

Research Article

Arun Ravichandran, Nicole E. Pashley, Brian Libgober, and Tirthankar Dasgupta*

Optimal allocation of sample size for randomization-based inference from 2^K factorial designs

<https://doi.org/10.1515/jci-2023-0046>

received July 03, 2023; accepted October 17, 2023

Abstract: Optimizing the allocation of units into treatment groups can help researchers improve the precision of causal estimators and decrease costs when running factorial experiments. However, existing optimal allocation results typically assume a super-population model and that the outcome data come from a known family of distributions. Instead, we focus on randomization-based causal inference for the finite-population setting, which does not require model specifications for the data or sampling assumptions. We propose exact theoretical solutions for optimal allocation in 2^K factorial experiments under complete randomization with A-, D-, and E-optimality criteria. We then extend this work to factorial designs with block randomization. We also derive results for optimal allocations when using cost-based constraints. To connect our theory to practice, we provide convenient integer-constrained programming solutions using a greedy optimization approach to find integer optimal allocation solutions for both complete and block randomizations. The proposed methods are demonstrated using two real-life factorial experiments conducted by social scientists.

Keywords: randomization, factorial design, optimum allocation, Neyman

MSC 2020: 62K05, 62K10, 62K15, 62D99

1 Introduction

Randomized 2^K factorial experiments are conducted to assess the marginal causal effects of K factors, each with two levels, along with their interactions on a response of interest. The two levels are often denoted as the “high level” and “low level” of the factor [1,2]. With K factors, there are 2^K unique treatment combinations to which units can be assigned. In the twentieth century, factorial designs have mostly been discussed in an industrial setting, whereas in recent times, there has been a lot of interest in their application in the social, behavioral and biomedical sciences and randomization-based inference from such designs (e.g., [3,4]).

Randomization-based inference is a useful methodology for drawing inference on causal effects of treatments in a finite population setting (e.g., [5,6]). A major advantage of randomization-based inference is that it applies even if the experimental units are not randomly sampled from a larger population, which is the case in most social science experiments [7,8]. The theory, methods, and applications of randomization-based inference for two-level factorial experiments with a completely randomized treatment assignment mechanism have been developed and discussed (e.g., [9,10]). Further, randomization-based inference from experiments with more general factorial structures and complex assignment mechanisms have been discussed in the study by

* **Corresponding author: Tirthankar Dasgupta**, Department of Statistics, Rutgers University, Piscataway, NJ, USA, e-mail: tirthankar.dasgupta@rutgers.edu

Arun Ravichandran, Nicole E. Pashley: Department of Statistics, Rutgers University, Piscataway, NJ, USA

Brian Libgober: Department of Political Science and Law, Northwestern University, Chicago, IL, USA

Mukerjee et al. [11]. Connections between regression- and randomization-based causal inference from factorial experiments have been studied by Zhao and Ding [12].

Despite the growing literature in this area, most of the recent research on randomization-based inference of factorial experiments has been confined to the analysis side. On the design side, the main focus has been on rerandomization [3,13,14], which generalizes the idea of blocking by predefining an acceptable criterion for randomization based on covariate balance between treatment groups. There have also been extensions to fractional and incomplete factorial designs [15] and to the use of screening steps [16]. However, the distribution of the total number of experimental units into the 2^K treatment groups has not received much attention. Balanced designs that assign an equal number of units to the treatment groups are often the default choice, but it is unclear whether they are the “best” design under different conditions. One work that does discuss how to allocate units to optimize precision of factorial estimators from the randomization-based perspective is the one by Blackwell et al. [17]. That work explores the advantages of Neyman allocation [18,19] by extending the two-stage adaptive design proposed in the study by Hahn et al. [20] to multiple treatment designs. Dai et al. [21] similarly explores Neyman allocation but within sequential designs.

To motivate this problem, we consider an education experiment from [22], conducted to assess the causal effects of two different interventions, a student support program (SSP) and a student fellowship program (SFP), on the academic performance of freshmen. This is a 2^2 factorial experiment in which each unit (freshman) can receive only one of four treatment combinations: control (neither of the two), SSP only, SFP only, and student fellowship and support program (SFSP) that indicated both interventions applied to a unit. The units were divided into two blocks based on their sex. Table 1 shows the allocation of units within each block to the four treatment combinations.

Clearly, the design is unbalanced, with the highest number of units assigned to control and the fewest to the treatment SFSP. Such an allocation, among other reasons, could be motivated by the budget for experimentation. The question we investigate in this article is the following: Under what assumptions, conditions, and requirements will such an allocation be the best possible one (in terms of being able to precisely answer scientific questions of interest)?

The problem of finding optimal designs in the context of model-based inference has been extensively studied in the twentieth century (see, e.g., [23]). In such settings, optimal designs depend on a postulated outcome model that may be linear or nonlinear. For example, for binary responses, optimal designs based on logistic models are likely to differ from those based on probit models and depend on unknown model parameters [24–26]. We aim to develop optimal designs that are tied to *model-free, randomization-based analysis* for finite and super populations. In addition to being robust to model assumptions, our approach works for continuous as well as binary outcomes as long as the finite-population estimand is well-defined.

This article is organized as follows: The next section introduces basic notation and estimands for factorial experiments using the potential outcomes framework. In Section 3, we derive optimal allocations of the N units in a population to different treatment combinations under three commonly used optimality criteria for a completely randomized design (CRD). In addition to theoretical results for exact optimal allocations, we also provide numerical algorithms for obtaining integer solutions. In Section 4, we extend our results for CRDs to the setting of randomized block designs (RBDs). In Section 5, we derive optimality results under cost constraints. Two different applications of the proposed methodology, the motivating education experiment and an audit experiment conducted to assess discrimination, are described in Section 6. We conclude with a discussion, including opportunities for future work, in Section 7.

Table 1: Allocation of units to treatment combinations

Sex	Control	SFP	SSP	SFSP
Female	574	150	142	82
Male	432	100	108	68
Total	1,006	250	250	150

2 2^K factorial experiments under the potential outcomes framework

Here, we introduce some key definitions and notations from the study by Dasgupta et al. [9]. Consider a 2^K experiment with N units, in which the levels of each of the K factors are denoted by 0 and 1. Each treatment combination is of the form $\mathbf{z}_j = (z_1, \dots, z_K)$, where $z_k \in \{0, 1\}$ for $k \in 1, \dots, K$. There are $J = 2^K$ treatment combinations arranged in lexicographic order $1, \dots, J$, where treatment combination \mathbf{z}_j is such that $j = 2^{K-1}z_1 + 2^{K-2}z_2 + \dots + z_K + 1$. In other words, (z_1, \dots, z_K) is a binary representation of integer $j - 1$. Thus, for example, in a 2^2 experiment, the four treatment combinations 00, 01, 10, and 11 are numbered as $j = 1, 2, 3$, and 4, respectively, and the 8th treatment combination in a 2^4 experiment is 0111. We will just refer to the treatment combination by its number (j) in notation below.

For $i = 1, \dots, N$, under the stable unit treatment value assumption (SUTVA) [27], the i th unit has $J = 2^K$ potential outcomes $Y_i(1), \dots, Y_i(J)$ corresponding to the J treatment combinations $\mathbf{z}_1, \dots, \mathbf{z}_J$. Let \mathbf{Y}_i denote the $J \times 1$ vector of potential outcomes for unit i . For unit i , the unit-level main effect of factor $k = 1, \dots, K$ is defined as the difference between the averages of potential outcomes for unit i for which the levels of factor k are at levels 0 vs 1. Mathematically, it is a contrast of the form $2^{-(K-1)}\lambda_k^T \mathbf{Y}_i = 2^{-(K-1)}\sum_{j=1}^J \lambda_{jk} Y_i(j)$, where \mathbf{x}^T denotes the transpose of vector \mathbf{x} , λ_k is a $J \times 1$ column-vector with coefficient λ_{jk} such that $\lambda_{jk} = -1$ if the level of factor k in j th treatment combination is 0, and $\lambda_{jk} = 1$ otherwise. For all $k = 1, \dots, K$, λ_k is a contrast vector, i.e., $\sum_{j=1}^J \lambda_{jk} = 0$.

Proceeding along the lines of the study by Dasgupta et al. [9], for unit i , we can define $\binom{K}{2}$ two-factor interactions, $\binom{K}{3}$ three-factor interactions, and finally one K -factor interaction as contrasts of the form $2^{-(K-1)}\lambda^T \mathbf{Y}_i$, where the contrast vector λ for any interaction can be derived by element-wise multiplication of the contrast vectors of the corresponding main effects λ_k , for factors involved in the interaction. Denoting the $J - 1 = 2^K - 1$ contrast vectors for the $J - 1$ unit-level factorial effects $\tau_{1i}, \dots, \tau_{J-1,i}$ by $\lambda_1, \dots, \lambda_{J-1}$, respectively, we define the $J \times J$ matrix as follows:

$$\mathbf{L} = (\lambda_0, \lambda_1, \dots, \lambda_{J-1}), \quad (1)$$

where λ_0 is the $J \times 1$ vector with all elements equal to one. We note that \mathbf{L} is an orthogonal matrix with $\mathbf{L}\mathbf{L}^T = \mathbf{L}^T\mathbf{L} = 2^{K-1}\mathbf{I}_J$, where \mathbf{I}_J denotes the identity matrix of order J . For $i = 1, \dots, N$, let $\boldsymbol{\tau}_i = (2\tau_{0i}, \tau_{1i}, \dots, \tau_{J-1,i})^T$, where τ_{0i} denotes the average of all potential outcomes for unit i . The linear transform between the vector of unit-level potential outcomes \mathbf{Y}_i and the vector of unit-level factorial effects $\boldsymbol{\tau}_i$ can be expressed as follows:

$$\boldsymbol{\tau}_i = 2^{-(K-1)}\mathbf{L}^T\mathbf{Y}_i. \quad (2)$$

Having defined unit-level factorial effects, we now move to their population-level counterparts. Let $\bar{\mathbf{Y}} = N^{-1}\sum_{i=1}^N \mathbf{Y}_i$ and $\boldsymbol{\tau} = N^{-1}\sum_{i=1}^N \boldsymbol{\tau}_i$, respectively, denote the $J \times 1$ vectors of average potential outcomes and the average factorial effects. Then, averaging equation (2) over $i = 1, \dots, N$, the vector of population-level factorial effects is given as follows:

$$\boldsymbol{\tau} = 2^{-(K-1)}\mathbf{L}^T\bar{\mathbf{Y}}. \quad (3)$$

Note that the first element of $\boldsymbol{\tau}$ is twice the average of all potential outcomes.

In a CRD, a pre-assigned number of units, N_j , are randomly assigned to treatment j . The experiment generates an $N \times 1$ vector of observed outcomes data from which the vector of factorial effects $\boldsymbol{\tau}$ can be unbiasedly estimated. We examine the properties of an unbiased estimator of $\boldsymbol{\tau}$ defined in the next section, with respect to its randomization distribution, and formulate the problem of optimally allocating the N units to the J treatment combinations in Section 3.

Remark 1. Incorporating the “intercept” (overall average) is a traditional practice in the analysis of factorial experiments (e.g., [28]). In addition to creating the matrix \mathbf{L} that is analogous to the “model matrix” in linear models, it has certain advantages in the context of randomization-based inference. Specifically, in the current

context, it leads to a factorization of the covariance matrix as in equation (5) of the article, which in turn leads to Lemma 1 of the Supplementary Material, providing the foundation for proving Theorem 1.

Remark 2. (A super population perspective) While the finite-population perspective does not depend on any hypothetical data generating process for the outcomes, alternative approaches assume that the potential outcomes are drawn from a, possibly hypothetical, super population. Assuming that $\mathbf{Y}_1, \dots, \mathbf{Y}_N$ are independent and identically distributed random vectors with $E[\mathbf{Y}_i] = \boldsymbol{\mu}$, factorial effects at a super population level are defined as follows:

$$\boldsymbol{\tau}^{\text{SP}} = 2^{-(K-1)} \mathbf{L}^T \boldsymbol{\mu}.$$

Ding et al. [29] discussed the conceptual and mathematical connections between finite- and super-population inference, showing that while the same estimator commonly used to estimate $\boldsymbol{\tau}$ unbiasedly is also an unbiased estimator of $\boldsymbol{\tau}^{\text{SP}}$, its sampling variances under the two perspectives are different.

3 Optimal designs for completely randomized experiments

In a randomized experiment with N_j units assigned to treatment combination $j \in \{1, \dots, J\}$, only one of the J potential outcomes is observed for unit i . This observed outcome is $y_i = Y_i(T_i)$ for $i = 1, \dots, N$, where T_i is the random treatment assignment variable for unit i taking value j if unit i receives treatment j . In a CRD, the joint probability distribution of (T_1, \dots, T_N) is

$$P[(T_1, \dots, T_N) = (t_1, \dots, t_N)] = \begin{cases} (N_1! \dots N_J!)/N!, & \text{if } \sum_{i=1}^N \mathbb{1}_{\{t_i=j\}} = N_j, \text{ for } j = 1, \dots, J, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathbb{1}_{\{A\}}$ denotes the indicator random variable for set A . Let $\bar{y}(j) = N_j^{-1} \sum_{i=1}^N \mathbb{1}_{\{t_i=j\}} Y_i(j)$ denote the average response for treatment j . Let $\bar{\mathbf{y}}$ denote the vector $(\bar{y}(1), \bar{y}(2), \dots, \bar{y}(J))^T$ of observed averages. Substituting $\bar{\mathbf{y}}$ in place of $\bar{\mathbf{Y}}$ in equation (3), we can unbiasedly estimate the vector of factorial effects as follows:

$$\hat{\boldsymbol{\tau}} = 2^{-(K-1)} \mathbf{L}^T \bar{\mathbf{y}}. \quad (4)$$

Lu [10] derived sampling properties of the estimator $\hat{\boldsymbol{\tau}}$ with respect to its randomization distribution for the general case of unequal N_1, \dots, N_J . Lu [10] showed that $\hat{\boldsymbol{\tau}}$ is an unbiased estimator of $\boldsymbol{\tau}$, and has the following finite-population covariance matrix:

$$\mathbf{V}_{\boldsymbol{\tau}} = \text{Var}(\hat{\boldsymbol{\tau}}) = \frac{1}{2^{2(K-1)}} \sum_{j=1}^{2^K} \frac{S_j^2}{N_j} \tilde{\boldsymbol{\lambda}}_j \tilde{\boldsymbol{\lambda}}_j^T - \frac{1}{N(N-1)} \sum_{i=1}^N (\boldsymbol{\tau}_i - \boldsymbol{\tau})(\boldsymbol{\tau}_i - \boldsymbol{\tau})^T, \quad (5)$$

where $\tilde{\boldsymbol{\lambda}}_j$ represents the transpose of row j of the model matrix \mathbf{L} defined in equation (1), $\boldsymbol{\tau}_i$ denotes the vector of unit-level factorial effects given by equation (2), $\boldsymbol{\tau}$ denotes the vector of population-level factorial effects given by equation (3), and

$$S_j^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i(j) - \bar{Y}(j))^2 \quad (6)$$

denotes the variance of all N potential outcomes for treatment j with divisor $N-1$, where $\bar{Y}(j) = N^{-1} \sum_{i=1}^N Y_i(j)$.

In the spirit of classical optimal designs [23], we can define a design optimality criterion as a functional of the matrix $\mathbf{V}_{\boldsymbol{\tau}}$ defined in equation (5). For example, the D-optimality criterion, which aims to minimize the determinant of the covariance matrix, or the A-optimality criterion, which aims to minimize the trace of the covariance matrix, or the E-optimality criterion, which aims to minimize the maximum eigenvalue of the covariance matrix, can be considered. However, the second term $1/(N(N-1)) \sum_{i=1}^N (\boldsymbol{\tau}_i - \boldsymbol{\tau})(\boldsymbol{\tau}_i - \boldsymbol{\tau})^T$, which is a

measure of heterogeneity of treatment effects, cannot be estimated from observed data, because none of the unit-level treatment effects τ_i are estimable without additional assumptions due to the missing potential outcomes. Because $1/(N(N-1))\sum_{i=1}^N(\tau_i - \tau)(\tau_i - \tau)^T$ is positive semi-definite, the first term of equation (5) can be considered as an upper bound of V_τ , which is attained under specific restrictions on the potential outcomes (e.g., treatment effect homogeneity). Thus, we propose optimizing a functional of the first term of equation (5), which in turn is equivalent to optimizing a functional of the positive definite matrix

$$\tilde{\mathbf{V}} = \sum_{j=1}^J \frac{S_j^2}{N_j} \tilde{\lambda}_j \tilde{\lambda}_j^T = \mathbf{L}^T \mathbf{A} \mathbf{L},$$

instead, where $\mathbf{A} = \text{diag}(S_1^2/N_1, \dots, S_J^2/N_J)$.

Another justification for using a functional of the matrix $\tilde{\mathbf{V}}$ as an optimality criterion comes from the super-population perspective mentioned in Section 2. Ding et al. [29] showed that the estimator $\hat{\tau}$ defined earlier is also an unbiased estimator of the super-population estimand τ^{SP} . Furthermore, extending their argument for a single factor with two levels to the case of 2^K factorial designs, if $V_j^2 = \text{Var}[Y_i(j)]$, $j = 1, \dots, J$, then a variance decomposition yields the following sampling variance of $\hat{\tau}$,

$$\text{Var}^{\text{SP}}(\hat{\tau}) = \frac{1}{2^{2(K-1)}} \sum_{j=1}^{2^K} \frac{V_j^2}{N_j} \tilde{\lambda}_j \tilde{\lambda}_j^T = \frac{1}{2^{2(K-1)}} \sum_{j=1}^{2^K} \frac{E(S_j^2)}{N_j} \tilde{\lambda}_j \tilde{\lambda}_j^T,$$

where Var^{SP} denotes variance over the random sampling from the super population and random assignment, and $V_j^2 = E(S_j^2)$ represents expectation of S_j^2 with respect to the distribution of the potential outcomes in the super population. This connection provides further motivation for the form of our optimization, but we focus on the finite-population setting going forward.

3.1 Exact optimal designs

The problem of finding an optimal design can be formulated as minimization of an appropriate functional $\psi(\tilde{\mathbf{V}})$ subject to the constraint $\sum_{j=1}^J N_j = N$ or equivalently as $\sum_{j=1}^J p_j = 1$ in terms of the proportions of units $p_j = N_j/N$ to be assigned to treatment combination j . As discussed earlier, we consider three widely used functionals in optimal design literature: the D-optimality criterion, where $\psi(\tilde{\mathbf{V}}) = |\tilde{\mathbf{V}}|$ and $|\cdot|$ refers to the determinant; the A-optimality criterion, where $\psi(\tilde{\mathbf{V}}) = \text{tr}(\tilde{\mathbf{V}})$; and the E-optimality criterion, where $\psi(\tilde{\mathbf{V}}) = \max\{\nu_1, \dots, \nu_J\}$ and ν_1, \dots, ν_J are the eigenvalues of $\tilde{\mathbf{V}}$. Theorem 1 proved in Supplementary Material S1, summarizes these optimality results.

Theorem 1. *Let N units be allocated to J treatment groups such that $p_j = N_j/N$ proportion of units receive treatment j . Then, the optimal allocation of N units to J treatment groups on the basis of covariance matrix $\tilde{\mathbf{V}}$ under*

- (a) *A-optimality is proportional to the finite-population standard deviations of potential outcomes in the corresponding treatment groups, i.e., $p_j = S_j/(\sum_j S_j)$.*
- (b) *D-optimality is a balanced assignment to all J treatment groups, i.e., $p_j = 1/J$.*
- (c) *E-optimality is proportional to the finite-population variances of potential outcomes in the corresponding treatment groups, i.e., $p_j = S_j^2/(\sum_j S_j^2)$.*

Remark 3. The optimality results in Theorem 1 are similar in spirit to the determination of optimal sample sizes in stratified survey sampling [19, Ch. 5]. The A-optimality result is the same as the Neyman allocation discussed in [17], where they motivate its use as reducing the identifiable portion of the finite-population variance of classical factorial estimators. Derivations of the A- and D-optimal designs are straightforward and

use the Lagrangian multiplier-based optimization technique. The proof of the E-optimality result uses the idea of perturbing the eigenvalues of a scaled identity matrix to show that the E-optimal design is indeed characterized by equal eigenvalues of the matrix $\tilde{\mathbf{V}}$.

Remark 4. In order to implement the results of Theorem 1, researchers need to “guess” the variances S_j^2 , $j = 1, \dots, J$, and substitute them into the expressions for p_j . This is similar to the application of optimal designs in nonlinear models, where optimal designs are actually “locally optimal” [30]. It is often a common practice to conduct pilot studies to obtain some preliminary estimates of the S_j^2 s, as done in finite-population survey sampling.

We now introduce two conditions associated with the matrix of potential outcomes under which Theorem 1 can be further simplified.

Condition 1. (Homoscedasticity) We call an $N \times J$ matrix of potential outcomes *homoscedastic* if each column has the same variance, i.e., $S_j^2 = S^2$ for $j = 1, \dots, J$.

Condition 2. (Strict additivity) Following the study by Dasgupta et al. [9], we call an $N \times J$ matrix of potential outcomes *strictly additive* if $Y_i(j) - Y_i(\tilde{j}) = \tau(j, \tilde{j})$ for all $j \neq \tilde{j} \in \{1, 2, \dots, J\}$. Potential outcomes satisfying this condition also satisfy Condition 1.

Corollary 1 of Theorem 1 is straightforward but useful.

Corollary 1. If the matrix of potential outcomes satisfies Condition 1, then the A-, D-, and E-optimal designs are all balanced designs with $p_j = 1/J$ for $j = 1, \dots, J$. Furthermore, under Condition 2, optimizations based on $\tilde{\mathbf{V}}$ and \mathbf{V}_τ are equivalent.

While Theorem 1 provides results on exact optimal designs in terms of proportions p_j , experimenters need integer solutions in terms of N_j s satisfying $\sum_j N_j = N$. For example, while the exact D-optimal design is balanced, N is not necessarily a multiple of 2^K , and the result does not provide a D-optimal allocation of, say, 69 units to the eight treatment combinations in a 2^3 factorial experiment. Thus, we need approximate integer solutions to the optimization problem in which additional constraints on the sample sizes assigned to specific treatment groups can also be introduced. Next, we discuss an integer programming approach to obtain such solutions.

1. Set $I \leftarrow \{1, \dots, J\}$.
2. Let $t = 0$. Initialize $N_j^{(t)} \leftarrow l_j$ for $j = 1, \dots, J$.
3. While $(\sum_{j=1}^J N_j^{(t)} \neq N \ \& \ I \neq \emptyset)$ do
 - for all $(j \in I)$, $\delta_j \leftarrow f_j(N_j^{(t)} + 1) - f_j(N_j^{(t)})$
 - choose $\mathcal{J} \leftarrow \operatorname{argmin}_{j \in I} \delta_j$
 - $j^* \leftarrow \min \mathcal{J}$
 - if $N_{j^*}^{(t)} + 1 \leq u_{j^*}$, then $N_{j^*}^{(t+1)} \leftarrow N_{j^*}^{(t)} + 1$, $t \leftarrow t + 1$
 else $I \leftarrow I \setminus \{j^*\}$
4. Return optimal (N_1, \dots, N_J) .

Figure 1: Greedy algorithm for separable functions.

3.2 Computation of integer optimal designs using an integer programming approach

Many sources in the integer programming literature address constrained optimization under integer space constraints (see, e.g., [31–34]). We adopt the methods proposed in the study by Friedrich et al. [35] that are designed for settings very similar to the ones we consider. Friedrich et al. [35] considered the following integer programming problem:

$$\min_{N_1, \dots, N_J} f(N_1, \dots, N_J) \text{ s.t. } \begin{cases} \sum_{j=1}^J N_j = N, \\ l_j \leq N_j \leq u_j \quad \forall j = 1, 2, \dots, J, \\ N_j \in \mathbb{Z}_+ \quad \forall j = 1, 2, \dots, J, \end{cases} \quad (7)$$

where \mathbb{Z}_+ is the set of positive integers, (l_j, u_j) are the lower and upper bound constraints on N_j , and $f: \mathbb{R}_+^J \rightarrow \mathbb{R}$ is a convex function. If $f(N_1, \dots, N_J)$ is separable (i.e., can be expressed as $\sum_{j=1}^J f_j(N_j)$) then the greedy algorithm in Figure 1 finds the globally optimal integer solution of the minimization problem given in equation (7) under some regularity conditions that we show in Supplementary Material S2.

From the proof of Theorem 1 in Supplementary Material S1, it follows that the A- and D-optimality criteria can, respectively, be expressed as $\sum_{j=1}^J (S_j^2/N_j)$ and $\sum_{j=1}^J (\log S_j^2/N_j)$. In Supplementary Material S2, we show that the conditions for global convergence of the greedy algorithm are met, and thus, convergence to the true integer optimal solution in the cases of A- and D-optimality are guaranteed if this algorithm is used. Hence, substitution of S_j^2/N_j and $\log S_j^2/N_j$ for $f_j(N_j)$ in the algorithm described in Figure 1 leads to optimal integer solutions for the A-optimality criterion and the D-optimality criterion, respectively. In our implementation of this algorithm, we take $l_j = 2$ and $u_j = N - 2J$ for $j = 1, \dots, J$ to guarantee at least two units are assigned to each treatment combination, allowing variance estimation within each treatment group.

We provide a modified greedy algorithm described in Figure 2 for solving the E-optimality problem since the optimization problem cannot be written in a separable form as in the aforementioned A- or D-optimality criterion. In this algorithm, we take $f_j(N_j) = S_j^2/N_j$ while keeping $l_j = 2$ and $u_j = N - 2J$ for $j = 1, \dots, J$. In Section 4.2, the ability of this greedy algorithm to find E-optimal solutions is demonstrated empirically for blocked designs, discussed next.

4 Optimal allocation for factorial experiments with blocks

Consider a block-randomized 2^K factorial design with H blocks. That is, units are pre-assigned membership to one of h blocks based on some similarity metric (we do not consider how to form blocks here). Let M_h denote

1. Set $I \leftarrow \{1, \dots, J\}$.
2. Let $t = 0$. Initialize $N_j^{(t)} \leftarrow l_j$ for $j = 1, \dots, J$.
3. While $(\sum_{j=1}^J N_j^{(t)} \neq N \ \& \ I \neq \emptyset)$ do
 - choose $\mathcal{J} \leftarrow \arg\max_{j \in I} f_j(N_j^{(t)})$
 - $j^* \leftarrow \min \mathcal{J}$
 - if $N_{j^*}^{(t)} + 1 \leq u_{j^*}$, then $N_{j^*}^{(t+1)} \leftarrow N_{j^*}^{(t)} + 1$, $t \leftarrow t + 1$
else $I \leftarrow I \setminus \{j^*\}$
4. Return optimal (N_1, \dots, N_J) .

Figure 2: Greedy algorithm for E-optimality.

the size of block h , $h = 1, \dots, H$. In addition, let $M_{h,j}$ be the number of units in block h assigned to the treatment j ($j = 1, \dots, J$). Finally, let $b_i(h)$, $i = 1, \dots, N$, be an indicator variable, taking the value 1 if unit i belongs to block h and 0 otherwise. Treatment assignment under a factorial RBD is equivalent to performing an independent CRD, as described in Section 3, within each block.

The population average treatment effect τ can be expressed as $\sum_h M_h \tau_h / N$, where τ_h is the block-level vector of factorial effects and its estimator $\hat{\tau}$ is a weighted average of $\hat{\tau}_h$, an unbiased estimator of τ_h defined in the same way as in equation (4) for block h . Extending equation (5) by noting the independence of the assignment to treatment across blocks, the covariance matrix of $\hat{\tau}_h$ in a factorial RBD can thus be obtained as follows:

$$\mathbf{V}_{\tau_h} = \text{Cov}(\hat{\tau}_h) = \frac{1}{2^{2(K-1)}} \sum_{j=1}^{2^K} \frac{1}{M_{h,j}} \tilde{\lambda}_j \tilde{\lambda}_j^T S_{h,j}^2 - \frac{1}{M_h(M_h - 1)} \sum_{i=1}^N b_i(h) (\tau_i - \tau_h)(\tau_i - \tau_h)^T, \quad (8)$$

where $\tilde{\lambda}_j$ represents the transpose of row j of the model matrix \mathbf{L} defined in equation (1) and $S_{h,j}^2$ denotes the variance of all M_h potential outcomes for units in block h under treatment j with divisor $M_h - 1$. The covariance matrix of $\hat{\tau}$ can then be expressed as $\text{Cov}(\sum_h M_h \hat{\tau}_h / N)$. Because the block-level treatment estimators $\hat{\tau}_h$ are independent across blocks, we have

$$\begin{aligned} \text{Cov}(\hat{\tau}) &= \sum_{h=1}^H \frac{M_h^2}{N^2} \mathbf{V}_{\tau_h} \\ &= \frac{1}{2^{2(K-1)}} \sum_{j=1}^{2^K} \tilde{\lambda}_j \tilde{\lambda}_j^T \left[\sum_{h=1}^H \frac{M_h^2}{N^2} \frac{S_{h,j}^2}{M_{h,j}} \right] - \sum_{h=1}^H \frac{M_h^2}{N^2} \frac{1}{M_h(M_h - 1)} \sum_{i=1}^N b_i(h) (\tau_i - \tau_h)(\tau_i - \tau_h)^T. \end{aligned}$$

Writing $S_{\text{blk},j}^2 = \sum_{h=1}^H (M_h^2 / N^2) (S_{h,j}^2 / M_{h,j})$ and proceeding along similar lines as in Section 3, we formulate a surrogate optimization problem to only optimize the first term, since the second term in the equation above is not identifiable. Then, choosing $M_{h,j}$ s to optimize some functional of $\text{Cov}(\hat{\tau})$ is equivalent to optimizing a functional of the matrix

$$\tilde{\mathbf{V}}_{\text{blk}} = \sum_{j=1}^{2^K} \tilde{\lambda}_j \tilde{\lambda}_j^T S_{\text{blk},j}^2 = \mathbf{L}^T \mathbf{A}_{\text{blk}} \mathbf{L}, \quad (9)$$

where $\mathbf{A}_{\text{blk}} = \text{diag}(S_{\text{blk},1}^2, \dots, S_{\text{blk},J}^2)$.

4.1 Exact optimal designs

The problem of finding an optimal design for RBDs can be formulated as minimization of an appropriate functional $\psi(\tilde{\mathbf{V}}_{\text{blk}})$ subject to the constraint $\sum_{j=1}^J M_{h,j} = M_h$ for each h or equivalently as $\sum_{j=1}^J p_{h,j} = 1$ in terms of the proportions of units to be assigned to treatment combination j in block h , $p_{h,j} = M_{h,j} / M_h$. While the A-optimality result is straightforward, finding exact D-optimal and E-optimal solutions in the setting with blocks is difficult without imposing restrictions on the potential outcomes. Before stating the optimality results, we first introduce two such restrictions that generalize Condition 1 to a block setting.

Condition 3. (Within-block homoscedasticity [WBH]) We call an $N \times J$ matrix of potential outcomes in H blocks to have *WBH* if within each block, all treatment columns have the same variance, i.e., within block h , $S_{h,j}^2 = S_h^2$ for $j = 1, \dots, J$.

Condition 4. (Between-block homoscedasticity [BBH]) We call an $N \times J$ matrix of potential outcomes in H blocks to have *BBH* if for each treatment column j , the variance of potential outcomes in each block is the same, i.e., $S_{h,j}^2 = S_{\cdot,j}^2$ for $h = 1, \dots, H$ and $j = 1, \dots, J$.

The following theorem now summarizes the optimality results for blocked designs.

Theorem 2. Let N units be distributed across H blocks such that there are M_h units in block h and $N = \sum_{h=1}^H M_h$. Let each set of M_h units be allocated to J treatment groups such that $p_{h,j}$ proportion of the units are allocated to treatment j in block h , and let $S_{h,j}^2$ be as defined in equation (8). Then, optimal allocation of M_h units to the J treatment groups on the basis of covariance matrix \tilde{V}_{blk} under different optimality criteria can be summarized as follows:

- The A-optimal allocation is the same as the A-optimal CRD allocation within each block, i.e., $p_{h,j} = S_{h,j} / (\sum_{j=1}^J S_{h,j})$ for each h .
- If either (or both) Condition 3 (WBH) or 4 (BBH) holds, the D-optimal allocation is the balanced assignment within each block, i.e., $p_{h,j} = 1/J$ for each h .
- If Condition 3 (WBH) holds, the E-optimal allocation is the balanced design within each block, i.e., $p_{h,j} = 1/J$ for each h .

Remark 5. Condition 2 for all N units implies both WBH and BBH. Consequently, by Theorem 2, the D- and E-optimal allocations for strictly additive potential outcomes in RBDs are balanced assignments within each block.

Remark 6. The A-optimal allocation under WBH is a balanced allocation within each block (same as D- and E-optimal allocations). However, under BBH, the A-optimal allocation is different from the D- and E-optimal allocations, and is proportional to the standard deviations of the treatments $S_{\cdot,j}$ that is constant across blocks.

4.2 Computation of integer optimal designs for factorial RBDs using an integer programming approach

As in the case of completely randomized factorial designs, whereas Theorem 2 provides results on exact optimal designs in terms of proportions $p_{h,j}$, experimenters need integer solutions in terms of $M_{h,j}$ s in which additional constraints on the sample sizes assigned to specific treatment groups can also be introduced.

Table 2: Summary of greedy algorithm solutions for E-optimality for $H = 2, K = 2$

S.No.	Case	Block size (M_h)	Variances ($S_{h,j}^2$)	Exhaustive search E-optimal solution	Greedy solution
1.	Equal blocks with equal variances	$\begin{bmatrix} 40 \\ 40 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \end{bmatrix}$	$\begin{bmatrix} 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \end{bmatrix}$
2.	Equal blocks with equal variances for all treatments within block	$\begin{bmatrix} 40 \\ 40 \end{bmatrix}$	$\begin{bmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \end{bmatrix}$	$\begin{bmatrix} 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \end{bmatrix}$
3.	Unequal blocks with equal variances across blocks for each treatment	$\begin{bmatrix} 40 \\ 20 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$	$\begin{bmatrix} 4 & 8 & 12 & 16 \\ 2 & 4 & 6 & 8 \end{bmatrix}$	$\begin{bmatrix} 4 & 8 & 12 & 16 \\ 2 & 4 & 6 & 8 \end{bmatrix}$
4.	Unequal blocks with equal variances but exact solution is noninteger	$\begin{bmatrix} 40 \\ 20 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & 5 \end{bmatrix}$	$\begin{bmatrix} 4 & 8 & 11 & 17 \\ 2 & 3 & 5 & 10 \\ 4 & 7 & 11 & 18 \\ 2 & 4 & 5 & 9 \\ 3 & 8 & 11 & 18 \\ 3 & 3 & 5 & 9 \\ 3 & 7 & 11 & 19 \\ 3 & 4 & 5 & 8 \end{bmatrix}$	$\begin{bmatrix} 4 & 7 & 11 & 18 \\ 2 & 4 & 5 & 9 \end{bmatrix}$
5.	Equal blocks with unequal variances	$\begin{bmatrix} 40 \\ 40 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{bmatrix}$	$\begin{bmatrix} 6 & 10 & 11 & 13 \\ 13 & 11 & 10 & 6 \\ 6 & 9 & 12 & 13 \\ 13 & 12 & 9 & 6 \end{bmatrix}$	$\begin{bmatrix} 6 & 9 & 12 & 13 \\ 13 & 12 & 9 & 6 \end{bmatrix}$

Furthermore, Theorem 2 provides D- and E-optimal solutions only under specific conditions like WBH and BBH. Thus, we discuss an integer programming approach to obtain integer solutions for settings covered and not covered by Theorem 2.

We can use the same algorithm in Figure 1 within each block to obtain the optimal integer solutions for A-optimality under the RBD by replacing the function $f_j(\cdot)$ by $f_{h,j}(M_{h,j}) = (M_h^2/N^2)(S_{h,j}^2/M_{h,j})$. For D- and E-optimality, we extend the greedy idea from the algorithm in Figure 2 with minor changes to the function $f_j(\cdot)$. We take $f_{h,j}(M_{h,j}) = \log(\sum_{h=1}^H (M_h^2/N^2)(S_{h,j}^2/M_{h,j}))$ for D-optimality and $f_{h,j}(M_{h,j}) = (M_h^2/N^2)(S_{h,j}^2/M_{h,j})$ for E-optimality. The exact algorithms taking the structure of the blocks into account are given in Supplementary Material S4. The main difference between the algorithm used in Section 3 and the one proposed here lies in the fact that now we have to allocate the next best unit at iteration t over an $H \times J$ matrix $((M_{h,j}^{(t)}))$ with upper bounds on row sums, rather than a vector $(N_1^{(t)}, \dots, N_J^{(t)})$ with an upper bound on the sum of the elements, as in the case of the CRD. Note that, by [35], the greedy algorithm finds the correct solution in the case of A-optimality for the block design. That guarantee, however, does not extend to the D- and E-optimality in the block case, due to the nature of the objective functions.

We now conduct an empirical exploration of the performances of the greedy algorithm in terms of its ability to find E-optimal solutions. Five different settings of 2^2 factorial designs in two blocks, each corresponding to a specific type of potential outcome matrix, are considered. Each setting is defined by the block sizes M_1 and M_2 , and the 4×2 matrix of variances $S_{h,j}^2$ as shown in columns 3 and 4 of Table 2, respectively.

The first setting considers blocks of equal sizes with potential outcomes satisfying Condition 2 (strict additivity), leading to an E-optimal design that is balanced within each block as per Remark 5. The second setting considers equal block sizes with potential outcomes satisfying Condition 3 (WBH), leading to a balanced design by Theorem 2. In this setting and the previous one, the exact optimal designs provide optimal integer solutions. This is not the case in the third setting, which considers unequal block sizes with potential outcomes satisfying Condition 4 (BBH) but not Condition 3 (WBH). Theorem 2 does not apply directly for E-optimality, but the greedy algorithm identifies the unique true E-optimal allocation determined by the exhaustive search. The fourth setting is similar to the third case above, but is one where the exhaustive search provides multiple solutions, identifying four different allocations, each of which is optimal. In this case, Theorem 2 does not apply directly, and the greedy algorithm identifies one of these solutions. The fifth setting satisfies neither Condition 3 (WBH) nor Condition 4 (BBH), and consequently Theorem 2 cannot provide an exact E-optimal solution. However, the greedy algorithm identifies one of the two (identified through exhaustive search) true optimal integer allocations. A quick note on our greedy algorithm is that, due to the nature of the algorithm in Figures S1 and S2 of Supplementary Material S4, ties are broken deterministically using minimum index when the greedy step returns more than one solution. Thus, our greedy solutions will always achieve the same solution for a given set of inputs without regard for the plurality of solutions (such as the ones in the fourth and fifth settings above).

A similar exploration performed for the D-optimal allocation (shown in Supplementary Material S3) provides evidence that the greedy algorithm can identify the true optimal integer solution when it is unique, and *one* of the true optimal solutions when multiple optimal integer solutions exist.

Table 3: Optimal π_j 's under cost constraints obtained from Theorem 3

Variance vector	Type of optimality	Cost vector (C_1, \dots, C_4)	
		(1, 1, 1, 1)	(0.1, 4, 4, 9)
(1, 1, 1, 1)	A	(0.250, 0.250, 0.250, 0.250)	(0.043, 0.273, 0.273, 0.410)
	D	(0.250, 0.250, 0.250, 0.250)	(0.250, 0.250, 0.250, 0.250)
	E	(0.250, 0.250, 0.250, 0.250)	(0.006, 0.234, 0.234, 0.526)
(1, 2, 3, 4)	A	(0.163, 0.230, 0.282, 0.325)	(0.025, 0.224, 0.275, 0.476)
	D	(0.250, 0.250, 0.250, 0.250)	(0.250, 0.250, 0.250, 0.250)
	E	(0.100, 0.200, 0.300, 0.400)	(0.002, 0.143, 0.214, 0.642)

5 Optimal allocation driven by cost constraints

So far, we have considered optimality criteria that are based on the covariance matrix of the estimated factorial effects, implicitly assuming that all treatment combinations are equally expensive (with respect to cost and/or time). However, such assumptions may not be true in many practical situations, and cost constraints can play an important role in determining optimal allocation. Thus, it is worthwhile to explore solutions to optimal allocation under cost constraints.

We consider the optimal allocation for 2^K factorial CRDs. Let the cost of assigning treatment combination j to one unit be $C_j > 0$, and the total available budget be C . In the new optimization problem, we replace the constraint $\sum_j N_j = N$ in the original problem described in Section 3.1 by the cost constraint $\sum_j C_j N_j \leq C$. The new optimization problem is, therefore:

$$\min_{N_j} \psi(\tilde{\mathbf{V}}) \text{ subject to } \sum_j C_j N_j \leq C, \quad (10)$$

where $\tilde{\mathbf{V}} = \sum_{j=1}^J \frac{S_j^2}{N_j} \tilde{\lambda}_j \tilde{\lambda}_j^T$ and $\psi(\tilde{\mathbf{V}})$ is a functional of $\tilde{\mathbf{V}}$. A straightforward approach to incorporate this new constraint into our previous setting is to re-write the constraint as $\sum_j \tilde{N}_j \leq C$, where $\tilde{N}_j = N_j C_j$ is the total cost for the suggested allocation to treatment arm j . Note that the total sample size N is no longer fixed before the optimization. Under this one-to-one transformation $\tilde{N}_j = C_j N_j$, the optimization problem in equation (10) is equivalent to minimizing the objective function over \tilde{N}_j [36] and can be written as follows:

$$\begin{aligned} \min_{\tilde{N}_j} \psi(\tilde{\mathbf{V}}) &= \min_{\tilde{N}_j} \psi \left(\sum_{j=1}^J \frac{S_j^2}{\tilde{N}_j} \tilde{\lambda}_j \tilde{\lambda}_j^T \right) \\ &= \min_{\tilde{N}_j} \psi \left(\sum_{j=1}^J \frac{S_j^2}{\tilde{N}_j / C_j} \tilde{\lambda}_j \tilde{\lambda}_j^T \right), \\ &\text{subject to } \sum_j \tilde{N}_j \leq C. \end{aligned}$$

Because the optimal solution of the above optimization problem is attained at $\sum_j \tilde{N}_j = C$, the inequality constraint can be replaced by the equality constraint. Then, proceeding along the lines of Theorem 1, one can obtain the cost for the optimal allocation to treatment arm j as $\tilde{N}_j \propto S_j \sqrt{(C_j)}$, $\tilde{N}_j = C/J$ and $\tilde{N}_j \propto S_j^2 C_j$ as the A-, D-, and E-optimal solutions. These results are formalized in terms of the optimal proportion of the budget allocated to each treatment arm, which can be used to determine the number of units to assign to each treatment arm, in Theorem 3.

Theorem 3. Let C be the total budget for the whole experiment and the cost of allocating one experimental unit to treatment j be $C_j > 0$. Let $\pi_j = C_j N_j / C$ denote the proportion of the total budget assigned to treatment j with $\sum_j \pi_j \leq 1$. Then, the

- (a) A-optimal cost-based allocation to the J treatment groups on the basis of covariance matrix $\tilde{\mathbf{V}}$ is $\pi_j = (S_j \sqrt{C_j}) / (\sum_j S_j \sqrt{C_j})$.
- (b) D-optimal cost-based allocation to the J treatment groups on the basis of covariance matrix $\tilde{\mathbf{V}}$ is $\pi_j = 1/J$.
- (c) E-optimal cost-based allocation to the J treatment groups on the basis of covariance matrix $\tilde{\mathbf{V}}$ is $\pi_j = (S_j^2 C_j) / (\sum_j S_j^2 C_j)$.

Remark 7. Theorem 3 can be extended to the case of block designs along the lines of Theorem 2.

Remark 8. If for $j = 1, \dots, J$, the costs C_j in Theorem 3 are the same and equal to C_0 , then the constraint $\sum_j C_j N_j \leq C$ reduces to $\sum_j N_j \leq N$, where $N = C/C_0$. In addition, $\pi_j = C_0 N_j / C \equiv p_j$. Thus, the optimization problem becomes the same as the one in Theorem 1, making it a special case of Theorem 3.

We use an example to demonstrate the applicability of Theorem 3. Consider a 2^2 factorial design, with $C = 100$, per unit cost vector $(C_1, \dots, C_4) = (0.1, 4, 4, 9)$. This setup represents many common scenarios where treatment arm 1 represents the control group 00 and involves a per-unit cost that is negligible compared to the ones with at least one active treatment. On the other hand, treatment arm 4 has both treatments at active level and involves the highest cost. Table 3 shows the A-, D-, and E- optimal proportions of total cost π_j s for two different vectors of variances (S_1^2, \dots, S_4^2) . In one setting, we take the vector as $(1, 1, 1, 1)$ and in another, set it to $(1, 2, 3, 4)$. For the sake of completeness, we also add a column of equal cost $(1, 1, 1, 1)$, under which the p_j 's of Theorem 1 and π_j 's of Theorem 3 become identical, as explained in Remark 8. Thus, the optimal allocations in the first column of Table 3 can also be derived from Theorem 1 with $N = C = 100$.

One can obtain the A-, D-, and E-optimal allocations of N_j s by substituting the optimal π_j s from Theorem 3 into $N_j = (C\pi_j)/C_j$. However, rounding these optimal N_j s into nearest integers may lead to violation of the constraint $\sum_j C_j N_j \leq C$. To avoid such possibilities, one can consider the optimal values of $\lfloor (C\pi_j)/C_j \rfloor$ as approximate integer solutions, where $\lfloor x \rfloor$ denotes the largest integer contained in x .

Researchers may decide to impose an additional constraint on the optimization problem (5) that forces the sum of N_j s to be exactly equal to a predetermined N . Such a problem would give optimal allocation under fixed N , unlike Theorem 3. However, imposing this additional constraint may force the set of feasible solutions to the optimization problem to be empty. For example, suppose for all j , $C_j > C/N$. Then, $\sum_j C_j N_j > \sum_j (C/N) N_j$, which exceeds the allowable cost C if the restriction $\sum_j N_j = N$ is imposed. Thus, additional conditions are necessary to guarantee that the feasible set is nonempty. Obtaining closed-form solutions under such conditions may not be straightforward, and one may need to rely on numerical methods to obtain such solutions.

6 Applications in real experiments

In this section, we demonstrate applications of the results and algorithms developed to two real-life experiments. First, we revisit the education example from [22] described in Section 1. Second, we discuss a pilot audit experiment reported in Libgober's work [37] conducted to identify how perceptions of race, gender, and affluence affect access to lawyers and demonstrate how the proposed methodology can be used to design follow-up experiments in similar populations.

6.1 Education experiment

In the experiment described in [22], the authors used a CRD to allocate the $N = 1,656$ units to the 2^2 treatments. Theorem 1 can directly inform us of the optimal allocation without costs, but there is more structure that we can exploit. For instance, there are potentially two blocks of experimental units or subjects representing female (block 1) and male (block 2) students, with block sizes $M_1 = 948$ and $M_2 = 708$, which can be used to improve their design. Theorem 2 will give us the optimal designs in this case.

Assuming that there is no prior information about the variances of potential outcomes (grade point average (GPA) after year 1), we assume that the variances are equal within and across blocks (Conditions 3 and 4). Then, optimal allocations under both CRD and RBD, from Theorems 1 and 2, respectively, are shown in Tables 4 and 5.

Table 4: A-, D-, and E-optimal allocations under CRD assuming Condition 1

Treatment combination				
	00	01	10	11
$N = 1,656$	414	414	414	414

Table 5: A-, D-, and E-optimal allocations under RBD assuming Conditions 3 and 4

Block	Block size	Treatment combination			
		00	01	10	11
1	$M_1 = 948$	237	237	237	237
2	$M_2 = 708$	177	177	177	177

Now let us consider a hypothetical situation where the number of units N is not prespecified, but there is a budget constraint that depends on the costs associated with the four treatment combinations in this experiment. The treatment combinations 01 (SFP but not SSP) and 10 (SSP but not SFP), each involve cost associated with one of two programs. Angrist et al. [22] reported about \$5,000 for individual students that were allocated to treatment 10 (SSP). Per unit cost for treatment combination 01(SFP) is not mentioned, but if we assume a similar cost as with SSP, then, we can infer that the cost to allocate a student to the treatment combination 11 (SFSP) would be the sum total of the individual costs (\$10,000). The control, representing the treatment combination 00, is possibly the cheapest to allocate units to, because it would involve only administrative cost, which we assume to be \$500. Then, under the original allocation (1,106, 250, 250, 150) in the actual experiment as shown in Table 1, the cost of the experiment would be approximately \$4.5 million. Assuming this amount to be our budget constraint C , the A-, D-, and E-optimal allocations for two different variance vectors obtained from Theorem 3 are shown in Table 6. The first row shows the allocation of the proportions π_j 's of the total budget to the four treatment arms, and the second row shows the corresponding approximate integer solution for N_j as $\lfloor C\pi_j/C_j \rfloor$.

Remark 9. From the education experiment results of [22], we can determine the pooled variance estimates across the four treatment combinations to be (0.92, 0.83, 0.92, 0.94). Given how the variances are nearly the same across treatment groups, the first variance vector (1, 1, 1, 1) is just the basic “equal variance of treatments” assumption informed by the variance estimates of the treatments. This allows us to compare the effect of the costs on the treatment allocation that can possibly justify the disproportionate allocation in education experiment data. For instance, the A-optimal allocation is very similar to the original allocation in the education experiment with most units allocated to the control treatment. The second variance vector (1, 2, 2, 2) then establishes how disproportionate variance structure along with the cost vector affects treatment allocation. It is important to note that N_j , the number of units allocated to treatment j , is unconstrained in this setting.

Table 6: Optimal allocations (π_j and $N_j = \lfloor C\pi_j/C_j \rfloor$) with cost vector (500, 5,000, 5,000, 10,000) and total budget of $C = 4.5$ million

Variance vector (S_1^2, \dots, S_4^2)	Type of optimality		Treatment combination			
			00	01	10	11
(1,1,1,1)	A	π_j	0.085	0.268	0.268	0.379
		N_j	762	241	241	170
	D	π_j	0.25	0.25	0.25	0.25
		N_j	2250	225	225	112
	E	π_j	0.024	0.244	0.244	0.488
		N_j	219	219	219	219
(1,2,2,2)	A	π_j	0.062	0.275	0.275	0.389
		N_j	553	247	247	174
	D	π_j	0.25	0.25	0.25	0.25
		N_j	2250	225	225	112
	E	π_j	0.012	0.245	0.245	0.494
		N_j	111	222	222	222

Table 7: Estimated variances for different treatment groups

Experiment	000	001	010	011	100	101	110	111
Replicate I	0.15	0.15	0.15	0.20	0.27	0.15	0.27	0.27
Replicate II	0.27	0.24	0.20	0.20	0.20	0.27	0.27	0.15
Pooled	0.21	0.20	0.18	0.20	0.23	0.21	0.27	0.21

Contrasting this setting with the previous one shows us that the D-optimal allocation of units does not change with the variance vector, while the E-optimal allocation of units typically follows the proportion in the variances. The A-optimal allocation is affected by both, as given by Theorem 3.

Remark 10. One of the goals of introducing cost-based allocation was to see if we can justify the original allocation in this way. Based on our results, we do not see any of the optimal allocations aligning exactly, but are able to hypothesize the conditions under which the original allocation could have been a valid (optimal) allocation. For example, when the cost vector is as in Section 6.1, A-optimal allocation comes closest to the original allocation.

6.2 Audit experiment

Libgober [37] reported an audit study in which the experimental units were 96 lawyers randomly selected from lawyers in California with a certification in criminal law. Each lawyer in the experiment received an email about a routine “driving under influence” case (a very common criminal matter). The email template suggested that the person sending the email was (i) either white or black (with a racially distinctive name being used to influence perceived race), (ii) either female or male (again cued via the email sender’s name), and (iii) either relatively affluent or relatively lower-income based on a description of client’s earnings. Thus, this experiment had a 2^3 factorial structure. The response was recorded as a binary outcome taking value 1 if there was a response to the email and 0 otherwise. The experiment was replicated with 96 additional lawyers after a certain period of time. The estimated variances s_j^2 for $j = 1, \dots, 8$ treatment groups for the individual replicates and their pooled values are shown in Table 7.

If another completely randomized experiment is planned with lawyers selected from a similar pool with a sample size of 192, then based on the pooled estimated variances shown in Table 7, we can apply Theorem 1 to obtain the optimal designs given in Table 8.

Now assume for illustration that (contrary to fact) instead of two replicates, the original experiment was conducted in two blocks, each block representing one type of lawyer (e.g., criminal and divorce), and suppose we want to obtain optimal allocations within each block for a future experiment. Further suppose that the variance estimates in row j of Table 7 represent estimates of the variances $S_{h,j}^2$ in block $j = 1, 2$ and block 2, respectively. Then, we can directly use part (a) of Theorem 2 to derive the A-optimal design. However, neither WBH nor BBH appears to hold, and we cannot apply parts (b) and (c) of Theorem 2. Conveniently, we can obtain the D- and E-optimal designs using the greedy search algorithm proposed in Section 4.2 (Table 9).

Table 8: Optimal allocations for future CRD

Optimality	000	001	010	011	100	101	110	111
A	24	23	22	23	25	24	27	24
D	24	24	24	24	24	24	24	24
E	24	22	20	22	26	24	30	24

Table 9: Optimal allocations for future RBD

Optimality	Block	000	001	010	011	100	101	110	111
A	I	11	11	10	12	14	10	14	14
	II	13	13	12	11	11	13	13	10
D	I	11	11	12	13	13	10	12	14
	II	13	13	13	12	11	13	11	10
E	I	10	10	10	12	15	10	16	13
	II	13	12	10	11	12	13	15	10

7 Discussion

In this article, we consider optimal allocations of a finite population of experimental units to different treatment combinations of a 2^K factorial experiment under the potential outcomes model. Rather than invoking the standard assumption in the mainstream optimal design literature that outcome data come from a known family of distributions, our work revolves around randomization-based causal inference for the finite-population setting. We find that for 2^K factorial designs with a completely randomized treatment assignment mechanism, D-optimal solutions are always balanced designs, while A- and E-optimal solutions are proportional to finite-population standard deviations and finite-population variances of the treatment groups, respectively. For blocked designs, our solution does not admit a closed form for D- or E-optimality without imposing specific restrictions on the potential outcomes, but the A-optimal allocation is equivalent to finding the A-optimal solution within each block. Convenient integer-constrained programming solutions using a greedy optimization approach to find integer optimal allocation solutions for both complete and block randomization are proposed. Optimal allocations are also derived under cost constraints.

While there is a large literature on model-based optimal designs, to the best of our knowledge, such designs have had very limited development for randomization-based inference for finite populations. The ideas explored and results developed in this article exploit the connection between finite-population sampling and experimental design. This recondite connection has recently been emphasized, explored, and utilized in various contexts by several researchers in causal inference, as discussed in the study by Mukerjee et al. [11]. This article attempts to further strengthen the bridge between finite-population survey sampling and experimental design by utilizing ideas from proportional and optimal allocation for stratified sampling in the context of optimal designs. While the optimal solutions are derived for a finite-population setting, they are readily applicable to a super-population setting without making any assumptions about the probability distribution of the outcome variable.

A question that practitioners may ask is which optimal design should be chosen for a given experiment. The answer would depend on the research goal of the experimenter. As our results have shown, strong assumptions like strict additivity lead to equivalence of A-, D-, and E-optimal designs. However, under treatment effect heterogeneity, different criteria will lead to different allocations. Both A- and D-optimality criteria are associated with quality of estimated causal effects – whereas A-optimality minimizes the average variance of estimators, the D-optimality criterion minimizes the volume of the confidence ellipsoid around the parameters. Some researchers (e.g., [38]) have argued that in model-based settings, A-optimal designs exhibit better performance than D-optimal designs when the objective is screening of active effects from inactive ones. On the other hand, when the goal is to draw the most precise inference on the vector of estimated causal effects, D-optimal design may be a better choice. The goal of the E-optimal design is to minimize the maximum variance of all possible normalized linear combinations of estimated treatment effects. Thus, the E-optimal design is useful when a large number of linear combinations of factorial effects are of interest. The E-optimal allocation, being a minimax strategy, is likely to provide a more conservative solution to the inference problem, but as shown by some researchers (e.g., [39]) in other contexts, the E-optimal solution may be less sensitive to incorrect prior information or assumptions about potential outcomes in comparison to A- and D-optimal designs. However, more investigation is required along these lines in the randomization-based setting.

In typical factorial experiments, all factorial effects are not equally important. An interesting question is, whether such relative importance can be taken into consideration while formulating the optimal design problem. Let us first consider the case of the A-optimality criterion, which minimizes the total (or average) variance of the estimated factorial effects. Because each factorial effect has the same variance $\sum_j S_j^2/N_j$, using a weighted sum of variances that puts more weight on some factorial effects (e.g., main effects) compared to others (e.g., higher-order interactions) does not change the optimal solution. This is also intuitive because, due to the equality of variances of each estimated factorial effect, the A-optimal design guarantees the same optimal inference for their marginal distributions and is thus insensitive to the level of importance attached to each effect. The question, however, is complicated for D- and E- optimal designs, which consider the joint distribution of estimated factorial effects. For instance, the D-optimality criterion represents the volume of the concentration ellipsoid constructed from the vector of estimated factorial effects, and it may not be obvious how such joint inference can be weighted with the relative importance of each individual dimension. Thinking along the lines of traditional factorial design literature, one can assume away some higher-order effects that are apriori known to be unimportant, thereby defining the optimality criterion as a functional of a submatrix of the covariance matrix \tilde{V} . However, while such an approach will lead to a computationally feasible solution, an arbitrary submatrix of \tilde{V} may not render itself to a spectral decomposition like Lemma 1 in the Supplementary Material that leads to nice closed-form solutions to the optimization problems. Using fractional factorial designs (e.g., [15]) in which important effects are completely aliased with unimportant higher-order effects, preserving the orthogonal basis decomposition structure is another alternative and may permit more straightforward extension of the results of this article.

The work presented in this article can be extended in several directions. One limitation of the proposed approach lies in the fact that the correlation among the potential outcomes under different treatment combinations is unidentifiable from the data, forcing us to ignore one term in the covariance matrix of estimated factorial effects while formulating the optimization problem. Basse and Airolidi [40] proposed a model-based approach to overcome this problem in two-armed experiments, in which information on the correlation among the outcomes is available pre-intervention. Such an idea may be extended to the setting of factorial experiments.

In addition, a natural extension of the randomization-based framework of causal inference is the Bayesian framework, in which the potential outcomes are assumed to follow a hierarchical probabilistic model containing hyperparameters with assumed prior distributions. The Bayesian framework proposed in the study by Dasgupta et al. [9] for drawing both super-population and finite-population causal inference from 2^K factorial designs can be utilized to obtain Bayesian optimal designs according to different criteria proposed in literature (e.g., [41]).

Another setting that has gained a lot of attention in recent times is when SUTVA is violated, for example, in the presence of interference between units. Extending the proposed results to such settings is a challenging, yet rewarding problem.

Finally, in certain situations, it is possible that instead of the traditional factorial effects defined by equation (3), the experimenter is interested in other contrasts of the treatment means or more general factorial effects. One such natural choice of contrast is one that compares the outcome of the control group with the average of all other groups that have at least one treatment. Optimal allocations under such a reformulated optimization problem would be an interesting problem to study.

Acknowledgement: We are grateful to the two reviewers whose comments and suggestions were instrumental in significantly improving the quality of the manuscript.

Funding information: This research was partially supported by the National Science Foundation grant SES 2217522. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Ethical approval: The conducted research is not related to either human or animal use.

Informed consent: Not applicable.

Data availability statement: All data necessary for reproducing the results in the paper are included in the paper [and its supplementary information files]. The original data analyzed in Section 6.1 are available from the following webpage: <https://economics.mit.edu/people/faculty/joshangrist/angrist-data-archive>. The original data analyzed in Section 6.2 are available from the corresponding author upon request. All R codes can be obtained from <https://github.com/arrrunn/2KFactOptimal>.

References

- [1] Fisher RA. The design of experiments. Edinburgh: Oliver and Boyd; 1935.
- [2] Yates F. The design and analysis of factorial experiments. Harpenden, England: Imperial Bureau of Soil Science; 1937. p. 35.
- [3] Branson Z, Dasgupta T, Rubin DB. Improving covariate balance in 2^K factorial designs via rerandomization with an application to a New York City Department of Education High School Study. *Ann Appl Stat.* 2016;10(4):1958–76. doi: 10.1214/16-AOAS959.
- [4] Egami N, Imai K. Causal interaction in factorial experiments: application to conjoint analysis. *J Am Stat Assoc.* 2019;114(526):529–40.
- [5] Freedman DA. Statistical models for causation: what inferential leverage do they provide? *Evaluat Rev.* 2006;30(6):691–713.
- [6] Freedman DA. On regression adjustments to experimental data. *Adv Appl Math.* 2008 February;40(2):180–93.
- [7] Abadie A, Athey S, Imbens GW, Wooldridge JM. Sampling-based versus design-based uncertainty in regression analysis. *Econometrica.* 2020 January;88:265–96.
- [8] Olsen R, Orr L, Bell S, Stuart E. External validity in policy evaluations that choose sites purposively. *J Policy Anal Manag.* 2013;32(1):107–21.
- [9] Dasgupta T, Pillai N, Rubin DR. Causal inference for 2^K factorial designs by using potential outcomes. *J R Stat Soc (Ser B).* 2015 September;77(4):727–53.
- [10] Lu J. On randomization-based and regression-based inferences for 2^K factorial designs. *Stat Probabil Lett.* 2016;112(C):72–8.
- [11] Mukerjee R, Dasgupta T, Rubin DB. Using standard tools from finite population sampling to improve causal inference for complex experiments. *J Am Stat Assoc.* 2018;113(522):868–81.
- [12] Zhao A, Ding P. Regression-based causal inference with factorial experiments: estimands, model specifications and design-based properties. *Biometrika.* 2022;109(3):799–815.
- [13] Li X, Ding P, Rubin DB. Rerandomization in 2^K factorial experiments. *Ann Stat.* 2020;48(1):43–63. doi: 10.1214/18-AOS1790.
- [14] Morgan KL, Rubin DB. Rerandomization to improve covariate balance in experiments. *Ann Stat.* 2012 April;40(2):1263–82.
- [15] Pashley NE, Bind MAC. Causal inference for multiple treatments using fractional factorial designs. *Canad J Stat.* 2023 June;51(2):444–68.
- [16] Shi L, Wang J, Ding P. Forward screening and post-screening inference in factorial designs. 2023. arXiv: <http://arXiv.org/abs/arXiv:230112045>.
- [17] Blackwell M, Pashley NE, Valentino D. Batch adaptive designs to improve efficiency in social science experiments. 2022. https://www.mattblackwell.org/files/papers/batch_adaptive.pdf.
- [18] Neyman J. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *J R Stat Soc.* 1934;97(4):558–625.
- [19] Cochran WG. Sampling techniques, 3rd Edition. New York: John Wiley & Sons; 1977.
- [20] Hahn J, Hirano K, Karlan D. Adaptive experimental design using the propensity score. *J Business Econ Stat.* 2011 January;29(1):96–108. doi: 10.1198/jbes.2009.08161.
- [21] Dai J, Gradu P, Harshaw C. Clip-OGD: an experimental design for adaptive Neyman allocation in sequential experiments. 2023. arXiv: <http://arXiv.org/abs/arXiv:230517187>.
- [22] Angrist J, Lang D, Oreopoulos P. Incentives and services for college achievement: evidence from a randomized trial. *Am Econ J Appl Econ.* 2009 January;1(1):136–63.
- [23] Atkinson A, Donev A, Tobias R. Optimum experimental designs, with SAS. Oxford: Oxford University Press; 2007.

- [24] Yang J, Mandal A, Majumdar D. Optimal designs for two-level factorial experiments with binary response. *Stat Sinica*. 2012;22:885–907.
- [25] Yang J, Mandal A. D-optimal factorial designs under generalized linear models. *Commun Stat*. 2015;44(9):2264–77.
- [26] Yang J, Mandal A, Majumdar D. Optimal designs for 2^k factorial experiments with binary response. *Stat Sinica*. 2016 January;26(1):385–411.
- [27] Rubin DB. Randomization analysis of experimental data: The Fisher randomization test comment. *J Am Stat Assoc*. 1980;75(371):591–3.
- [28] Wu CF, Hamada MS. Experiments: planning, analysis, and optimization. 2nd Edition. Wiley; 2009.
- [29] Ding P, Li X, Miratrix L. Bridging finite and super population causal inference. *J Causal Infer*. 2017;5:20160027.
- [30] Chernoff H. Locally optimal designs for estimating parameters. *Ann Math Stat*. 1953 December;24(4):586–602.
- [31] Nemhauser G, Wolsey L. Integer and combinatorial optimization. New York: John Wiley & Sons; 1988.
- [32] Schrijver A. Theory of linear and integer programming. Chichester, UK: John Wiley & Sons; 1998.
- [33] Khan MGM. Mathematical programming in sampling. PhD thesis. India: Aligarh Muslim University; 1995. <http://hdl.handle.net/10603/51752>.
- [34] Sofi N, Ahmad A, Maqbool DS, Ahmad B. A branch and bound approach to optimal allocation in stratified sampling. *Math Theory Model*. 2016;6(4):20–6.
- [35] Friedrich U, Münnich R, de Vries S, Wagner M. Fast integer-valued algorithms for optimal allocations under constraints in stratified sampling. *Comput Stat Data Anal*. 2015 December;92:1–12.
- [36] Boyd S, Vandenberghe L. Convex optimization. Cambridge, UK: Cambridge University Press; 2004.
- [37] Libgober B. Getting a Lawyer While Black: A Field Experiment. *Lewis Clark Law Rev*. 2020;24(1):53–108.
- [38] Jones B, Allen-Moyer K, Goos P. A-optimal versus D-optimal design of screening experiments. *J Quality Technol*. 2021;53(4):369–82.
- [39] Wong WK. Comparing robust properties of A-, D- E- and G-optimal designs. *Comput Stat Data Anal*. 1994 November;18(4):441–8.
- [40] Basse G, Airolidi E. Model-assisted design of experiments in the presence of network-correlated outcomes. *Biometrika*. 2018 December;105(4):849–58.
- [41] Chaloner K, Verdinelli I. Bayesian experimental design: a review. *Stat Sci*. 1995 August;10(3):273–304.