Online Appendix for "A Theory of Experimenters: Robustness, Randomization, and Balance"

By Abhijit V. Banerjee and Sylvain Chassang and Sergio Montero and Erik Snowberg

A. Balance as Subjective Expected Utility

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

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

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

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

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

Given a treatment assignment $(\tau_i)_{i\in\{1,\dots,N\}}$, let $\overline{x}^\tau \equiv \frac{2}{N}\sum_{i=1}^N x_i \mathbf{1}_{\tau_i=\tau} \in \mathbb{R}^k$ and $\phi \equiv \frac{2}{N}\sum_{i=1}^N (-1)^{1-\tau_i}\varepsilon_i$. We make the (asymptotically correct) assumption that ϕ is normally distributed with variance σ_ϕ^2. Under the empirical success rule, the subjective expected utility of the decision-maker under prior h_0 is

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

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

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

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

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

B. 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 IV do.

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

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

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

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

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

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

This differs significantly from a t-statistic: mean treatment effect $\hat{p}^1 - \hat{p}^0$ is

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

scaled by the difference of variances, rather than the sum of standard deviations. In particular, risk aversion means that the decision-maker values certainty (a small variance in outcomes) as well as a higher mean outcome. Greater standard deviation σ_0 makes action $a = 0$ less attractive, not more.

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

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

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

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

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

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

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

C. 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:

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

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

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

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

- we can pick $\mathcal{N}_0$ such that $\operatorname{argmax}_{e \in E} v(h, e) \subset \operatorname{argmax}_{e \in E} v(h_0, e)$ for all priors $h \in \mathcal{N}_0$;
- we can assume that for any experiment e, a fixed policy rule $\alpha(y, e)$ is uniquely optimal for all priors $h \in \mathcal{N}_0$.

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

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

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

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

where $\alpha^*_{h_0} \in \mathcal{A}$ denotes an optimal policy rule under h_0. Suppose $e \neq e'$ are both optimal under h_0. As $\mathcal{Y}$ is finite, by breaking indifferences in favor of one policy

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

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

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

and $f^\theta(p)^{\tau'}_{x'} = p^{\tau'}_{x'}$ for $(x', \tau') \neq (x, \tau)$. Let h^θ_1 be the distribution of transformed state $f^\theta(p)$ under h_1. As θ approaches 0, h^θ_1 approaches h_1 under the statistical distance. Notice that h^θ_1 garbles h_1 at (x, τ) alone, and does not change the expected performance of decision rules that depend on assignments $(x', \tau') \neq (x, \tau)$. Hence, it does not affect the value of experiment e'. We now show it must change the value of experiment e.

Let $v^\theta(e) \equiv \mathbb{E}_{h^\theta_1}\left[p^0 + \alpha^*_{h_1}(e, y)(p^1 - p^0)\middle| e\right]$ denote the value of experiment e for the fixed policy rule $\alpha^*_{h_1}$, evaluated at prior h^θ_1. For θ close to 0, the fact that the optimal policy does not change for priors in $\mathcal{N}_1$ implies that $v^\theta(e) = v(h^\theta_1, e)$. Note that $v^\theta(e)$ is a polynomial in θ. We show it is not constant. As e is optimal and all data points are valuable under h_1, sampling the pair (x, τ) is strictly valuable. Hence, there exists θ close enough to 1 such that $v^\theta(e) < v^0(e)$. As a non-zero polynomial has finitely many zeros, there exists θ arbitrarily close to 0 such that $v(h^\theta_1, e) \neq v(h_1, e) = v(h_1, e') = v(h^\theta_1, e')$. This proves the induction step.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

for all $\alpha \in \mathcal{A}$,

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

■

Proof of Proposition 3: Consider the set of optimal experiments

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

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

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

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

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

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

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

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

which implies Proposition 3. ■

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

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

experimental assignments $\{e_k,\ k \in \{1,\dots,K\}\}$, losses $L(p)$ from first-best satisfy

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

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

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

The mapping $x \mapsto x\exp(-Nx^2/4)$ is maximized at $x = \sqrt{\frac{2}{N}}$. As $K \geq 2$, we have $2\sqrt{\frac{\ln K}{N}} > \sqrt{\frac{2}{N}}$, which implies that both terms on the right-hand side of (2) are maximized at $p^1 - p^0 = 2\sqrt{\frac{\ln K}{N}}$. This implies that $L(p) \leq 2\sqrt{\frac{\ln K}{N}}$. An identical reasoning applies in the case where $p^1 - p^0 < 0$. ■

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

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

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

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

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

utility is

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

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

We now turn to point (ii). We use Proposition 7—established below—which implies that there exist randomized experiments leading to optimal decisions up to a penalty of order $O(1/\sqrt{N})$. This implies that the decision-maker can guarantee herself a payoff greater than $-O(1/\sqrt{N})$. We show this is not true when the decision-maker implements a deterministic experiment e. For $d \in (-1/2, 1/2)$, let $p(d)$ denote the state such that

$$\begin{array}{lll} p^0_x = \frac{1}{2} + d, & p^1_x = \frac{1}{2} & \text{if } \tau_x = 1; \\ p^0_x = \frac{1}{2}, & p^1_x = \frac{1}{2} - d & \text{if } \tau_x = 0. \end{array}$$

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, $\Delta^1_p = \overline{p}^1 - \overline{p}^0 = -d$. Hence, under prior h_e, regardless of the action a taken by the decision-maker, there is probability 0.5 that $\Delta^a_p = -\nu$ and probability 0.5 that $\Delta^a_p = +\nu$. As $w(p, a)$ is locally strictly concave in Δ^a_p around $\Delta^a_p = 0$, it follows that expected payoff from running experiment e under h_e is bounded below 0. This implies that for N large enough, randomized experiments are strictly optimal. ■

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

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

D. Simulations

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

D.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 covari-

ates x, Assumption 1 does not hold, and the losses from running a deterministic experiment maximizing balance vanish as the sample size grows large.

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

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

Under correct belief h_0, each parameter μ_τ is independently drawn according to a five-dimensional truncated normal: $\mu_\tau \sim \prod_{k=1}^{5} \mathcal{N}(0,1)_{|[-3,3]}$. The set of adversarial priors H consists of all doctrinaire priors corresponding to fixed values $\mu_\tau \in [-3,3]^5$. We denote by α^* the Bayes optimal policy rule under this model.

We consider $\mathcal{E} = \mathcal{E}_{\text{RCT}}$ or $\mathcal{E} = \mathcal{E}_K$, the rerandomized experiment 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 experimental treatment assignments.

- *Bayes loss given Bayes optimal policy rule:*

$$\mathbb{E}_{h_0,x,\mathcal{E}}\left[\max_{a\in\{0,1\}} u(p,a) - u(p,\alpha^*)\right]. \tag{1}$$

- *Loss under worst prior given Bayes optimal policy rule:*

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

- *Loss under worst prior, and worst assignment τ, given Bayes optimal policy rule:*

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

As Figure D1 shows, the ex ante Bayes expected loss (1) is essentially identical under randomization and rerandomization. Loss measure (2) chooses the prior that maximizes first-best losses 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 (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

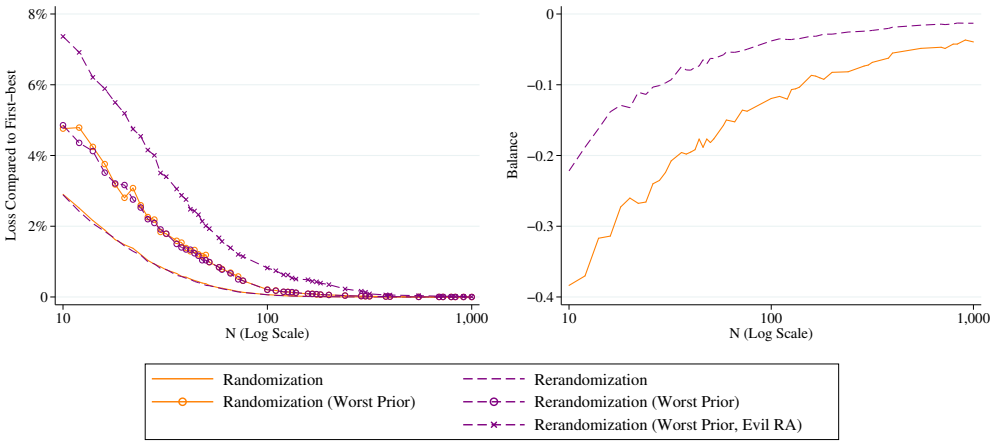

FIGURE D1. RERANDOMIZATION SUBSTANTIALLY INCREASES BALANCE WITH NO COST TO ROBUSTNESS.

substantially increases expected losses. However, it is important to note even under this highly unrealistic scenario—the evil RA must know the data-generating process—losses are under 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 D2 shows that balance improves substantially with K, especially for the first 20 rerandomizations, while worst-prior expected losses remain essentially flat.

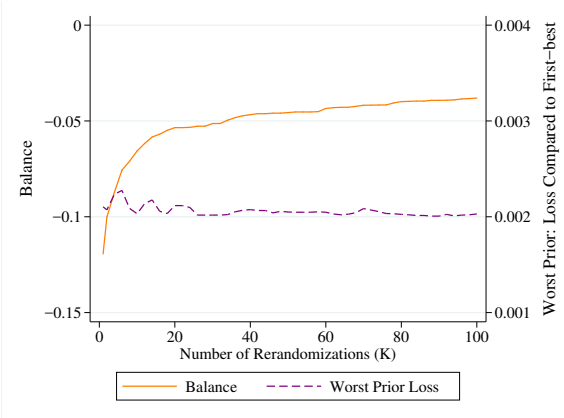

FIGURE D2. RERANDOMIZATION INCREASES BALANCE WITH NO ROBUSTNESS COST WITH FIXED N.

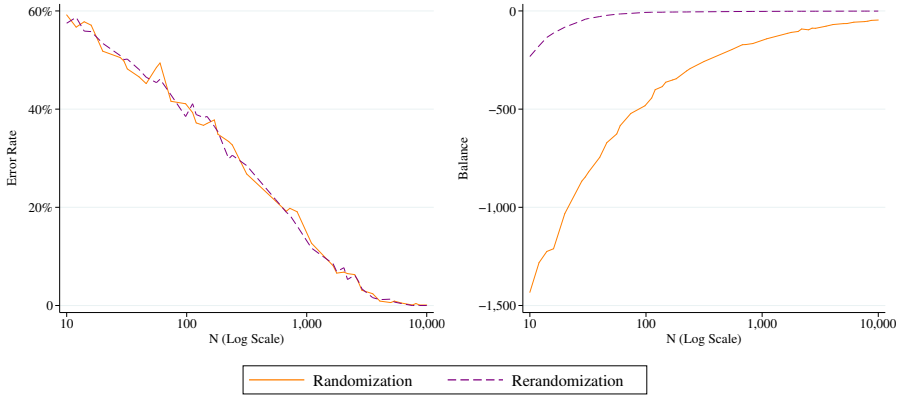

FIGURE D3. RERANDOMIZATION SUBSTANTIALLY INCREASES BALANCE WITH NO COST TO ROBUSTNESS.

D.2. Poorly Behaved Treatment Effects

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

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

Specifically, we set $X = \{1, 2, \ldots, 10^4\}$ and

$$p_x^0 = \begin{cases} \frac{1}{5} & \text{if } x \text{ is odd,} \\ \frac{1}{2} & \text{if } x \text{ is even,} \end{cases} \qquad p_x^1 = \begin{cases} \frac{4}{5} & \text{if } x \text{ is odd,} \\ \frac{1}{4} & \text{if } x \text{ is even.} \end{cases}$$

Even covariates are twice as likely as odd covariates so that, on aggregate, $u(p, 1) = \frac{13}{30} > \frac{2}{5} = u(p, 0)$, and 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 D3 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

Blackwell, David. 1951. "Comparison of Experiments." In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Vol. 1, 93–102. Berkeley, CA:University of California Press.

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.

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

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

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.

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