A Theory of Experimenters*

Abhijit V. Banerjee
Massachusetts Institute of Technology

Sylvain Chassang
New York University

Sergio Montero
University of Rochester

Erik Snowberg
University of British Columbia

April 17, 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 enough and robustness is an important concern. We apply our model to shed light on the practice of rerandomization. We show that rerandomization creates a tradeoff between subjective and robust performance, but that losses in robust performance guarantees grow very slowly with the number of randomizations. This suggests that moderate levels of rerandomization usefully expand the set of acceptable compromises between subjective performance and robustness. We argue that targeting a fixed quantile of balance is a safer approach than targeting an absolute balance objective.

JEL Classification: C90, D78, D81
Keywords: experiment design, robustness, ambiguity aversion, 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, 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.

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, Glennerster and Kremer, 2008; Bruhn and McKenzie, 2009; Deaton, 2010; Imbens, 2010; Humphreys, de la Sierra and Van der Windt, 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 decision making under uncertainty. A key modeling step needed to account for randomization is to allow for ambiguity aversion on the part of the decision maker. 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.

[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.

The paper reports two main sets of results. The first set of results shows that RCTs can be optimal for a decision-maker facing an adversarial audience, and clarifies the circumstances in which this is the case. If the decision-maker places non-zero weight on satisfying her adversarial audience, then, for sufficiently large sample sizes, it is always strictly optimal for the decision-maker to use a RCT. Indeed, as the sample size grows large, RCTs permit robust prior-free inference. 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 introduction of new technologies inside a firm—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. We study the common experimental practice of rerandomizing to improve covariate balance between treatment and control groups. Rerandomization draws multiple treatment assignments, then chooses the one that maximizes balance. 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 succinctly captured in our framework. Successive rerandomizations improve balance, a common objective of experimenters that 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 when the number of rerandomizations grows linearly or sublinearly 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 extend under regret-minimization. In fact, regret aversion lets us simplify technical assumptions, as well as rationalize the practice of null-hypothesis statistical testing (NHST; Tetenov, 2012).[4]

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 NHST in decision-making. Section 6 contains several discussions: of other possible rational-

[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.

[4]A related paper by three of the authors provides a non-technical discussion of our results on randomization, and a limited discussion of our results on rerandomization (Banerjee, Chassang and Snowberg, 2017).

izations of randomization, of the positive implications of our theory, of practical implications for rerandomization, and of possible directions for future research.[5]

2 Model

We first lay out the problem of experimental design, similar to Banerjee, Chassang and Snowberg (2017). We then specify the preferences and beliefs of the decision-maker—the main difference between our work and prior theories of experimentation.

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 large number of 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.

We assume that the set of covariates X is finite, but large. For simplicity, we assume that all individuals are unique. As a result, the mass of individuals with a given type x in the population is $1/|X|$. This assumption is consistent with the Neyman-Rubin (Neyman, 1923; Rubin, 1974) "potential outcomes" framework. Each person i is unique and cannot be observed in both the treated and untreated state.

The probability of success given treatment τ and covariate x is denoted by $p_x^\tau \equiv \text{Prob}(y^\tau = 1|x)$. Outcomes are i.i.d. conditional on 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. Note that state-space P is compact, convex,

[5]In addition, Online Appendix A clarifies how regret aversion can rationalize the use of NHST. Proofs are contained in Online Appendix B. Online Appendix C presents simulations.

and finite-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 would 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 may cause problems when the distribution of z conditional on x changes rapidly: individuals with similar values of x may have very different values of z.[6] In such a setting, the mapping $x \mapsto p_x^\tau$ becomes discontinuous in x.

Experiments and strategies. To maximize her odds of making the correct policy choice, the decision-maker can run an experiment on N participants. We take as given the participants' covariates $(x_i)_{i \in \{1,\dots,N\}}$ and keep them fixed throughout the paper. For simplicity, we assume that the sample is representative of the underlying set of types, so that $N = |X|$ and $\{x_i\}_{i \in \{1,\cdots,N\}} = X$. This lets us abstract away from questions of sampling—which often is not under the control of experimenters—and focus on treatment assignment. Note that as a result, when we consider large sample sizes N, we consider environments where both N and $|X|$ are large.[7]

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

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

[7] Our main results extend under random sampling provided that: (i) the set of covariates $x \in X$ is sampled with full support, and (ii) the sample size N is not too large compared to X. Specifically, Assumption 1 (below) can be satisfied whenever $N \leq 2\rho|X|$, where $\rho \in (0,1)$ is fixed.

$i \in \{1, \cdots, N\}$ 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.[8]

The standard RCT. A special case of interest is the standard RCT, assigning 50% of participants to treatment $\tau = 1$.[9] 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 in the paper stem from random assignment, not random policy-making.

2.2 Preferences and Beliefs

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:

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

[8] Following standard notation, $u(p, \alpha) \equiv \mathbb{E}_\alpha[u(p, a)]$.

[9] 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.

where H is a convex 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), but parameterized to emphasize subjective expected utility maximization as a special case ($\lambda = 1$) and facilitate comparative statics.[10] 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.

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 $h \in H$. If $\lambda = 1$, the decision-maker does not care about the audience and we recover the standard subjective expected utility, or *Bayesian* model. An alternative interpretation is that the decision-maker is constrained by a heterogeneous audience whose members have veto power and share a common reservation utility. In this interpretation, the ratio $\frac{1-\lambda}{\lambda}$ can be interpreted as the Lagrange multiplier placed on audience members' individual rationality constraint. 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 $\lambda h_0 + (1-\lambda)h$. Going forward, we use this audience interpretation in discussions, but it is equally appropriate to view decision problem (DP) as a model of *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 $(p^\tau_x)_{x,\tau \in e}$ for pairs of covariate and treatment (x, τ) sampled by experiment e is not sufficient for efficient decision-making under prior h. 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:

[10] Note in particular that it coincides with the standard maxmin representation with objective $\min_{h \in H_0} \mathbb{E}_{h,\mathcal{E}}[u(p, \alpha(e, y))]$, where $H_0 \equiv \{\lambda h_0 + (1-\lambda)h | h \in H\}$.

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*

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

Note that, for any experiment e, $(p_x^\tau)_{x,\tau \in e}$ is an upper bound for the information generated by experiment e in the sense of Blackwell (1953). 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 complete failures. If this were the case, nature would minimize the decision-maker's welfare using the prior $p^1 = p^0 = 0$, making learning and experimentation moot. As shown in Section 5, Assumption 1 can be disposed of if the decision-maker exhibits regret aversion rather than ambiguity aversion.

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. For simplicity, assume that $|X|/2$ is an integer. 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, we show that for any deterministic experiment e there exists a prior $h \in H$ such that decision-making is bounded away from first-best. Since $N = |X|$, 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 $\xi = 1/8$.[11] □

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 $|X| = N > N/2$). If not, all pairs (x,τ) could be sampled by a deterministic experiment. Second, pernicious priors are such that non-sampled pairs (x,τ) create significant residual uncertainty about the correct treatment. As the deterministic experiment e provides no information on 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 one is optimal depends on sample size N and the weight λ given to subjective performance of the policy decision $\mathbb{E}_{h_0} u(a,p)$ versus the worst-case performance guarantee $\min_{h\in H} \mathbb{E}_h u(a,p)$.

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

[11] 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'.

generically every prior h_0 for 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—$d(h, h') \equiv \sup_{A \subset P \atop A \text{ meas.}} |h(A) - h'(A)|$—sometimes described as *the* statistical distance.

In order to rule out trivial indifferences between experiments, we assume that the decision-maker values every data-point, that is, reducing the sample size reduces the quality of decision-making under prior h_0. This holds automatically if there is some cost of acquiring data-points, and the sample size N is chosen optimally by the decision-maker. Because sample size N is exogenously given, this is an assumption in our case.

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 $E_{-1} = \{e_{-i} \mid i \in \{1, \ldots, N\}, e \in E\}$ 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 $\Delta(\{0,1\})$.

Assumption 2 (all data points are valuable)**.** *All data points are valuable if and only if*

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

The following holds.

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 satisfying Assumption 2, there exists $\underline{\lambda} \in (0,1)$ and a deterministic experiment e^* such that, for all $\lambda > \underline{\lambda}$, e^* is uniquely optimal.

This result strengthens the observation that deterministic experiments are weakly optimal for a Bayesian decision-maker. When all data points are valuable, a deterministic experiment is generically strictly optimal.[12] This remains true if the decision-maker puts a small, but

[12]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.

non-zero, weight $1-\lambda$ on the preferences of her audience.

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. To simplify the presentation of results, let $H_0 = \{\lambda h_0 + (1-\lambda)h \mid h \in H\}$ denote the set of mixtures between the experimenter's prior h_0 and the audience's priors $h \in H$. Note that $\lambda \mathbb{E}_{h_0} u(p,a) + (1-\lambda)\min_{h\in H} \mathbb{E}_h u(p,a) = \min_{h\in H_0} \mathbb{E}_h u(p,a)$.

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:*

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

$$\max_{\alpha\in\mathcal{A}} \min_{h\in H_0} \mathbb{E}_{h,\mathcal{E}^*}\left[u(p,\alpha(e,y))\right] \geq \min_{h\in H_0} \mathbb{E}_h\left[\max_{a\in\{0,1\}} u(p,a)\right] - \sqrt{\frac{2}{N}}.$$

(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 $1/\sqrt{N}$. Note that this performance bound is 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. Note that this is not an asymptotic result: threshold $\underline{N}$ can be defined as an explicit function of ξ.

The traditional interpretation of an ambiguity-averse decision-maker's problem as a zero-sum game played against nature helps clarify 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 it 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).[13] In our context, for large sample size N, the Bayesian component of preferences is close to the first-best regardless of the prior. Thus, the subjective benefits from running a deterministic experiment are small. In contrast, under Assumption 1, the impact of randomization on robust payoff guarantees is bounded away from 0.

We note that randomization can also be interpreted as 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 particularly adversarial priors.[14] Randomization provides ex ante performance guarantees even under the most skeptical priors. Approximate randomization does not.

[13] 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.

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

3.3 RCTs as a Rule of Thumb

A well known corollary 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* $h \in \Delta(P)$*,*

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

This result implies that a standard RCT guarantees approximately optimal decision-making for both Bayesian and ambiguity-averse decision-makers.[15] Importantly, the decision-maker does not even need to specify her own preferences. When sample N is large enough, a RCT is near optimal for all parameters λ and set 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.

This corollary supports an interpretation of standard RCTs as a simple rule-of-thumb that captures the bulk of the value from randomization. When choosing from a restricted set of choices—say between deterministic experiments and a standard RCT—a decision-maker with a large sample and preferences (DP) will prefer the standard RCT over any deterministic experiment. Indeed, the decision-maker would not be willing to spend significant resources in order to run the unrestricted optimal experiment under (DP), as opposed to a standard RCT.[16]

While RCTs have many useful attributes, in practice they also have some faults. Specifically, they can fail to generate samples whose covariates are balanced. We explore this issue, and rerandomization as a potential remedy, in the next section.

[15] Note that Corollary 1 holds for any arbitrary set of covariates X, even if $|X| \gg N$, provided covariates are sampled representatively. Because outcomes are bounded, it is not possible for rare—and therefore hard to sample—covariate realizations to have a substantial impact on payoffs. Balance in expectation is sufficient to guarantee approximately optimal decision-making.

[16] A revealed-preference way to make this point would be to let the decision-maker trade off the share of people receiving her preferred treatment a versus the ability to run her preferred experiment.

4 Rerandomization

As noted above, a standard RCT is only a near-optimal solution to (DP). 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 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 tradeoff 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.

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 used in the standard RCT: $\alpha_{\text{RCT}}(e, y) = \mathbf{1}_{\overline{y}^1 \geq \overline{y}^0}$.

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 assignment $e_k = (\tau_{i,k})_{i\in\{1,\dots,N\}}$ an exchangeable assignment such that 50% of participants receive treatment $\tau = 1$.
2. Select an assignment $e_K^* \in \operatorname{argmax}_{e\in\{e_1,\cdots,e_K\}} \mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e, y))]$ that maximizes the decision-making value of experiment e under prior h_0.
3. Run experiment e_K^*, generating outcomes y_K.
4. Choose a policy $a \in \{0, 1\}$ according to α_{RCT}.

Balance-targeting rerandomization protocols fix a minimum target level for $\mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e, y))]$, and draw assignments e until the target is achieved or exceeded (Morgan and Rubin, 2012). This contrasts with K-rerandomization, which guarantees an upper bound K to the number of sample draws. As shown below, open-ended draws may sometimes lead to significant robustness losses.

4.2 Balance as Subjective Expected Utility

In practice, experimenters engaging in rerandomization select assignment e_K^* according to a specific statistical balance objective capturing the distance between treatment and control samples. A prominent objective is to minimize 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).[17]

This objective is a special case of subjective expected utility maximization for a natural class of priors h_0, which we now describe. Assume the decision-maker believes outcomes

[17]The Mahalanobis distance between two arbitrary vectors $v_1,\ v_2 \in \mathbb{R}^k$ is given by $(v_1 - v_2)^\top \mathsf{cov}(\mathbf{x})^{-1}(v_1 - v_2)$, in which $\mathbf{x} = (x_i)_{i\in\{1,\dots,N\}}$. 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 $\mathbf{x}$.

$y_i^\tau \in \{0,1\}$ are distributed according to the linear model

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

in which zs are the underlying determinants of outcomes ys, I_k is the k-dimensional identity matrix, and $(\varepsilon_i)_{i\in\{1,\ldots,N\}}$ are independent, mean-zero error terms. Although terms z_i are unobservable, they are assumed to be a transformation of observables x_i: $x_i = \mathsf{cov}(\mathbf{x})^{1/2} z_i$. This implies that

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

Given a treatment assignment $(\tau_i)_{i\in\{1,\ldots,N\}}$, 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-maker under prior h_0 is

$$\mathbb{E}\left\langle \beta, \overline{x}^0 \right\rangle + \mathbb{E}_\Delta[\Delta \times \mathrm{Prob}(\Delta + \left\langle \beta, \overline{x}^1 - \overline{x}^0 \right\rangle + \phi \geq 0)].$$

It is immediate that this expression is decreasing in the variance of $\langle \beta, \overline{x}^1 - \overline{x}^0 \rangle$:

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

which is precisely the Mahalanobis distance between the mean of covariates in the treatment and control groups, and is trivially minimized when $\overline{x}^0 = \overline{x}^1$. Under this prior, the assignment $(\tau_i)_{i\in\{1,\ldots N\}}$ that maximizes subjective expected utility minimizes the Mahalanobis distance of covariate means between the treatment and control groups.

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

of rerandomization.

4.3 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. Note this holds for *any* preference over deterministic assignments used to select e_K^* among K assignments.

Remark 1. $\mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e_K^*, y))]$ *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). Rerandomization with such a balance objective leads to significantly more precise estimates of treatment effects when outcomes come from a linear Gaussian model. More generally, Bugni, Canay and Shaikh (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.

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

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

Proposition 3. *There exists $\underline{\rho} > 0$ such that for all N, if $K \geq 2^N$, then*

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

where ξ is defined in Assumption 1.

Intuitively, when K is sufficiently large, the assignment is essentially deterministic. Proposition 2 implies that this precludes first-best performance guarantees. Proposition 3 raises concerns about a popular class of rerandomization schemes that sets a fixed balance target, and rerandomizes 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.

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

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

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 also holds for *any* preference over deterministic assignments used to select e_K^* among K assignments, similar to Remark 1. 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 preferences of stakeholders.

[19] We also note that this bound is non-asymptotic and conservative.

4.4 Direct Approaches to Balance Involve Similar Tradeoffs

We conclude this section by highlighting that the tradeoffs of K-rerandomization apply to 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 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]

The potential issue here is that such a 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 will come at a cost to robustness bounded away from 0. 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 small, causing significant losses of robustness.

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 implies the following corollary.

[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 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 $p_{E_\dagger}$.

Corollary 2. *For all $h \in H$,*

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

This implies that a matching 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. In terms of the partial stratification procedure above, it suggests that in order to avoid robustness losses, one needs to accept lower levels of balance on other dimensions if one wants to stratify on a few dimensions.

Quantile targeting rerandomization. This analysis helps qualify the procedure, endorsed by Morgan and Rubin (2012), that calls for rerandomizing until a desired level of balance is achieved. If the experimenter sets an excessively high balance target, set $E_\dagger$ will be small, leading to the possibility of non-negligible losses in robustness. K-rerandomization protects against such robustness losses by limiting the number of draws, rather than letting them grow to meet some balance threshold. Corollary 2 suggests specifying balance targets as a specific quantile of balance. 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 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) in data-driven decision-making is also puzzling for standard models. Recent advances, however, have shown that NHST can be rationalized using reference-dependent preferences (Tetenov, 2012). We establish that such reference-dependent preferences are also compatible with our framework. The version of reference-dependence we use nests regret aversion, and has the additional benefit of allowing

us to dispense with Assumption 1. In particular, consider payoffs

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

where $\Delta^a \equiv u(p, a) - u(p, 1-a)$. These payoffs depend on both the policy choice a made by the decision-maker, and the performance difference Δ^a between this choice and the alternative.[22] Crucially, the κ parameters—which only affect preferences when the decision-maker makes a mistake, that is $\Delta^a < 0$—imply that the possibility of a success of size Δ^a cannot offset an equal possibility of a mistake of the same size.[23] When $\kappa_0 = \kappa_1$ then regret about a mistake of fixed 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 believes strongly enough that $p^1 > p^0$ in order to reject the null hypothesis.

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

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

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

These preferences allow us to dispense with Assumption 1 (limited extrapolation). As noted earlier, limited extrapolation rules out the doctrinaire prior setting $p^1 = p^0 = 0$. This would be the unconstrained worst-case prior chosen by nature under (DP), and no learning is possible at these priors. Regret aversion implies that nature would no longer select such a prior. The worst-case prior is now one where the decision-maker can make a mistake, otherwise there is nothing to regret. As a result, under regret aversion, information

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

[23] These parameters are similar to loss aversion in the classic reference-dependent preferences of Prospect Theory (Kahneman and Tversky, 1979).

[24] We define $w(\Delta^\alpha, \alpha) \equiv \mathbb{E}_\alpha[w(\Delta^a, a)]$.

is valuable even under the unconstrained worst-case prior.

Our main results extend to this class of preferences:

Proposition 5 (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 6 (rerandomization)**.** *There exists $M > 0$ such that for every prior $h \in H$,*

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

There are two superficial differences between these results and those presented earlier. First, part (i) of Proposition 5 features weak optimality, rather than generic strict optimality. This is for simplicity. The result that random assignment is strictly optimal for large sample size N continues to hold. Second, Proposition 6 features a coefficient M. As shown in the proof, M is a function of parameters κ_0, κ_1.

6 Discussion

Our ambition is to provide a decision-theoretic framework for experimental design that can clarify debates about experimental practice. We take as a starting point that any such framework must be able to rationalize a key experimental practice: randomization. In this last section, we discuss alternative rationales for randomization, and further implications of our framework.

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.

The behavior of many real-life experimenters, however, indicates a strict preference for randomization. Careful randomization is 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. In addition, experimenters do obtain covariates, and frequently use them to determine experimental assignment by rerandomizing or stratifying. The fact that decision-makers are willing to sacrifice some balance for the sake of randomization highlights that they are not merely indifferent between assignments: they have strict preferences for randomization.

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, and wants to limit the ways in which an interested experimenter could bias findings. This concern about moral hazard is not present in our framework. 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. For example, a Bayesian regulator would mandate specific experimental assignments once a sample is drawn.[25] In addition, in many cases, experimenters are entirely trusted by the audience. Experimenters could lie about randomizing and simply pick their preferred experimental assignment with little possibility of detection. The fact that they choose to randomize

[25]The regulator would also likely have preferences over the specific sample, but we do not deal with the problem of drawing an optimal sample in this work.

suggests that they intrinsically respect their audience's concern for robustness. This choice is also consistent with the preferences of (DP) expressing internal doubt.

6.2 Deterministic versus Randomized Experimentation

As shown in Section 3, our framework rationalizes either Bayesian 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.

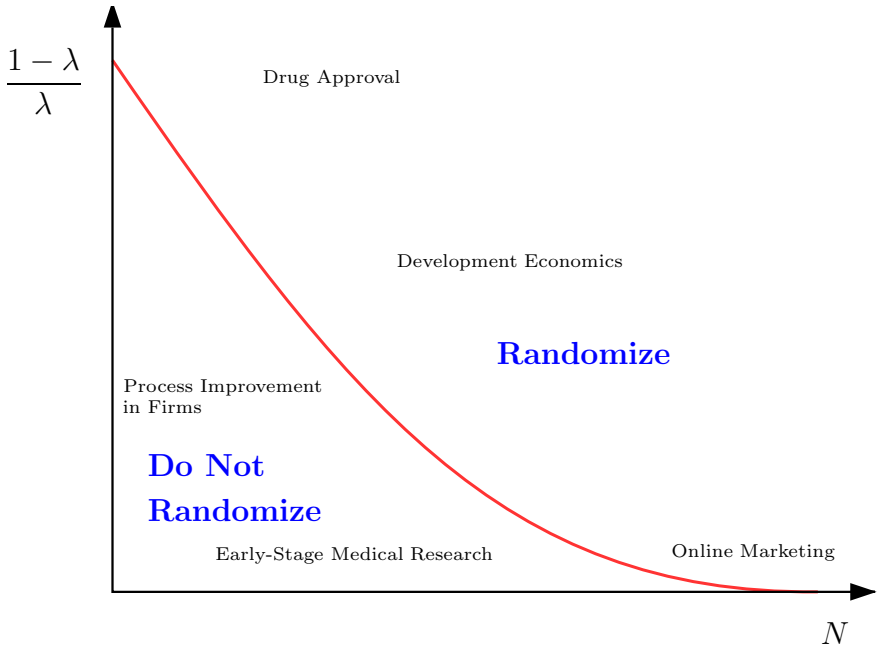
When sample points are scarce (N is small), or when the decision-maker does not put much weight on satisfying her audience (λ is close to 1), the optimal experiment is deterministic, driven by prior h_0. The experimenter assigns treatment and control to maximize the decision-making value of the data collected. For example, in the context of process improvement, a firm testing out new machinery may assign treatment to its best performing teams 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 usually occurs in one of two distinct cases. First, 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. The second case is when experimenting with participants who have conditions known to rapidly result in severe disability or death.[26] 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 must instead satisfy a sufficiently adversarial audience, or when

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

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



she has a sufficiently large sample, she will randomize. The former is the case in 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

This subsection discusses how our results may be used to inform the practice of rerandomization. Our propositions clarify the tradeoffs inherent in both randomization and rerandomization, however, optimal policies ultimately depend on the preferences of decision-makers. Any guideline we offer ultimately reflects our own subjective judgements.

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. Proposition 4 and Corollary 2 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 $\sqrt{\frac{\ln(N)}{N}}$. In addition to this linear upper bound, one may consider applying a constant bound. One set of guidelines that strikes us as an attractive compromise is to set $K \leq \min\{100, N\}$ (Banerjee, Chassang and Snowberg, 2017). Indeed, note that with probability $1 - 0.95^K$ the K-rerandomized assignment is in the 95^{th} quantile of most balanced assignments. Whenever $N \geq 100$ this event has probability greater than 99%.

The second approach, described in Section 4.4, is to set a quantile of balance as a 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, Corollary 2 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 eschew 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, our rerandomization results do not rely on the fact that selected experimental assignment e_K^* maximizes the particular subjective performance objective we use throughout. These results 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/or 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 further care about targeting treatment to those they believe will benefit the most, or may simply dislike 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; pick the assignment e_K^* from that group of K randomizations that maximizes the preferences of stakeholders; (iii) run the experiment ; (iv) take policy decision $a = \mathbf{1}_{\overline{y}_1 \geq \overline{y}_0}$.

We note that in this description, the preferences of the stakeholder are expressed ex ante, although Proposition 4 would continues to hold even if stakeholders picked the assignment ex post without pre-specifying their preferences. There are two reasons for this. First, formally specifying the preferences of stakeholders permits randomization inference tests

(Fisher, 1935; Morgan and Rubin, 2012; Young, 2016). Given the process for assignment a statistician can simulate the distribution of treatment effects that would be observed if we had $p_x^1 = p_x^0$, for all $x \in X$. This procedure can be used to calculate exact p-values, and/or to 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. For this reason, we believe it is important for all parties (including stakeholders) to agree on a process ex ante, and limit the scope for ex post disagreement over assignments.

6.4 Future Directions

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, provided all collected data is made available. The first difficulty is to explain why we rely on low-dimensional statistics rather than simply published data. This likely requires a model of bounded rationality. Altogether we believe that treating existing experimental practices as a form of revealed preferences provides an opportunity to improve our models of learning, experimentation, and communication.

Online Appendix—Not Intended for Publication

A Rationalizing 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 outcome p^a of action $a \in \{0, 1\}$.[1] A risk-neutral Bayesian decision-maker solving $\max_{a\in\{0,1\}} \mathbb{E}[p^a]$ will simply take action $a = 1$ if and only if $\hat{p}^1 - \hat{p}^0 > 0$. However, the t-statistic for a given treatment effect 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. In particular there always exists σ_0 large enough to cause the decision-maker to stick with $a = 0$ regardless of the estimated treatment effect.

Risk aversion over policy outcomes. A natural supposition 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

[1] 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 $\overline{y}$ and $\hat{p}$ are equal.

follows that

$$\mathbb{E}\left[\Gamma(\hat{p}^1)\right] \geq \mathbb{E}\left[\Gamma(\hat{p}^0)\right] \iff \frac{2N}{-\Gamma''}\frac{\Gamma(\hat{p}_1)-\Gamma(\hat{p}_0)}{\sigma_1^2-\sigma_0^2} = \gamma\frac{\hat{p}_1-\hat{p}_0}{\sigma_1^2-\sigma_0^2} > 1$$

with $\gamma = \frac{2N\Gamma'(\hat{p})}{-\Gamma''}$ for $\hat{p} \in [\hat{p}_0, \hat{p}_1]$.

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. Indeed, risk aversion means that the decision-maker values certainty (a small variance in outcomes) as well as a higher mean outcome. In addition, payoffs depend on the variances of parameter estimates, rather than their standard deviations.

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). Recall that for all $a \in \{0,1\}$, $p^a \equiv \frac{1}{|X|}\sum_{x\in X} p_x^a$ denotes the mean outcome under policy a. As in Section 5, consider a decision-maker who seeks to solve

$$\max_{a\in\{0,1\}} \mathbb{E}[w(\Delta^a, a)], \tag{A.1}$$

in which $\Delta^a \equiv p^a - p^{1-a}$ and $w(\Delta^a, a) \equiv \Delta^a \times (1 + \kappa_a \mathbf{1}_{\Delta^a<0})$ with $0 < \kappa_0 \leq \kappa_1$, as defined in Section 5.

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

Proof of Lemma A.1: Let $\bar{t} \equiv \frac{p^1-p^1}{\sqrt{\sigma_0^2+\sigma_1^2}}$. As $p^1 - p^0 \sim \mathcal{N}(\hat{p}_1 - \hat{p}_0, \sigma_0^2 + \sigma_1^2)$, 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:

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

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 $\bar{t}$. This implies that $\mathbb{E}_{\bar{t}}\left[w\big(\bar{t},1\big) - w\big(-\bar{t},0\big)\big|\, t=0\right] < 0$, so that $t^* > 0$. ■

B 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:

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

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 $\text{supp}\,\mathcal{E} \subset \underset{e\in E}{\text{argmax}}\, v(h_0,e)$.

To prove that deterministic experiments are generically strictly optimal when all datapoints are valuable, we show that $\text{argmax}_{e\in E}\, v(h_0,e)$ is generically a singleton. 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.

To show that this set is dense in the space of priors for which all data points are valuable, fix a neighborhood N_0 of h_0 and assume that all data-points are valuable under h_0. The proof is by induction on the number of optimal experiments in $\operatorname{argmax}_{e \in E} v(h_0, e)$. We show that, if there are n such optimal experiments, then there exists a prior $h \in N_0$ such that there are at most $n - 1$ optimal experiments in $\operatorname{argmax}_{e \in E} v(h, e)$.

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 $h_1 \in N_0$ and a neighborhood $N_1 \subset N_0$ of h_1 such that for all $h \in N_1$ the optimal policies $\alpha^*_h(e, y)$ and $\alpha^*_h(e', y)$ are unique and respectively equal to $\alpha^*_{h_1}(e, y)$ and $\alpha^*_{h_1}(e', y)$. This implies that, for all $h \in H_1$, $v(h, e) = \mathbb{E}_h\left[Q^e(p)\right]$ and $v(h, e') = \mathbb{E}_h\left[Q^{e'}(p)\right]$, where Q^e and $Q^{e'}$ are polynomials in p of order less than 2 in p^τ_x as (x, τ) is sampled at most once. Furthermore, h_1 and N_1 can be chosen so that, for all $h \in N_1$, $\operatorname{argmax}_{\tilde{e} \in E} v(h, \tilde{e}) \subset \operatorname{argmax}_{\tilde{e} \in E} v(h_0, \tilde{e})$, and all data-points are valuable.

If e is not optimal under h_1, this concludes the inductive step. Consider the case where e is 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 be the distribution of $f^\theta(p)$ under h_1. Notice that h^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 show that it must change the value of experiment e.

Let $v^\theta(e) \equiv \mathbb{E}_{h^\theta}\left[p^0 + \alpha^*_{h_1}(e, y)(p^1 - p^0)\middle| e\right]$ denote the value of policy rule $\alpha^*_{h_1}$ at h^θ. For θ close to 0, the fact that the optimal policy does not change implies that $v^\theta(e) = v(h^\theta, e)$. For all $\theta \in [0, 1]$, $v^\theta(e) = \mathbb{E}_{h^\theta}[Q^e(p)]$, and $v^\theta(e)$ is a polynomial in θ. Since e is optimal and all data points are valuable under h_1, the pair (x, τ) is strictly valuable. Hence, there exists θ close enough to 1 such that $v^\theta(e) < v^0(e)$. In contrast, for all $\theta \in [0, 1]$, $v(h^\theta, e') = v(h_1, e')$. Thus, there exists θ close enough to 1 such that $v(h^\theta, e) \neq v(h^\theta, e')$. This implies the required induction step.

The case where λ is below but close to 1 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

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

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 McDiarmid's inequality (McDiarmid, 1989). A small difficulty is that assignment $e = (\tau_i)_{i \in \{1, \ldots, N\}}$ is exchangeable but not i.i.d. since 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 $\{1, \ldots, N\}$, 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 McDiarmid's inequality to this sum of $N/2$ independent terms, we obtain that

$$\begin{aligned} \text{Prob}_{p,\mathcal{E}_{\text{RCT}}}(\overline{y}^1 - \overline{y}^0 \leq 0) &= \text{Prob}_{p,\mathcal{E}_{\text{RCT}}}\left(\overline{y}^0 - \overline{y}^1 - (p^0 - p^1) \geq (p^1 - p^0)\right) \\ &\leq \exp\left(-\frac{1}{8} N (p^1 - p^0)^2\right), \end{aligned}$$

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

$$L(p) \leq \sqrt{\frac{4\exp(-1)}{N}} \leq \sqrt{\frac{2}{N}}.$$

An analogous argument holds in the case where $p^1 - p^0 \leq 0$. Hence, given any $h \in H_0$,

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

where we defined H_0 as $H_0 = \{\lambda h_0 + (1-\lambda)h \mid h \in H\}$.

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 $(p_x^\tau)_{x,\tau \in e}$. 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 a deterministic experiment $e_E \in \text{argmax}_{e \in \text{supp } E} \mathbb{E}_{h_0}[u(p, \alpha_{\text{RCT}}(e, y))]$. The number of experiments that assign treatment to $N/2$ participants out of N is necessarily less than 2^N. Hence the probability that the kth rerandomized trial selects e_E is at least $\rho \equiv 1 - \left(1 - 2^{-N}\right)^K$. For $K \geq 2^N$,

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

Hence, there exists $\underline{\rho} > 0$ such that, for all N, rerandomized experiment $\mathcal{E}_K$ selects experiment e_E with probability at least $\underline{\rho}$. By Assumption 1, this implies that

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

■

Proof of Proposition 4: 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 $\mathcal{E}_K$. Note that the manner in which e_K^* is selected from experimental assignments $\{e_k,\ k \in \{1, \ldots, K\}\}$ is unimportant.

In the case where $p^1 - p^0 > 0$, losses $L(p)$ from first best take the form

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

The proof of Proposition 2 shows that $\text{Prob}_{p,e_k}(\overline{y}_{1,k} - \overline{y}_{0,k} \leq 0) \leq \exp\left(-N(p^1 - p^0)^2/8\right)$. We have that $K\exp(-N(p^1 - p^0)^2/8) \leq 1 \iff p^1 - p^0 \geq 2\sqrt{\frac{2\ln K}{N}}$, which implies that

$$L(p) \leq \begin{cases} p^1 - p^0 & \text{if} \quad p^1 - p^0 < 2\sqrt{\frac{2\ln K}{N}}, \\ K(p^1 - p^0)\exp(-N(p^1 - p^0)^2/8) & \text{if} \quad p^1 - p^0 \geq 2\sqrt{\frac{2\ln K}{N}}. \end{cases} \tag{B.2}$$

The mapping $x \mapsto x\exp(-Nx^2/8)$ is maximized at $x = 2\sqrt{\frac{1}{N}}$. As $K \geq 2$, we have $2\sqrt{\frac{2\ln K}{N}} > 2\sqrt{\frac{1}{N}}$, which implies that both terms on the right-hand side of (B.2) are maximized at $p^1 - p^0 = 2\sqrt{\frac{2\ln K}{N}}$. This implies that indeed $L(p) \leq 2\sqrt{\frac{2\ln K}{N}}$. An identical reasoning applies in the case where $p^1 - p^0 < 0$. ■

Proof of Corollary 2: 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,

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

where the last inequality uses Proposition 4. ■

Proof of Proposition 5: 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

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

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

We now turn to point (ii). We know from Proposition 6 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

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

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, $\Delta^1 = \overline{p}^1 - \overline{p}^0 = -d$. Hence, regardless of the action a taken by the decision-maker, there is probability 0.5 that $\Delta^a = -\nu$ and probability 0.5 that $\Delta^a = +\nu$. As w_a is locally strictly concave around 0, it follows that expected payoffs from running a deterministic experiments are bounded below 0.

This implies that for N large enough, randomized experiments are strictly optimal. ■

Proof of Proposition 6: The proof is closely related to that of Proposition 4. Consider first the case where $\Delta^1 \equiv p^1 - p^0 > 0$. Then the efficiency loss compared to first-best is equal to

$$L = \mathbb{E}_h[w(\Delta^1, 1) - w(\Delta^1, \alpha_{\text{RCT}})|\Delta^1 > 0].$$

Since $w(\Delta, a)$ is Lipschitz-continuous in Δ and $w(0, a) = 0$, there exists $M > 0$ such that

$$L \leq M\mathbb{E}_h[(1 - \alpha_{\text{RCT}})\Delta^1|\Delta^1 > 0].$$

The proof of Proposition 4 implies that $\mathbb{E}_h[(1-\alpha_{\text{RCT}})\Delta^1 \ |\Delta^1 > 0]$ is bounded above by $\sqrt{\frac{2\log K}{N}}$. An identical argument holds in the case of $\Delta^1 < 0$, which yields Proposition 6. ■

C Simulations: The Tradeoffs of Rerandomization

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

C.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. These are mapped to outcomes according to a five-dimensional unknown parameter $\mu \in \mathbb{R}^5$:

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

Under correct belief h_0, parameter μ is drawn according to a five-dimensional truncated normal: $\mu \sim \prod_{k=1}^5 \mathcal{N}(0,1)_{|[-2,2]}$. The set of adversarial priors H consists of all doctrinaire priors corresponding to a fixed $\mu \in [-2,2]^5$. We denote by τ^* and α the Bayes optimal assignment of treatment and policy choice 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*

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

- *Loss under worst prior given Bayes optimal assignment*

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

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

$$\max_{\mu}\mathbb{E}_{x}\max_{\tau}\mathbb{E}\left[\max_{a\in\{0,1\}} u(p,a) - u(p,\alpha)\right]. \tag{C.3}$$

The ex ante Bayes expected loss (C.1) is essentially identical under randomization and rerandomization. Loss (C.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 (C.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 $N \geq 300$.

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

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

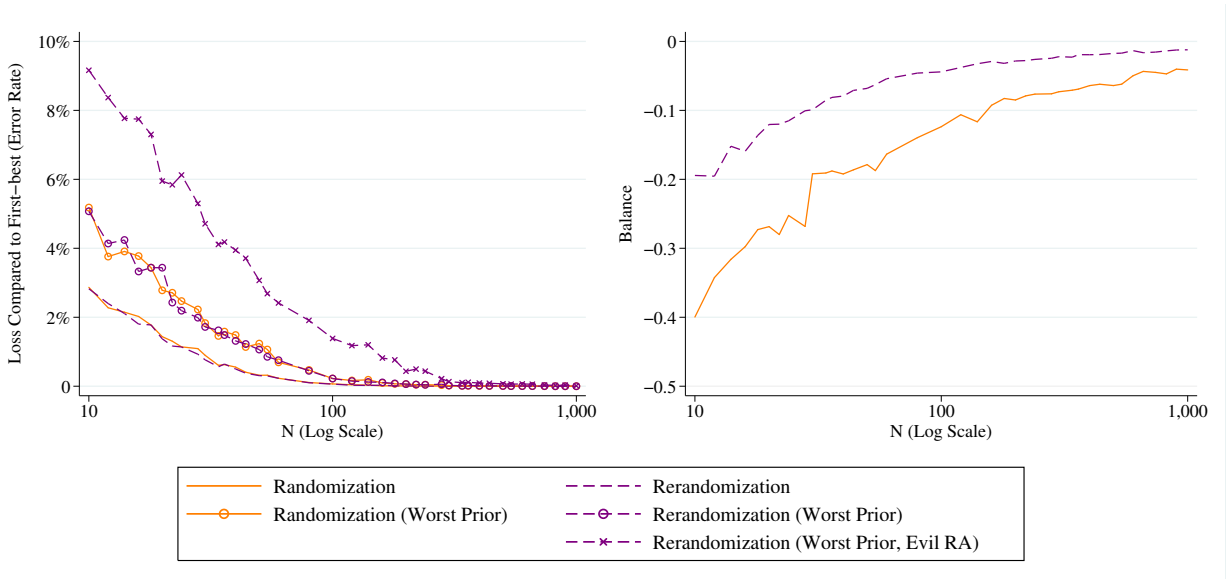
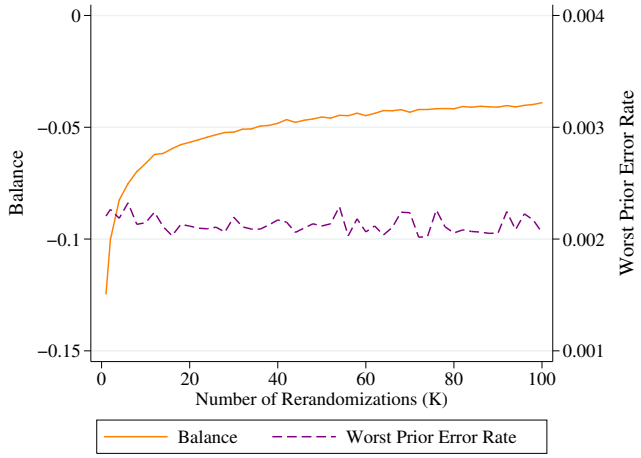


Figure C.2: Rerandomization increases balance with no robustness cost with fixed N.

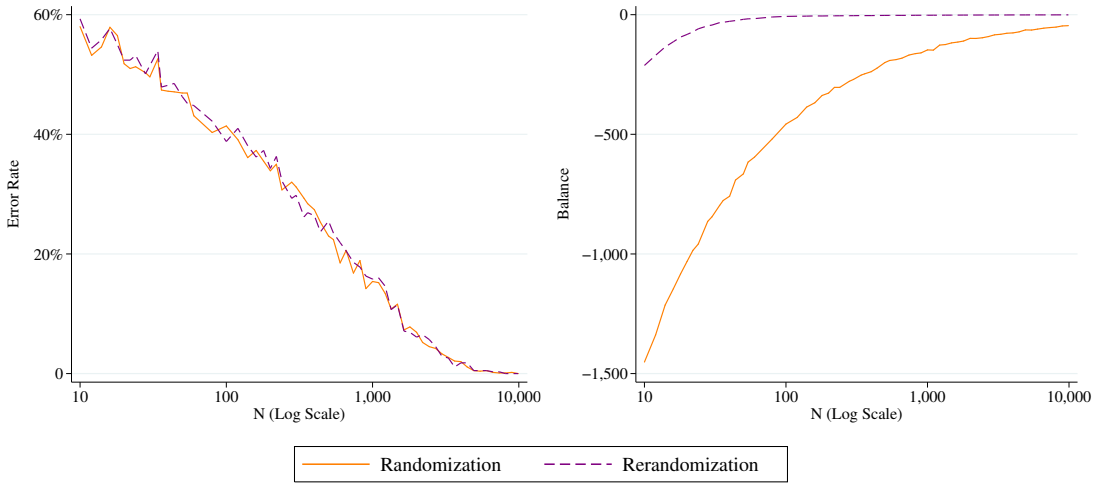


C.2 Poorly Behaved Treatment Effects

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

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



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.

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 C.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.

References

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

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

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

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

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

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

Bergemann, Dirk, and Juuso Välimäki. 2002. "Information Acquisition and Efficient Mechanism Design." *Econometrica*, 70(3): 1007–1033.

Bergemann, Dirk, and Juuso Välimäki. 2006. "Bandit Problems." Cowles Foundation paper no. 1292.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

McDiarmid, Colin. 1989. "On the Method of Bounded Differences." *Surveys in Combinatorics*, 141(1): 148–188.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Weerahandi, Samaradasa, and James V. Zidek. 1983. "Elements of Multi-Bayesian Decision Theory." *The Annals of Statistics*, 11(4): 1032–1046.

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

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