables," <https://www.math.hkbu.edu.hk/UniformDesign/>.
- [41] C. S. Cheng, "Orthogonal arrays with variable numbers of symbols," *Annals of Statistics*, vol. 8, no. 2, pp. 447–453, 1980.
- [42] X. Mi, K. M. Eskridge, V. George, and D. Wang, "Structural equation modeling of gene-environment interactions in coronary heart disease," *Annals of Human Genetics*, vol. 75, no. 2, pp. 255–265, 2011.
- [43] B. T. Campbell, P. S. Baenziger, K. S. Gill et al., "Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat," *Crop Science*, vol. 43, no. 4, pp. 1493–1505, 2003.
- [44] P. Dhungana, K. M. Eskridge, P. S. Baenziger, B. T. Campbell, K. S. Gill, and I. Dweikat, "Analysis of genotype-by-environment interaction in wheat using a structural equation model and chromosome substitution lines," *Crop Science*, vol. 47, no. 2, pp. 477–484, 2007.
- [45] G. Casella and R. L. Berger, *Statistical Inference*, Thomson Learning, Pacific Grove, CA, USA, 2001.