**Title**
# On the Equivalence of Constrained and Compound Optimal Designs



**Authors**
R. Dennis Cook
Weng-Kee Wong

**Publication Date**
2011-10-24

# On the Equivalence of Constrained and Compound Optimal Designs

R. Dennis Cook and Weng Kee Wong*

Constrained and compound optimal designs represent two well-known methods for dealing with multiple objectives in optimal design as reflected by two functionals $\phi_1$ and $\phi_2$ on the space of information matrices. A constrained optimal design is constructed by optimizing $\phi_2$ subject to a constraint on $\phi_1$, and a compound design is found by optimizing a weighted average of the functionals $\phi = \lambda\phi_1 + (1-\lambda)\phi_2$, $0 \le \lambda \le 1$. We show that these two approaches to handling multiple objectives are equivalent.

KEY WORDS: D optimality; Efficiency; Information Matrix; Large sample design.

## 1. INTRODUCTION

Consider the problem of selecting an experimental design to furnish information on the linear model $y = f^T(x)\beta + \varepsilon$, where $y$ is the observable response variable, $\beta$ is a $p \times 1$ vector of unknown parameters, the errors are assumed to be independent and identically distributed with mean 0 and constant variance $\sigma^2$, and $f$ is a known continuous function of the $q \times 1$ vector of design variables $x$ that is constrained to lie in a compact subset $\chi$ of $R^q$. An experimental design can be described as a probability mass function $\xi$ that places total mass on a finite collection of $s$ points in the design region $\chi$. In small sample design, $n\xi(x_i)$ is required to be an integer for $i = 1, \ldots, s$, where $n$ is the sample size. In large sample design, this integral restriction is not imposed. Our considerations here are limited to the construction of large sample designs.

Virtually all optimality criteria characterize the worth of a design through a concave functional $\phi$ that depends only on the $p \times p$ information matrix

$$M(\xi) = \sum_{i=1}^{s} \xi(x_i) f(x_i) f^T(x_i). \qquad (1)$$

Such criteria $\phi$ are defined on the space of information matrices $\mathcal{M}$ and are chosen to reflect the goals of the experiment. We shall usually denote $\phi(M(\xi))$ by $\phi(\xi)$ for notational convenience. An optimal design problem then reduces to the problem of finding the design that maximizes $\phi$ over $\mathcal{M}$. Perhaps the most common criterion is $D$ optimality, which corresponds to $\phi(\xi) = \log|M(\xi)|$. Another is A optimality, $\phi(\xi) = -\operatorname{tr} M(\xi)^{-1}$. (For additional background on optimal design, see Fedorov 1972; Pazman 1986; and Silvey 1980.) Reviews of recent developments in optimal design are available in articles by Atkinson (1988) and Pukelsheim (1987).

This general approach to experimental design has been criticized because it apparently reflects an overly myopic view: that the various practical goals of a design cannot be incorporated by choosing one of the usual objective functions. Box and Draper (1975) gave a comprehensive list of 14 criteria that should be considered in any design problem. Several of these criteria by themselves would be difficult to represent adequately by a single objective function, to say nothing of the elusive task of combining all 14.

There seem to be two standard methods in the literature for molding optimal designs to satisfy multiple objectives: constrained and compound optimal design. We briefly review these methods in Sections 2 and 3. In Section 4 we show that these approaches are essentially equivalent. We provide examples in Section 5 and concluding comments in Section 6.

## 2. CONSTRAINED OPTIMAL DESIGN

Suppose that the experiment must satisfy two distinct objectives as implemented by the concave functions $\phi_1$ and $\phi_2$, representing primary and secondary criteria that are both defined on $\mathcal{M}$. To balance these competing objectives, a design $\xi$ could be selected to do as well as possible on the secondary objective $\phi_2$ provided that it does not stray too far from the optimal design $\xi^1$ under $\phi_1$ alone. These ideas can be implemented by selecting a design $\xi$ to maximize $\phi_2$ subject to the constraint that $\phi_1(\xi) \ge c$, where $c$ is a user-selected constant. Let $\Delta_c$ denote the set of all such constrained optimal designs for a specific choice of $c$, $\Delta_c = \{\xi \mid \xi \text{ maximizes } \phi_2$ subject to the constraint $\phi_1(\xi) \ge c\}$. It is easily verified that $\Delta_c$ is convex. Generally, designs in $\Delta_c$ can have distinct information matrices. When this happens, it seems sensible to add the further constraint that the required design $\xi_c$ satisfy $\xi_c = \operatorname{argmax}_{\Delta_c} \phi_1(\xi)$, where the maximum is taken over all designs in $\Delta_c$. In other words, if there is a choice between designs in $\Delta_c$, then we select the design that does the best on the primary criterion $\phi_1$.

In what follows, any design with a subscript "$c$" indicates a constrained optimal design.

To fix ideas, we next turn to a simple example. Consider a situation where primary interest is in precise parameter estimation for the model $y = f_1^T(x)\beta_1 + \varepsilon$, and secondary interest is in detecting lack of fit by using the expanded model $y = f^T(x)\beta + \varepsilon = f_1^T(x)\beta_1 + f_2^T(x)\beta_2 + \varepsilon$. Let $p$, $p_1$, and $p_2$ denote the dimensions of $\beta$, $\beta_1$, and $\beta_2$. The information matrix corresponding to $f$ is denoted by $M$ and $D = M^{-1}$. Partition $M = (M_{ij})$ and $D = (D_{ij})$ $i, j = 1, 2$, so that $M_{11}$ is the information matrix for $f_1$.

* R. Dennis Cook is Professor, Department of Applied Statistics, University of Minnesota, St. Paul, MN 55108. Weng Kee Wong is Assistant Professor, Department of Biostatistics, University of California, Los Angeles, CA 90024. The authors thank V. V. Fedorov who provided helpful comments on an earlier version of this work, which was supported in part by National Science Grant DMS-9212413. They also thank the referees and an associate editor who provided useful comments on an earlier version of this manuscript.


Journal of the American Statistical Association
June 1994, Vol. 89, No. 426, Theory and Methods

There are several general ways of proceeding in this example. One is to select a design for discriminating between the rival models. But discrimination is not the primary goal in this example, and it is well known that designs for precise estimation of $\beta_2$ can be inefficient for estimation of $\beta_1$. An alternative is to use sequential designs that first concentrate observations on model discrimination and then, as the model becomes known, gradually transfer effort to parameter estimation (see, for example, Hill, Hunter, and Witchern 1968). This may be an effective strategy when sequential experimentation is possible.

Alternatively, we could ask to do as well as possible for estimating $\beta_2$ subject to the restriction that a specified minimum efficiency be maintained for the primary model. Specifically, let

$$\phi_1(\xi) = \log\left(\frac{|M_{11}(\xi)|}{|M_{11}(\xi^1)|}\right)^{1/p_1} \tag{2}$$

denote the log efficiency of the design $\xi$ relative to the optimal design $\xi^1$ for $\beta_1$ in the primary model and let

$$\phi_2(\xi) = \log\left(\frac{|D_{22}^{-1}(\xi)|}{|D_{22}^{-1}(\xi^2)|}\right)^{1/p_2} \tag{3}$$

denote the log efficiency of the design $\xi$ relative to the optimal design $\xi^2$ for $\beta_2$ in the expanded model. Then the objective is to maximize $\phi_2$ subject to the constraint that $\phi_1(\xi) \geq c$, where $c \in [-\infty, 0]$. In this example, $\phi_1$ and $\phi_2$ represent the same type of functional: They are both log determinants, a condition that is unnecessary in general.

Alternatively, this example could have been formulated by interchanging the roles of $\phi_1$ and $\phi_2$ so that $\phi_2$ is now the primary criterion. Further, if we consider polynomial regression and set $f^T(x) = (1, x, x^2, \ldots, x^p)$ and $f_2(x) = x^p$, then the constrained problem reduced to finding the $C$-restricted $D$-optimal design suggested in the pioneering paper by Stigler (1971). An algorithm for constructing these designs was provided by Mikulecka (1983); see also the work of Pazman (1986). Using a technique involving canonical moments, Studden (1982) investigated design construction under a generalization of Stigler's criterion.

More recently, Lee (1987, 1988) introduced several constrained optimality criteria and provided necessary and sufficient conditions for a design to be optimal. Lee provided a review of the literature in the area and further motivation for considering constrained problems, mentioning that constrained designs are an alternative to compound designs. Böhning (1981) discussed algorithms for the construction of constrained designs.

## 3. COMPOUND OPTIMAL DESIGNS

As an alternative to constrained design, we could consider compound designs based on a weighted average of the functionals $\phi_1$ and $\phi_2$,

$$\phi(\xi \mid \lambda) = \lambda\phi_1(\xi) + (1 - \lambda)\phi_2(\xi), \tag{4}$$

where $0 \leq \lambda \leq 1$ is a user-selected constant. This is essentially the approach suggested by Läuter (1974, 1976). When used in the context of competing models, this approach seems to be based on the rationale that we would like our design to be reasonable no matter which model is finally selected. Cook and Nachtsheim (1982) adapted Läuter's approach for linear optimality criteria. Dette (1990) investigated compound optimal designs for polynomial regression.

Compound designs can be constructed by using standard methodology. Equivalence theorems for specific versions of (4) are available from Läuter (1974, 1976). General equivalence theory, as described by Fedorov (1980, thm. 1) and Kiefer (1974), can often be easily adapted to (4) for use in practice. Similar comments apply to standard design algorithms. The user-selected parameter $\lambda$ is somewhat annoying, however. It is intended to reflect the relative importance of the criteria $\phi_1$ and $\phi_2$, but the connection with the efficiency of a design is unclear. Different choices for the scales of the criteria are also a practical complication. Does choosing $\lambda = \frac{1}{2}$ necessarily imply equal interest in $\phi_1$ and $\phi_2$, for example? For these reasons, it can be difficult to implement (4) in an informed fashion.

In what follows, any design with a subscript "$\lambda$" indicates a compound optimal design maximizing (4). When we want to indicate a particular value, we use the notational form $\xi_{\lambda=.5}$, which indicates a compound optimal design with $\lambda = .5$.

## 4. EQUIVALENCE OF CONSTRAINED AND COMPOUND OPTIMAL DESIGN

In this section we argue that the constrained and compound design problems described in Sections 2 and 3 are essentially equivalent.

### 4.1 Assumptions

We assume that the primary and secondary criteria—$\phi_1$ and $\phi_2$—are monotonic functions of the design efficiencies relative to the optimal designs under each criterion alone and that they are both defined on a space of information matrices $\mathcal{M}$. Letting $\mathcal{M}^+$ denote the set of nonsingular information matrices, the constrained design problem is

$$\text{maximize } \phi_2(M(\xi)) \text{ over } M(\xi) \in \mathcal{M}^+ \quad \text{subject to } \phi_1(M(\xi)) \geq c, \tag{5}$$

where for definiteness we use the range of $c$ corresponding to (2), $c \in (-\infty, 0)$. The extreme points $c = 0$ and $c = -\infty$ are uninteresting special cases. If $c = -\infty$, for example, (5) is no longer constrained. The compound design problem is to maximize (4) over $M(\xi) \in \mathcal{M}^+$ for $\lambda \in (0, 1)$. We assume that solutions to these problems exist and are in $\mathcal{M}^+$. Otherwise, in the terminology of Pazman (1986, p. 76), our intent is that $\phi(\xi \mid \lambda)$ should be a "global" criterion for $0 < \lambda < 1$, whereas $\phi_1$ and $\phi_2$ may be "partial" criteria. Specifically, we impose the following additional conditions:

1. $\phi(\xi \mid \lambda)$ is assumed to be concave on $\mathcal{M}$ and strictly concave on $\mathcal{M}^+$ for $\lambda \in (0, 1)$. Consequently, the information matrices of designs optimal under (4) are unique for $0 < \lambda < 1$.
2. $\phi(\xi \mid \lambda)$ is assumed to be continuous in $\mathcal{M}$, whereas $\phi_1$ and $\phi_2$ are assumed to be continuous in $\mathcal{M}^+$. Conditions

necessary to ensure continuity have been given by Pazman (1986, p. 109).

### 4.2 Design Characterizations

The following lemmas are instrumental in understanding the relationship between the constrained design problem (5) and the compound design problem (4). The justifications are available in the Appendix.

*Lemma 1.* For $\lambda \in (0, 1)$, let $\xi_\lambda$ maximize the functional $\phi(\xi \mid \lambda)$ as defined in (4) and let $c_\lambda = \phi_1(\xi_\lambda)$, the primary design criterion evaluated at $\xi_\lambda$. Then $\xi_\lambda$ maximizes $\phi_2(\xi)$ subject to the constraint $\phi_1(\xi) \geq c_\lambda$.

The central consequence of Lemma 1 is that the set of designs $\{\xi_\lambda \mid 0 < \lambda < 1\}$ is a subset of the union over $c$ of $\Delta_c$. In other words, every solution of the compound optimization problem (4) is also a solution to the constrained optimization problem for some $c$. Beginning with the constrained problem (5), if a $\lambda$ exists such that $c = \phi_1(\xi_\lambda)$ and we can find it, a constrained optimal design can be constructed by using (4). For this idea to work, however, we need to know something of how $\phi_1(\xi_\lambda)$ and $\phi_2(\xi_\lambda)$ behave as a function of $\lambda$.

*Lemma 2.* $\phi(\xi_\lambda \mid \lambda)$ and $\phi_1(\xi_\lambda)$ are continuous functions of $\lambda \in (0, 1)$.

*Lemma 3.* $\phi_1(\xi_\lambda)$ is a nondecreasing function of $\lambda$ on $[0, 1]$.

*Lemma 4.* $\phi_2(\xi_\lambda)$ is a nonincreasing function of $\lambda$ on $[0, 1]$.

The behavior of $\phi_1(\xi_\lambda)$ and $\phi_2(\xi_\lambda)$ as described in Lemmas 3 and 4 is not surprising. Clearly, when faced with two competing objective functions, any gain in $\phi_1$ has to be met by a corresponding sacrifice in terms of $\phi_2$. The next lemma is the central result connecting constrained optimal designs with compound optimal designs according to the functional $\phi$. Recall that $\xi_{\lambda=0}$ and $\xi_{\lambda=1}$ denote compound optimal designs for $\lambda = 0$ and $\lambda = 1$; that is, $\xi_{\lambda=1} = \xi^1$ and $\xi_{\lambda=0} = \xi^2$.

*Lemma 5.* Assume that

$$\lim_{\lambda \uparrow 1} \phi_1(M(\xi_\lambda)) = \phi_1(M(\xi_{\lambda=1}))$$

$$\text{and} \quad \lim_{\lambda \downarrow 0} \phi_1(M(\xi_\lambda)) = \phi_1(M(\xi_{\lambda=0})). \quad (6)$$

Then every solution $\xi_c$ to the constrained design problem (5) for $c \in (-\infty, 0)$ corresponds to a solution to the compound design problem (4) for some $\lambda$.

It follows from Lemmas 1 and 5 that constrained optimal designs are compound optimal designs and vice versa. Some consequences of this are illustrated in the next section.

## 5. EXAMPLES

In this section we apply the results from the previous section in three examples. Our approach is a graphical one that may be useful in practice. The first example has been described by Stigler (1972), Studden (1982), and Lee (1988). We choose it to show that the same solution to the problem can be found graphically without resorting to the theory of canonical moments, as did Studden (1982). In both examples we construct compound optimal designs $\xi_\lambda$ and then use the results of Section 4 to find the constraint bound $c$ as a function of $\lambda$ for the equivalent constrained design problem. Instead of working with constraints of the form $\phi_1 \geq c$, however, we characterize the constraint corresponding to a particular value of $\lambda$ in terms of design efficiency. This is helpful, because in practice it is usually easiest to work in a scale that has the same meaning from problem to problem. Accordingly, let $E_k(\xi)$ denote the efficiency of $\xi$ relative to the optimal design $\xi^k$ under $\phi_k$ alone, $k = 1, 2$. Explicit definitions of $E_k$ will be given for the functionals used in the examples that follow. Because $\phi_k$ is a monotonic function of $E_k$, the constrained optimal design problem can be reformulated: maximize $\phi_2(\xi)$ subject to $E_1(\xi) \geq e_1$, where $e_1 \in (0, 1)$.

Because all examples assume a polynomial regression model of degree $p$, it will be helpful to denote the information matrix for the design $\xi$ by $M_p(\xi)$ and denote its variance function at the point $x$ by $d_p(x, \xi) = f_p^T(x) M_p^{-1}(\xi) f_p(x)$, where $f_p^T(x) = (1, x, x^2, \ldots, x^p)$.

*Example 1.* Consider linear and quadratic polynomial regression models defined on $\chi = [-1, 1]$. Assume that primary interest is in detecting lack of fit through precise estimation of the quadratic coefficient in the quadratic model and that secondary interest is in precise estimation of the parameters in the linear model. This situation can be quantified by setting

$$\phi_1(\xi) = -\frac{b^T M_2^{-1}(\xi) b}{4} = -[E_1(\xi)]^{-1} \quad (7)$$

and

$$\phi_2(\xi) = \log |M_1(\xi)|^{1/2} = \log E_2(\xi), \quad (8)$$

where $b^T = (0, 0, 1)$. Both $\phi_1$ and $\phi_2$ are concave functionals and are expressed as monotonic functions of design efficiency, because $b^T M_2^{-1}(\xi^1) b = 4$ and $|M_1(\xi^2)| = 1$. When a lower bound is placed on $\phi_1(\xi)$, the solution to this problem is Stigler's (1971) $C$-restricted $D$-optimal design.

It is easily seen that the compound functional $\phi(\xi \mid \lambda)$ is concave, and for $\lambda > 0$ it takes the value $-\infty$ whenever $\xi$ is supported by less than three points. The optimal compound design $\xi_\lambda$ is supported symmetrically at $\pm 1$ and 0. It is well known that $\xi_{\lambda=0}(1) = \frac{1}{2}$ and $\xi_{\lambda=1}(1) = \frac{1}{4}$, where $\xi_\lambda(x)$ denotes the design mass at $x$, and thus the optimal designs for $\phi_1$ and $\phi_2$ do not have identical support. Using these results, we find that

$$\xi_\lambda(1) = \tfrac{1}{4} + \tfrac{1}{4}(1 - E_1(\xi_\lambda))^{1/2} \quad (9)$$

and

$$[E_2(\xi_\lambda)]^2 = \tfrac{1}{2} + \tfrac{1}{2}(1 - E_1(\xi_\lambda))^{1/2}. \quad (10)$$

Equation (9) shows how the mass at 1 changes as a function of $\lambda$ and provides an implicit connection with Stigler's $C$-restricted $D$-optimal designs: Stigler's $C = 4/E_1(\xi_\lambda)$ (see also Lee 1988, ex. 5.1). Clearly, in this example the set of compound optimal designs (9) for $\lambda \in [0, 1]$ generates the set of $C$-restricted $D$-optimal designs for Stigler's $C \geq 4$. From (10) it follows that $E_2$ is a decreasing function of $E_1$, as expected.

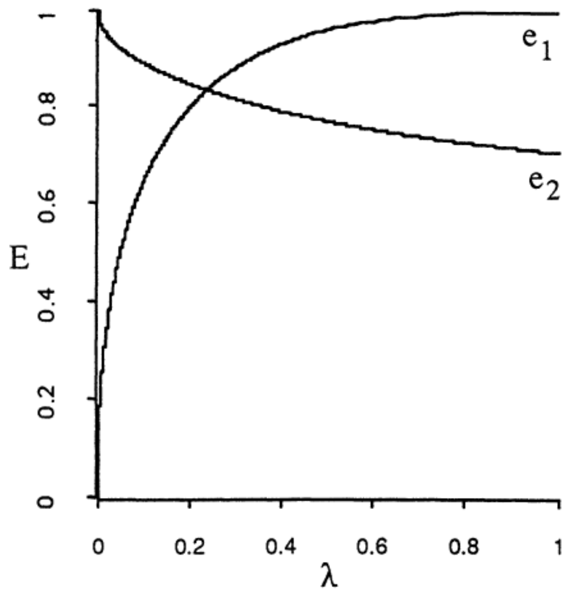


Figure 1. Efficiencies E Versus λ for Example 1.

Further, using (9), we obtain the values given in the first two lines of table 1 in Studden (1982).

To connect explicitly the constrained design problems and the compound design problems, the relationship between a constraint expressed as $E_1(\xi) \geq e_1$ (or $E_2(\xi) \geq e_2$) and the corresponding value of λ must be established. Using results from Fedorov (1980, thm. 1), it can be shown that $\xi_\lambda$ maximizes $\phi(\xi | \lambda)$ if and only if

$$2(1-\lambda)d_1(x, \xi_\lambda) + \lambda\{b^T M_2^{-1}(\xi_\lambda) f_2(x)\}^2 - 4(1-\lambda) - \lambda b^T M_2^{-1}(\xi_\lambda) b \leq 0 \quad (11)$$

for all $x$ in $[-1, 1]$. Further (11) becomes an equality at the support points for $\xi_\lambda$. In that case, substituting (9) into (11) yields

$$\lambda = \frac{e_1^2}{4 + 4(1-e_1)^{1/2} - 4e_1 + e_1^2}. \quad (12)$$

Figure 1, constructed using (10) and (12), shows the relationship between optimal designs with efficiency constraints and compound optimal designs. For example, a design that maximizes $\phi_2$ subject to the constraint $E_1(\xi) \geq .6$ can be found by maximizing $\phi(\xi | \lambda)$ with $\lambda \approx .1$. Figure 1 contains useful information on the interpretation of λ as well. In particular, it might be felt that setting $\lambda = .5$ would yield a design in which equal interest is placed on the two criteria. But from Figure 1, the compound design problem with $\lambda = .5$ is equivalent to the constrained problem in which we maximize $\phi_2$ subject to the constraint that $E_1(\xi) \geq .96$. The resulting constrained design has $E_2(\xi_{\lambda=.5}) \approx .78$. In terms of the efficiencies, placing equal interest on the two criteria would seem to require $\lambda = .25$, because at that point $E_1(\xi_\lambda) \approx E_2(\xi_\lambda) \approx .84$. Finally, reconstructing the plot in Figure 1 so that the horizontal axis is $1 - \lambda$ rather than λ provides the corresponding plot for maximizing $\phi_1$ subject to a constraint on $\phi_2$.

*Example 2.* For the simple linear regression model $f_1^T(x) = (1, x)$ on $\chi = [-1, 1]$, consider balancing $A$ optimality with precise estimation of the response at the point $z = .75$:

$$\phi_1(\xi) = -d_1(z, \xi)/d_1(z, \xi^1) = -[E_1(\xi)]^{-1}$$

and

$$\phi_2(\xi) = -\text{tr}\, M_1^{-1}(\xi)/\text{tr}\, M_1^{-1}(\xi^2) = -[E_2(\xi)]^{-1}.$$

The design $\xi^1$ is optimal for $\phi_1$ and has the minimum variance possible for a fitted value at the point $z$. This minimum variance is the same as that obtained under the design that places mass 1 at $z$. The design $\xi^2$ is the $A$-optimal design. Both $\xi^1$ and $\xi^2$ are supported at $\pm 1$, with the masses at 1 being $\frac{7}{8}$ and $\frac{1}{2}$. Thus $d_1(z, \xi^1) = 1$ and $\text{tr}[M_1^{-1}(\xi^2)] = 2$.

Define $w(x) = (9x^2 - 25x + 16)^{1/2}$ and $g(x) = (3x + 4 - w(x))/8$ for $0 \leq x \leq 1$. Then

$$\xi_\lambda(1) = g(E_1(\xi_\lambda)), \quad (13)$$

which is the analog of (9) for this example. Next, again from Fedorov (1980, thm. 1) $\xi_\lambda$ satisfies

$$\lambda(f_1^T(x) M_1^{-1}(\xi_\lambda) f_1(z))^2 + (1-\lambda) f_1(x) M_1^{-2}(\xi_\lambda) f_1(x) = \lambda d_1(z, \xi_\lambda) + (1-\lambda)\text{tr}\, M_1^{-1}(\xi_\lambda)$$

at the support points $x = \pm 1$. Substituting (13) into this equation yields

$$\lambda = \frac{8(w(e_1) - 3e_1)}{34g(e_1) - 17 - 48g^2(e_1)},$$

which is the analog of (12) for $e_1 \geq .64$. From this we constructed Figure 2. The relationships in Figure 2 are, of course, qualitatively similar to those in Figure 1. But note that the value of λ at which the efficiencies are equal is much larger than that for Figure 1. Generally, the interpretation of λ depends heavily on the functionals involved. Useful inter-

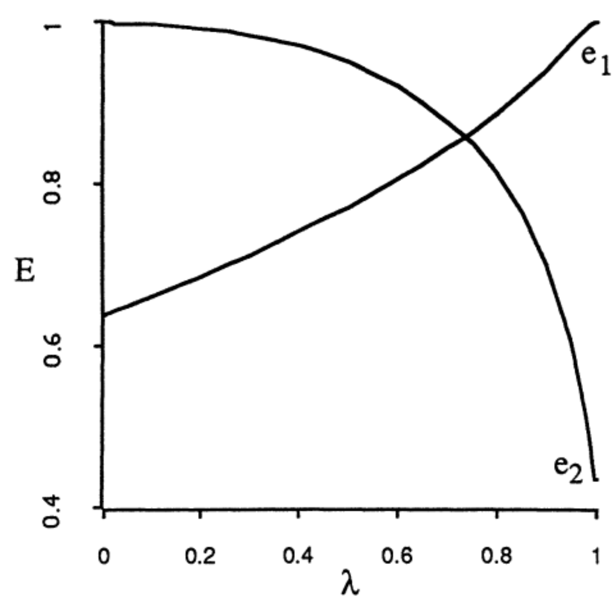


Figure 2. Efficiencies E Versus λ for Example 2.

pretations can always be obtained by reference to the corresponding constrained problem, however.

*Example 3.* This final brief example is intended to illustrate what can happen in situations where the optimal designs for $\phi_1$ and $\phi_2$ have distinct supports.

Suppose that $\chi = [-1, 1]$, $f^T(x) = (1, x)$ and the error variance at $x$ is $\sigma^2(x) = (2 - x^2)^{-1}$. The criteria of interest correspond to $A$-optimal and $D$-optimal designs,

$$\phi_1(\xi) = -\frac{\operatorname{tr} M_1^{-1}(\xi)}{\operatorname{tr} M_1^{-1}(\xi^1)} \quad \text{and} \quad \phi_2(\xi) = \frac{1}{2}\ln\left\{\frac{|M_1(\xi)|}{|M_1(\xi^2)|}\right\},$$

where $\xi^1$ and $\xi^2$ are the optimal designs under $\phi_1$ and $\phi_2$, as defined previously. Using standard design theory, it is easily verified that $\xi^1$ is equally supported at $\pm a$, with $a = (\sqrt{3} - 1)^{1/2}$, and $\xi^2$ is equally supported at $\pm d$, with $d = (2/3)^{1/2}$. The compound optimal design is supported equally at $\pm b$, where $b$ is related to $\lambda$ by

$$\lambda = \frac{4b^2 - 8b^4 + 3b^6}{3b^6 - 4\sqrt{3}b^4 + (20 - 8\sqrt{3})b^2 - 8(2 - \sqrt{3})}.$$

The support point $b$ varies between the minimum $d = (2/3)^{1/2}$ at $\lambda = 0$ and the maximum $a = (\sqrt{3} - 1)^{1/2}$ at $\lambda = 1$. Graphs of the efficiency can now be constructed as in Examples 1 and 2, but the general interpretation is the same.

## 6. DISCUSSION

Stigler (1971) mentioned in passing that a useful characterization theorem for a $C$-restricted $D$-optimal design appears difficult because of the mathematical complexity of the problem. One advantage in our approach is now apparent, because constrained designs are characterized in terms of compound designs.

By simultaneously plotting the two efficiency functions as in Figures 1 and 2, the experimenter can contrast the behavior of the two competing optimal designs. The graphs are useful for understanding the trade-offs. A steep efficiency graph of $\phi_1$, for example, will generally mean that the optimal design for $\phi$ is rather sensitive to small departures from $\xi^1$. On the other hand, a relatively flat efficiency curve indicates that the design could be allowed to stray from the optimum without suffering serious harm to its efficiency. In the extreme case, when $\phi_1$ and $\phi_2$ have the same optimal design, both the efficiency functions will coincide and $\xi^1 = \xi^2 = \xi_\lambda$ for all $\lambda \in [0, 1]$. The examples of Section 5 are based on closed-form expressions for the compound optimal designs. Such expressions are unnecessary in practice, because algorithms are available, and thus graphs analogous to Figures 1 and 2 can be constructed without relying on analytic derivations.

When several functionals $\phi_i(\xi)$, $i = 1, \ldots, m$ are involved, one extension of the constrained design problem is to find the design that maximizes $\phi_m$ subject to the constraints $\phi_i \geq c_i$, $i = 1, \ldots, m - 1$. The corresponding compound design problem is to find $\xi_\lambda = \arg\max \phi(\xi \mid \lambda)$, where $\lambda = (\lambda_i)$ is now a vector of nonnegative weights adding to 1 and $\phi(\xi \mid \lambda) = \Sigma\lambda_i\phi_i(\xi)$. It is not difficult to see that Lemma 1 extends immediately to this case: $\xi_\lambda$ maximizes $\phi_m(\xi)$ subject to the constraints $\phi_i(\xi) \geq \phi_i(\xi_\lambda)$, $i = 1, \ldots, m - 1$. Thus any optimal compound design is again a constrained optimal design for particular constraint bounds. The remaining results in Section 4 do not extend so easily, but Merlise Clyde (personal communication) has established a multiple-constraint version of Lemma 5 by adapting a Lagrangian argument, with the $\lambda$'s in a compound functional being essentially scaled Lagrange multipliers. One difficult part of a multiple-constraint problem seems to be choosing the constraint constants $c_i$ so that the set of candidate designs is nonempty. Another is developing graphical displays that allow trade-offs to be understood, as Figures 1 and 2 do for problems with a single constraint. Work on these problems is underway.

In the meantime, practically useful results for multiple criteria might be obtained by pairwise application of the results in this article. For definiteness, suppose that three criteria are represented in order of importance by the functionals $\delta_1(\xi), \ldots, \delta_3(\xi)$. We continue to use $\phi_1(\xi)$ and $\phi_2(\xi)$ to represent the primary and secondary criteria, at each pairwise comparison. Begin with the two most important criteria by setting $\phi_1 = \delta_1$ and $\phi_2 = \delta_2$, and then develop a compound design, say $\xi^{12}$ with corresponding $\lambda = \lambda_{12}$. Next, set $\phi_1 = \lambda_{12}\delta_1(\xi) + (1 - \lambda_{12})\delta_2(\xi)$ and $\phi_2(\xi) = \delta_3(\xi)$ and select another compromise design, say $\xi^{123}$, that hopefully will be adequate. There is clearly a lot of flexibility in this general idea. Flexibility could be an advantage in some problems, because it allows for an understanding of the trade-offs between criteria of relatively similar importance.

## APPENDIX: JUSTIFICATIONS

*Lemma 1.* Assume that $\xi_\lambda$ does not maximize $\phi_2$ subject to the constraint $\phi_1 \geq c_\lambda$. Then there exists a design $\xi^*$ such that $\phi_2(\xi^*) > \phi_2(\xi_\lambda)$ and $\phi_1(\xi^*) \geq c_\lambda$. Substituting $\xi^*$ into $\phi$,

$$\phi(\xi^* \mid \lambda) = (1 - \lambda)\phi_2(\xi^*) + \lambda\phi_1(\xi^*) \geq (1 - \lambda)\phi_2(\xi^*) + \lambda\phi_1(\xi_\lambda) > (1 - \lambda)\phi_2(\xi_\lambda) + \lambda\phi_1(\xi_\lambda) = \phi(\xi_\lambda \mid \lambda),$$

which contradicts the optimality of $\xi_\lambda$.

*Lemma 2.* $\phi(\xi \mid \lambda)$ is trivially a continuous function of $\lambda$ and, by assumption, is also continuous on $\mathcal{M}$. From this and the compactness of $\mathcal{M}$, it follows from Pshenichny (1971, p. 71) that $\phi(\xi_\lambda \mid \lambda)$ is a continuous function of $\lambda$.

Because $M(\xi_\lambda) \in \mathcal{M}^+$ for $\lambda \in (0, 1)$ and $\phi_1(\xi)$ is continuous on $\mathcal{M}^+$, it follows that $\phi_1(\xi_\lambda)$ is continuous in $\lambda$.

*Lemma 3.* Let $\lambda_2 > \lambda_1$. Then we need to show that $\phi_1(\xi_{\lambda_2}) \geq \phi_1(\xi_{\lambda_1})$. Because $\xi_{\lambda_1}$ maximizes $\phi(\xi \mid \lambda_1)$,

$$(1 - \lambda_1)\phi_2(\xi_{\lambda_1}) + \lambda_1\phi_1(\xi_{\lambda_1}) \geq (1 - \lambda_1)\phi_2(\xi_{\lambda_2}) + \lambda_1\phi_1(\xi_{\lambda_2}),$$

which implies that

$$(1 - \lambda_1)\{\phi_2(\xi_{\lambda_1}) - \phi_2(\xi_{\lambda_2})\} + \lambda_1\{\phi_1(\xi_{\lambda_1}) - \phi_1(\xi_{\lambda_2})\} \geq 0. \quad \text{(A.1)}$$

Similarly,

$$(1 - \lambda_2)\{\phi_2(\xi_{\lambda_2}) - \phi_2(\xi_{\lambda_1})\} + \lambda_2\{\phi_1(\xi_{\lambda_2}) - \phi_1(\xi_{\lambda_1})\} \geq 0. \quad \text{(A.2)}$$

If $\lambda_2 = 1$, the conclusion follows immediately from (A.2). Otherwise, dividing both sides of (A.1) and (A.2) by $(1 - \lambda_1)$ and $(1 - \lambda_2)$ and adding the resulting equations gives the desired conclusion, $\{\phi_1(\xi_{\lambda_2}) - \phi_1(\xi_{\lambda_1})\} \geq 0$.

*Lemma 4.* From (A.1) and Lemma 3, we have

$$(1-\lambda_1)\{\phi_2(\xi_{\lambda_1})-\phi_2(\xi_{\lambda_2})\}+\lambda_1\{\phi_1(\xi_{\lambda_2})-\phi_1(\xi_{\lambda_1})\}\geq 0,$$

which implies the desired conclusion.

*Lemma 5.* We need to show that for each $c \in (-\infty, 0)$, the constrained optimal design $\xi_c$ is also a solution to the compound problem. Let $c^* = \phi_1(\xi_c)$. By definition, $c^* \geq c$ and $\phi_2(\xi_c) \geq \phi_2(\xi)$ for all $\xi$ satisfying $\phi_1(\xi) \geq c$.

Clearly, $c^* \leq \phi_1(\xi_{\lambda=1})$, because $\xi_{\lambda=1}$ maximizes $\phi_1$. Further, if $\phi_1(\xi_{\lambda=0}) \geq c$, then the optimal design for $\phi_2$ alone satisfies the constraint, $\xi_c = \xi_{\lambda=0}$, and thus $c^* = \phi_1(\xi_{\lambda=0})$. And if $\phi_1(\xi_{\lambda=0}) < c$, then $c^* > \phi_1(\xi_{\lambda=0})$. Together, these inequalities imply that $\phi_1(\xi_{\lambda=0}) \leq c^* \leq \phi_1(\xi_{\lambda=1})$.

The conclusion now follows immediately if $c^* = \phi_1(\xi_{\lambda=1})$ or $c^* = \phi_1(\xi_{\lambda=0})$. For example, if $c^* = \phi_1(\xi_{\lambda=0})$, then $\xi_c$ maximizes $\phi_2$, and thus it corresponds to a compound design with $\lambda = 0$.

Assume that $\phi_1(\xi_{\lambda=0}) < c^* < \phi_1(\xi_{\lambda=1})$. By the continuity of $\phi_1(\xi_\lambda)$ and (6), there exists $\lambda^*$ such that $c^* = \phi_1(\xi_{\lambda^*})$, $0 < \lambda^* < 1$. Because

$$(1-\lambda^*)\phi_2(\xi_{\lambda^*})+\lambda^*\phi_1(\xi_{\lambda^*})\geq(1-\lambda^*)\phi_2(\xi_c)+\lambda^*\phi_1(\xi_c)$$

and $\phi_1(\xi_{\lambda^*}) = \phi_1(\xi_c)$, it follows that $\phi_2(\xi_{\lambda^*}) \geq \phi_2(\xi_c)$. But this contradicts the optimality of $\xi_c$ unless equality holds, in which case we have

$$\begin{aligned}\phi(\xi_c\,|\,\lambda=\lambda^*)&=(1-\lambda^*)\phi_2(\xi_c)+\lambda^*\phi_1(\xi_c)\\&=(1-\lambda^*)\phi_2(\xi_{\lambda^*})+\lambda^*\phi_1(\xi_{\lambda^*})\geq\phi(\xi\,|\,\lambda=\lambda^*)\end{aligned}$$

for all $\xi$.

*[Received April 1992. Revised February 1993.]*

## REFERENCES


Atkinson, A. C. (1988), "Recent Developments in the Methods of Optimum and Related Experimental Designs," *International Statistical Review,* 56, 99–116.

Böhning, D. (1981), "On the Construction of Optimal Experimental Designs: A Penalty Approach," *Mathematische Operationsforschung und Statistik, Series Statistics,* 12, 487–495.

Box, G. E. P., and Draper, N. R. (1975), "Robust Design," *Biometrika,* 62, 347–352.

Cook, R. D., and Nachtsheim, C. (1982), "Model Robust, Linear-Optimal Designs," *Technometrics,* 24, 49–55.

Dette, H. (1990), "A Generalization of $D$- and $D_1$-Optimal Designs in Polynomial Regression," *The Annals of Statistics,* 18, 1784–1804.

Fedorov, V. V. (1972), *Theory of Optimal Experiments,* New York: Academic Press.

——— (1980), "Convex Design Theory," *Mathematische Operationsforschung und Statistik, Series Statistics,* 11, 403–413.

Hill, W. J., Hunter, W. G., and Witchern, D. W. (1968), "A Joint Design Criterion for the Dual Problem of Model Discrimination and Parameter Estimation," *Technometrics,* 10, 145–160.

Kiefer, J. (1974), "General Equivalence Theory for Optimal Designs," *The Annals of Statistics,* 2, 849–879.

Läuter, E. (1974), "Experimental Planning in a Class of Models," *Mathematische Operationsforschung und Statistik,* 5, 673–708.

——— (1976), "Optimal Multipurpose Designs for Regression Models," *Mathematische Operationsforschung und Statistik,* 7, 51–68.

Lee, C. M. S. (1987), "Constrained Optimal Designs for Regression Models." *Communications in Statistics, Part A—Theory and Methods,* 16, 765–783.

——— (1988), "Constrained Optimal Design," Journal of Statistical Planning and Inference, 18, 377–389.

Mikulecka, J. (1983), "On Hybrid Experimental Design," *Kybernetika,* 19, 1–14.

Pazman, A. (1986), *Foundations of Optimum Experimental Design,* Boston: D. Reidel.

Pshenichny, B. N. (1971), "Necessary Conditions for an Extremum," New York: Marcel Dekker.

Pukelsheim, F. (1987), "Information Increasing Orderings in Experimental Design Theory," *International Statistical Review,* 56, 203–219.

Silvey, S. D. (1980), *Optimal Design,* London: Chapman and Hall.

Stigler, S. M. (1971), "Optimal Experimental Design for Polynomial Regression," *Journal of the American Statistical Association,* 66, 311–318.

Studden, W. J. (1982), "Some Robust-Type $D$-Optimal Designs in Polynomial Regression," *Journal of the American Statistical Association,* 77, 916–921.