A Theory of Experimenters: Robustness, Randomization, and Balance*,†

Abhijit V. Banerjee
Massachusetts Institute of Technology

Sylvain Chassang
New York University

Sergio Montero
University of Rochester

Erik Snowberg
University of British Columbia

October 6, 2019

Abstract

This paper studies the problem of experiment design by an ambiguity-averse decision-maker who trades off subjective expected performance against robust performance guarantees. This framework accounts for real-world experimenters' preference for randomization. It also clarifies the circumstances in which randomization is optimal: when the available sample size is large and robustness is an important concern. We apply our model to shed light on the practice of rerandomization, used to improve balance across treatment and control groups. We show that rerandomization creates a tradeoff between subjective performance and robust performance guarantees. However, robust performance guarantees diminish very slowly with the number of rerandomizations. This suggests that moderate levels of rerandomization usefully expand the set of acceptable compromises between subjective performance and robustness. Targeting a fixed quantile of balance is safer than targeting an absolute balance objective.

JEL Classification: C90, D78, D81
Keywords: experiment design, robustness, ambiguity aversion, balance, randomization, rerandomization

*We are grateful to David Ahn, Angus Deaton, Pascaline Dupas, Jeff Ely, Guido Imbens, Thomas Lemieux, Charles Manski, Pablo Montagnes, Marco Ottaviani, Bruno Strulovici, Alexey Tetenov, Duncan Thomas, Chris Udry, several helpful referees, as well as audience members at Bocconi, the Cowles Econometric Conference (2016), the ESSET Meeting at Gerzensee (2017), Emory, INSEAD, MIT, the NBER Development Economics Summer Meeting (2016), the North-American Econometric Society Meeting (2016), Penn State, the SISL Conference at Caltech (2017), UBC, UC Berkeley, the UCL/Cemmap workshop on Econometrics for Public Policy (2016), and UPenn for several helpful discussions. Chassang and Snowberg gratefully acknowledge the support of NSF grant SES-1156154.

†Software implementing the treatment allocation protocols discussed in this paper is available at `https://github.com/sylvaingchassang/rct`.

1 Introduction

The proliferation of experiments in academia, business, and public policy has been accompanied by spirited debates about best practices for experiment design and the analysis of experimental results. Topics of debate include pre-registration of experiment designs, pre-analysis plans, the pros and cons of rerandomization, clustering, stratification, and statistical significance testing (Duflo et al., 2008; Bruhn and McKenzie, 2009; Deaton, 2010; Imbens, 2010; Humphreys et al., 2013; Olken, 2015; Athey and Imbens, 2017; Benjamin et al., 2018). At the heart of these debates are different—usually implicit—models of both knowledge generation and how people interpret experimental evidence. Moreover, classical models of experimentation—in which decision-makers are subjective expected utility maximizers—cannot explain the strong preference for randomized controlled trials (RCTs) expressed by experimenters (Kasy, 2016; Savage, 1954).[1]

This paper seeks to help clarify these debates. We propose an instrumental model of experimentation in which a decision-maker collects experimental data in order to improve decisions under uncertainty. A key step, which allows us to account for randomization, is to model the decision-maker as ambiguity averse. Specifically, we use the maxmin framework of Gilboa and Schmeidler (1989), reformulated as the problem of a Bayesian decision-maker seeking to satisfy an adversarial audience. Examples of adversarial audiences abound: the Food and Drug Administration for drug trials, seminar audiences and journal referees for research papers, and governments or NGOs for public policy proposals.[2] This intuitive reinterpretation of Gilboa and Schmeidler (1989) (also known as the ε-contamination model; Huber, 1964) facilitates meaningful comparative statics.

The paper reports two main sets of results. The first set shows that randomized ex-

[1] RCTs are mixed strategies over experimental assignments. As a result they can never be strictly optimal for a subjective expected utility maximizer. For seminal contributions to the economic literature on experimentation and information acquisition, see Rothschild (1974); Grossman and Stiglitz (1980); Aghion et al. (1991); Bergemann and Välimäki (1996); Persico (2000); Bergemann and Välimäki (2002, 2006).

[2] The audience may also be seen as a stand-in for the decision-maker's self-doubt.

periments can be strictly optimal for a decision-maker facing an adversarial audience, and clarifies the circumstances in which this is the case. If the decision-maker cares sufficiently about the adversarial audience, and the sample size is sufficiently large, then it is strictly optimal for the decision-maker to use a randomized experiment. Specifically, all deterministic experiments are bounded away from the first-best, while a standard RCT approaches first-best up to a loss that is vanishing in the sample size. On the other hand, deterministic experiments are generically strictly optimal when the sample size is small and the decision-maker puts sufficiently high weight on her own subjective expected utility.

This set of results accords with the observed heterogeneity in experimental practice. Randomized experiments tend to be used by decision-makers who put a high value on convincing an adversarial audience—for example, scientists and pharmaceutical companies—or when the decision-maker can afford large samples—for example, A/B testing in online marketing. When data points are few and the decision-maker puts little weight on satisfying an adversarial audience—for example, the CEO of a privately held firm deciding whether to adopt a new production technology—optimal experiments are deterministic and optimize the subjective informational value of each data point.

Our second set of results applies our model to an open issue in experiment design: whether or not to rerandomize to improve covariate balance between treatment and control groups. Rerandomization draws multiple treatment assignments, then chooses the one that maximizes a prespecified balance objective. For example, a medical researcher may want to ensure that treatment and control groups are similar in terms of gender, age, race, and baseline health variables such as blood pressure and weight (Morgan and Rubin, 2012). Despite the ease of using rerandomization to improve balance, researchers are concerned that it may adversely affect the reliability of findings (Bruhn and McKenzie, 2009).

We show that the tradeoffs at the heart of rerandomization are concisely captured in our framework. Successive rerandomizations improve balance, which can be expressed through the subjective expected utility component of preferences. However, rerandomization reduces

the robust performance guarantees offered by RCTs. In the extreme case where the allocation is rerandomized until perfect balance is achieved, the allocation is effectively deterministic and worst-case performance guarantees are bounded away from first-best, even for large samples. In contrast, losses against the first-best are vanishing in sample size N provided the number of rerandomizations is polynomial in N. We clarify the potential robustness costs of alternative procedures used to achieve balance: setting a balance target and rerandomizing until it is met (Morgan and Rubin, 2012); or equivalently, selecting an assignment uniformly from a constrained set of assignments achieving a pre-specified balance objective.

Our framework builds on a long line of work in statistics, starting with Wald (1950). Variants of Wald's framework have been used in economics and econometrics to study questions of identification and model uncertainty (Gilboa and Schmeidler, 1989; Hansen and Sargent, 2001; Manski, 2004, 2009; Marinacci, 2015; Kitagawa and Tetenov, 2018).[3] Different approaches to model uncertainty—specifically ambiguity aversion or regret aversion—remain an ongoing subject of debate in this literature. Ambiguity aversion has more attractive normative foundations (see, for example, Marinacci, 2015), while regret aversion is more directly prescriptive. We state our main results under ambiguity aversion but show that they also hold when using regret aversion. In fact, the latter framework allows to dispense with certain technical assumptions, as well as rationalize the practice of null-hypothesis statistical testing (NHST; Tetenov, 2012). In a related paper, Banerjee et al. (2017), three of the authors provides a non-technical discussion of our results on randomization, and a limited discussion of our results on rerandomization.

The paper is structured as follows. Section 2 introduces our framework. Section 3 delineates the forces that determine whether running a randomized or deterministic experiment is optimal. Section 4 studies the tradeoffs involved in rerandomization. Section 5 shows that our results extend to reference-dependent preferences, better suited to explain the use of

[3]This paper is also related to the dormant literature in multi-Bayesian statistical decision theory (Weerahandi and Zidek, 1981, 1983). In these models, Bayesians with conflicting preferences adopt random decision rules, rather than randomized experiments.

NHST in decision-making. Section 6 contains several discussions: of other possible rationalizations of randomization, of the positive implications of our theory, of practical implications for rerandomization, and of possible directions for future research. Appendix A extends our analysis to dynamic settings where participants arrive over time, and experimenters must choose treatment assignments in real time, without knowing what the ultimate sample of covariates will be.[4]

2 Model

In this section, we first lay out the problem of experiment design along lines similar to Banerjee et al. (2017). We then specify the preferences and beliefs of the decision-maker.

2.1 A Framework for Studying Experiment Design

Decisions and payoffs. A decision-maker chooses whether or not to provide a treatment $\tau \in \{0, 1\}$ to a population of heterogeneous individuals. For simplicity, we assume that the final policy choice $a \in \{0, 1\}$ is all-or-nothing and sets $\tau = a$ for all individuals. Potential outcomes for individual i with treatment status τ are random variables $y_i^\tau \in \{0, 1\}$; $y = 1$ is referred to as a success. Each individual has observable covariates $x_i \in X$ that affect the distribution of outcomes y_i. Covariates x_i are uniformly distributed in X.

The probability of success, given treatment τ and covariate x, is denoted by $p_x^\tau \equiv \text{Prob}(y^\tau = 1|x)$. The state of the world is described by the finite-dimensional vector $p = (p_x^0, p_x^1)_{x \in X} \in [0, 1]^{2X} \equiv P$ of success probabilities p_x^τ conditional on treatment status $\tau \in \{0, 1\}$ and covariate x. Outcomes y_i^τ for different individuals are drawn independently given state of the world p. Note that state-space P is compact, convex, and finite-

[4]Online Appendix B clarifies that a traditional balance objective—minimizing the Mahalanobis distance between mean covariates across treatment and control samples—coincides with subjective expected utility maximization under specific priors. Online Appendix C clarifies how regret aversion can rationalize the use of NHST. Proofs are contained in Online Appendix D. Online Appendix E presents simulations.

dimensional. Given a state p and a policy decision $a \in \{0, 1\}$, the decision-maker's payoff $u(p, a)$ is

$$u(p, a) \equiv \mathbb{E}_p y^a = \frac{1}{|X|} \sum_{x \in X} p_x^a.$$

Although covariates x are observable, our framework is consistent with unobservable characteristics. Denoting unobserved characteristics by z, and given a joint distribution F over (x, z) we have $p_x^\tau = \int p_{x,z}^\tau dF(z|x)$. When x and z are embedded in $\mathbb{R}^k$, our framework captures traditional concerns over omitted variables by allowing the mapping $x \mapsto p_x^\tau$ from observed covariates to outcomes to be discontinuous. Even if $p_{x,z}^\tau$ is continuous in (x, z), omitted variables cause $x \mapsto p_x^\tau$ to become discontinuous when the distribution of z conditional on x changes rapidly. That is, when individuals with similar values of x may have very different values of z.[5]

Experiments and strategies. To maximize her odds of making the correct policy choice, the decision-maker can run an experiment on N participants, which are a representative sample of a broader population. The decision-maker chooses a treatment status τ_i for each participant i, observes outcomes $y_i^{\tau_i}$, and then makes a policy decision a that applies to the entire broader population.

We simplify away the problem of representative sampling by taking as given the covariates $(x_i)_{i \in \{1,\ldots,N\}}$ of experimental participants and assuming that they are *exactly* representative of the underlying set of types. That is, we assume that $N = |X|$ and $\{x_i\}_{i \in \{1,\cdots,N\}} = X$. This means that random sampling is not needed to ensure that experimental participants are representative of the overall population. This assumption is consistent with the Neyman-Rubin "potential outcomes" framework (Neyman, 1923; Rubin, 1974). Each person i is unique and cannot be observed in both the treated and untreated state. This assumption simplifies exposition and focuses on randomness in treatment assignment as the only random component of experiment design. Under this exact sampling assumption, the broader population can be

[5]That is, $F(z|x)$ is not continuous in x.

thought of as identical copies of the experimental sample.[6]

Given covariates $(x_i)_{i\in\{1,\cdots,N\}}$, an experimental assignment is a tuple $e = (\tau_i)_{i\in\{1,\ldots,N\}} \in \{0,1\}^N \equiv E$. Experiment e generates outcome data $y = (y_i)_{i\in\{1,\ldots,N\}} \in \{0,1\}^N \equiv \mathcal{Y}$, with y_is independent realizations of $y_i^{\tau_i}$ given (x_i, τ_i). We say that a covariate and treatment pair $(x,\tau) \in X \times \{0,1\}$ is *sampled* by experiment e—denoted by $x, \tau \in e$—if and only if there exists $i \in \{1, \cdots, N\}$ such that $(x_i, \tau_i) = (x, \tau)$.

The decision-maker's strategy consists of both an experimental design $\mathcal{E} \in \Delta(E)$, which is a mixed strategy over experimental assignments e, and an allocation rule $\alpha : E \times \mathcal{Y} \to \Delta(\{0,1\})$, which maps experimental data (e, y) to policy decisions $a \in \{0,1\}$. We denote by $\mathcal{A}$ the set of such mappings α.[7]

The standard RCT. For simplicity, we assume that $N/2$ is an integer throughout the paper. An experiment design of interest is the standard RCT, assigning 50% of participants to treatment $\tau = 1$.[8] It corresponds to the strategy $(\mathcal{E}_{\text{RCT}}, \alpha_{\text{RCT}})$:

- $\mathcal{E}_{\text{RCT}}$ draws an exchangeable profile $(\tau_i)_{i\in\{1,\ldots,N\}} \in \{0,1\}^N$ of treatment assignments such that $\sum_{i=1}^N \tau_i = N/2$,
- Policy a is chosen according to the *empirical success* rule: $\alpha_{\text{RCT}}(e, y) \equiv \mathbf{1}_{\overline{y}^1 \geq \overline{y}^0}$, where $\overline{y}^\tau \equiv \frac{2}{N}\sum_{i=1}^N y_i \mathbf{1}_{\tau_i=\tau}$ is the mean outcome for participants with treatment status τ.

Note that policy choice α_{RCT} is deterministic conditional on experimental outcomes: performance guarantees for RCTs established here stem from random assignment, not random policy-making.

2.2 Preferences and Beliefs

[6] Our results extend when $N < |X|$ and experimental participants are randomly sampled from X according to the underlying distribution of covariates.

[7] Following standard notation, for any $\sigma \in \Delta(\{0,1\})$, $u(p,\sigma) \equiv \mathbb{E}_{a\sim\sigma}[u(p,a)]$.

[8] The analysis extends to experiments assigning any fixed share $\pi \in (0,1)$ to treatment. Bounds use the reduced sample $N' = 2\min\{\pi, 1-\pi\}N$ obtained by dropping excess data points in the larger subsample.

The experiment designer's preferences and beliefs are at the heart of our analysis. As noted in the Introduction, subjective expected utility maximization does not yield strict preferences for randomization. Ambiguity aversion does, under conditions explored below.

Preferences. We consider an ambiguity-averse decision-maker that chooses strategy $(\mathcal{E}, \alpha)$ to maximize:

$$\lambda \mathbb{E}_{h_0,\mathcal{E}}[u(p, \alpha(e, y))] + (1 - \lambda) \min_{h \in H} \mathbb{E}_{h,\mathcal{E}}[u(p, \alpha(e, y))]. \tag{DP}$$

where H is a set of priors $h \in \Delta(P)$ over states $p \in P$. These preferences coincide with the standard maxmin model of Gilboa and Schmeidler (1989): the decision-maker's objective can be expressed as $\min_{h \in H_0} \mathbb{E}_{h,\mathcal{E}}[u(p, \alpha(e, y))]$, with $H_0 \equiv \{\lambda h_0 + (1 - \lambda)h' | h' \in H\}$. This parameterization, known as the ε-contamination model of Huber (1964, 2010), is popular in the robust statistics literature, with λ serving as a tuning parameter. It lets us emphasize subjective expected utility maximization as a special case ($\lambda = 1$), and facilitates comparative statics.

Representation (DP) of the standard maxmin model admits an intuitive interpretation: a subjective-expected-utility maximizing decision-maker, with prior h_0, faces an adversarial audience of Bayesian stakeholders with non-common priors $h \in H$. Note that heterogeneous priors within the audience are essential: if the audience entertained a common prior h, then the decision problem would reduce to subjective expected utility maximization for the mixed prior $\lambda h_0 + (1 - \lambda)h$. If $\lambda = 1$, the decision-maker does not care about the audience, and we recover the standard subjective expected utility, or *Bayesian* model. Throughout the paper, we interpret results in terms of an adversarial audience, but it is equally appropriate to view decision problem (DP) as a model of the experimenter's *internal* doubt.

Beliefs. Our main assumption concerns the set of priors H entertained by audience members. We assume that for any given experiment e, there exists a welfare-minimizing prior h such that knowing, without error, the success rates $(p_x^\tau)_{x,\tau \in e}$ for pairs (x, τ) sampled by ex-

periment e is not sufficient for efficient decision-making. Denoting the expected probability of success given policy $a \in \{0,1\}$ by $p^a \equiv \frac{1}{|X|}\sum_{x\in X} p^a_x$, the assumption is:

Assumption 1 (limited extrapolation)**.** *There exists $\xi > 0$ such that, for all $e \in E$, there exists a prior $h \in \arg\min_{h\in H} \mathbb{E}_h(\max_{a\in\{0,1\}} p^a)$ such that*

$$\mathbb{E}_h\left[\max_{a\in\{0,1\}} p^a - \max_{a\in\{0,1\}} \mathbb{E}_h\left[p^a|(p^\tau_x)_{x,\tau\in e}\right]\right] \geq \xi.$$

Note that, for any experiment e, $(p^\tau_x)_{x,\tau\in e}$ is an upper bound, in the sense of Blackwell (1953), of the information generated by experiment e. $\mathbb{E}_h\left[p^a|(p^\tau_x)_{x,\tau\in e}\right]$ is the expected payoff of action a from the perspective of a decision-maker with initial prior h after observing the true success rate p^τ_x for each pair (x,τ) sampled by the experiment. For any deterministic experiment e, the limited extrapolation assumption implies there exists a pernicious prior h such that data from sampled pairs $(x,\tau) \in e$ does not permit extrapolation to the optimal policy for the entire population.

A key implication of Assumption 1 is that audience members cannot be arbitrarily pessimistic. They cannot be certain that both treatment and control would result in certain failure. If this were the case, nature would minimize the decision-maker's welfare using the prior $p^1 = p^0 = 0$, which makes experimentation worthless.

It is useful to point out two specific features of our modeling exercise. The first is that any comparative static exercise in which the sample size N grows, while Assumption 1 is maintained for $\xi > 0$ fixed, also requires type space X to grow with N. If a finite type space X was held fixed while N grew then all pairs (x,τ) could be sampled by a deterministic experiment when N is sufficiently large. The second feature is that Assumption 1 becomes unnecessary in an alternative model in which the decision-maker exhibits regret aversion, rather than ambiguity aversion, as described in Section 5. A key advantage of regret aversion is that the priors that minimize the decision-maker's welfare must still allow for the possibility of learning. Otherwise there is nothing to regret. The main disadvantage of regret aversion

is that it may lead to intransitivity for some choice problems (Marinacci, 2015).

The following lemma establishes that the set of priors satisfying Assumption 1 is non-empty. The proof is by construction: it contains an example that is useful for understanding the content of the limited extrapolation assumption.

Lemma 1 (existence and example)**.** *There exists a set of priors H that satisfies limited extrapolation for parameter $\xi = \frac{1}{8}$.*

Proof. Consider all pairs (X', τ), with $X' \subset X$ satisfying $|X'| = |X|/2$, and $\tau \in \{0, 1\}$. Such a pair specifies a set of covariates and a single treatment status. For each pair (X', τ), consider the prior $h_{X',\tau}$ such that if $x \notin X'$ or $\tilde{\tau} \neq \tau$ then $p_x^{\tilde{\tau}} = \frac{1}{2}$. If $x \in X'$ and $\tilde{\tau} = \tau$, then the prior over this pair is given by $p_x^{\tilde{\tau}} = \frac{1+\sigma}{2}$ where σ takes values 1 and -1 with equal probability. Note that σ takes a single value for all such pairs $(x, \tilde{\tau})$. Set H is the set of such priors $h_{X',\tau}$.

Set H satisfies limited extrapolation. First, the expected value $\mathbb{E}_h(\max_{a\in\{0,1\}} p^a)$ is constant and equal to $5/8$ for all $h \in H$. This implies that $\operatorname{argmin}_{h\in H} \mathbb{E}_h(\max_{a\in\{0,1\}} p^a) = H$. Next, for any deterministic experiment e there exists a prior $h \in H$ such that decision-making is bounded away from first-best. As $N = |X|$, there must exist X' and $\tau \in \{0, 1\}$ such that no pair in $X' \times \{\tau\}$ is sampled by deterministic experiment e. Thus, under the corresponding belief $h_{X',\tau}$, data generated by experiment e carries no information about the sign of σ. As a result, for all $\tau \in \{0, 1\}$, $\max_{a\in\{0,1\}} p^a - p^\tau = \frac{1}{4}$ with probability $1/2$. Thus, limited extrapolation holds for $\xi = 1/8$.[9] □

Two points raised by the proof are worth highlighting. First, in order to ensure that Assumption 1 is non-empty, $|X|$ needs to be larger than $N/2$—we assume $|X| = N$. Second, pernicious priors are such that non-sampled pairs (x, τ) create significant residual uncertainty

[9]Note that H in this example will continue to satisfy Assumption 1 even if it is expanded by adding other priors as long as $\mathbb{E}_{h'}(\max_{a\in\{0,1\}} p^a) \geq 5/8$ for each added prior h'.

about the correct treatment. As deterministic experiment e provides no information about these pairs, this uncertainty is not resolved by the experiment.

3 Optimal Design and Randomization

This section shows that both deterministic and randomized experiments can be strictly optimal for a decision-maker solving (DP). Which type of experiment is optimal depends on the sample size N and the weight λ given to subjective performance $\mathbb{E}_{h_0} u(a,p)$, versus worst-case performance guarantees $\min_{h\in H} \mathbb{E}_h u(a,p)$.

3.1 Bayesian Experimentation

When $\lambda = 1$, the decision-maker is a standard subjective expected utility maximizer. In this case, it is known that deterministic experiments are *weakly* optimal. In fact, we show that for generically every prior h_0 under which all data points are valuable, deterministic experiments are *strictly* optimal. We use the topological version of genericity—that is, open, dense sets—under the total variation distance over distributions—$d(h,h') \equiv \sup_{\substack{A\subset P \\ A \text{ meas.}}} |h(A) - h'(A)|$—sometimes described as *the* statistical distance.

We assume that the decision-maker values every data point. Formally, for any deterministic experiment $e \in E$, let e_{-i} denote the experiment in which sample point (x_i, τ_i) has been removed. Let $E_{-1} = \{e_{-i} \mid i \in \{1,\dots,N\}, e \in E\}$ denote the set of experiments with samples of length $N-1$. We denote by $\mathcal{A}_{-1}$ the set of policy rules mapping an experiment $e \in E_{-1}$ and outcomes $y \in \{0,1\}^{N-1}$ to $\Delta(\{0,1\})$.

Definition 1. *We say that all data points are valuable under prior h_0 whenever*

$$\max_{\alpha\in\mathcal{A}} \max_{e\in E} \mathbb{E}_{h_0}[u(p,\alpha(e,y))|e] > \max_{\alpha\in\mathcal{A}_{-1}} \max_{e\in E_{-1}} \mathbb{E}_{h_0}[u(p,\alpha(e,y))|e].$$

That is, when all data points are valuable under prior h_0, reducing the sample size reduces the quality of decision-making. This holds automatically if there is some cost of acquiring data points and the sample size N is chosen optimally by the decision-maker. However, as N is exogenously given here, we assume all data points are valuable. This rules out trivial indifferences between experiments.

Proposition 1 (near-Bayesians do not randomize)**.** *For every prior h_0, if $\lambda = 1$, then there exists a deterministic experiment e^* solving* (DP)*.*

For generically every prior h_0 such that all data points are valuable, there exists $\underline{\lambda} \in (0,1)$ and a deterministic experiment e^* such that, for all $\lambda > \underline{\lambda}$, e^* is uniquely optimal.

That is, when all data points are valuable, a deterministic experiment is generically strictly optimal when $\lambda = 1$, and also when the decision-maker puts a small, but non-zero, weight $1 - \lambda > 0$ on the preferences of her audience.[10] In recent work, Kasy (2016) uses a similar result to argue that RCTs may be suboptimal. Instead, we believe that Proposition 1 shows the limits of subjective expected utility maximization as a positive model of experimenters.

3.2 Adversarial Experimentation

We now examine the case where the experimenter cares about her audience's preferences. We reiterate that the type space X must grow with N in order to maintain Assumption 1 for a fixed value $\xi > 0$, and that $H_0 \equiv \{\lambda h_0 + (1-\lambda)h \mid h \in H\}$.

Proposition 2. *Take weight $\lambda \in (0,1)$ and parameter $\xi > 0$ in Assumption 1 as given. There exists $\underline{N}$ such that for all $N \geq \underline{N}$, and all decision problems* (DP) *satisfying Assumption 1 with parameter ξ, optimal experiments are randomized. More precisely, the following hold:*

[10] Here genericity applies relative to the set of priors under which all data points are valuable. Note, however, that this is an open set under the statistical distance. Hence, the set of priors satisfying the second half of Proposition 1 is in fact open within the set of all priors.

(i) For any N, any optimal experiment design $\mathcal{E}^*$ satisfies

$$\max_{\alpha\in\mathcal{A}} \min_{h\in H_0} \mathbb{E}_{h,\mathcal{E}^*}\left[u(p,\alpha(e,y))\right] \geq \min_{h\in H_0} \mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p,a)\right] - \sqrt{\frac{1}{N}}.$$

(ii) All deterministic experiments $e \in E$ are bounded away from first-best:

$$\forall e \in E, \quad \max_{\alpha\in\mathcal{A}} \min_{h\in H_0} \mathbb{E}_{h,e}\left[u(p,\alpha(e,y))\right] < \min_{h\in H_0} \mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p,a)\right] - (1-\lambda)\xi.$$

Point (i) shows that the efficiency loss of the optimal experiment compared to the first-best decision is bounded above by a term of order $1/\sqrt{N}$. This performance bound is potentially attained by several experiment designs, including the standard RCT: $(\mathcal{E}_{\text{RCT}}, \alpha_{\text{RCT}})$. Point (ii) shows that the loss from a deterministic experiment is bounded below by $(1-\lambda)\xi$, where ξ is defined by Assumption 1. Hence, taking ξ as given, there exists $\underline{N}$ such that for all $N > \underline{N}$ the optimal experiment is not deterministic; it is randomized. Note that this is not an asymptotic result: threshold $\underline{N}$ is the lowest value such that $\sqrt{\frac{1}{\underline{N}}} < (1-\lambda)\xi$.

The traditional interpretation of an ambiguity-averse decision-maker's problem as a zero-sum game played against nature can be used to understand the value of randomization. The decision-maker selects an experimental design $\mathcal{E}$ and a decision rule α, while nature picks the prior h to minimize the decision-maker's payoff. If there is a known pattern in the choice of experimental assignments, nature can exploit the pattern to lower the decision-maker's payoff. Randomization eliminates patterns that nature can exploit. This is related to the fact that ambiguity-averse agents may have preferences for randomization even if they exhibit risk-aversion over known lotteries (Saito, 2015).[11] In our context, for large sample size N, the Bayesian component of preferences is close to the first-best for any prior. Thus, the subjective benefits of running a deterministic experiment are small. In contrast, under Assumption 1, the impact of randomization on robust payoff guarantees is bounded away from 0.

[11] A key modeling choice here is that nature cannot observe the outcome of the decision-maker's randomization before picking a prior. Kasy (2016) assumes that nature observes the outcome of the experimenter's randomization and then picks a prior, which renders randomization useless.

Randomization can also be seen as a way to get parties with divergent priors to agree on a design. Such stakeholders need not be satisfied by any given deterministic experiment: there may always exist a prior under which the experiment's design is badly flawed. Running a randomized experiment guarantees each stakeholder that, in expectation, the final decision will be close to optimal from her perspective.

Regardless of the interpretation, our results emphasize the importance of actual randomization. Assignments that are only "nearly" random, such as assignment based on time of day (see Green and Tusicisny, 2012, for a critique), or alphabetical order (as in the case of Miguel and Kremer, 2004; see Deaton, 2010 for a critique), remain problematic under some adversarial priors.[12] Randomization provides ex ante performance guarantees even under the most skeptical priors. Approximate randomization does not.

3.3 Standard RCTs as a Rule of Thumb

A well known corollary of the proof of Proposition 2 (see, for example, Manski, 2004) is that the standard RCT $(\mathcal{E}_{\text{RCT}}, \alpha_{\text{RCT}})$ leads to near optimal decisions under every prior.

Corollary 1. *Experimentation policy* $(\mathcal{E}_{\text{RCT}}, \alpha_{\text{RCT}})$ *is such that for all priors* $h \in \Delta(P)$*,*

$$\mathbb{E}_{h,\mathcal{E}_{\text{RCT}}}\left[u(p, \alpha_{\text{RCT}}(e, y))\right] \geq \mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p, a)\right] - \sqrt{\frac{1}{N}}.$$

This result implies that a standard RCT guarantees approximately optimal decision-making for both Bayesian and ambiguity-averse decision-makers.[13] Importantly, the decision-maker does not even need to specify her own preferences. When sample N is large enough, a

[12]Specifically, Miguel and Kremer (2004) stratify by administrative unit before assigning schools to three treatment groups according to alphabetical order. A later paper argues that these precautions allowed for valid inference (Baird et al., 2016). Still, the Deaton (2010) critique provides an example of the challenge an adversarial prior can pose for non-randomized designs.

[13]Note that Corollary 1 holds even in models where $|X|$ is arbitrarily large compared to N, provided covariates are sampled in a representative way. As outcomes are bounded, it is not possible for rare—and therefore hard to sample—covariate realizations to have a substantial impact on payoffs.

RCT is near optimal for all parameters λ and sets of priors H. Note also that the corollary holds using the deterministic empirical success rule $\alpha_{\text{RCT}}(e, y)$. This shows that the robustness of RCTs stems from random assignment, not random policy-making.

Together with Proposition 2, Corollary 1 makes the following predictions about experimenter preferences when the sample size is large and robustness is a concern. First, experimenters strongly prefer standard RCTs over any deterministic experiment.[14] Second, while the standard RCT need not be exactly optimal, the marginal improvement from using the exactly optimal random design is small compared to the improvement from using the standard RCT over a deterministic design. Note that while standard RCTs are also near optimal under a subjective expected utility model, such a model does not rationalize the strong preferences for randomization that experimenters exhibit in practice. In economics, randomization frequently requires costly negotiations with implementation partners. Furthermore, randomization comes at the cost of balance, which experimenters clearly care about, as revealed by covariate collection efforts.

4 Rerandomization

As noted above, a standard RCT is only a near-optimal solution to (DP). With small probability, it may result in unbalanced assignments that provide little real opportunity for learning under reasonable priors. To improve balance, treatment assignment needs to be non-exchangeable: if someone with particular covariates is assigned to treatment, then a participant with similar covariates should be assigned to the control group.

In practice, experimenters sometimes resolve this difficulty through rerandomization: they repeatedly randomize until they obtain an assignment that satisfies their balance objective. As Bruhn and McKenzie (2009) highlight, this practice is widespread, poorly theorized

[14] Cardinal measures of the intensity of experimenter preferences could be obtained by letting the experimenter trade-off the choice of her preferred experiment design against making policy decision a at random with some probability.

and—in principle—a substantial deviation from one-shot randomization.

Our framework is well suited to clarify the tradeoffs involved in rerandomization: it improves the subjective value of experiments at some cost to robustness. We argue that, used in moderation, rerandomization provides a simple way to trade off the subjective value of an experiment design for tolerable losses in robustness. Rerandomization captures what may be an important aspect of optimal solutions to (DP)—it correlates assignment across participants. It provides an expanded class of rules-of-thumb that lets experimenters trade off subjective performance versus robust performance guarantees, while ensuring near-optimal performance along both objectives.

We discuss two approaches to rerandomization: K-rerandomization, which allows for a fixed number K of rerandomizations, and balance-targeting rerandomization, which sets a specific balance target for the assignment, and draws independent assignments until that target is reached.

4.1 K-rerandomized Experiments

K-rerandomized experiments extend the standard RCT. As in our previous analysis, we fix the policy rule α to be the empirical success rule: $\alpha_{\text{RCT}}(e, y) = \mathbf{1}_{\overline{y}^1 \geq \overline{y}^0}$.

Given a profile of covariates $(x_i)_{i\in\{1,\dots,N\}}$, K-rerandomized experiment design $\mathcal{E}_K$ takes the following form:

1. Independently draw K assignments $\{e_1, \cdots, e_K\}$ with each $e_k = (\tau_{i,k})_{i\in\{1,\dots,N\}}$ an exchangeable assignment such that 50% of participants receive treatment $\tau = 1$.

2. Select an assignment $e_K^* \in \text{argmax}_{e\in\{e_1,\cdots,e_K\}} \mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e, y))]$ that maximizes the decision-making value of experiment e under prior h_0.[15]

3. Run experiment e_K^*, generating outcomes y_K.

4. Choose a policy $a \in \{0, 1\}$ according to α_{RCT}.

[15] An assignment is selected with uniform probability if multiple assignments maximize subjective performance.

More generally, one could replace the objective $\mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e, y))]$ in stage 2 by any utility function $B : e \in E \mapsto B(e) \in \mathbb{R}$ defined directly over experimental assignments. Propositions 3 and 4 (below) hold for this more general class of experiment designs.

We discuss two practical aspects of rerandomization in Appendices. First, Appendix A extends our analysis of rerandomization to sequential settings where participants arrive one after the other, and treatment must be chosen in real time. This makes balance more difficult to achieve. Second, Online Appendix B describes priors h_0 under which maximizing subjective expected utility term $\mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e, y))]$ coincides with a standard balance objective: minimizing the Malahanobis distance between mean covariates across the treated and control groups (Rubin, 1980; Cochrane and Rubin, 1973; Rubin, 1979; Morgan and Rubin, 2012).[16]

The remainder of this section uses our framework to clarify the potential costs and benefits of rerandomization.

4.2 The Tradeoff of Rerandomization

The benefit of rerandomization. The benefit of rerandomization is immediately apparent: it increases the subjective expected performance of the implemented design.[17]

Remark 1. *Subjective performance* $\mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e^*_K, y))]$*, in which the expectation is taken given a realization for* e^*_K*, first-order stochastically dominates* $\mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e^*_{K-1}, y))]$*.*

In the special case where prior h_0 rationalizes minimizing the Mahalanobis distance between mean covariates, this observation coincides with the analysis of Morgan and Rubin (2012). In that work, rerandomization leads to significantly more precise estimates of treatment effects when outcomes come from a linear Gaussian model. More generally, Bugni

[16]The Mahalanobis distance defined by covariates $\mathbf{x} = (x_i)_{i \in \{1,\dots,N\}}$ (with $x_i \in \mathbb{R}^k$) between two arbitrary vectors $v_1,\ v_2 \in \mathbb{R}^k$ is given by $(v_1 - v_2)^\top \mathsf{cov}(\mathbf{x})^{-1}(v_1 - v_2)$. This is equivalent to the standard Euclidean distance in the rotated and rescaled coordinate system in which the sample covariance matrix of covariates $\mathbf{x}$ becomes the identity matrix—that is, the Euclidean distance in the basis defined by the principal components of $\mathbf{x}$.

[17]This holds for *any* preference over deterministic assignments used to select e^*_K among K assignments.

et al. (2018) show that balance from symmetric stratification procedures also substantially increases precision.

The cost of rerandomization. Although rerandomization provides clear benefits in terms of improving the subjective value of experiments, there are common, but vague, concerns about its potential costs. As Bruhn and McKenzie (2009) document, this leads many researchers to rerandomize, but not report the fact that they did. Our framework clarifies the issue by showing that rerandomization can indeed reduce robustness. However this cost grows very slowly with the number of rerandomizations.

We assume that the number of experimental assignments e maximizing $\mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e, y))]$ is bounded above by a finite number, independently of N.[18]

Proposition 3. *There exists $\underline{\rho} > 0$ such that, for all N and H satisfying Assumption 1 with parameter $\xi > 0$, if $K \geq 2^N$, then*

$$\max_{\alpha} \min_{h \in H} \mathbb{E}_{h, \mathcal{E}_K} \left[u(p, \alpha(e, y)) \right] < \min_{h \in H} \mathbb{E}_h \left[\max_{a \in \{0,1\}} u(p, a) \right] - \underline{\rho} \xi.$$

Intuitively, when K is sufficiently large, the experimental assignment is essentially deterministic. Proposition 2 implies that this precludes first-best performance guarantees. As a result, Proposition 3 encourages caution towards balance-targeting rerandomization schemes that set a fixed balance target, and rerandomize until that target is reached. If very few assignments satisfy this target, then such rerandomization algorithms cause non-vanishing robustness losses.

Still, the number of rerandomizations K necessary to cause non-vanishing robustness losses is exponential in the sample size. This suggests that a moderate number of rerandomizations may have little impact on robustness. This is indeed the case.

[18] In the case where the balance objective coincides with the Mahalanobis distance, there will generically be two optimal assignments as treatment and control status can be switched without changing balance.

Proposition 4. *Given $K \geq 2$, consider K-rerandomized experiment $\mathcal{E}_K$. For all $h \in \Delta(P)$,*

$$\mathbb{E}_{h,\mathcal{E}_K}[u(p, \alpha_{\text{RCT}}(e, y))] \geq \mathbb{E}_h\left[\max_{a \in \{0,1\}} u(p, a)\right] - 2\sqrt{\frac{\ln K}{N}}.$$

Proposition 4 clarifies that additional losses from rerandomization are of order $\sqrt{\ln K}$.[19] When $K = O(N)$, this implies that the K-rerandomized experiment still approaches the first-best as N grows large. Still, whether this potential loss in robustness is worthwhile depends on experimenter preferences. We propose what we believe are reasonable guidelines for practice in Section 6, but emphasize that they are ultimately subjective.

Remarkably, Proposition 4 holds for *any* preference over deterministic assignments used to select e^*_K among K assignments. As we discuss in Section 6, this suggests that K-rerandomized experiment designs could be used to trade off robustness with objectives other than standard statistical balance, including the preferences of stakeholders.

4.3 Direct Approaches to Balance Involve Similar Tradeoffs

We conclude this section by showing that understanding K-rerandomization helps us understand other, more direct, approaches to balance.

Consider the following experiment design, denoted by $\mathcal{E}_{E_\dagger}$. First, the experimenter defines a set of acceptable assignments $E_\dagger \subset E$. For example, the set of assignments that are perfectly balanced on a few discrete-valued dimensions of covariates $x \in \mathbb{R}^k$, and tolerably balanced on the remaining dimensions. Second, an assignment e is drawn from $E_\dagger$ with uniform probability, and implemented. This overall design could be implemented by stratifying on these few dimensions—achieving perfect balance on those dimensions—and then drawing stratified samples until they satisfy the overall balance objective.[20]

[19] This bound is non-asymptotic and conservative.

[20] Stratification is also sometimes referred to as "blocking." It usually involves dividing a sample along some particular dimension, say gender, and then assigning treatment through draws without replacement within each gender. In this way, the treatment and control groups will be exactly balanced on this dimension. In our

The potential issue here is that set $E_\dagger$ may be quite small. If this is the case, then the experiment is near deterministic, and the argument of Proposition 3 applies: balance constraints reduce performance guarantees by an amount that does not vanish as N gets large. This is especially problematic if set $E_\dagger$ is difficult to compute explicitly, so that the decision-maker is not naturally aware of how demanding a balance target is. The bound K in K-rerandomization mechanically ensures that the experimenter does not seek to achieve excessively difficult balance objectives. Intuitively, if it takes many draws of random samples to reach the acceptable set $E_\dagger$, then it is likely that the set $E_\dagger$ is excessively small.

This intuition can be used to obtain a lower bound on the performance of experiment design $\mathcal{E}_{E_\dagger}$. Let $p_{E_\dagger}$ denote the probability that a uniformly chosen assignment $e \in E$ belongs to $E_\dagger$.[21] Proposition 4 can be used to establish the following efficiency bound.

Proposition 5. *For all $h \in \Delta(P)$,*

$$\mathbb{E}_{h,\mathcal{E}_{E_\dagger}}\left[u(p, \alpha_{\text{RCT}}(e,y))\right] \geq \mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p,a)\right] - \min_{K\geq 2}\left[2\sqrt{\frac{\ln K}{N}} + (1-p_{E_\dagger})^K\right].$$

This implies that a procedure seeking to achieve a given balance objective will come at a limited cost to robustness if an acceptable assignment can be reached with high probability within a small number of rerandomizations.

Quantile-targeting rerandomization. This analysis helps qualify the balance-targeting rerandomization procedure endorsed by Morgan and Rubin (2012). Proposition 5 suggests specifying balance targets as a quantile of balance among all possible assignments. This makes the size of set $E_\dagger$ explicit and ensures that the balance objective picked by the experimenter does not cause significant losses in robustness. For instance, if the experimenter

framework, stratification on all dimensions is not possible, as individuals are unique. However, stratification is possible along a few discrete-valued dimensions of covariates $x \in \mathbb{R}^k$, for instance gender.

[21] Computing $p_{E_\dagger}$ may be difficult for complex balance criteria. In that case, Monte Carlo simulations can be used to compute an estimate of $p_{E_\dagger}$.

targets the 95^{th} percentile of balance, then $p_{E_\dagger} = 5\%$ and robustness losses due to rerandomization vanish as sample size N grows large.

5 Reference Dependence and Regret Aversion

The use of null-hypothesis statistical testing (NHST) is another common practice that standard models fail to predict. Recent work, however, shows that NHST can be rationalized if the decision-maker exhibits reference-dependence (Tetenov, 2012). We establish that reference-dependent preferences are compatible with our framework.

The version of reference-dependence we use nests regret aversion as a special case. It has the additional benefit of allowing us to dispense with Assumption 1. In particular, consider payoffs

$$w(p, a) = \Delta_p^a \times (1 + \kappa_a \mathbf{1}_{\Delta_p^a < 0}), \quad \text{with} \quad 0 < \kappa_0 \leq \kappa_1,$$

where $\Delta_p^a \equiv p^a - p^{1-a}$. These payoffs depend on both the policy choice a made by the decision-maker, and the performance difference Δ_p^a between this choice and the alternative.[22] Crucially, the parameters κ_0, κ_1—which only affect preferences when the decision-maker makes a mistake ($\Delta_p^a < 0$)—imply that the possibility of a success of size Δ_p^a cannot offset an equal possibility of a mistake of the same size.[23] When $\kappa_0 = \kappa_1$, regret about a mistake of a given magnitude is the same whether a equals 0 or 1. This is the standard formulation of regret aversion in the statistical decision theory literature. On the other hand, when $\kappa_0 < \kappa_1$, mistakes made when the decision-maker chooses $a = 1$ are given more weight than those made when the decision-maker chooses $a = 0$. Thus, $a = 1$ will only be chosen if the decision-maker is sufficiently certain that $p^1 > p^0$.

[22] Appendix C shows that these preferences rationalize NHST as an optimal decision rule.

[23] The decision-maker exhibits loss aversion, consistent with Prospect Theory (Kahneman and Tversky, 1979).

We consider the variant of (DP) in which the decision-maker seeks to maximize

$$\lambda \mathbb{E}_{h_0,\mathcal{E}}[w(p,\alpha)] + (1-\lambda) \min_{h \in H} \mathbb{E}_{h,\mathcal{E}}[w(p,\alpha)] \tag{DP$'$}$$

where $H = \Delta(P)$ is now the set of all possible priors over states $p \in P$.[24]

Note that we no longer impose Assumption 1 (limited extrapolation). The set of priors H is simply the set of all priors over P. As noted earlier, Assumption 1 is needed to rule out the doctrinaire prior in which $p^1 = p^0 = 0$. This would be the unconstrained worst-case prior chosen by nature under problem (DP), and no learning is possible at these priors. Nature does not select such a prior under regret aversion. Under the worst-case prior for problem (DP$'$), one policy must be better than the other. Otherwise the decision-maker cannot make a mistake, and there is nothing to regret. As a result, under regret aversion, information is valuable even under the unconstrained worst-case prior.

In this environment, analogs of our earlier propositions hold:

Proposition 6 (randomization)**.** *Consider a decision-maker solving Problem* (DP$'$)*:*

(i) Whenever $\lambda = 1$, running a deterministic experiment is weakly optimal.

(ii) For every h_0 and every $\lambda \in (0,1)$, as N becomes arbitrarily large, deterministic experiments remain bounded away from efficiency, and randomized experiments are strictly optimal.

Proposition 7 (rerandomization)**.** *There exists $M > 0$ such that for every prior $h \in H$ and $K \geq 1$,*

$$\mathbb{E}_{h,e\sim\mathcal{E}_K,a\sim\alpha_{\text{RCT}}}[w(p,a)] \geq \mathbb{E}_h\left[\max_{a\in\{0,1\}} w(p,a)\right] - M\sqrt{\frac{\ln(K+1)}{N}}.$$

[24] Alternatively, for any $\underline{p} < 1/2 < \overline{p}$, we could consider the set of priors $H = \Delta([\underline{p},\overline{p}]^{2X})$

6 Discussion

The ambition of this paper is to provide a decision-theoretic framework for experimental design that can clarify debates about experimental practice. Our primary requirement is that any such framework should rationalize revealed preferences for randomization. In this last section, we discuss alternative rationales for randomization and formulate suggestions for practical experiment design.

6.1 Alternative Rationales for Randomization

An alternative rationale for randomization is that experimenters are indifferent over experimental assignment, and may as well randomize. This indifference could be ascribed to a lack of opinion over the variation in treatment effects associated with different covariates. Alternatively, experimenters may have such opinions, but find the collection of covariate information too costly.

We believe that this rationale is not consistent with the preferences revealed by the behavior real-life experimenters. First, experimenters indicate strong preferences against deterministic assignments and in favor of randomized designs, including, but not limited to, standard RCTs. Ensuring proper randomization is often costly, and experimenters are concerned when they engage in any behavior that strays from this gold standard. This is especially true in economics, where experimenters often invest considerable time and energy trying to convince implementation partners to randomize. Second, experimenters are not indifferent about balance, and reveal mild preferences for rerandomized assignments over standard RCTs. Indeed, experimenters do obtain covariates, and frequently use them to determine experimental assignment by rerandomizing or stratifying. Concerns over balance typically vanish as the sample size grows large. Our model captures all these facts: deterministic assignments are unattractive even when the sample size is large; standard RCTs are a sizeable improvement, although they may be improved by enhancing balance; the value of

this last improvement vanishes as N gets large.

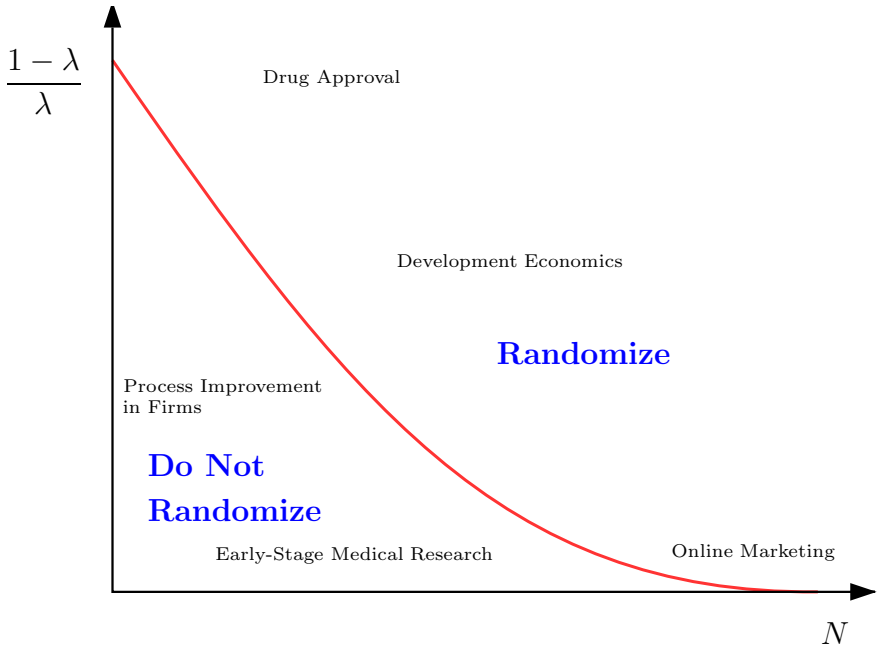
Another potential avenue to rationalize randomization is to view it as a delegation rule in a principal-agent setting: the audience wants data to be collected, but wants to limit the ways in which an interested experimenter could bias findings. This concern about moral hazard is not present in our framework: we believe that experimenters and their audience find randomization valuable even in the absence of incentive issues. In other words, we believe that a team of researchers and policy designers sharing the same objectives, but not necessarily the same beliefs, would find randomization a valuable tool. While we certainly believe that some experimenters are interested in obtaining specific outcomes—for example, pharmaceutical companies—it is still difficult to explain randomization without introducing heterogeneous priors. A Bayesian regulator would prefer mandating specific experimental assignments, for example, assignments maximizing balance.

6.2 Deterministic versus Randomized Experimentation

As shown in Section 3, our framework predicts either deterministic experimentation or randomized experimentation depending on the sample size N and the weight the decision-maker places on satisfying her audience λ. Here we describe how these findings relate to stylized facts about experimentation, as summarized in Figure 1. Throughout, we maintain Assumption 1 for a fixed parameter $\xi > 0$ (alternatively, we could assume the reference-dependent preferences of Section 5). Under this maintained assumption, changes in sample size N are accompanied with changes in the set of type X and priors H.

When sample points are scarce (N is small), or when the decision-maker does not put much weight on satisfying her audience (λ is close to 1), the optimal experiment is deterministic, driven by prior h_0. The experimenter assigns treatment and control to maximize the decision-making value of the data collected. For example, in the context of process improvement, a firm testing out new machinery may assign treatment to its best performing teams

Figure 1: When to randomize? Theory matches broad patterns.



so that the evaluation is not muddled by operational errors. When refining a new product, firms test their prototypes with selected alpha users.

Similarly, early-stage medical research often does not feature randomization. This occurs in Phase I trials, which examine the toxicity of a new compound. These trials recruit healthy volunteers, which allows adverse reactions to be cleanly ascribed to treatment rather than to any underlying disease condition. Sample sizes in Phase I trials also tend to be relatively small. Deterministic experimentation also occurs when studying patients who have conditions known to rapidly result in severe disability or death.[25] Any improvement can be reliably assigned to the treatment. In our framework, the reasoning underlying these two cases is the same: relatively little disagreement between the experimenter and the adversarial audience leads to deterministic designs maximizing the experimenter's subjective expected utility.

When the decision-maker cares sufficiently about satisfying her adversarial audience, or when she has a sufficiently large sample, she will randomize. The former is the case in

[25]For recent examples, see Harmon (2016) and Yuhas (2016).

scientific research, especially when the experiment has significant policy implications—for example, in development economics. The latter is the case for firms doing A/B testing online. Although the firm only needs to convince itself of the effectiveness of a particular ad, there are so many observations available that randomization is used to effectively address internal concerns about robustness. This is also the case for later-stage medical research seeking regulatory approval: government regulators, doctors, patients, and investors form the adversarial audience for pharmaceutical and medical device companies.

6.3 The Practice of Rerandomization

We now discuss how our results may be used to inform the practice of rerandomization. We emphasize that our results apply to environments where the set of covariates is known at the time of treatment assignment. We extend our results to the case of sequential treatment assignment in Appendix A.

Proposition 2 and Corollary 1 establish that: 1) randomization is a key aspect of any compromise between subjective expected utility and robust payoff guarantees; and 2) the standard RCT is an acceptable compromise, providing approximately optimal performance for each objective. We interpret Propositions 3 and 4 as expanding the range of such compromises. We emphasize two insights from our analysis that seem most useful for practice: first, how to set balance targets so that they do not inadvertently cause severe losses in robustness, and second, how to use rerandomization to allow stakeholders to express preferences.

Setting balance targets properly. Our analysis identifies the following issue with the practice of rerandomization: excessively ambitious absolute balance objectives may lead to severe losses in robustness. Propositions 4 and 5 suggest two ways to address the issue by changing the metric used to set balance targets.

The first approach is to use K-rerandomization, with K less than sample size N. This ensures losses on the order of $\sqrt{\frac{\ln(N)}{N}}$. A rule of thumb that strikes us as a (subjectively)

attractive compromise is to set $K \leq \min\{100, N\}$ (Banerjee et al., 2017). Indeed, note that with probability $1 - 0.95^K$ the K-rerandomized assignment is in the 95^{th} quantile of most balanced assignments. For any $K \geq 100$ this event has probability greater than 99%.

The second approach, described in Section 4.3, is to set a quantile of balance as an objective, for instance the 95^{th} quantile of balance among all assignments $e \in E$, and rerandomize until that target is achieved. For any such fixed quantile, Proposition 5 ensures that losses against the first-best vanish as sample size N grows large. Concretely, an approximate way to generate such an assignment would be to draw a large set of independent, exchangeable assignments, and choose one with uniform probability among the 5% most balanced.

In both cases, the idea is to avoid setting absolute balance goals, so that rerandomization does not inadvertently lead to an extremely reduced set of possible assignments. Expressing rerandomization objectives using either a bound on the number of rerandomizations, or by setting a quantile of balance, clarifies how "selected" possible assignments are. While the optimal degree of rerandomization is ultimately subjective, we think these guidelines—setting $K = \min\{N, 100\}$ or targeting the 95^{th} quantile of balance—offer a reasonable compromise between subjective performance and robust performance guarantees.

Expressing stakeholder preferences. As noted in the text surrounding Remark 1 and Proposition 4, these results do not rely on the fact that selected experimental assignment e_K^* maximizes a specific subjective performance objective. They hold regardless of the way the final assignment e_K^* is selected, provided it is selected among K uniformly-drawn random assignments.

This means that one can use a K-rerandomized design as a way to let stakeholders and implementation partners express preferences, albeit in a constrained way. Regulators, funders, or the communities from which experimental participants are drawn often have equity concerns and distributional preferences. They may care about targeting treatment to those they believe will benefit the most, or may simply dislike the lack of control inherent

to randomization and wish to exert some control over assignment. The ability to at least *partially* accommodate the preferences of stakeholders, by using their preferences to select an assignment among K options, may help build goodwill and ensure cooperation.

Concretely, such a protocol would take the following form: (i) let stakeholders express preferences over assignments; (ii) draw K random assignments; (iii) pick the assignment e_K^* from that group of K randomizations that maximizes the preferences of stakeholders; (iv) run the experiment ; (v) take policy decision $a = \mathbf{1}_{\overline{y}_1 \geq \overline{y}_0}$.

We note that in this description, the preferences of stakeholders are expressed ex ante, although Proposition 4 would continues to hold even if stakeholders picked the assignment ex post without pre-specifying their preferences. Stating preferences ex ante has two benefits. First, formally specifying the preferences of stakeholders permits randomization inference tests (Fisher, 1935; Morgan and Rubin, 2012; Young, 2016). That is, given the process for assignment, a statistician can simulate the distribution of treatment effects that would be observed if $p_x^1 = p_x^0$, for all $x \in X$. This procedure can be used to calculate exact p-values, and infer standard errors. Second, randomization helps parties with differing priors agree on a process ex ante, but not ex post: if audience members start looking into the details of the realized assignment, someone may well find issue with it.

6.4 Future Directions

We believe that debates over the proper way to do empirical research are an opportunity to both improve, and put to gainful use, economists' models of decision-making. Other possible uses of our framework include understanding subgroup analysis and multiple-hypothesis testing. We believe these factors can be introduced in our framework by considering more complex policies that tailor treatment to the covariates of participants. Pre-analysis plans—that is, the practice of pre-specifying which statistics of the data will be reported—also form a challenge to traditional models of experimentation. This practice is difficult to rationalize

using a Bayesian framework as long as all collected data is made available. A related, but more fundamental, difficulty is explaining the reliance on low-dimensional statistics, rather than just publishing data. We believe that models of rational inattention provide an attractive path forward.

Appendix

A Sequential Treatment Assignment

In many settings, subjects arrive over time and treatment assignment has to be made sequentially. This is particularly true in settings where recruitment takes time (for example, medical trials), or in settings where the cost of interacting with subjects is large (for example, economic development experiments occurring in hard to reach rural areas). In such settings, it may not be practical to first learn the set of covariates X in the full sample and assign treatments afterwards. In this appendix we show that much of our analysis holds in such environments. We propose a variant of K-rerandomization appropriate for sequential assignment.

Sequential treatment assignment has received attention from statisticians. For example, Efron (1971) advocates biased coin designs to ensure that treatment and control are sampled at similar rates, without generating significant bias in treatment assignment. More recently, Atkinson (1982, 2002) extends the approach to the design of sequentially balanced experiments when participants have covariates. Guiteras et al. (2016) provides simulations and details from field implementations highlighting the value of such designs in practice. To our knowledge, this literature does not address the impact of such designs on robust decision-making.

Framework. N experimental participants are labeled by their arrival period i. Each participant $i \in \{1, \cdots, N\}$ exhibits an observable covariate $d_i \in D$, where D is finite. We denote by $x_i = (i, d_i)$ covariates augmented with the time of arrival of participants. In principle, outcomes $y_i \in \{0, 1\}$ may depend on the time of arrival. Let $X = \{x_i\}_{i \in \{1, \cdots, N\}}$ denote the final set of covariates. Note that X is a random variable. We denote by $\mathcal{X} = \text{supp } X$ the set of potential sets of covariates X. A state of the world is a pair (X, p), where p takes the form $(p_x^\tau)_{\tau \in \{0,1\},\, x \in X}$. A prior h now denotes a prior over the pair (X, p). As in Section 2, the payoff of a decision-maker taking decision a at state (X, p) is $u(X, p, a) = \frac{1}{|X|} \sum_{x \in X} p_x^a$.

We continue to denote by $\tau_i \in \{0, 1\}$ the treatment assignment of the i^{th} participant. A design history s_n in period $n \leq N$ is a tuple $(x_i, \tau_i, x_n)_{i \in \{1, \cdots, n-1\}}$. An experimental assignment function $e \in E$ is a mapping from design histories s_n to treatment assignments $\tau_n \in \{0, 1\}$. A randomized experiment $\mathcal{E}$ is a distribution over experimental designs $e \in E$. The final realized experimental assignment is described by history $s_{N+1} = (x_i, \tau_i)_{i \in \{1, \cdots, N\}}$.

A policy rule α is a mapping from the final experiment design s_{N+1} and realized outcome data $y = (y_1, \cdots, y_N)$ to policy choices $\Delta(\{0, 1\})$. The decision-maker evaluates choices over experiment designs $\mathcal{E}$ and policy rules α according to criterion

$$\lambda \mathbb{E}_{h_0, \mathcal{E}, \alpha}[u(X, p, a)] + (1 - \lambda) \min_{h \in H} \mathbb{E}_{h, \mathcal{E}, \alpha}[u(X, p, a)], \tag{DP†}$$

where h_0 is a prior over (X, p), and H is a set of priors h over (X, p). For simplicity, we assume that H takes the form $H = \{\delta_X \times h_p \,|\, X \in \mathcal{X}, h_p \in H_p\}$, where δ_X is the Dirac mass at X, and $H_p \subset \Delta\left([0, 1]^{2N}\right)$ is a set of priors over p satisfying Assumption 1 for some fixed $\xi > 0$.

The benchmark sequential RCT design $\mathcal{E}_{\text{sRCT}}$ takes the following form:

- For each history s_n, $\mathcal{E}_{\text{sRCT}}$ assigns treatment $\tau = 1$ with probability 0.5.
- Policy a is chosen according to the empirical success rule: $\alpha_{\text{RCT}}(e, y) \equiv \mathbf{1}_{\overline{y}^1 \geq \overline{y}^0}$ defined

in Section 2.

Benchmark results. The following results hold:

(i) If $\lambda = 1$, there exists a deterministic experiment design e that maximizes objective (DP†).

The decision-maker's value at any history s_n satisfies the Bellman equation

$$V(s_n) = \max_{\tau \in \{0,1\}} \mathbb{E}_{h_0} V(s_{n+1})$$
$$\text{with } V(s_{N+1}) = \mathbb{E}_{h_0, s_{N+1}} \left[\max_{a \in \{0,1\}} \mathbb{E}_{h_0}[u(X, p, a) | s_{N+1}, y] \right]$$

At any history s_n the decision-maker picks τ_n solving the first equation above. As she is an expected utility maximizer, there is no advantage to random treatment choice. A deterministic contingent plan is optimal.

(ii) If $\lambda > 0$ and Assumption 1 holds for $\xi > 0$ fixed, then any deterministic experiment is bounded away from the first-best by a term greater than $(1 - \lambda)\xi$.

The result follows from the fact deterministic sequential experiment designs form a subset of the deterministic experiment designs studied in Section 2. As a result, payoffs under deterministic sequential experiment designs are bounded above by payoffs under deterministic experiments.

(iii) Sequential trial design $\mathcal{E}_{\text{sRCT}}$ yields payoffs within $O\left(\sqrt{\frac{1}{N}}\right)$ of first-best. The proof is identical to that of Corollary 1.

Sequential K-rerandomization. It is possible to provide a sequential version of the K-rerandomization process described Section 4. The design, denoted by $\mathcal{E}_{sK}$, takes as given a value function W defined over states s_n and period n assignments τ_n. We provide a few possibilities below. Assignments are defined as follows:

1. Draw K sequential assignment mappings $(e_1, \cdots, e_K) \in E \times \cdots \times E$ according to K sequential RCT designs $(\mathcal{E}^1_{\text{sRCT}}, \cdots, \mathcal{E}^K_{\text{sRCT}})$. Note that each individual design $\mathcal{E}^k_{\text{sRCT}}$ assigns treatment independently across covariates, but the realizations of the assignment mappings themselves may be correlated across values $k \in \{1, \cdots, K\}$. We provide examples below (in particular, see Figure A.1).

2. For each history s_n, we define the set of feasible treatments $T(s_n)$ at s_n as

$$T(s_n) = \{e(x_n) | e \in \{e_1, \cdots, e_K\} \; s.t. \; \forall n' < n, \; e(x_{n'}) = \tau_{n'}\}.$$

   $T(s_n)$ is the set of assignments $e_1(x_n), \cdots, e_K(x_n)$, for mappings e_k whose assignments for past covariates $x_{n'}$ are consistent with realized past assignments τ'_n (with $n' < n$).

   Treatment τ_n is chosen from $\text{argmax}_{\tau \in T(s_n)} W(s_n, \tau)$. Indifferences are resolved with a uniform draw.

3. Policy rule α is the empirical success rule $\alpha_{\text{RCT}} \equiv \mathbf{1}_{\overline{y}^1 \geq \overline{y}^0}$.

The proof of Proposition 4 implies that this experiment design guarantees payoffs close to the first-best, up to losses of order $\sqrt{\frac{\ln K}{N}}$. Indeed, the proof of Proposition 4 only relies on the fact that each assignment $\mathcal{E}^k_{\text{sRCT}}$ for $k \in \{1, \cdots, K\}$ provides robust payoff guarantees. Proposition 4 holds regardless of the correlation between assignment mappings.

There are several ways to define value function W. Given a prior μ_0 over X, and a balance function over final assignments $B(s_{N+1})$, one can define W via the Bellman equation

$$\begin{aligned} V(s_{N+1}) &= B(s_{N+1}) \\ V(s_n) &= \max_{\tau_n \in T(s_n)} \mathbb{E}_{\mu_0}[V(s_{n+1}) | s_n, \tau_n] \\ W(s_n, \tau_n) &= \mathbb{E}_{\mu_0}[V(s_{n+1}) | s_n, \tau_n] \end{aligned}$$

In practice, it may be convenient to use the sample empirical prior estimated over past

covariates to predict future covariates at each history. Concretely, for any n, $\widehat{\mu}_n$ is the belief over future covariates $((n+1, d_{n+1}), \cdots, (N, d_N))$ obtained by resampling covariates $(d_1, \cdots, d_n)$ with replacement. At every n we define V_n by

$$V_n(s_{N+1}) = B(s_{N+1})$$
$$\forall n' \geq n, V_n(s_{n'}) = \max_{\tau_{n'} \in T(s_{n'})} \mathbb{E}_{\widehat{\mu}_n}[V_n(s_{n'+1})|s_{n'}, \tau_{n'}]$$

and define $W(s_n, \tau) = \mathbb{E}_{\widehat{\mu}_n}[V_n(s_{n+1})|s_n, \tau]$.

The reason that one may want to correlate assignments $e_1, \cdots, e_K$ is to optimally delay selecting the ultimate assignment. For instance, if $K = 2$, it may be beneficial to draw two assignments e_1 and e_2 that are identical up to participant $N/2$, and independent after participant $N/2$. That is, for all $d \in D$, $e_1((i,d)) = e_2((i,d))$ if $i \leq N/2$, and $e_1((i,d))$ independent of $e_2((i,d))$ if $i > N/2$. This allows the experimenter to first learn the distribution of covariates z in the population before picking the continuation assignment that maximizes expected balance.

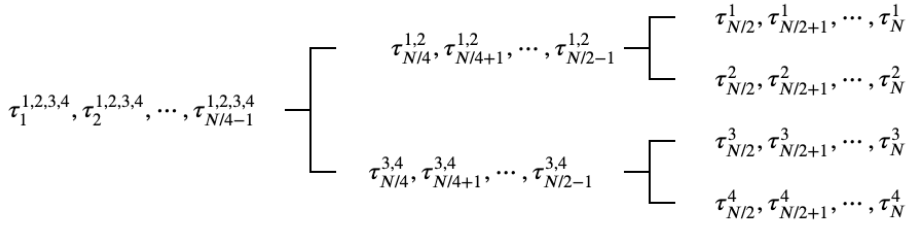


Figure A.1: Correlated assignments

More generally, if the number of randomizations is $K = C^G$, with C and G two integers, one could select the final assignment e in G steps, with choices made at times $n_1, \cdots, n_G$. At each period n_g, the decision-maker chooses between C independently drawn assignment mappings for covariate realized between time n_g and $n_{g+1} - 1$. In that case, action set $T(s_n)$ is a singleton for all periods $n \notin \{n_1, \cdots, n_G\}$. Figure A.1 illustrates such correlated assignments in a setting with $K = 4$, $G = 2$, $C = 2$, $n_1 = N/4$, and $n_2 = N/2$.

References

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien, "Optimal Learning by Experimentation," *The Review of Economic Studies*, 1991, *58* (4), 621–654.

Athey, Susan and Guido Imbens, "The Econometrics of Randomized Experiments," in Esther Duflo and Abhijit Banerjee, eds., *Handbook of Field Experiments*, Vol. 1, Elsevier, 2017, pp. 73–140.

Atkinson, Anthony C, "Optimum biased coin designs for sequential clinical trials with prognostic factors," *Biometrika*, 1982, *69* (1), 61–67.

_ , "The comparison of designs for sequential clinical trials with covariate information," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2002, *165* (2), 349–373.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel, "Worms at work: Long-run impacts of a child health investment," *The quarterly journal of economics*, 2016, *131* (4), 1637–1680.

Banerjee, Abhijit, Sylvain Chassang, and Erik Snowberg, "Decision Theoretic Approaches to Experiment Design and External Validity," in Esther Duflo and Abhijit Banerjee, eds., *Handbook of Field Experiments*, Vol. 1, Elsevier, 2017, pp. 141–174.

Benjamin, Daniel, James Berger, Magnus Johannesson, Brian Nosek, Eric-Jan Wagenmakers, Richard Berk, Kenneth Bollen, Björn Brembs, Lawrence Brown, and Colin Camerer, et al., "Redefine Statistical Significance," *Nature Human Behaviour*, 2018, *2* (1), 6–10.

Bergemann, Dirk and Juuso Välimäki, "Learning and Strategic Pricing," *Econometrica*, September 1996, *64* (5), 1125–1149.

_ **and** _ , "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 2002, *70* (3), 1007–1033.

_ **and** _ , "Bandit Problems," 2006. Cowles Foundation paper no. 1292.

Blackwell, David, "Comparison of Experiments," in "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability," Vol. 1 1951, pp. 93–102.

_ , "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 1953, *24* (2), 265–272.

Bruhn, Miriam and David McKenzie, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 200–232.

Bugni, Federico A., Ivan A. Canay, and Azeem M. Shaikh, "Inference under Covariate-Adaptive Randomization," *Journal of the American Statistical Association*, 2018, *113* (524), 1784–1796.

Cochrane, William G. and Donald B. Rubin, "Controlling Bias in Observational Studies," *Sankhyā: The Indian Journal of Statistics, Series A*, 1973, *35* (4), 417–446.

Deaton, Angus, "Instruments, Randomization, and Learning about Development," *Journal of Economic Literature*, June 2010, *48* (2), 424–455.

Duflo, Esther, Rachel Glennerster, and Michael Kremer, "Using Randomization in Development Economics Research: A Tool Kit," in T. Paul Schultz and John Strauss, eds., *Handbook of Development Economics, Vol. 4*, Amsterdam: Elsevier, 2008, pp. 3895–3962.

Efron, Bradley, "Forcing a sequential experiment to be balanced," *Biometrika*, 1971, *58* (3), 403–417.

Fisher, Ronald Aylmer, *The Design of Experiments*, Edinburgh and London: Oliver & Boyd, 1935.

Gilboa, Itzhak and David Schmeidler, "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics*, 1989, *18* (2), 141–153.

Green, Donald P and Andrej Tusicisny, "Statistical Analysis of Results from Laboratory Studies in Experimental Economics: A Critique of Current Practice," 2012. Columbia University, *mimeo.*

Grossman, Sanford J and Joseph E Stiglitz, "On the Impossibility of Informationally Efficient Markets," *The American Economic Review*, June 1980, *70* (3), 393–408.

Guiteras, Raymond P, David I Levine, and Thomas H Polley, "The pursuit of balance in sequential randomized trials," *Development Engineering*, 2016, *1*, 12–25.

Hansen, Lars Peter and Thomas J. Sargent, "Robust Control and Model Uncertainty," *American Economic Review*, 2001, *91* (2), 60–66.

Harmon, Amy, "After Long Fight, Drug Gives Sudden Reprieve," *The New York Times*, Februrary 22, 2016. Published online at: http://www.nytimes.com/2010/02/23/health/research/23trial.html.

Hoeffding, W., "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association*, 1963, *58* (301), 13–30.

Huber, Peter J., "Robust Estimation of a Location Parameter," *The Annals of Mathematical Statistics*, 1964, *35* (1), 73–101.

Huber, Peter J, "Robust Statistics," in Miodrag Lovric, ed., *International Encyclopedia of Statistical Science*, Berlin: Springer, 2010, pp. 1248–1251.

Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter Van der Windt, "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration," *Political Analysis*, 2013, *21* (1), 1–20.

Imbens, Guido W, "Better LATE than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature*, June 2010, *48* (2), 399–423.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, March 1979, *47* (2), 263–292.

Kasy, Maximilian, "Why Experimenters Might not Always Want to Randomize, and what they Could Do Instead," *Political Analysis*, 2016, *24* (3), 324–338.

Kitagawa, Toru and Aleksey Tetenov, "Who should be Treated? Empirical Welfare Maximization Methods for Treatment Choice," *Econometrica*, 2018, *86* (2), 591–616.

Manski, Charles F., "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 2004, *72* (4), 1221–1246.

_ , "The 2009 Lawrence R. Klein Lecture: Diversified Treatment under Ambiguity," *International Economic Review*, 2009, *50* (4), 1013–1041.

Marinacci, Massimo, "Model Uncertainty," *Journal of the European Economic Association*, 2015, *13* (6), 1022–1100.

Miguel, Edward and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, January 2004, *72* (1), 159–217.

Morgan, Kari Lock and Donald B. Rubin, "Rerandomization to Improve Covariate Balance in Experiments," *The Annals of Statistics*, 2012, *40* (2), 1263–1282.

Neyman, Jerzy, "On the Application of Probability theory to Agricultural Experiments. Essay on principles. Section 9. Translated by Dorota Dabrowska, and Terence Speed [1990]," *Statistical Science*, 1923, pp. 465–472.

Olken, Benjamin A., "Promises and Perils of Pre-Analysis Plans," *The Journal of Economic Perspectives*, 2015, *29* (3), 61–80.

Persico, Nicola, "Information Acquisition in Auctions," *Econometrica*, 2000, *68* (1), 135–148.

Rothschild, Michael, "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 1974, *9* (2), 185–202.

Rubin, Donald B, "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 1974, *66* (5), 688–701.

Rubin, Donald B., "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies," *Journal of the American Statistical Association*, 1979, *74* (366a), 318–328.

_ , "Bias Reduction using Mahalanobis-metric Matching," *Biometrics*, 1980, *36* (2), 293–298.

Saito, Kota, "Preferences for Flexibility and Randomization under Uncertainty," *The American Economic Review*, 2015, *105* (3), 1246–1271.

Savage, Leonard J., *The Foundations of Statistics*, Courier Corporation, 1954.

Tetenov, Aleksey, "Statistical Treatment Choice Based on Asymmetric Minimax Regret Criteria," *Journal of Econometrics*, 2012, *166* (1), 157–165.

Wald, Abraham, *Statistical Decision Functions*, New York: Wiley, 1950.

Weerahandi, Samaradasa and James V. Zidek, "Multi-Bayesian Statistical Decision Theory," *Journal of the Royal Statistical Society. Series A (General)*, 1981, *144* (1), 85–93.

_ **and** _ , "Elements of Multi-Bayesian Decision Theory," *The Annals of Statistics*, 1983, *11* (4), 1032–1046.

Young, Alwyn, "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," 2016. *mimeo.*, London School of Economics.

Yuhas, Alan, "Cancer Researchers Claim 'Extraordinary Results' using T-cell Therapy," *The Guardian*, February 15, 2016. Published online at: https://www.theguardian.com/science/2016/feb/15/cancer-extraordinary-results-t-cell-therapy-research-clinical-trials.

Online Appendix—Not Intended for Publication

B Balance as Subjective Expected Utility

In this appendix we clarify that under an appropriate prior, maximizing subjective performance coincides with a traditional balance objective: minimizing the Malahanobis distance—defined by sample covariates $(x_i)_{i\in\{1,\dots,N\}}$ with $x_i \in \mathbb{R}^k$—between mean covariates across the treated and control groups (Rubin, 1980; Cochrane and Rubin, 1973; Rubin, 1979; Morgan and Rubin, 2012). To simplify the analysis, we allow for Gaussian priors, although they do not satisfy the framework of Section 2.[1]

Prior h_0 is generated as follow. Assume the decision-maker believes outcomes y_i^τ are distributed according to the linear model

$$y_i^\tau = \tau_i \Delta + b^\intercal z_i + \varepsilon_i \quad \text{with} \quad \Delta \sim F_\Delta\,, \quad b \sim \mathcal{N}(0, I_k)$$

in which zs are the underlying determinants of outcomes ys, I_k is the k-dimensional identity matrix, and $(\varepsilon_i)_{i\in\{1,\dots,N\}}$ are independent, mean-zero error terms. Although terms z_i are unobservable, they are assumed to be a linear transformation of observables x_i, so that $x_i = Mz_i$, with M invertible. This implies that

$$y_i^\tau = \tau_i \Delta + \beta^\intercal x_i + \varepsilon_i\,, \quad \text{with} \quad \beta \sim \mathcal{N}(0, \mathsf{cov}(\mathbf{x})^{-1}).$$

Given a treatment assignment $(\tau_i)_{i\in\{1,\dots,N\}}$, let $\overline{x}^\tau \equiv \frac{2}{N}\sum_{i=1}^N x_i \mathbf{1}_{\tau_i=\tau} \in \mathbb{R}^k$ and $\phi \equiv \frac{2}{N}\sum_{i=1}^N (-1)^{1-\tau_i}\varepsilon_i$. We make the (asymptotically correct) assumption that ϕ is normally distributed with variance σ_ϕ^2. Under the empirical success rule, the subjective expected utility of the decision-

[1] Our analysis extends to such environments provided the mean and variance of possible outcome distributions are bounded.

maker under prior h_0 is

$$\mathbb{E}\beta^\intercal \overline{x}^0 + \mathbb{E}_\Delta[\Delta \times \text{Prob}(\Delta + \beta^\intercal(\overline{x}^1 - \overline{x}^0) + \phi \geq 0)].$$

This expression is decreasing in the variance of $\beta^\intercal(\overline{x}^1 - \overline{x}^0)$:

$$\mathsf{var}(\beta^\intercal(\overline{x}^1 - \overline{x}^0)) = (\overline{x}^1 - \overline{x}^0)^\top \mathsf{cov}(\mathbf{x})^{-1}(\overline{x}^1 - \overline{x}^0),$$

which is precisely the Mahalanobis distance between the mean of covariates in the treatment and control groups. Under this prior, the assignment $(\tau_i)_{i\in\{1,\dots N\}}$ that maximizes subjective expected utility minimizes the Mahalanobis distance between covariate means across the treatment and control groups.

C Null-Hypothesis Statistical Testing

Decision problem (DP) does not rationalize null-hypothesis statistical testing (NHST, using t- or z-statistics favoring implementation of the null treatment $a = 0$), a mainstay of experimental practice. In that decision problem, the raw treatment effect—that is, the difference in average outcomes—is sufficient for near-optimal decision-making. This appendix clarifies that other standard preferences (including risk aversion over treatment effects) do not rationalize NHST, while the reference-dependent preferences introduced in Section 5 do.

Ambiguity aversion does not play a role in this argument, so we consider a decision-maker with independent Gaussian posteriors $\mathcal{N}(\hat{p}^a, \frac{\sigma_a^2}{N})$ over the mean outcomes p^a of actions $a \in \{0, 1\}$.[2] A risk-neutral Bayesian decision-maker solving $\max_{a\in\{0,1\}} \mathbb{E}[p^a]$, expectations being taken under the posterior $\mathcal{N}(\hat{p}^a, \frac{\sigma_a^2}{N})$, will take action $a = 1$ if and only if $\hat{p}^1 - \hat{p}^0 > 0$.

[2]Parameters $\hat{p}^a$ and σ_a^2/N could be derived from a standard Gaussian learning model with diffuse priors. Under such a model $\hat{p}_a$ would be equal to the sample average of outcomes y following treatment $\tau = a$.

However, the t-statistic for a given treatment effect $\hat{p}^1 - \hat{p}^0$ is given by

$$t \equiv \sqrt{N} \frac{\hat{p}^1 - \hat{p}^0}{\sqrt{\sigma_0^2 + \sigma_1^2}}.$$

Thus, decision rules that choose $a = 1$ if and only if $t > \underline{t} > 0$ (where $\underline{t}$ is a fixed-threshold) are suboptimal. Indeed, since $\bar{t} > 0$, for any given estimated treatment effect $\hat{p}^1 - \hat{p}^0$, there always exists σ_0 large enough such that $t < \bar{t}$. As a result the decision-maker sticks with $a = 0$ regardless of the estimated treatment effect.

Risk aversion over policy outcomes. A natural hypothesis is that risk aversion may drive the reliance on hypothesis testing using t-statistics. However, this is not the case. To show this, we assume (w.l.o.g.) that $\sigma_0 < \sigma_1$, and consider a decision-maker who wants to solve $\max_{a \in \{0,1\}} \mathbb{E}\left[\Gamma(p^a)\right]$ where Γ is quadratic and concave. As $\mathbb{E}\left[\Gamma(p^a)\right] = \Gamma(\hat{p}^a) + \frac{1}{2}\Gamma'' \frac{\sigma_a^2}{N}$ it follows that

$$\mathbb{E}\left[\Gamma(p^1)\right] \geq \mathbb{E}\left[\Gamma(p^0)\right] \iff \frac{2N}{-\Gamma''} \frac{\Gamma(\hat{p}^1) - \Gamma(\hat{p}^0)}{\sigma_1^2 - \sigma_0^2} = \gamma \frac{\hat{p}^1 - \hat{p}^0}{\sigma_1^2 - \sigma_0^2} > 1$$

with $\gamma = \frac{2N\Gamma'(\tilde{p})}{-\Gamma''}$ for some value $\tilde{p} \in [\hat{p}^0, \hat{p}^1]$.

This differs significantly from a t-statistic: mean treatment effect $\hat{p}^1 - \hat{p}^0$ is scaled by the difference of variances, rather than the sum of standard deviations. In particular, risk aversion means that the decision-maker values certainty (a small variance in outcomes) as well as a higher mean outcome. Greater standard deviation σ_0 makes action $a = 0$ less attractive, not more.

Reference-dependent preferences. The asymmetric treatment of the null and alternative hypotheses suggests that one must resort to reference-dependent preferences to motivate NHST using t-statistics (see Tetenov, 2012). As in Section 5, consider a decision-maker who

seeks to solve

$$\max_{a\in\{0,1\}} \mathbb{E}[w(p,a)], \tag{C.1}$$

where $w(p,a) \equiv \Delta_p^a \times (1+\kappa_a \mathbf{1}_{\Delta_p^a<0})$ with $\Delta_p^a \equiv p^a - p^{1-a}$ and $0<\kappa_0\leq\kappa_1$.

Lemma C.1. *Consider a reference-dependent agent solving* (C.1)*, with* $\kappa_0<\kappa_1$*. The optimal-decision rule takes the form* $t>t^*$*, with* $t^*>0$*.*

Proof of Lemma C.1: Let $\bar{t} \equiv \sqrt{N}\frac{p^1-p^0}{\sqrt{\sigma_0^2+\sigma_1^2}}$. As $p^1-p^0 \sim \mathcal{N}\left(\hat{p}_1-\hat{p}_0, \frac{\sigma_0^2+\sigma_1^2}{N}\right)$, it follows that conditional on observing a t-statistic t, $\bar{t}\sim\mathcal{N}(t,1)$. Note that w is positively homogeneous of degree 1. Conditional on realized data, the decision-maker chooses $a=1$ if and only if:

$$\begin{aligned}
\mathbb{E}\left[w(\Delta_p^1,1)-w(\Delta_p^0,0)\right]>0 &\iff \mathbb{E}_{\bar{t}}\left[w\left(\bar{t}\sqrt{\frac{\sigma_0^2+\sigma_1^2}{N}},1\right)-w\left(-\bar{t}\sqrt{\frac{\sigma_0^2+\sigma_1^2}{N}},0\right)\middle|\, t\right]>0\\
&\iff \mathbb{E}_{\bar{t}}\left[w\big(\bar{t},1\big)-w\big(-\bar{t},0\big)\big|\, t\right]>0\\
&\Rightarrow t>t^*
\end{aligned}$$

for some value of t^*. Note that $w\big(\bar{t},1\big)-w\big(-\bar{t},0\big) = (2+\kappa_0)\bar{t}+(\kappa_1-\kappa_0)\bar{t}\mathbf{1}_{\bar{t}<0}$. As $\kappa_0<\kappa_1$ it follows that $w\big(\bar{t},1\big)-w\big(-\bar{t},0\big)$ is increasing and concave in $\bar{t}$, and strictly so around 0. As $\bar{t}$ has expectation equal to zero conditional on $t=0$, this implies that $\mathbb{E}_{\bar{t}}\left[w\big(\bar{t},1\big)-w\big(-\bar{t},0\big)\big|\, t=0\right] < w(0,1)-w(0,0)=0$, so that $t^*>0$. ■

D Proofs

Proof of Proposition 1: We begin by showing that deterministic experiments are always weakly optimal for a Bayesian decision-maker. The decision-maker's indirect utility from

running experiment $\mathcal{E}$ can be written as

$$\max_{\alpha \in \mathcal{A}} \mathbb{E}_{h_0,\mathcal{E}} \left[u(p, \alpha(e, y)) \right] = \sum_{e \in E} \mathcal{E}(e) v(h_0, e),$$

where $v(h_0, e)$ is the indirect utility from decision-making given realized experiment e:

$$v(h_0, e) \equiv \sum_{y \in \mathcal{Y}} \text{Prob}_{h_0,e}(y) \max_{a \in \{0,1\}} \mathbb{E}_{h_0,e} \left[u(p, a) | y \right]. \tag{D.1}$$

Any deterministic experiment e^* solving $\max_{e \in E} v(h_0, e)$ is optimal. More strongly, $\mathcal{E}$ solves (DP) if and only if supp $\mathcal{E} \subset \underset{e \in E}{\text{argmax}}\, v(h_0, e)$.

We now prove that deterministic experiments are generically strictly optimal when all data-points are valuable. We first consider the case where $\lambda = 1$. It is straightforward to show that the set of priors for which there exists a uniquely optimal deterministic experiment is open. Suppose e is uniquely optimal under h_0. As E is finite, there exists $\eta > 0$ such that $v(h_0, e) > v(h_0, e') + \eta$ for all $e' \neq e$. As v is continuous in h_0, there exists a neighborhood N_0 of h_0 such that $v(h, e) > v(h, e') + \eta/2$ for all $h \in N_0$ and $e' \neq e$. Hence, the set of priors for which there exists a uniquely optimal deterministic experiment is open.

We now show that the set of priors for which there exists a uniquely optimal experiment is dense in the space of priors for which all data points are valuable. The proof is by induction on the number of optimal experiments in $\text{argmax}_{e \in E} v(h_0, e)$. Fix a neighborhood $\mathcal{N}_0$ of h_0 such that all data-points are valuable under priors $h \in \mathcal{N}_0$. We show that, if there are n such optimal experiments, then there exists a prior $h \in \mathcal{N}_0$ such that there are at most $n-1$ optimal experiments in $\text{argmax}_{e \in E} v(h, e)$. The proof consists of two main steps. First, we establish that the following simplifying assumptions are without loss of generality:

- we can pick $\mathcal{N}_0$ such that $\text{argmax}_{e \in E} v(h, e) \subset \text{argmax}_{e \in E} v(h_0, e)$ for all priors $h \in \mathcal{N}_0$;

- we can assume that for any experiment e, a fixed policy rule $\alpha(y, e)$ is uniquely optimal for all priors $h \in \mathcal{N}_0$.

Second, given two experiments e, e' that are optimal at prior h_0, we exploit the fact that one pair (x, τ) must be sampled by e and not by e' to construct a family of priors in $\mathcal{N}_0$ that garble the information provided by pair (x, τ). Such priors change the value of experiment e, but not the value of experiment e', establishing the induction step.

Step 1: simplifications. The fact that we can pick $\mathcal{N}_0$ such that $\operatorname{argmax}_{e \in E} v(h, e) \subset \operatorname{argmax}_{e \in E} v(h_0, e)$ for all priors $h \in \mathcal{N}_0$ follows from the fact that E is finite, and $v(h, e)$ is continuous in h under the statistical distance.

The decision-maker's indirect utility from running experiment e can be rewritten as

$$v(h_0, e) = \mathbb{E}_{h_0}\left[p^0 + \alpha^*_{h_0}(e, y)(p^1 - p^0)\middle| e\right],$$

where $\alpha^*_{h_0} \in \mathcal{A}$ denotes an optimal policy rule under h_0. Suppose $e \neq e'$ are both optimal under h_0. As $\mathcal{Y}$ is finite, by breaking indifferences in favor of one policy (say $a = 1$), one can find $h_1 \in \mathcal{N}_0$ and a neighborhood $\mathcal{N}_1 \subset \mathcal{N}_0$ of h_1 such that for all $h \in \mathcal{N}_1$ the optimal policies $\alpha^*_h(e, y)$ and $\alpha^*_h(e', y)$ are unique and respectively equal to $\alpha^*_{h_1}(e, y)$ and $\alpha^*_{h_1}(e', y)$. Furthermore, h_1 and $\mathcal{N}_1$ can be chosen so that, for all $h \in \mathcal{N}_1$, $\operatorname{argmax}_{\tilde{e} \in E} v(h, \tilde{e}) \subset \operatorname{argmax}_{\tilde{e} \in E} v(h_0, \tilde{e})$, and all data-points are valuable. If e or e' is not optimal under h_1, this concludes the inductive step.

Step 2: targeted garbling. Consider the case where e and e' are optimal under h_1. The fact that $e' \neq e$ implies that there exists a pair (x, τ) that is sampled by e but not by e'. For $\theta \in [0, 1]$ and any state $p \in P$, let $f^\theta(p) \in P$ denote the state of the world such that

$$f^\theta(p)^\tau_x = (1 - \theta)p^\tau_x + \theta \mathbb{E}_{h_1}[p^\tau_x | (p^{\tau'}_{x'})_{(x', \tau') \neq (x, \tau)}]$$

and $f^\theta(p)^{\tau'}_{x'} = p^{\tau'}_{x'}$ for $(x', \tau') \neq (x, \tau)$. Let h^θ_1 be the distribution of transformed state $f^\theta(p)$

under h_1. As θ approaches 0, h_1^θ approaches h_1 under the statistical distance. Notice that h_1^θ garbles h_1 at (x,τ) alone, and does not change the expected performance of decision rules that depend on assignments $(x',\tau') \neq (x,\tau)$. Hence, it does not affect the value of experiment e'. We now show it must change the value of experiment e.

Let $v^\theta(e) \equiv \mathbb{E}_{h_1^\theta}\left[p^0 + \alpha^*_{h_1}(e,y)(p^1-p^0)\middle|e\right]$ denote the value of experiment e for the fixed policy rule $\alpha^*_{h_1}$, evaluated at prior h_1^θ. For θ close to 0, the fact that the optimal policy does not change for priors in $\mathcal{N}_1$ implies that $v^\theta(e) = v(h_1^\theta, e)$. Note that $v^\theta(e)$ is a polynomial in θ. We show it is not constant. As e is optimal and all data points are valuable under h_1, sampling the pair (x,τ) is strictly valuable. Hence, there exists θ close enough to 1 such that $v^\theta(e) < v^0(e)$. As a non-zero polynomial has finitely many zeros, there exists θ arbitrarily close to 0 such that $v(h_1^\theta, e) \neq v(h_1, e) = v(h_1, e') = v(h_1^\theta, e')$. This proves the induction step.

Finally, to conclude the proof of Proposition 1, we need to show that if a unique experiment is optimal at h_0 for $\lambda = 1$, then it is also uniquely optimal for λ below but close to 1. The result follows immediately from the continuity of objective (DP) in λ, and the fact that there are finitely many experiments. Any experiment that is strictly optimal for $\lambda = 1$ remains strictly optimal for λ close to 1. ■

Proof of Proposition 2: To establish point (i) and Corollary 1, we use the standard RCT $(\mathcal{E}_{\text{RCT}}, \alpha_{\text{RCT}})$. Losses $L(p)$ from first-best, given state of the world p, are defined as

$$L(p) \equiv \max_{a\in\{0,1\}} p^a - p^0 - (p^1 - p^0) \times \text{Prob}_{p,\mathcal{E}_{\text{RCT}}}(\overline{y}^1 - \overline{y}^0 > 0).$$

We show that for all $p \in P$, $L(p) \leq \sqrt{\frac{1}{N}}$. By symmetry, it suffices to treat the case where $p^1 - p^0 > 0$. In this case, we have $L(p) = (p^1 - p^0) \times \text{Prob}_{p,\mathcal{E}_{\text{RCT}}}(\overline{y}^1 - \overline{y}^0 \leq 0)$. We bound the probability of choosing the suboptimal policy using Hoeffding's inequality (Hoeffding,

1963). A small difficulty is that assignment $e = (\tau_i)_{i\in\{1,\dots,N\}}$ is exchangeable but not i.i.d. as, by construction, $\sum_{i=1}^{N} \tau_i = N/2$. For this reason we decompose the draw of exchangeable assignment e as: (1) a uniform draw of a pairing $Q = \{\{i, j\}\}$, such that for all i in $\{1, \dots, N\}$, there exists a unique pair $\{l, m\} \in Q$ such that $i \in \{l, m\}$; (2) independently across each pair $\{i, j\} \in Q$, draw an assignment $(\tau_i, \tau_j) \in \{(0, 1), (1, 0)\}$, with equal probabilities. Given a pairing Q, we have that

$$\overline{y}^1 - \overline{y}^0 = \frac{2}{N} \sum_{\{i,j\}\in Q} \tau_i(y_i^1 - y_j^0) + (1 - \tau_i)(y_j^1 - y_i^0).$$

Conditional on a pairing Q, variables $\tau_i(y_i^1 - y_j^0) + (1 - \tau_i)(y_j^1 - y_i^0)$ are independent across pairs and take values within $[-1, 1]$. Applying Hoeffding's inequality to this sum of $N/2$ independent terms, we obtain that

$$\begin{aligned}\text{Prob}_{p,\mathcal{E}_{\text{RCT}}}(\overline{y}^1 - \overline{y}^0 \leq 0) &= \text{Prob}_{p,\mathcal{E}_{\text{RCT}}}\left(\overline{y}^0 - \overline{y}^1 - (p^0 - p^1) \geq (p^1 - p^0)\right) \\ &\leq \exp\left(-\frac{1}{4}N(p^1 - p^0)^2\right).\end{aligned}$$

For any $a > 0$, the mapping $x \mapsto x \exp(-ax^2)$ is log-concave and maximized at $x = (2a)^{-1/2}$. This implies that

$$L(p) \leq \sqrt{\frac{2\exp(-1)}{N}} \leq \sqrt{\frac{1}{N}}.$$

An analogous argument holds in the case where $p^1 - p^0 \leq 0$. Hence, given any $h \in \Delta(P)$,

$$\mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p, a)\right] - \mathbb{E}_{h,\mathcal{E}_{\text{RCT}}}\left[u(p, \alpha_{\text{RCT}}(e, y))\right] \leq \sqrt{\frac{1}{N}}.$$

To establish point (ii), fix a deterministic experiment $e \in E$. By Assumption 1,

$$\max_{\alpha} \min_{h\in H} \mathbb{E}_{h,e}\left[u(p, \alpha(e, y))\right] \leq \min_{h\in H} \mathbb{E}_{h,e}\left[\max_{a\in\{0,1\}} \mathbb{E}_{h,e}\left[u(p, a)|(p_x^\tau)_{x,\tau\in e}\right]\right] \leq \min_{h\in H} \mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p, a)\right] - \xi$$

where the first inequality follows from the fact that experimental outcomes are a garbling of $(p_x^\tau)_{x,\tau\in e}$ — i.e. given $(p_x^\tau)_{x,\tau\in e}$ the decision-maker can simulate the outcome y of an experiment simply by drawing outcomes y_x^τ independently according to p_x^τ (see Blackwell, 1951, for a general treatment). This implies that for all $\alpha \in \mathcal{A}$,

$$\min_{h\in H_0} \mathbb{E}_h\left[u(p,\alpha)\right] \leq \min_{h\in H_0} \mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p,a)\right] - (1-\lambda)\xi.$$

■

Proof of Proposition 3: Consider the set of optimal experiments

$$E^* = \operatorname*{argmax}_{e\in E} \mathbb{E}_{h_0}[u(p,\alpha_{\text{RCT}}(e,y))].$$

By assumption its cardinal is bounded above independently of N. The number of experiments that assign treatment to $N/2$ participants out of N is necessarily less than 2^N. Hence the probability that a K-rerandomized trial selects $e \in E^*$ is at least $\rho \equiv 1 - \left(1-2^{-N}\right)^K$. For $K \geq 2^N$,

$$\rho \geq 1 - \exp\left(2^N \ln\left(1-2^{-N}\right)\right) \xrightarrow{N\to\infty} 1 - \exp(-1) > 0.$$

A sequence of strictly positive terms converging to a strictly positive number is bounded below by a strictly positive number. Hence, there exists $\rho' > 0$ such that, for all N, rerandomized experiment $\mathcal{E}_K$ selects an experiment $e \in E^*$ with probability at least ρ'.

For all policy rules $\alpha \in \mathcal{A}$, $h \in H$ we have that

$$\mathbb{E}_{h,\mathcal{E}_K}[u(p,\alpha(e,y))] \leq (1-\rho')\mathbb{E}_h(\max_{a\in\{0,1\}} u(p,a)) + \sum_{e\in E^*} \frac{\rho'}{|E^*|}\mathbb{E}_h\left(\max_{a\in\{0,1\}} \mathbb{E}_h(u(p,a)|(p_x^\tau)_{\tau,x\in e})\right).$$

By Assumption 1, this implies that for all $\alpha \in \mathcal{A}$

$$\min_{h\in H} \mathbb{E}_{h,\mathcal{E}_K}[u(p,\alpha(e,y))] \leq \min_{h\in H} \mathbb{E}_h \left[\max_{a\in\{0,1\}} u(p,a)\right] - \frac{\rho'}{|E^*|}\xi,$$

which implies Proposition 3. ■

Proof of Proposition 4: The proof that follows applies for any procedure used to pick experiment e_K^* among $(e_1, \cdots, e_K)$.

Denote by $(\overline{y}_{0,k}, \overline{y}_{1,k})$ the sample average of outcomes by treatment group for experiment e_k, and by $(\overline{y}_0^*, \overline{y}_1^*)$ the sample average of outcomes by treatment group for the experimental design e_K^* selected by rerandomized experiment $\mathcal{E}_K$. In the case where $p^1 - p^0 > 0$, regardless of the manner in which e_K^* is selected from experimental assignments $\{e_k,\ k \in \{1,\ldots,K\}\}$, losses $L(p)$ from first-best satisfy

$$\begin{aligned} L(p) &= (p^1 - p^0)\text{Prob}_{p,\mathcal{E}_K}(\overline{y}_1^* - \overline{y}_0^* \leq 0) \\ &\leq (p^1 - p^0)\text{Prob}_{p,\mathcal{E}_K}\left(\min_{k\in\{1,\ldots,K\}} \overline{y}_{1,k} - \overline{y}_{0,k} \leq 0\right) \\ &\leq (p^1 - p^0)\min\left\{1, \sum_{k=1}^{K} \text{Prob}_{p,e_k}(\overline{y}_{1,k} - \overline{y}_{0,k} \leq 0)\right\}. \end{aligned}$$

The proof of Proposition 2 shows that $\text{Prob}_{p,e_k}(\overline{y}_{1,k} - \overline{y}_{0,k} \leq 0) \leq \exp\left(-N(p^1 - p^0)^2/4\right)$. We have that $K\exp(-N(p^1 - p^0)^2/4) \leq 1 \iff p^1 - p^0 \geq 2\sqrt{\frac{\ln K}{N}}$, which implies that

$$L(p) \leq \begin{cases} p^1 - p^0 & \text{if} \quad p^1 - p^0 < 2\sqrt{\frac{\ln K}{N}}, \\ K(p^1 - p^0)\exp(-N(p^1 - p^0)^2/4) & \text{if} \quad p^1 - p^0 \geq 2\sqrt{\frac{\ln K}{N}}. \end{cases} \tag{D.2}$$

The mapping $x \mapsto x\exp(-Nx^2/4)$ is maximized at $x = \sqrt{\frac{2}{N}}$. As $K \geq 2$, we have $2\sqrt{\frac{\ln K}{N}} > \sqrt{\frac{2}{N}}$, which implies that both terms on the right-hand side of (D.2) are maxi-

mized at $p^1 - p^0 = 2\sqrt{\frac{\ln K}{N}}$. This implies that $L(p) \leq 2\sqrt{\frac{\ln K}{N}}$. An identical reasoning applies in the case where $p^1 - p^0 < 0$. ■

Proof of Proposition 5: Consider the generalized K-rerandomized experiment $\mathcal{E}_K$ such that the selected experiment e_K^* is chosen to maximize objective function $B(e) \equiv \mathbf{1}_{e \in E_\dagger}$. Proposition 4 applies as is.

Experiment design $\mathcal{E}_K$ is equivalent to running experiment $\mathcal{E}_{E_\dagger}$ (that is, picking uniformly from $E_\dagger$) with probability $1 - (1 - p_{E_\dagger})^K$ and experiment $\mathcal{E}_{E \setminus E_\dagger}$ with probability $(1 - p_{E_\dagger})^K$. As u takes values in $[0, 1]$ this implies that for all h, and $K \geq 2$,

$$\begin{aligned}
&\mathbb{E}_{h,\mathcal{E}_K}\left[u(p, \alpha_{\text{RCT}}(e, y))\right] \leq \left(1 - (1 - p_{E_\dagger})^K\right) \mathbb{E}_{h,\mathcal{E}_{E_\dagger}}\left[u(p, \alpha_{\text{RCT}}(e, y))\right] + (1 - p_{E_\dagger})^K \\
\Rightarrow &\mathbb{E}_{h,\mathcal{E}_{E_\dagger}}\left[u(p, \alpha_{\text{RCT}}(e, y))\right] \geq \mathbb{E}_h\left[\max_{a \in \{0,1\}} u(p, a)\right] - 2\sqrt{\frac{\ln K}{N}} - (1 - p_{E_\dagger})^K,
\end{aligned}$$

where the last inequality uses Proposition 4. Taking the maximum of the right-hand side over $K \geq 2$ concludes the proof. ■

Proof of Proposition 6: Point (i) follows from a reasoning similar to that of Proposition 1. For $\lambda = 1$, given an experiment $\mathcal{E}$, the decision-maker's indirect utility is

$$\max_{\alpha, \mathcal{E}} \mathbb{E}_{h_0}[w(p, \alpha)] = \sum_{e \in E} \mathcal{E}(e) W(h_0, e),$$

where $W(h_0, e) \equiv \sum_{y \in \mathcal{Y}} \text{Prob}(y|e) \max_{a \in \{0,1\}} \mathbb{E}_{p \sim h_0}\left[w(p, a)|e, y\right]$. Hence, an experiment $\mathcal{E}$ is optimal if and only if $\text{supp}\, \mathcal{E} \subset \arg\max_{e \in E} W(h_0, e)$.

We now turn to point (ii). We use Proposition 7—established below—which implies that there exist randomized experiments leading to optimal decisions up to a penalty of order $O(1/\sqrt{N})$. This implies that the decision-maker can guarantee herself a payoff greater than

$-O(1/\sqrt{N})$. We show this is not true when the decision-maker implements a deterministic experiment e. For $d \in (-1/2, 1/2)$, let $p(d)$ denote the state such that

$$\begin{aligned} p_x^0 = \tfrac{1}{2} + d, \quad p_x^1 = \tfrac{1}{2} \qquad & \text{if } \tau_x = 1; \\ p_x^0 = \tfrac{1}{2}, \qquad p_x^1 = \tfrac{1}{2} - d \quad & \text{if } \tau_x = 0. \end{aligned}$$

Consider the prior h_e that puts probability 0.5 on $p(d = \nu)$ and 0.5 on $p(d = -\nu)$ for $\nu \in (0, 1/2)$. By construction the information generated by experiment e is independent of whether $d = \nu$ or $d = -\nu$. In addition, $\Delta_p^1 = \overline{p}^1 - \overline{p}^0 = -d$. Hence, under prior h_e, regardless of the action a taken by the decision-maker, there is probability 0.5 that $\Delta_p^a = -\nu$ and probability 0.5 that $\Delta_p^a = +\nu$. As $w(p, a)$ is locally strictly concave in Δ_p^a around $\Delta_p^a = 0$, it follows that expected payoff from running experiment e under h_e is bounded below 0. This implies that for N large enough, randomized experiments are strictly optimal. ■

Proof of Proposition 7: The proof is closely related to that of Proposition 4. Consider first the case where $\Delta_p^1 \equiv p^1 - p^0 > 0$ so that the first-best action is $a = 1$. Given p, the efficiency loss compared to first-best is equal to $L(p) = \mathbb{E}_{h,\mathcal{E}_K}[w(\Delta_p^1, 1) - w(\Delta_p^{\alpha_{\text{RCT}}}, \alpha_{\text{RCT}})]$.

As $\Delta_p^1 > 0$, we have that $L(p) = (2 + \kappa_0)\text{Prob}_{h,\mathcal{E}_K}(\alpha_{\text{RCT}} = 0)\Delta_p^1$. The proofs of Propositions 2 and 4 imply that $\text{Prob}_{h,\mathcal{E}_K}(\alpha_{\text{RCT}} = 0)\Delta_p^1$ is bounded above by $2\sqrt{\frac{\ln(K+1)}{N}}$. An identical argument holds in the case of $\Delta^1 < 0$, which yields Proposition 7. ■

E Simulations

In this appendix, we use numerical simulations to highlight the tradeoffs of rerandomization.

E.1 Well-behaved Treatment Effects

We first consider an environment where treatment effects depend smoothly on covariates. We note that because treatment effects depend smoothly on covariates x, Assumption 1 does not hold, and the losses from running a deterministic experiment maximizing balance vanish as the sample size grows large.

Covariates $x \in \mathbb{R}^5$ are drawn i.i.d. according to $\prod_{k=1}^{5} U[0,1]$, a five-dimensional uniform distribution. For each treatment status $\tau \in \{0,1\}$, these are mapped to outcomes according to a five-dimensional unknown parameter $\mu_\tau \in \mathbb{R}^5$:

$$\text{Prob}(y_i = 1|x) = \frac{\exp(\mu_\tau \cdot x)}{1 + \exp(\mu_\tau \cdot x)}.$$

Under correct belief h_0, each parameter μ_τ is independently drawn according to a five-dimensional truncated normal: $\mu_\tau \sim \prod_{k=1}^{5} \mathcal{N}(0,1)_{|[-2,2]}$. The set of adversarial priors H consists of all doctrinaire priors corresponding to fixed values $\mu_\tau \in [-2,2]^5$. We denote by e^* and α the Bayes optimal experimental assignment and policy rule under this model.

We consider the rerandomized experiment $\mathcal{E}_K$, with K following the rule of thumb $K = \min\{N, 100\}$. We report balance—captured by the negative of the L_2 norm between mean characteristics across treatment and control—as well as losses compared to first-best under various priors, and sample selection criteria.

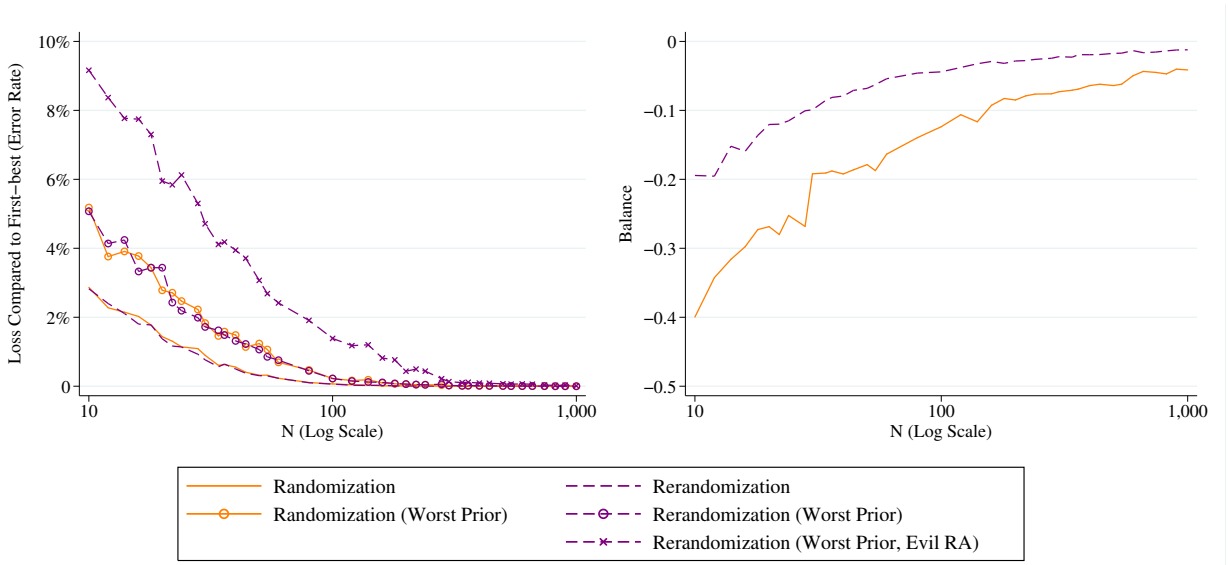
- *Bayes loss given Bayes optimal assignment*

$$\mathbb{E}_{\mu,x,\tau^*}\left[\max_{a\in\{0,1\}} u(p,a) - u(p,\alpha)\right]; \tag{E.1}$$

- *Loss under worst prior given Bayes optimal assignment*

$$\max_{\mu_0,\,\mu_1} \mathbb{E}_{x,\tau^*}\left[\max_{a\in\{0,1\}} u(p,a) - u(p,\alpha)\right]; \tag{E.2}$$

Figure E.1: Rerandomization substantially increase balance with no cost to robustness.

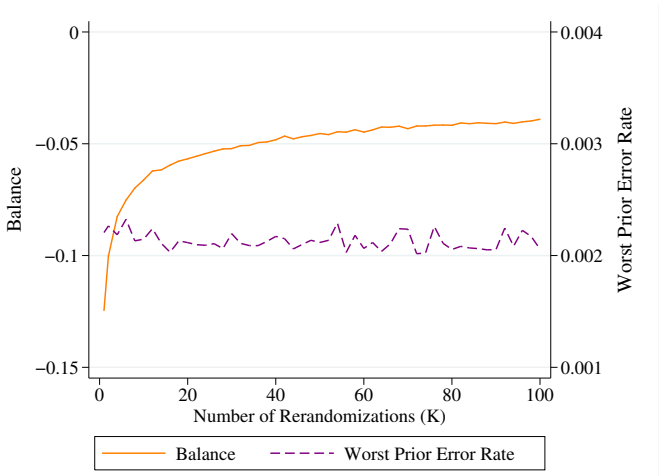


- *Loss under worst prior, and worst assignment τ*

$$\max_{\mu_0,\,\mu_1} \mathbb{E}_x \max_{\tau} \mathbb{E}\left[\max_{a\in\{0,1\}} u(p,a) - u(p,\alpha)\right]. \tag{E.3}$$

As Figure E.1 shows, the ex ante Bayes expected loss (E.1) is essentially identical under randomization and rerandomization. Loss (E.2) chooses the prior that maximizes the error rate given the experimental strategy $\mathcal{E}$ of the experimenter. While this is substantially higher than the Bayes expected loss—as one might anticipate—it is not substantially different between randomization and rerandomization. Finally, loss measure (E.3) stacks the deck against the experimenter, and assumes that the experimenter has an "evil RA" who chooses the experimental assignment τ from e_K that maximizes the expected loss. This has no application in the case of randomization, but in the case of rerandomization it substantially increases error rates. However, it is important to note even under this highly unrealistic scenario—the evil RA must know the data-generating process—the error rate is about one-tenth of 1% for $N \geq 300$.

Figure E.2: Rerandomization increases balance with no robustness cost with fixed N.



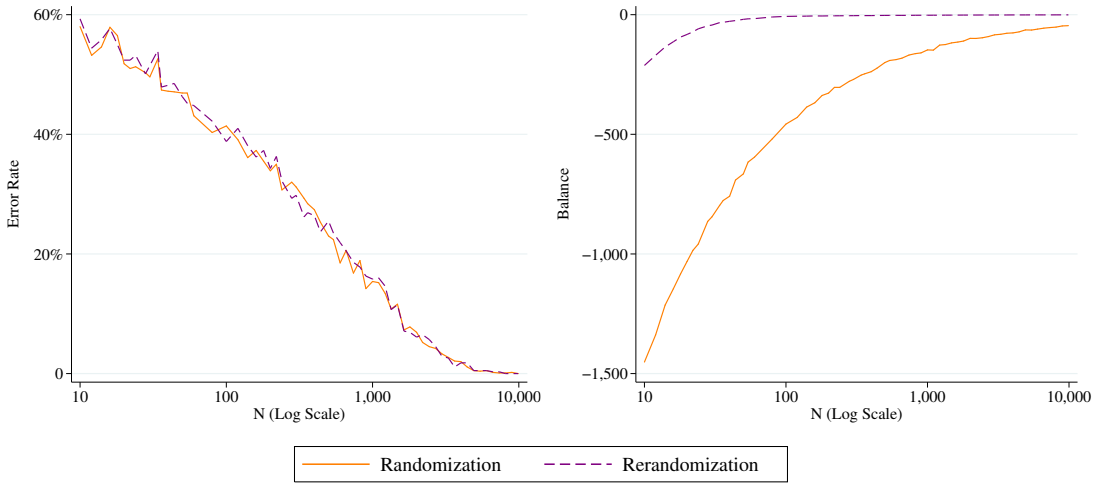
We also evaluate losses for N fixed at 100 while varying the number of rerandomizations K. Figure E.2 shows that balance improves substantially with K, especially for the first 20 rerandomizations, while the error rate is essentially flat.

E.2 Poorly Behaved Treatment Effects

We now consider the impact of rerandomization in a specific state of the world p such that a natural balance objective fails to improve the quality of decision-making.

Specifically, the environment is as follows. Covariates are on the real line $x \in X \subset \mathbb{R}$ and the balance objective is to minimize the distance between the mean of each treatment group: $B(e) = -|\overline{x}^1 - \overline{x}^0|$. The difficulty here is that treatment effects are very jagged as a function of x, so that balance with respect to $\overline{x}^1$ and $\overline{x}^0$ does not help identify treatment effects. Natural, deterministic assignments achieving a high balance objective will result in non-vanishing efficiency losses.

Figure E.3: Rerandomization substantially increase balance with no cost to robustness.



Specifically we set $X = \{1, 2, \ldots, 10,000\}$ and

$$p_x^0 = \begin{cases} \frac{1}{5} & \text{if} \quad x \mod 6 \in \{2,4\} \\ \frac{1}{2} & \text{if} \quad x \mod 6 \notin \{2,4\} \end{cases} \qquad p_x^1 = \begin{cases} \frac{4}{5} & \text{if} \quad x \mod 6 \in \{2,4\} \\ \frac{1}{4} & \text{if} \quad x \mod 6 \notin \{2,4\} \end{cases}.$$

On aggregate, $u(p,1) \simeq \frac{13}{30} > \frac{2}{5} \simeq u(p,0)$, so that treatment $(a = 1)$ is beneficial.

For this specific state, the aspect of covariates that balance seeks to improve is unrelated to treatment effects. In fact, a natural matching algorithm systematically assigning consecutive xs to treatment and control (starting with treatment) results in an experimental assignment that does not lead to the efficient decision. Figure E.3 examines the error rates and balance of randomization and rerandomization. Both schemes yield the same error rate. However, once again, rerandomization substantially improves the balance of the samples. This is particularly true for small and moderate sample sizes. This is not useful for this particular state of the world, but may be valuable at states where treatment effects are better behaved as a function of x.