Supporting Information for:

Adaptive Experimental Design:

Prospects and Applications in Political Science

Molly Offer-Westort, Alexander Coppock, and Donald P. Green

February 11, 2021

A Algorithms

We suppose that K arms have unknown success rates $\theta_1, \ldots, \theta_K$, following their respective Bernoulli distributions, with likelihoods

$$f_{X_1|\Theta_1}(x_1|\theta_1), \ldots, f_{X_K|\Theta_K}(x_K|\theta_K).$$

Posteriors follow Beta distributions with parameters $\alpha_{k,t}, \beta_{k,t}$.

Algorithm 1: Batch-wise Thompson sampling

1: Initialize priors such that $(\alpha_{k,1} = 1, \beta_{k,1} = 1)$ for $k = 1, \ldots, K$.

For periods $t = 1, \ldots, T$:

2: Calculate $p_{k,t} = P\left[\Theta_k = \max_k\{\Theta_1, \ldots, \Theta_K\} | (\alpha_{1,t}, \beta_{1,t}), \ldots, (\alpha_{K,t}, \beta_{K,t})\right]$ for $k = 1, \ldots, K$.[1]

3: Sample n observations, assigning treatment with probabilities $(p_{1,t}, \ldots, p_{K,t})$.

4: Update posteriors, for $k = 1, \ldots, K$:

$$\alpha_{k,t+1} = \alpha_{k,t} + \# \text{ successes observed for arm } k \text{ in period } t,$$
$$\beta_{k,t+1} = \beta_{k,t} + \# \text{ failures observed for arm } k \text{ in period } t.$$

Algorithm 2: Control-augmented Thompson sampling

1: Let C index the control arm. Initialize priors such that $(\alpha_{k,1} = 1, \beta_{k,1} = 1)$ for $k \neq C$.

For periods $t = 1, \ldots, T$:

2: Calculate $p_{k,t}$ as above in step 2, excluding C.

[1]See C.2 for a worked mathematical example; alternatively, the relative probabilities can be determined by sampling from the implied posterior distributions.

3: Retrieve the "current best arm," and calculate the difference between the cumulative sample assigned to that arm and the control arm:

$$b = \underset{k}{\text{argmax}}\ p_{k,t},$$

$$d = n_{b,t} - n_{C,t}.$$

4: Calculate the proportion of the next batch needed for the control to match the cumulative sample of the "best" arm, up to a maximum of Z_t of the batch, where $Z_t \in (0, 1)$ may be fixed, or may be data adaptive:

$$q = \min(\max(d/n, 0), Z_t).$$

5: The probability of assignment to the control condition is a combination of the allocation to match the cumulative sample of the control to the current best arm, and R_t of the remaining probability, for $R_t \in (0, 1)$. This value may be fixed, or may also be data adaptive:

$$\tilde{p}_{C,t} = q + R_t * (1 - q).$$

6: The treatment arms are assigned according to their posterior probabilities, scaled to the remaining sampling probability, for $k \neq C$:

$$\tilde{p}_{k,t} = p_{k,t} * (1 - R_t) * (1 - q).$$

7: Sample n observations, assigning treatment with probabilities $(\tilde{p}_{1,t}, \ldots, \tilde{p}_{C,t}, \ldots, \tilde{p}_{K,t})$.

8: Update posteriors, for $k \neq C$:

$$\alpha_{k,t+1} = \alpha_{k,t} + \#\text{ successes observed for arm } k \text{ in period } t,$$

$$\beta_{k,t+1} = \beta_{k,t} + \#\text{ failures observed for arm } k \text{ in period } t.$$

In simulations and experiments, we set Z_t in step 4 as fixed at .90, and R_t in steps 5 and 6 as fixed at 1/K. However, these values may be revised by the researcher; in particular, in relatively small samples, we may want to allow the control arm to "catch up" to the best treatment arm to facilitate approximate balance, even as the arm we identify as "best" changes across batches. After early fluctuations, researchers may wish to let Z_t approach zero and R_t approach a fixed value, so that the portion of the batch allocated to allow the control to catch up (step 4) shrinks to zero, and the remaining probability assigned to the control (step 5) approaches R. The optimal rate of decay in particular for small sample performance, however, will depend on the context of the experiment.

Furthermore, for simplicity, we have assumed equal variance of outcomes under the best treatment arm and the control arm, in which case the optimal allocation of treatment is balanced between the best treatment arm and the control. Letting Z_t approach zero and R_t approach 2, our algorithm will approach this optimal allocation. However, if the variances are not equal, under, e.g., a Hájek-style estimator (Hájek 1971), variance of the estimator can be minimized by assigning treatment proportional to the relative standard deviations of the two arms. Under a Horvitz-Thompson-style estimator, (Horvitz and Thompson 1952), with simple random assignment, optimal assignment is a function of the raw moments.[2] Robbins (1952) proposes a two-stage sequential design to learn these values and assign treatment according to them; Dimmery (2018) extends this approach to an algorithm that learns and updates continuously. Here, R_t could be learned adaptively in a similar manner.

[2]We thank Peter Aronow for discussion on these points, among many others.

B Estimation and theory

B.1 Alternative estimation procedures

For estimation and hypothesis testing, we have assumed that there is a unique best arm. When this is not the case, hypothesis testing becomes more challenging; more general methods for inference on adaptively collected data have been proposed by Hadad et al. (2019), Deshpande et al. (2020) and Zhang et al. (2020). The estimators described here rely on martingale central limit theorems for asymptotic consistency.

Hadad et al. (2019) propose an augmented version of the IPW estimator, which allows for an optional conditional means model with *evaluation weights* that, under specified conditions, achieves asymptotic normality for adaptively weighted estimates even in the no-signal setting. When there *is* a unique best arm, however, these necessary conditions for asymptotic normality of IPW-type estimators may hold in the absence of evaluation weights.

Deshpande et al. (2020) propose a *W-decorrelation* estimator, under which they augment the standard OLS estimate with a decorrelation matrix. This matrix is a function of a tuned regularization parameter, which trades off bias and variance.

Zhang et al. (2020) propose a "batched" OLS hypothesis testing procedure, noting the asymptotic normality of OLS estimates within batches. They construct a test statistic by combining batch-wise t-statistics and compare it to a simulated null distribution in order to obtain a p-value. The authors demonstrate that this testing procedure has reliable Type-1 error control in small samples and is robust to non-stationarity.

Even with a unique best arm, our simulations and those presented elsewhere have demonstrated the possibility that in small samples, even robust confidence intervals may undercover. The "sufficient" experiment size for valid coverage will depend on the value of the best arm relative to the other arms, or the signal-to-noise ratio (Zhang et al. 2020). When in doubt, researchers have several options: simulations tailored to a specific application may be helpful in determining whether undercoverage is a concern; to ensure valid coverage for small samples, researchers may use conservative methods, such as uniformly valid confidence sequences (Howard et al. 2018); or they may use the "batched" OLS procedure proposed by Zhang et al. (2020).

B.2 Finite-n unbiasedness of the Horvitz-Thompson estimator under adaptive assignment

We consider estimation using a Horvitz-Thompson-style estimator (Horvitz and Thompson 1952). We demonstrate the unbiasedness of the estimator here, but see Bowden and Trippa (2017) for further investigation of this and other estimators in the adaptive setting. Let S_i be the state, defined by the history of treatment and outcomes observed prior to assigning treatment to observation i. This is to account for the differences in treatment assignment probabilities across runs of the experiment. In batch-wise designs, this history will be the same for all observations within a given batch.

For treatment K_i, define $\pi_i(k; S_i) = \Pr[K_i = k | S_i]$.

$$\hat{\mu}_k^{HT} = \frac{1}{N} \sum_{i=1}^{N} Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)}$$

By design, $\pi_i(k; S_i) > 0, \forall i, k$.

Under the potential outcomes framework: $\mathrm{E}\left[\hat{\mu}_k^{HT}\right] = \mathrm{E}\left[Y_i(k)\right].$

We require only independence of potential outcomes and treatment conditional on history,

$Y_i(k) \perp\!\!\!\perp K_i|S_i$, which is given by the experimental design.

$$\begin{aligned}
\mathrm{E}\left[\hat{\mu}_k^{HT}\right] &= \mathrm{E}\left[\frac{1}{N}\sum_{i=1}^{N} Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)}\right] \\
&= \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left[Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)}\right].
\end{aligned}$$

Considering the i^{th} unit, by the Law of Iterated Expectations,

$$\mathrm{E}\left[Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)}\right] = \mathrm{E}\left[\mathrm{E}\left[Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)}\middle| S_i\right]\right]$$

Taking the interior term, $\mathrm{E}\left[Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)}\middle| S_i\right]$, by definition,

$$= \mathrm{E}\left[\frac{Y_i \mathbb{1}\{K_i = k\}}{\Pr[K_i = k|S_i]}\middle| S_i\right]$$

By the potential outcomes model,

$$= \mathrm{E}\left[Y_i(k) \times \frac{\mathbb{1}\{K_i = k\}}{\Pr[K_i = k|S_i]}\middle| S_i\right]$$

And because $Y_i \perp\!\!\!\perp K_i|S_i$,

$$\begin{aligned}
&= \mathrm{E}\left[Y_i(k)|S_i\right] \times \mathrm{E}\left[\frac{\mathbb{1}\{K_i = k\}}{\Pr[K_i = k|S_i]}\middle| S_i\right] \\
&= \mathrm{E}\left[Y_i(k)|S_i\right] \times \frac{\mathrm{E}\left[\mathbb{1}\{K_i = k\}|S_i\right]}{\Pr[K_i = k|S_i]} \\
&= \mathrm{E}\left[Y_i(k)|S_i\right] \times \frac{\Pr[K_i = k|S_i]}{\Pr[K_i = k|S_i]} \\
&= \mathrm{E}\left[Y_i(k)|S_i\right].
\end{aligned}$$

Then returning to the Law of Iterated Expectations from above,

$$\mathrm{E}\left[Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)}\right] = \mathrm{E}\left[\mathrm{E}\left[Y_i(k)|S_i\right]\right] = \mathrm{E}\left[Y_i(k)\right].$$

B.3 Theoretical properties of Control-augmented Thompson sampling

The objective of our algorithm is to optimally allocate observations to treatment and control conditions for estimating the average treatment effect with respect to the best treatment arm and a pre-specified control condition. We note that there are tradeoffs: by using an adaptive algorithm as discussed here, we prioritize assignment and estimation for the best arm, at the expense of sub-optimal arms. When we use the control-augmented version of the adaptive algorithm, we are further reducing some of the sample assigned to the best arm, and estimate the best arm mean somewhat less precisely in favor of targeting the difference between the best arm and the control.

We will consider the Horvitz-Thompson estimator, presented above. For simplicity we assume here that batches are of size 1. Performance of adaptive algorithms with larger batches is addressed in Jun et al. (2016) and Perchet et al. (2016).

The Horvitz-Thompson estimator of the ATE for arm k, τ_k (letting 0 signify the control condition) is,

$$\hat{\tau}_k^{HT} = \frac{1}{N}\sum_{i=1}^{N} Y_i \frac{\mathbb{1}\{K_i = k\}}{\pi_i(k; S_i)} - \frac{1}{N}\sum_{i=1}^{N} Y_i \frac{\mathbb{1}\{K_i = 0\}}{\pi_i(0; S_i)}.$$

We require again independence of potential outcomes and treatment conditional on history, $Y_i(k) \perp\!\!\!\perp K_i|S_i$, which is given by the experimental design. Let us assume constant means of the form $\Pr[Y(k) = 1]$ for all arms with the associated Bernoulli variances. If we knew which treatment arm was best and the arms' means ex-ante, we would simply fix probabilities at the start of the experiment to minimize variance. However, if this is not the case, we must conduct some amount of exploration to also learn which arm is best.

For the adaptive experiment, the probabilities $\pi_i(k; S_i)$ are functions of the random variable S_i. Conditional on history, the variance can be expressed as,

$$\text{Var}\left[\hat{\tau}_k^{HT}\middle|S_i\right] = \frac{1}{N^2}\sum_{i=1}^{N}\text{Var}\left[Y_i(k)\frac{\mathbb{1}\{K_i=k\}}{\pi_i(k;S_i)}\middle|S_i\right] + \frac{1}{N^2}\sum_{i=1}^{N}\text{Var}\left[Y_i(0)\frac{\mathbb{1}\{K_i=0\}}{\pi_i(0;S_i)}\middle|S_i\right]$$

Because we can treat the probabilities as fixed, conditional on history,

$$=\frac{1}{N^2}\sum_{i=1}^{N}\frac{\text{Var}\left[Y_i(k)\mathbb{1}\{K_i=k\}|S_i\right]}{\Pr[K_i=k|S_i]^2} + \frac{1}{N^2}\sum_{i=1}^{N}\frac{\text{Var}\left[Y_i(0)\mathbb{1}\{K_i=0\}|S_i\right]}{\Pr[K_i=0|S_i]^2}$$

The $Y_i(k)$ and $\mathbb{1}\{K_i = k\}$ terms are independent binomials, and the product of two binomials is also a binomial, and so we express the variance of their product accordingly.

$$\begin{aligned}=&\frac{1}{N^2}\sum_{i=1}^{N}\frac{\Pr[Y_i(k)=1|S_i]\Pr[K_i=k|S_i]\times\Big(1-\Pr[Y_i(k)=1|S_i]\Pr[K_i=k|S_i]\Big)}{\Pr[K_i=k|S_i]^2}\\ &+\frac{1}{N^2}\sum_{i=1}^{N}\frac{\Pr[Y_i(0)=1|S_i]\Pr[K_i=0|S_i]\times\Big(1-\Pr[Y_i(0)=1|S_i]\Pr[K_i=0|S_i]\Big)}{\Pr[K_i=0|S_i]^2}\end{aligned}$$

Simplifying, we cancel terms and drop the conditioning on S_i for the Y_i

$$\begin{aligned}=&\frac{1}{N^2}\sum_{i=1}^{N}\frac{\Pr[Y_i(k)=1]\times\Big(1-\Pr[Y_i(k)=1]\Pr[K_i=k|S_i]\Big)}{\Pr[K_i=k|S_i]}\\ &+\frac{1}{N^2}\sum_{i=1}^{N}\frac{\Pr[Y_i(0)=1]\times\Big(1-\Pr[Y_i(0)=1]\Pr[K_i=0|S_i]\Big)}{\Pr[K_i=0|S_i]}.\end{aligned}$$

To minimize this variance, we should like each of the $\pi_i(k^*; S_i)$ for the best arm k^* and $\pi_i(0; S_i)$ terms to be proportional to the relative square root of the success rates.

Under the static design, as the size of the experiment grows, regardless of the probability of being best that we associate with a given arm, we will never exploit that arm more than any other, and will assign all arms with probability $1/(K+1)$. Consequently, in the adaptive design, if we are able to approach $\pi_i(k^*; S_i)$ and $\pi_i(0; S_i)$ equal to 1/2, with appropriate bounding, the asymptotic variance will be smaller than under the static case for $K > 1$.

Of course, performance in finite samples will depend on the rate at which we approach these assignment probabilities, which will be a function of both the number of arms and the distribution of outcomes under the various arms, here parameterized by the mean. We may also wish to impose probability floors (similar to Dimakopoulou et al. 2017; Zhang et al. 2020) to avoid extreme weights. As we see in simulations in SI D.2, the performance of the control-augmented algorithm relative to the static algorithm in finite samples will be better when the difference between the best arm mean and the second best arm mean is larger.

To provide some understanding for the rate at which we will approach optimal assignment, we refer to Agrawal and Goyal (2017)'s regret bounds for standard online Thompson sampling with Beta priors, where expected total regret is the difference in total expected rewards under optimal assignment as compared to assignment under a given algorithm. Optimal regret is zero. They show regret as a function of the expected sum of assignments to suboptimal arms, multiplied by the difference between the best arm mean and the mean under ams $k \neq k^*$. For our objectives, we should like to see the component that is expected sum of assignments to suboptimal arms to grow more slowly than N. Extrapolating from the proof of Theorem 1.2, we can see the expected sum of all plays of suboptimal arms under Thompson sampling is $O(\ln(N))$, as compared to under a random design, where it is $O(N)$.

Finally, we have used Thompson sampling throughout for its heuristic interpretation and well-studied performance. However, our control-augmented approach could be paired with many other adaptive algorithms for best arm selection. Indeed, since Thompson sampling is designed to minimize regret and not speed best arm identification, it may not approach the optimal arm assignment at the fastest rate. Russo (2016) proposes alternative algorithms which will be optimal in this goal with appropriate parameter tuning.

C Worked Examples

C.1 Beta distribution as conjugate prior to the Binomial distribution

For a review of Bayesian inference, see Wasserman (2013, chapter 11). The below case is given as example 11.1 in the text. Consider a random variable X, which follows a Bernoulli distribution parameterized by an unknown Θ. We propose a uniform prior over the possible values of Θ, i.e., the parameter is distributed Beta$(1, 1)$, with density

$$f_\Theta(\theta) = 1,\ 0 \leq \theta \leq 1.$$

Our data consists of n i.i.d. observations, $X^{\{n\}} = (X_1, \ldots X_n)$. Using Bayes rule, the distribution of the parameter Θ given the data $X^{\{n\}}$ is,

$$f_{\Theta|X^{\{n\}}}(\theta|x^{\{n\}}) = \frac{f_{X^{\{n\}}|\Theta}(x^{\{n\}}|\theta) f_\Theta(\theta)}{\int f_{X^{\{n\}}|\Theta}(x^{\{n\}}|\theta) f_\Theta(\theta) d\theta}.$$

Plugging in the total likelihood and the prior,

$$f_{\Theta|X^{\{n\}}}(\theta|x^{\{n\}}) \propto \theta^{\left(\sum_{i=1}^n x_i\right)} (1-\theta)^{\left(n - \sum_{i=1}^n x_i\right)},\quad 0 < \theta < 1.$$

The posterior follows a Beta distribution, with parameter values $\alpha = \sum_{i=1}^n x_i + 1$, and $\beta = n - \sum_{i=1}^n x_i + 1$. That is, the α parameter is the number of observed successes plus 1, and the β parameter is the number of observed failures plus 1.

C.2 Naive estimation under Thompson sampling

As an example, consider an experiment where we are comparing two treatment arms, with success rate Θ_1 and Θ_2. The true, but unknown values of these parameters are both 0.5.

Again we set the prior for both arms as uniform over the parameter space, i.e., Beta$(1,1)$. We require one observation from each arm in the first period. We then assign treatment to one observation each subsequent period, with assignment probabilities proportional to the probability that each arm is best, updating after each period.

Let $n_{k,t}$ be the number of trials observed for arm k up to and including period t, and let $X_k^{\{n_{k,t}\}} = (X_{[1]k}, \ldots, X_{[n_{k,t}]k})$ be the vector of responses under treatment arm k observed up until and including time t. Thus at time t, posteriors follow Beta distributions with parameters $\alpha_{k,t} = \sum_{i=1}^{n_{k,t}} x_{[i]k} + 1$ and $\beta_{k,t} = n_{k,t} - \sum_{i=1}^{n_{k,t}} x_{[i]k} + 1$.

Suppose that in the first period, we observe one success from arm one and one failure from arm two. The posterior for Θ_1 is now Beta$(2,1)$, while the posterior for Θ_2 is Beta$(1,2)$.

Based on these posteriors, we calculate the probability that each arm is best.

$$
\begin{aligned}
&P\left(\Theta_1 \geq \Theta_2 \middle| X_1^{\{n_{1,t}\}} = x_1^{\{n_{1,t}\}}, X_2^{\{n_{2,t}\}} = x_2^{\{n_{2,t}\}}\right) \\
&\qquad = \int_{\theta_1=0}^{1} \int_{\theta_2=0}^{\theta_1} f_{\Theta_2|X_2^{\{n_{2,t}\}}}(\theta_2|x_2^{\{n_{2,t}\}}) f_{\Theta_1|X_1^{\{n_{1,t}\}}}(\theta_1|x_1^{\{n_{1,t}\}}) d\theta_2 d\theta_1 \\
&\qquad = \int_{\theta_1=0}^{1} c_{1,t}{\theta_1}^{\alpha_{1,t}-1}(1-\theta_1)^{\beta_{1,t}-1} \int_{\theta_2=0}^{\theta_1} c_{2,t}{\theta_2}^{\alpha_{1,t}-1}(1-\theta_2)^{\beta_{1,t}-1} d\theta_2 d\theta_1
\end{aligned}
$$

where $c_{k,t}$ represents the normalization constant for the Beta distribution, $\dfrac{\Gamma(\alpha_{k,t}+\beta_{k,t})}{\Gamma(\alpha_{k,t})\Gamma(\beta_{k,t})}$.

At time $t = 1$,

$$
\begin{aligned}
&= \frac{\Gamma(3)}{\Gamma(2)\Gamma(1)} \frac{\Gamma(3)}{\Gamma(1)\Gamma(2)} \int_{\theta_1=0}^{1} {\theta_1}^1 (1-\theta_1)^0 \int_{\theta_2=0}^{\theta_1} {\theta_2}^0 (1-\theta_2)^1 d\theta_2 d\theta_1 \\
&= 4 \int_{\theta_1=0}^{1} \theta_1 \left(\int_{\theta_2=0}^{\theta_1} (1-\theta_2) d\theta_2 \right) d\theta_1 \\
&= 4 \int_{\theta_1=0}^{1} \theta_1 \left(\theta_1 - \frac{{\theta_1}^2}{2} \right) d\theta_1 \\
&= \frac{5}{6}.
\end{aligned}
$$

That is, having seen one success from arm one and one failure from arm two, we find that the probability that arm one is the best arm is 5/6, and the probability that arm two is the best is 1/6.

In period two, we would assign arm one again with probability 5/6 and arm two with probability 1/6. This preferential assignment probability conditional on observed success or failure introduces a source of bias into estimation. To see this, consider all possible states we could arrive at after two periods, represented in Table C.1. A state is defined by the number of times we see each treatment and the number of successes we observe under each treatment,

$$s = \{n_{1,2}, n_{2,2}, |x_1^{\{n_{1,2}\}}|, |x_2^{\{n_{2,2}\}}|\}.$$

The sample mean for arm k after two periods is defined as,

$$\widehat{\theta}_k = \frac{\sum_{i=1}^{n_{k,2}} x_{[i]k}}{n_{k,2}}.$$

If we were to arrive at all states with equal probability, the sample means for each arm would coincide with the true means in expectation. However, we do not see each state with equal probability. If we take this into account, the sample mean we would observe for either

Table C.1: Possible States after Two Periods

$n_{1,2}$	$n_{2,2}$	$\lvert x_1^{\{n_{1,2}\}}\rvert$	$\lvert x_2^{\{n_{2,2}\}}\rvert$	$P(\Theta_1 \geq \Theta_2 \vert\, s)$	$P(s)$
2	1	2	1	0.60	0.063
2	1	1	1	0.30	0.083
1	2	1	1	0.70	0.083
1	2	1	2	0.40	0.063
2	1	2	0	0.90	0.104
2	1	1	0	0.70	0.167
1	2	1	0	0.90	0.021
2	1	0	1	0.10	0.021
1	2	0	2	0.10	0.104
1	2	0	1	0.30	0.167
2	1	0	0	0.40	0.063
1	2	0	0	0.60	0.063

arm is, in expectation, only 0.458.

$$\begin{aligned} \mathrm{E}\left[\widehat{\theta}_k\right] &= \sum_s \widehat{\theta}_k | s \cdot P(s) \\ &\approx 0.458 \end{aligned}$$

If we condition on the best observed arm, however, the expected sample mean of the best arm is 0.708, although the true parameter value for either arm is 0.5.

$$\begin{aligned} \mathrm{E}\left[\widehat{\theta}_{\max}\right] &= \sum_s \widehat{\theta}_{\max} | s \cdot P(s) \\ &\approx 0.708 \end{aligned}$$

As we observe more outcomes under each arm, we would expect the sample mean to regress to the true mean. However, the differing sampling probabilities imply that arms that initially under-perform by chance will converge at a slower rate than those that initially over-perform by chance.

By using inverse probability weights, however, we are able to fully account for the different sampling probabilities, and the IPW estimate for either arm is 0.5 in expectation.

C.3 Control-augmented batch-wise Thompson sampling

Consider an experiment where we are comparing two treatment arms, with success rates Θ_{T1} and Θ_{T2}, and a control arm, with success rate Θ_C. The true, but unknown values of treatment parameters are 0.75 and 0.5 respectively, and the control parameter value is 0.25. Again we set the prior for all arms as uniform over the parameter space, i.e., Beta$(1,1)$.

We will consider a batched design, with nine observations per period. In the first period, we require three observations from each arm. We then assign treatment according to control-augmented Thompson sampling in the two following periods.

Suppose in the first period we observe three successes for arm T1, one success and two failures for arm T2, and three failures for arm C. The posteriors are then Beta$(4,1)$ for Θ_{T1}, Beta$(2,3)$ for Θ_{T2}, and Beta$(1,4)$ for Θ_C.

We then calculate the posterior probability that arm T1 is the best treatment arm as,

$$\begin{aligned}
&P\left(\Theta_{T1} \geq \Theta_{T2} \middle| X_{T1}^{\{3\}} = x_{T1}^{\{3\}}, X_{T2}^{\{3\}} = x_{T2}^{\{3\}}\right) \\
&\qquad = \frac{\Gamma(5)}{\Gamma(4)\Gamma(1)} \frac{\Gamma(5)}{\Gamma(2)\Gamma(3)} \int_{\theta_{T1}=0}^{1} \theta_{T1}^{3} \int_{\theta_{T2}=0}^{\theta_{T1}} \theta_{T2}(1-\theta_{T2})^2 d\theta_{T2} d\theta_{T1} \\
&\qquad = 0.929.
\end{aligned}$$

The posterior probability for T2 must then be 0.071.

The next step in our algorithm is to identify the treatment condition with the largest posterior probability, which here, is clearly T1, and to find the difference in cumulative sample between that arm and the control condition. As we have fixed our sample for this first period, that difference is 0.

We then assign the control probability to be $\tilde{p}_{C,2} = 1/3$. For the treatment arms, the probabilities are proportional to the posterior probabilities scaled so that all probabilities sum to 1:

$$\tilde{p}_{T1,2} = 0.929 * \frac{2}{3} = 0.619,$$
$$\tilde{p}_{T2,2} = 0.071 * \frac{2}{3} = 0.048.$$

Suppose that in period two based on the sampling probabilities ($\tilde{p}_{T1,2},\ \tilde{p}_{T2,2},\ \tilde{p}_{2,C}$), from our period size of nine, we assign six observations to arm T1, one to arm T2, and two to the control. We observe three successes and three failures for T1, one success for T2, and one success and one failure for the control.

The posterior probabilities for arms T1 and T2 are then 0.713 and 0.287 respectively. (Note that we must take into account the *cumulative* successes and failures). The treatment condition with the largest posterior probability is still T1. However, we now have a difference in the cumulative sample between the best arm and the control condition; we have assigned a cumulative sample of nine to arm T1, and five to the control, for a difference of four.

The proportion of the next batch needed for the control to match the cumulative sample of the best arm is 4/9 (below the probability cap of .9). So the control probability is 4/9 plus 1/3 of the remaining probability,

$$\tilde{p}_{C,3} = \frac{4}{9} + \frac{1}{3} * \frac{5}{9} = 0.630,$$

For the treatment conditions, the probabilities are proportional to the posterior probabilities, scaled so that all probabilities sum to 1:

$$\tilde{p}_{T1,3} = 0.713 * \frac{2}{3} * \frac{5}{9} = 0.264$$
$$\tilde{p}_{T2,3} = 0.287 * \frac{2}{3} * \frac{5}{9} = 0.106.$$

We would then use these sampling probabilities in the third and final period.

D Additional Simulations

D.1 Varying number of batches

We replicate simulations presented in Table 1 with 1,000 total observations and varying number of batches. Most of the gains are realized by five total batches, with diminishing marginal returns thereafter.

In practice, the frequency of updating and batch size will depend on the experimental setting. For online survey experiments, it may be more convenient to update each day, and batch size will relate to response rates on a given platform. In clinical trials or field experiments, the time between intervention and outcome measurement may be longer, requiring larger batches and fewer updates when the duration of the entire experiment is limited. When there is flexibility around the frequency of updating and batch size, the researcher may consider tradeoffs in costs to decreasing batch size with improvements in algorithm performance (Perchet et al. 2016; Gao et al. 2019).

Figure D.1: Iterated Simulation Statistics, Varying Number of Batches, Arm selection and coverage

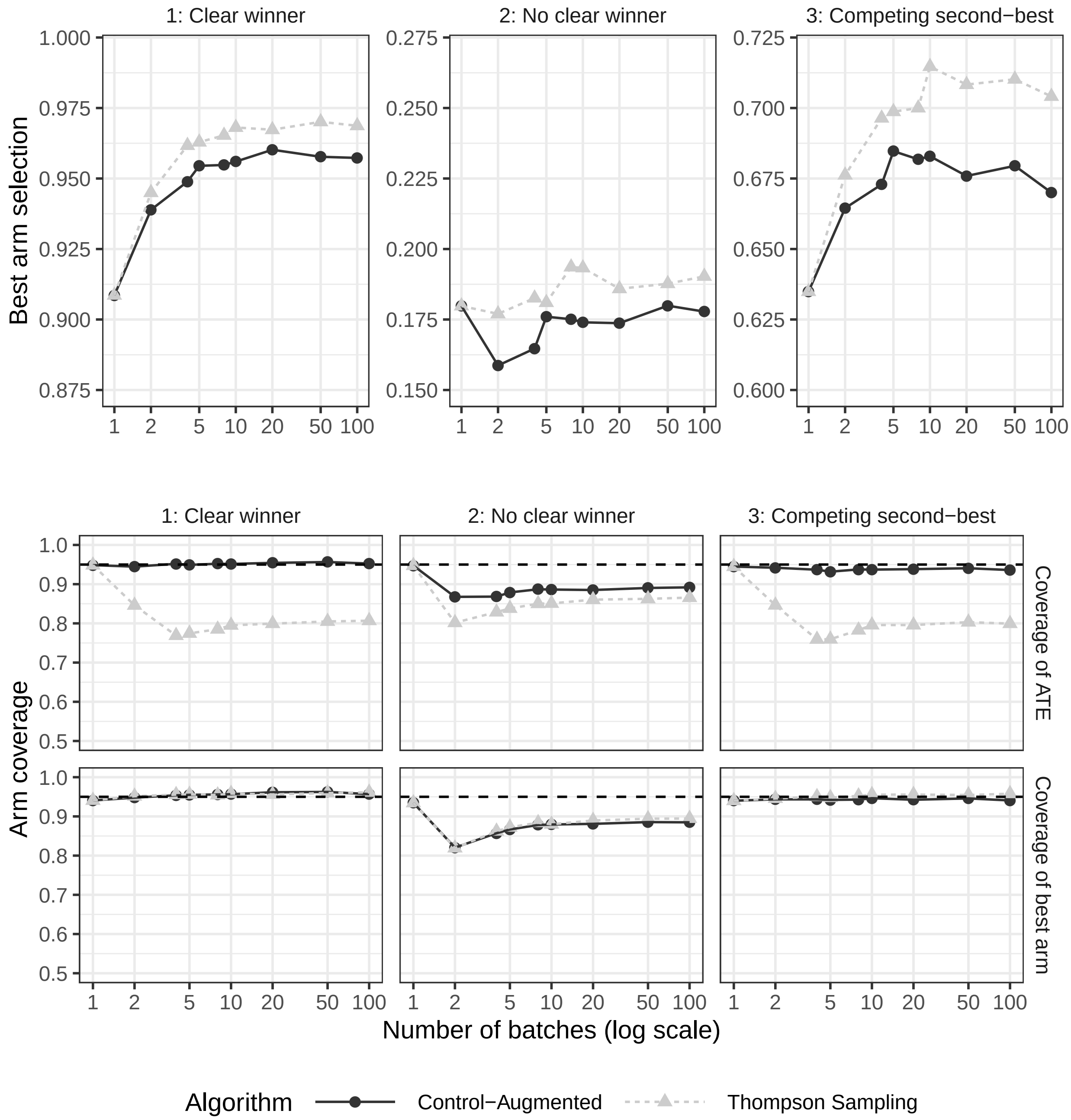


Figure D.2: Iterated Simulation Statistics, Varying Number of Batches, RMSE

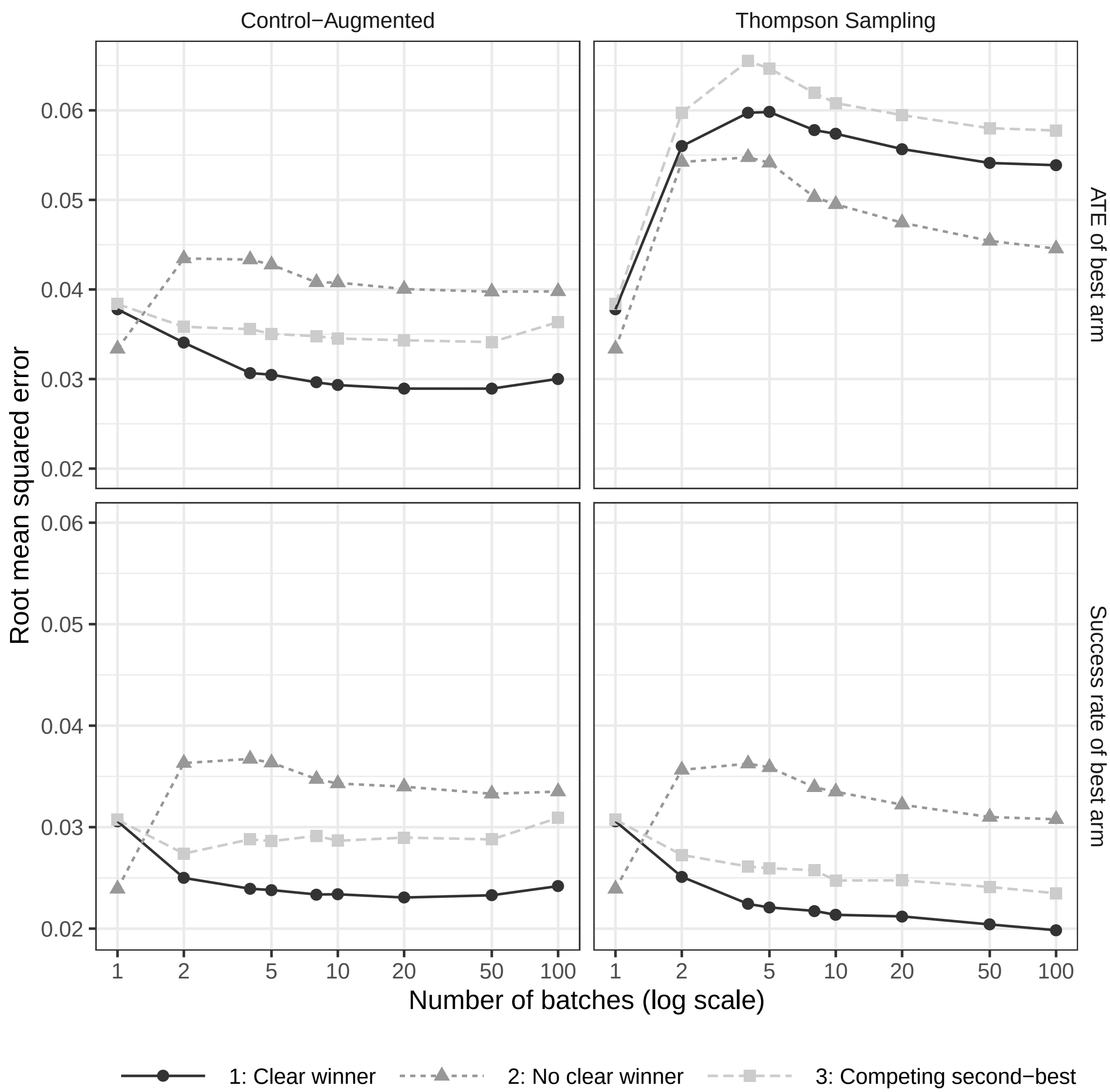


Table D.2: Iterated Simulation Statistics, Varying Number of Batches

Design				RMSE		Coverage	
Assignment algorithm	Case	Batches	Best arm selected	Best arm	ATE	Best arm	ATE
	1: Clear winner	1	0.909	0.031	0.038	0.941	0.949
TS		2	0.945	0.025	0.056	0.952	0.846
		4	0.962	0.022	0.060	0.955	0.769
		5	0.963	0.022	0.060	0.956	0.775
		8	0.965	0.022	0.058	0.955	0.785
		10	0.968	0.021	0.057	0.958	0.795
		20	0.967	0.021	0.056	0.956	0.799
		50	0.970	0.020	0.054	0.960	0.805
		100	0.969	0.020	0.054	0.961	0.807
	2: No clear winner	1	0.180	0.024	0.033	0.935	0.947
		2	0.177	0.036	0.054	0.819	0.802
		4	0.183	0.036	0.055	0.862	0.829
		5	0.181	0.036	0.054	0.872	0.838
		8	0.194	0.034	0.050	0.884	0.850
		10	0.193	0.033	0.050	0.880	0.851
		20	0.186	0.032	0.047	0.890	0.861
		50	0.188	0.031	0.045	0.894	0.863
		100	0.190	0.031	0.045	0.895	0.866
	3: Competing second best	1	0.635	0.031	0.038	0.940	0.945
		2	0.676	0.027	0.060	0.946	0.846
		4	0.696	0.026	0.066	0.950	0.759
		5	0.699	0.026	0.065	0.948	0.760
		8	0.700	0.026	0.062	0.952	0.783
		10	0.715	0.025	0.061	0.956	0.796
		20	0.708	0.025	0.059	0.956	0.796
		50	0.710	0.024	0.058	0.955	0.803
		100	0.704	0.023	0.058	0.958	0.799
TS,	1: Clear winner	2	0.939	0.025	0.034	0.948	0.945
Control-Augmented		4	0.949	0.024	0.031	0.954	0.952
		5	0.955	0.024	0.030	0.955	0.949
		8	0.955	0.023	0.030	0.956	0.952
		10	0.956	0.023	0.029	0.957	0.952
		20	0.960	0.023	0.029	0.962	0.954
		50	0.958	0.023	0.029	0.962	0.956
		100	0.957	0.024	0.030	0.957	0.953
	2: No clear winner	2	0.159	0.036	0.043	0.821	0.868
		4	0.165	0.037	0.043	0.857	0.868
		5	0.176	0.036	0.043	0.867	0.879
		8	0.175	0.035	0.041	0.878	0.887
		10	0.174	0.034	0.041	0.879	0.886
		20	0.174	0.034	0.040	0.881	0.885
		50	0.180	0.033	0.040	0.886	0.890
		100	0.178	0.034	0.040	0.885	0.892
	3: Competing second best	2	0.665	0.027	0.036	0.944	0.942
		4	0.673	0.029	0.036	0.943	0.937
		5	0.685	0.029	0.035	0.942	0.932
		8	0.682	0.029	0.035	0.943	0.938
		10	0.683	0.029	0.035	0.946	0.937
		20	0.676	0.029	0.034	0.943	0.938
		50	0.680	0.029	0.034	0.946	0.941
		100	0.670	0.031	0.036	0.941	0.936

Note: Assignment algorithms are batch-wise Thompson sampling (TS) and control-augmented batch-wise Thompson sampling (TS, Control-Augmented); the balanced static design is equivalent to the single-batch design. "Best arm selected" column presents the portion of simulations under which the true best arm was selected. RMSE is average root mean squared error of the estimate of the mean of the true best arm, and the average treatment effect of the true best arm relative to the control. Coverage is with respect to 95% confidence intervals of the estimate. In all cases one of the inferior arms with a true success rate of 0.10 is selected as the control comparison for estimating the Average Treatment Effect.

D.2 Varying value of best arm

We additionally replicate the simulations with 1,000 total observations, 100 observations per batch, varying value of the best arm, holding fixed the value of eight other arms at 0.10.

Figure D.3: Iterated Simulation Statistics, Varying Value of Best Arm

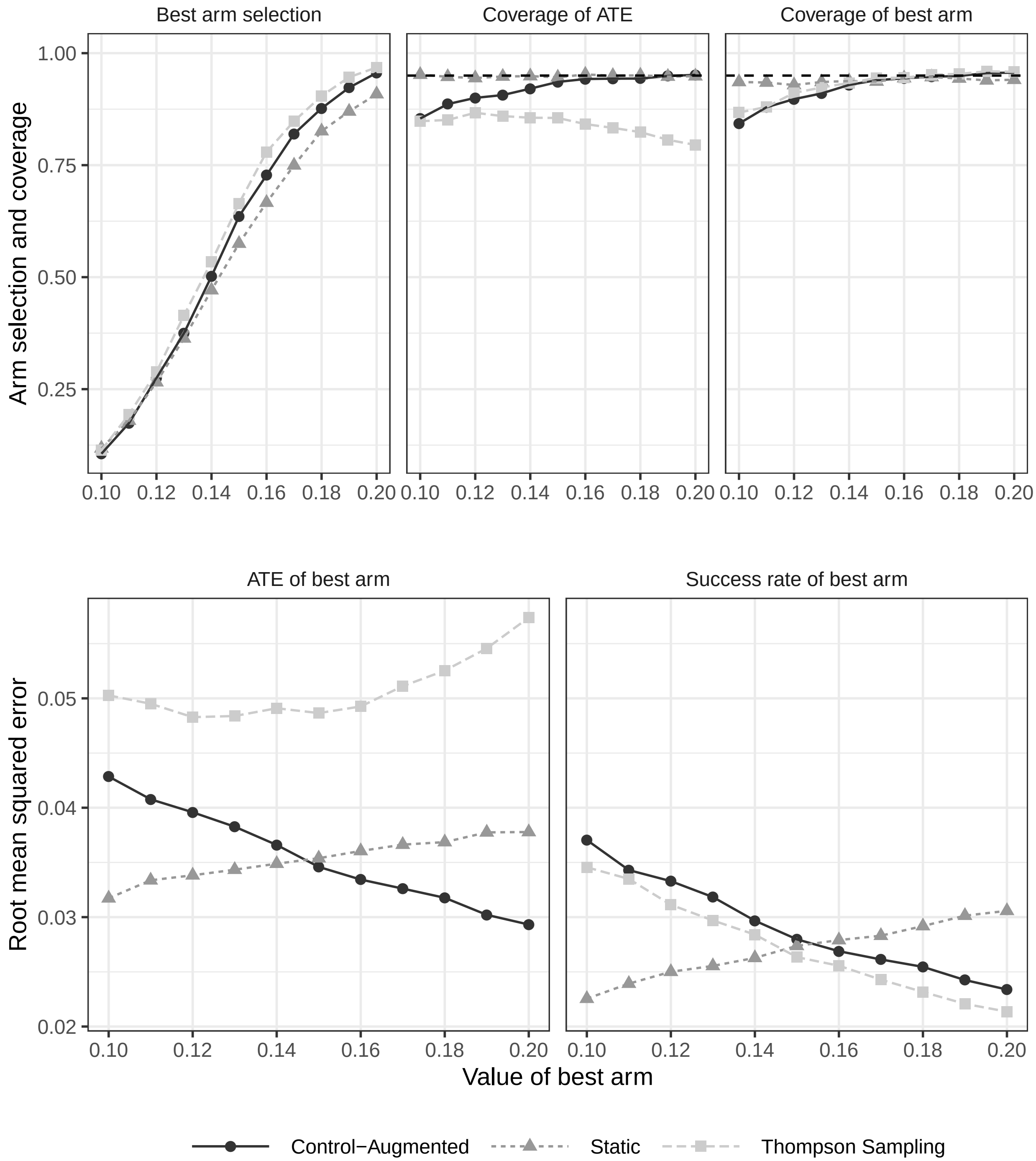


Table D.3: Iterated Simulation Statistics, Varying Value of Best Arm

Design			RMSE		Coverage	
Assignment algorithm	Best arm value	Best arm selected	Best arm	ATE	Best arm	ATE
TS	0.10	0.113	0.035	0.050	0.867	0.848
	0.11	0.193	0.033	0.050	0.880	0.851
	0.12	0.289	0.031	0.048	0.911	0.867
	0.13	0.415	0.030	0.048	0.923	0.860
	0.14	0.534	0.028	0.049	0.934	0.856
	0.15	0.665	0.026	0.049	0.945	0.855
	0.16	0.779	0.026	0.049	0.945	0.842
	0.17	0.848	0.024	0.051	0.952	0.833
	0.18	0.905	0.023	0.053	0.954	0.824
	0.19	0.946	0.022	0.055	0.960	0.807
	0.20	0.968	0.021	0.057	0.958	0.795
Static	0.10	0.118	0.023	0.032	0.936	0.952
	0.11	0.180	0.024	0.033	0.935	0.947
	0.12	0.266	0.025	0.034	0.929	0.945
	0.13	0.363	0.026	0.034	0.936	0.948
	0.14	0.471	0.026	0.035	0.938	0.949
	0.15	0.575	0.027	0.035	0.937	0.946
	0.16	0.666	0.028	0.036	0.944	0.953
	0.17	0.750	0.028	0.037	0.947	0.950
	0.18	0.826	0.029	0.037	0.944	0.951
	0.19	0.870	0.030	0.038	0.939	0.948
	0.20	0.909	0.031	0.038	0.941	0.949
TS,	0.10	0.106	0.037	0.043	0.843	0.854
Control-Augmented	0.11	0.174	0.034	0.041	0.879	0.886
	0.12	0.275	0.033	0.040	0.897	0.900
	0.13	0.375	0.032	0.038	0.910	0.907
	0.14	0.502	0.030	0.037	0.929	0.921
	0.15	0.636	0.028	0.035	0.940	0.936
	0.16	0.728	0.027	0.033	0.944	0.942
	0.17	0.820	0.026	0.033	0.947	0.943
	0.18	0.877	0.025	0.032	0.949	0.943
	0.19	0.923	0.024	0.030	0.956	0.949
	0.20	0.956	0.023	0.029	0.957	0.952

Note: Assignment algorithms are batch-wise Thompson sampling (TS), balanced static design (Static), and control-augmented batch-wise Thompson sampling (TS, Control-Augmented). "Best arm selected" column presents the portion of simulations under which the true best arm was selected. RMSE is average root mean squared error of the estimate of the mean of the true best arm, and the average treatment effect of the true best arm relative to the control. Coverage is with respect to 95% confidence intervals of the estimate. In all cases one of the inferior arms with a true success rate of 0.10 is selected as the control comparison for estimating the Average Treatment Effect.

D.3 Varying first batch size

One practical question that often arises in adaptive designs is how much sample to allocate to the first batch, prior to the start of adaptation. We conducted a series of simulations with 1,000 total observations, varying the size of the first batch, fixing the total number of batches at 10.

We note that for RMSE of the best arm and for best arm selection, there may be an initial benefit to increasing batch size in the cases where there is an identifiable best arm; increasing the first batch size may reduce chance under-assignment to the true best arm in early periods. That is, by extending the first period of exploration, we avoid exploiting the wrong arms. After this initial benefit, however, enforcing more exploration in the first period comes at the cost of exploiting the true best arm.

The point at which we see decreasing returns depends on the parameter values of the different arms. When there is no clear winner, however, we may generally be better off with larger first batch sizes, all the way up to a static design—where the first batch size is the full experiment, as the adaptive algorithms do not sufficiently exploit the best arm. We do not see returns to exploitation in cases where we are not able to able to identify the best arm to exploit.

Finally, when it comes to RMSE of the ATE with respect to the best arm, for the control-augmented algorithm, we see the same trends as for the best-arm estimate. Regarding standard Thompson sampling, however, increasing first batch size improves RMSE for the ATE, as we are forcing a larger sample to be collected for the control arm.

Figure D.4: Iterated Simulation Statistics, Varying Size of First Batch, Arm Selection and Coverage

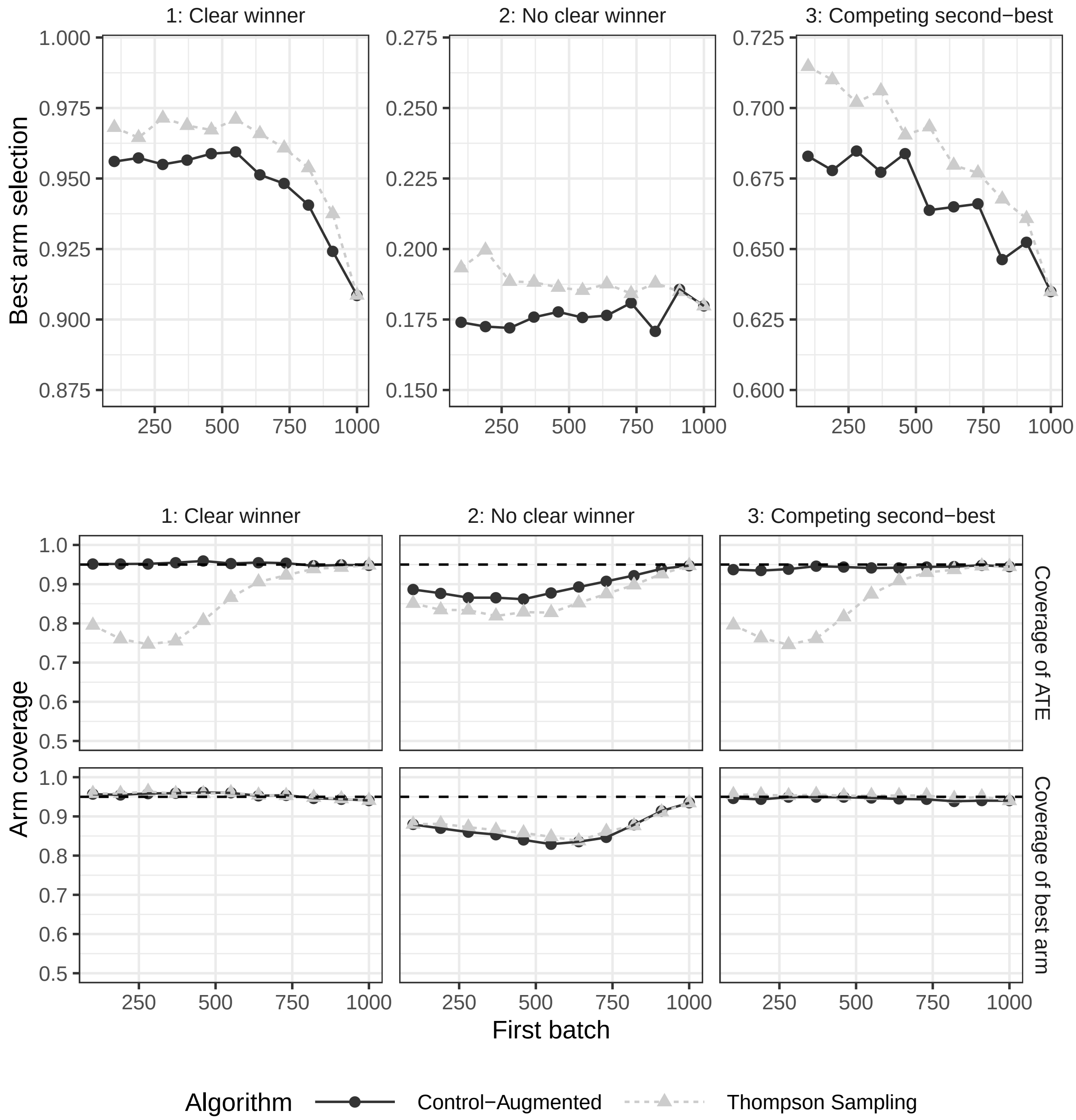


Figure D.5: Iterated Simulation Statistics, Varying Size of First Batch, RMSE

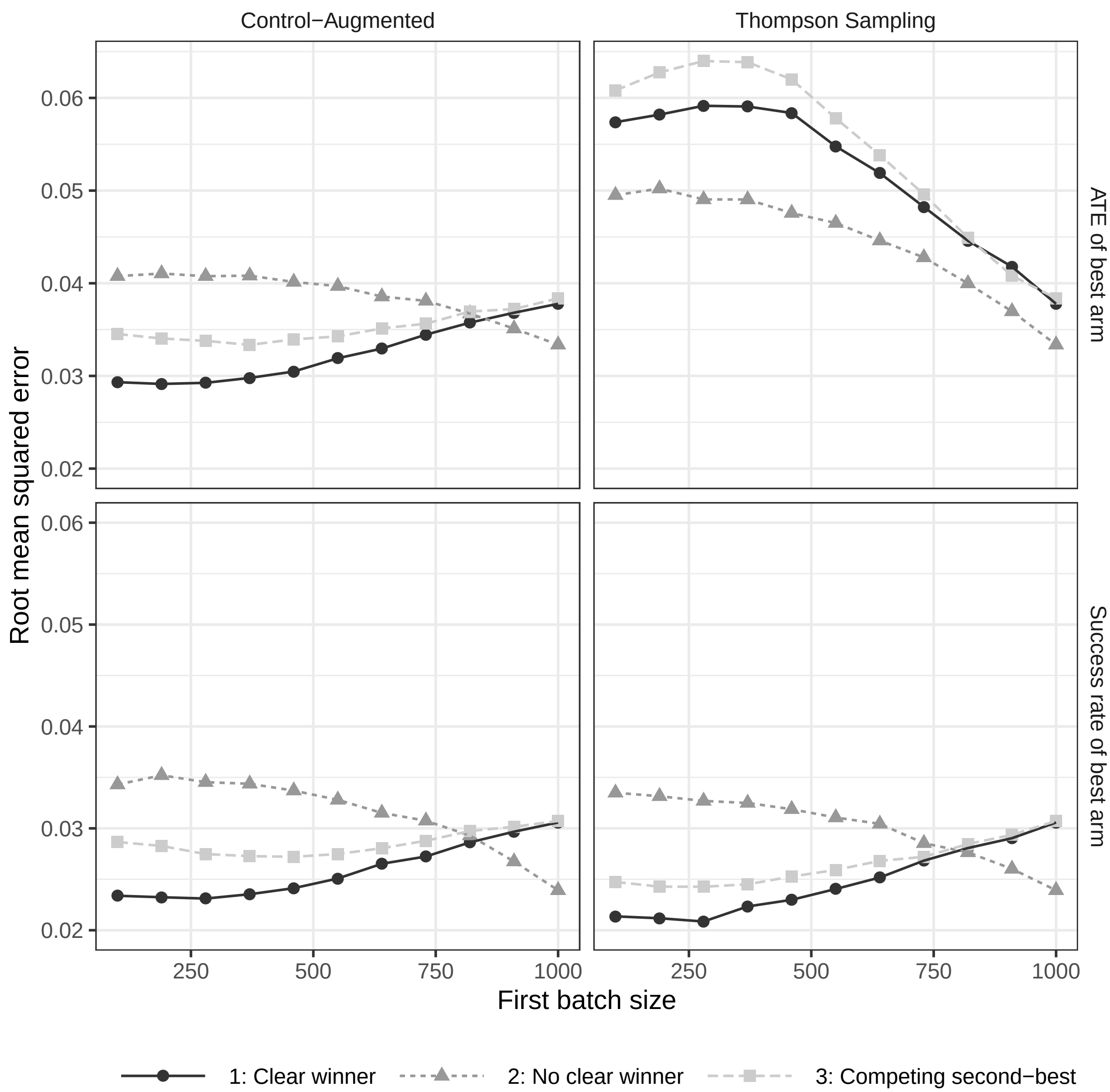


Table D.4: Iterated Simulation Statistics, Varying First Batch Size: Thompson Sampling

Assignment algorithm	Design: Case	First batch size	Best arm selected	RMSE: Best arm	RMSE: ATE	Coverage: Best arm	Coverage: ATE
	1: Clear winner	100	0.968	0.021	0.057	0.958	0.795
TS		190	0.965	0.021	0.058	0.958	0.760
		280	0.971	0.021	0.059	0.964	0.747
		370	0.969	0.022	0.059	0.958	0.755
		460	0.967	0.023	0.058	0.958	0.807
		550	0.971	0.024	0.055	0.960	0.866
		640	0.966	0.025	0.052	0.954	0.906
		730	0.961	0.027	0.048	0.953	0.923
		820	0.954	0.028	0.045	0.948	0.939
		910	0.938	0.029	0.042	0.945	0.943
	2: No clear winner	100	0.193	0.033	0.050	0.880	0.851
		190	0.200	0.033	0.050	0.881	0.835
		280	0.189	0.033	0.049	0.873	0.834
		370	0.188	0.032	0.049	0.865	0.819
		460	0.186	0.032	0.048	0.858	0.829
		550	0.185	0.031	0.046	0.848	0.827
		640	0.188	0.030	0.045	0.837	0.852
		730	0.184	0.029	0.043	0.862	0.875
		820	0.188	0.028	0.040	0.876	0.898
		910	0.185	0.026	0.037	0.911	0.926
	3: Competing second best	100	0.715	0.025	0.061	0.956	0.796
		190	0.710	0.024	0.063	0.955	0.763
		280	0.702	0.024	0.064	0.953	0.746
		370	0.706	0.025	0.064	0.956	0.761
		460	0.690	0.025	0.062	0.953	0.817
		550	0.693	0.026	0.058	0.953	0.874
		640	0.680	0.027	0.054	0.953	0.909
		730	0.677	0.027	0.050	0.953	0.930
		820	0.668	0.028	0.045	0.946	0.937
		910	0.661	0.029	0.041	0.950	0.946

Table D.5: Iterated Simulation Statistics, Varying First Batch Size, Control Augmented

Design				RMSE		Coverage	
Assignment algorithm	Case	First batch size	Best arm selected	Best arm	ATE	Best arm	ATE
TS, Control-Augmented	1: Clear winner	100	0.956	0.023	0.029	0.957	0.952
		190	0.957	0.023	0.029	0.955	0.951
		280	0.955	0.023	0.029	0.959	0.952
		370	0.957	0.024	0.030	0.959	0.955
		460	0.959	0.024	0.030	0.961	0.959
		550	0.959	0.025	0.032	0.960	0.953
		640	0.951	0.027	0.033	0.953	0.955
		730	0.948	0.027	0.034	0.954	0.954
		820	0.941	0.029	0.036	0.947	0.947
		910	0.924	0.030	0.037	0.944	0.949
	2: No clear winner	100	0.174	0.034	0.041	0.879	0.886
		190	0.173	0.035	0.041	0.870	0.877
		280	0.172	0.035	0.041	0.860	0.866
		370	0.176	0.034	0.041	0.854	0.866
		460	0.178	0.034	0.040	0.840	0.862
		550	0.176	0.033	0.040	0.829	0.877
		640	0.176	0.032	0.039	0.835	0.893
		730	0.181	0.031	0.038	0.846	0.907
		820	0.171	0.029	0.037	0.879	0.922
		910	0.186	0.027	0.035	0.914	0.939
	3: Competing second best	100	0.683	0.029	0.035	0.946	0.937
		190	0.678	0.028	0.034	0.943	0.935
		280	0.685	0.027	0.034	0.949	0.938
		370	0.677	0.027	0.033	0.949	0.946
		460	0.684	0.027	0.034	0.949	0.944
		550	0.664	0.027	0.034	0.947	0.941
		640	0.665	0.028	0.035	0.944	0.942
		730	0.666	0.029	0.036	0.943	0.944
		820	0.646	0.030	0.037	0.939	0.945
		910	0.652	0.030	0.037	0.940	0.948

Note: Assignment algorithms are batch-wise Thompson sampling (TS) and control-augmented batch-wise Thompson sampling (TS, Control-Augmented). "Best arm selected" column presents the portion of simulations under which the true best arm was selected. RMSE is average root mean squared error of the estimate of the mean of the true best arm, and the average treatment effect of the true best arm relative to the control. Coverage is with respect to 95% confidence intervals of the estimate. In all cases one of the inferior arms with a true success rate of 0.10 is selected as the control comparison for estimating the Average Treatment Effect.

E Additional Information, Study One

Table E.6: Minimum Wage Rates as of June, 2018

State	Minimum wage
Alabama	$7.25
Alaska	$9.84
Arizona	$10.50
Arkansas	$8.50
California	$11.00
Colorado	$10.20
Connecticut	$10.10
Delaware	$8.25
Florida	$8.25
Georgia	$7.25
Hawaii	$10.10
Idaho	$7.25
Illinois	$8.25
Indiana	$7.25
Iowa	$7.25
Kansas	$7.25
Kentucky	$7.25
Louisiana	$7.25
Maine	$10.00
Maryland	$9.25
Massachusetts	$11.00
Michigan	$9.25
Minnesota	$9.65
Mississippi	$7.25
Missouri	$7.85
Montana	$8.30
Nebraska	$9.00
Nevada	$8.25
New Hampshire	$7.25
New Jersey	$8.60
New Mexico	$7.50
New York	$10.40
North Carolina	$7.25
North Dakota	$7.25
Ohio	$8.30
Oklahoma	$7.25
Oregon	$10.25
Pennsylvania	$7.25
Rhode Island	$10.10
South Carolina	$7.25
South Dakota	$8.85
Tennessee	$7.25
Texas	$7.25
Utah	$7.25
Vermont	$10.50
Virginia	$7.25
Washington	$11.50
West Virginia	$8.75
Wisconsin	$7.25
Wyoming	$7.25

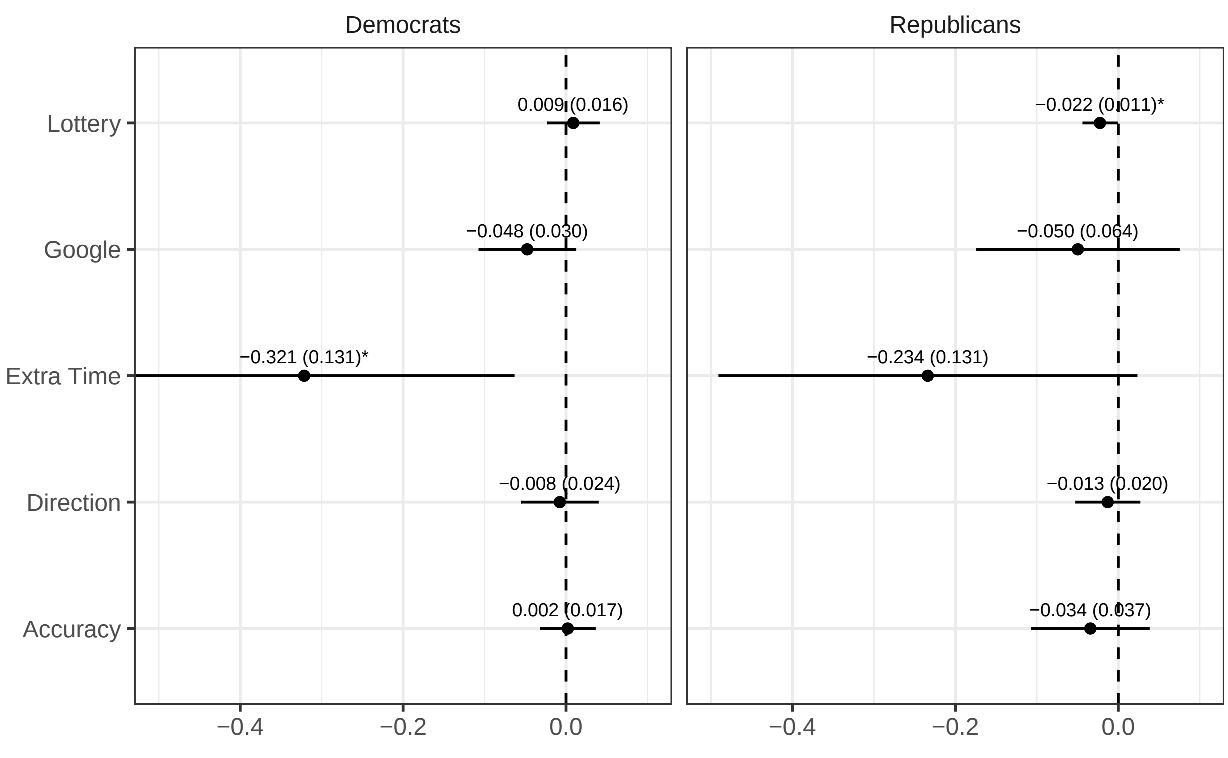
Source: `https://en.wikipedia.org/wiki/Minimum_wage_in_the_United_States`

F Additional Analyses, Study Two

F.1 Response rates

We consider response rates by party and treatment condition in Figure F.6. Control means are 0.956 and 0.977 for Democrats and Republicans, respectively. For the Lottery, Direction, and Accuracy conditions, response rates are very close to those under the control condition. Response rates under the Google condition are about 5 percentage points lower for both Democrats and Republicans; response rates under the Extra Time condition are 32 and 23 percentage points lower for Democrats and Republicans respectively.

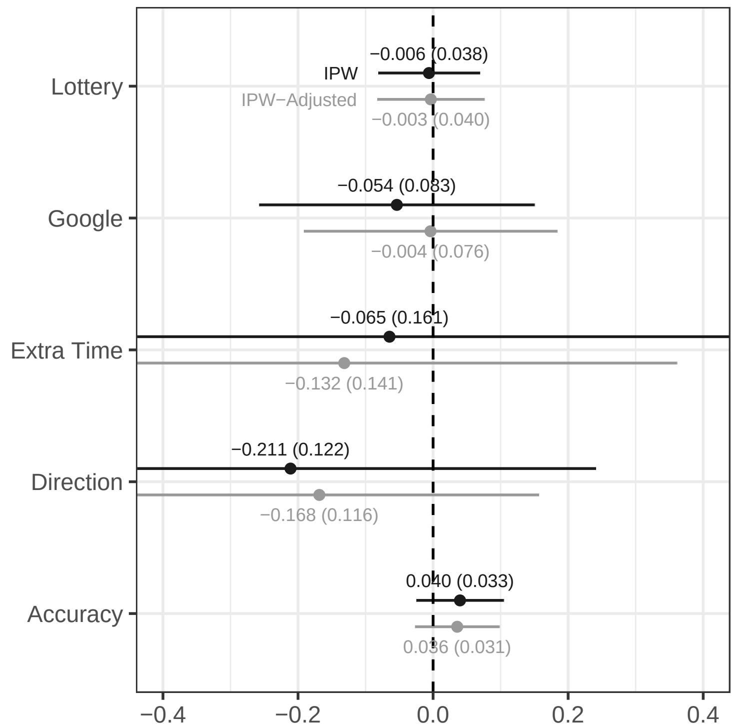
Figure F.6: Study Two, Response Rates



Note: *p<0.05. Estimates are with respect to rates of response in the control condition, and are inverse-probability weighted. Standard errors are heteroskedasticity consistent HC2 corrected.

F.2 Analysis for Independents

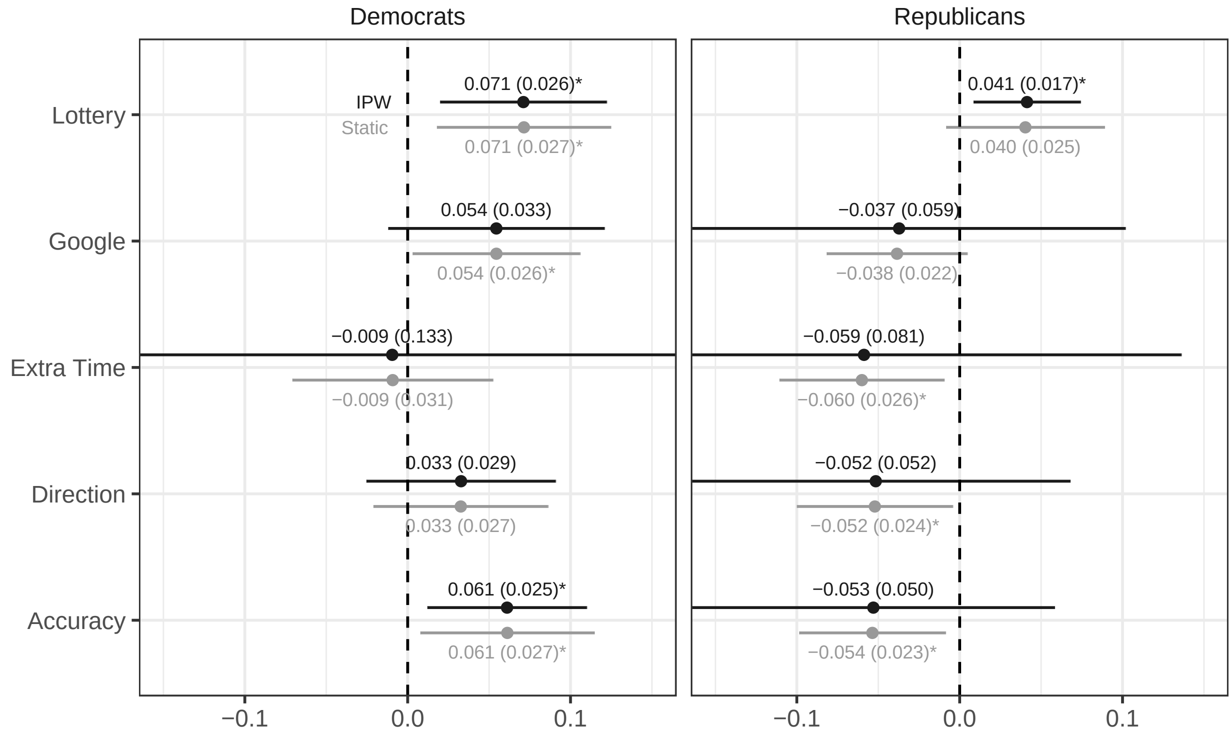
Figure F.7: Study Two, Mean Correct Responses, Independents



Note: *p<0.05. Estimates are inverse probability weighted. Standard errors are bias-reduced linearization CR2 adjusted, clustered at the subject level.

F.3 Comparison with alternative design

Figure F.8: Study Two, Mean Correct Responses, Simulated Comparisons



Note: *p<0.05. "IPW" estimates are produced from the original study and are inverse probability weighted. Standard errors for these estimates are bias-reduced linearization CR2 adjusted, clustered at the subject level. "Static" estimates are produced by sampling with replacement from observed responses, with sampling weights as the inverse probability of treatment weights, to produce a hypothetical static trial of the same size as the adaptive trial. These estimates are bootstrapped 1,000 times, and confidence intervals are produced under the assumption that bootstrap replicates follow a t-distribution.

F.4 Downstream analyses

We asked two additional follow-up questions later in the survey, to evaluate downstream effects on current evaluations of and optimism about the economy. These questions are:

> **Follow-up 1:** "What do you think about the state of the economy these days in the United States?"
> Response options are "Very good," "Good," "Neither good nor bad," "Bad," and "Very bad." Outcomes are coded as 1 if the response is "Very good" or "Good," and 0 otherwise, with NAs left as is.

> **Follow-up 2:** "What about the next 12 months? Do you expect the economy, in the country as a whole, to…"
> Response options are "Get better," "Stay about the same," and "Get worse." Outcomes are coded as 1 if the response is "Get better," and 0 otherwise, with NAs left as is.

We exclude from this analysis those treated with Google and Extra Time conditions, as the response rates under these two treatments are much lower than the others (see Figure F.6). We include all models with and without covariate adjustment.

Table F.7: Study Two, Downstream Effects of Treatment

	Dependent variable:							
	Follow-up 1				Follow-up 2			
	Democrats		Republicans		Democrats		Republicans	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lottery	−0.005	−0.002	−0.005	0.003	0.019	0.047	0.018	0.015
	(0.044)	(0.043)	(0.030)	(0.029)	(0.032)	(0.031)	(0.034)	(0.034)
Accuracy	−0.008	−0.013	−0.067	−0.066	0.021	0.043	−0.181	−0.174
	(0.042)	(0.041)	(0.093)	(0.079)	(0.032)	(0.032)	(0.101)	(0.094)
Direction	−0.034	−0.044	0.058	0.055	0.022	0.049	−0.112	−0.123
	(0.047)	(0.046)	(0.059)	(0.064)	(0.038)	(0.037)	(0.106)	(0.090)
Covariate adjusted	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1,143	1,143	1,124	1,124	1,143	1,143	1,120	1,120
R^2	0.001	0.072	0.009	0.174	0.001	0.100	0.025	0.172

Note: *p<0.05. Estimates are inverse probability weighted. Standard errors are heteroskedasticity consistent $HC2$ corrected.

G Study Three: An Adaptive Conjoint Trial

We consider an additional setting where arms are composed of factorial components as in a conjoint experiment, where multiple dimensions of treatments are varied. Combinations of factors can quickly result in a large number of arms that may be unwieldy for exploration. In the case of binary rewards, the reward distribution can be modeled by a probit regression on the respective factorial components (see, e.g., Shahriari et al. 2016; for more general cases, see Filippi et al. 2010). Modeling assumptions, such as the stipulation that the factors exert main effects without cross-factor interactions, allow us to pool information across arms sharing common components and to estimate success probabilities for arms which we have not observed. Such model-based approaches can be used to select from among a large number of possible treatment profiles.

Here, we consider ballot measures composed of four elements, each addressing an aspect of campaign finance: personal limits, corporate limits, public funding, and disclosures for individual campaign contributions. Each element may take on one of several levels, such that the combination of factors with $4 \times 3 \times 4 \times 4$ levels results in 192 unique experimental arms. As with the minimum wage and right-to-work measures, the outcome is defined as a success if a subject responds that they would vote in favor of the measure. We are interested in identifying the experimental arm representing the combination of factors that is associated with the highest success rate.

The parameters of interest are a vector $\boldsymbol{\gamma}$, which we estimate as the coefficients in our model,

$$P(x=1) = \Phi\left(\omega_0 + \sum_{j=1}^{4}\sum_{\ell=1}^{L_j} \gamma_{j\ell} D_{j\ell}\right),$$

where Φ represents the standard normal cumulative distribution function, and the parameter vector $\boldsymbol{\gamma} = (\omega_0, \gamma_{1,1}, \ldots, \gamma_{4,3})$ includes an intercept and coefficients for dummy variables, $D_{j\ell}$, indexed over the four factors, j, and the levels within each factor, ℓ. We assume a uniform prior over the parameter vector.[3] Alternative approaches to modeling could account for interactions through regularization.

For this study, we recruited a convenience sample of 979 subjects from Lucid, collected as a target of 100 responses for each of 10 waves. Our study ran from November 26, 2018 to December 7, 2018. We paid \$1 for each survey response; participants were compensated in money, points, or other rewards depending on how they were enrolled in the Lucid subject pool. Due to constraints in survey implementation at the time of the experiment, we were not able to prevent Lucid subjects from taking the survey on multiple days, but we were able to identify them and remove them from ex-post analysis for each wave. While we include 10 unique waves in our design and analysis, data collection for some waves was extended to a second day to facilitate identification of re-sampled subjects and to augment the sample to account for these subjects.

We separately ran a conjoint experiment with the same features using a static design, in which all factor levels were presented with equal probability. For this experiment, we recruited a population from MTurk from November 4 to 12, 2018. As the MTurk population was recruited as part of a larger multi-wave study, we randomly select a subsample of 979 subjects for comparison to the adaptive design. MTurk subjects were also paid \$1 for their participation.

[3]The prior is an improper prior and can not be integrated, as it is uniform over all real numbers. However, this does not cause problems when sampling from the posterior, which is well-defined.

Design

Both the static and adaptive experiments follow a conjoint design with four sets of attributes. The full text of all treatment levels is presented in Table G.8, where the first level of each attribute represents the status quo at the time of the experiment. Subjects were shown two measures, with assignment conducted independently. (Bansak et al. 2018 demonstrate that conjoint designs are robust to assigning subjects multiple choice tasks.) The outcome measure is response to the question, "If you were casting a ballot tomorrow, would you vote yes or no on this constitutional amendment?" with response categories "Yes," "No," and "Undecided." Our primary outcome of interest is binary: an indicator for whether the response was "Yes."

In the first period, we assign treatment to a subset of conditions to facilitate estimation of main effects. Following the first wave of data collection, we then update this prior by modeling success rate for each of the conditions using probit regression with a main effects only model.

We simulate 10,000 draws from the posterior distribution, and compute the predicted success probability in each of the 192 treatment conditions for each draw, and then calculate the probability that each condition is "best" as the proportion of draws under which the arm had the highest predicted success rate. On the subsequent wave, sampling of each condition is conducted in proportion to this probability.

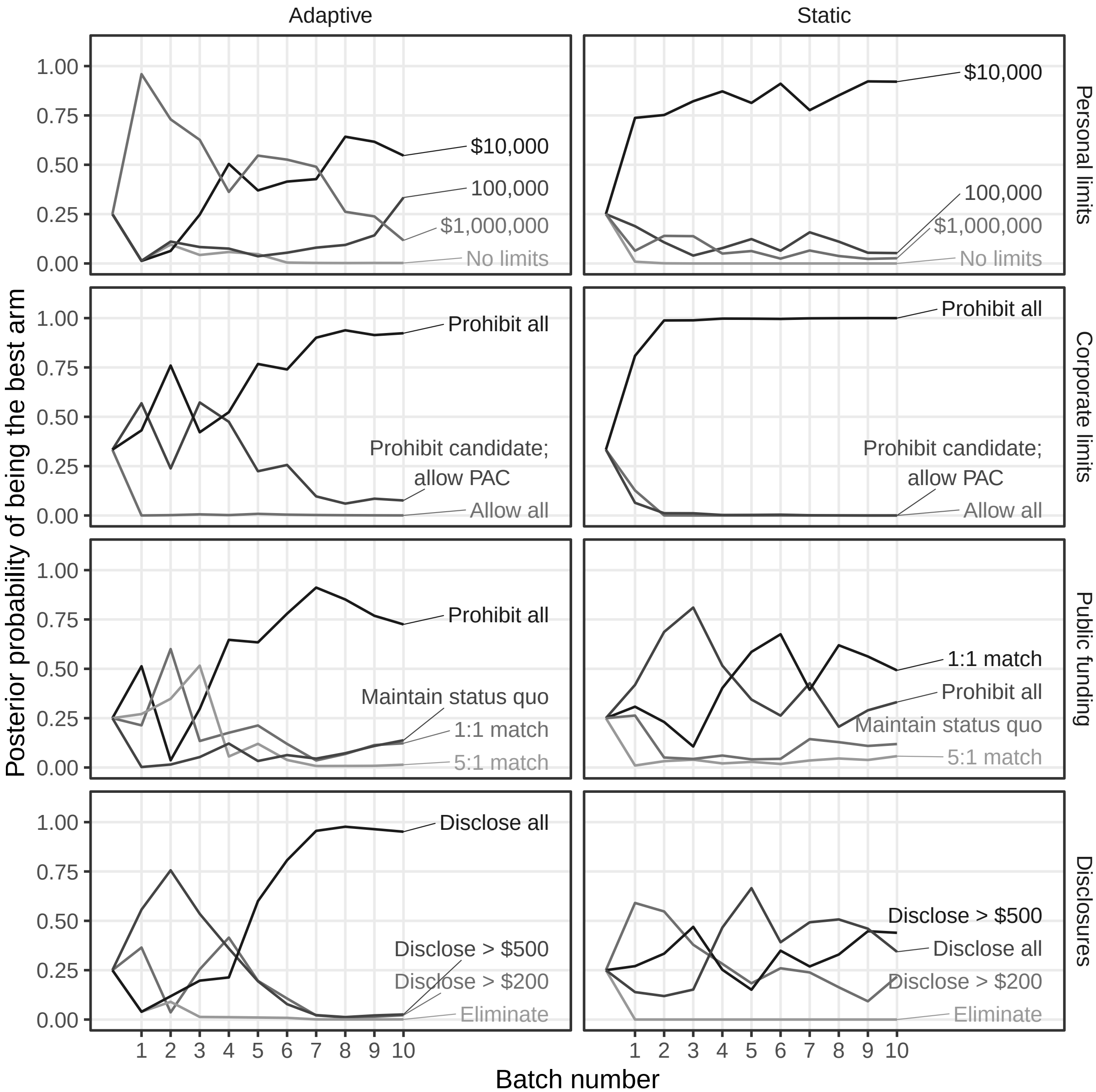
Results

The benefits of the adaptive design become evident when considering the development of probabilities of being best over time. For each attribute, we may consider probabilities of being best for each level marginalizing over the other attributes, as we have assumed a main-effects only model. Under this assumption, the adaptive conjoint design effectively reduces

Table G.8: Study Three, Treatments and Outcome Measures

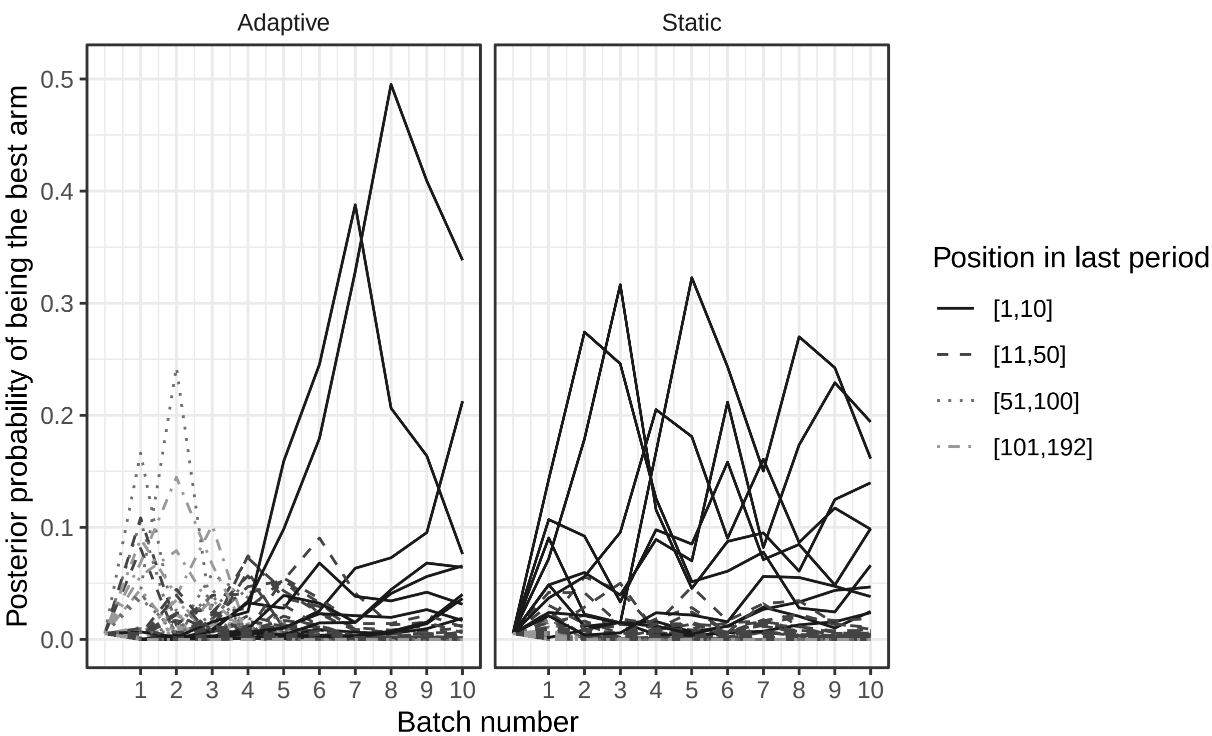
	Personal Limits	Corporate Limits	Public Funding	Disclosures
Question Text	Consider an amendment to the Constitution of the United States on the topic of campaign finance. This amendment includes the following provisions: **[ballot measure text].** If you were casting a ballot tomorrow, would you vote yes or no on this constitutional amendment? [Yes; No; Undecided]			
Level 1 [Status quo]	maintain no limits on how much an individual may contribute in aggregate to all candidates in a calendar year	maintain the prohibition on corporate contributions to candidates, while allowing contributions to political action committees and independent expenditures	maintain the public funding option for presidential elections, and introduce no new reforms for public funding for other federal office	maintain requirements for disclosures for contributions above the current federal limit, $200 per election cycle
Level 2	set the limit on how much an individual may contribute in aggregate to all candidates at $1 million per calendar year	prohibit corporations from providing financial support to candidates for office, either through direct contributions to candidates, contributions to political action committees, or independent expenditures	prohibit Congress from passing a bill to establish public funding of candidates	eliminate all disclosure requirements
Level 3	set the limit on how much an individual may contribute in aggregate to all candidates at $100,000 per calendar year	allow corporations to provide financial support to candidates for office, either through direct contributions to candidates, contributions to political action committees, or independent expenditures	make available public funds for candidates for federal office decided by a November general election; funds would be provided at a rate of $1 in matching funds for every $1 raised through small (under $175) donations from constituents, in exchange for the candidate agreeing to campaign spending limits and increased financial oversight	require that every campaign contribution be disclosed, no matter how small
Level 4	set the limit on how much an individual may contribute in aggregate to all candidates at $10,000 per calendar year		make available public funds for candidates for federal office decided by a November general election; funds would be provided at a rate of $5 in matching funds for every $1 raised through small (under $175) donations from constituents, in exchange for the candidate agreeing to campaign spending limits and increased financial oversight	require disclosures for contributions above $500 per election cycle

Figure G.9: Study Three, Overtime Posterior Probabilities by Attribute



Note: Marginal posterior probabilities updated after each day's data collection.

Figure G.10: Study Three, Overtime Posterior Probabilities, Joint



Note: Joint posterior probabilities updated after each day's data collection.

to four separate adaptive experiments, each running over the course of the study. Considering the left panel in Figure G.9, for the adaptive design a clear winner has emerged in the corporate limits, public funding, and disclosures attributes, and a likely winner has emerged for personal limits. From these, we would gather that the most preferred measure profile would propose personal contribution limits of only $10,000, would prohibit all corporate contributions as well as public funding, and would institute required disclosures of all contributions. Considering the right panel in Figure G.9, we would infer under the static design that the most preferred profile measure would also propose personal contribution limits of only $10,000 and would prohibit all corporate contributions, but would support $1 to $1 in matching funds, and would require disclosures above $500.

Figure G.10 presents the probability of being best for each condition as the joint probabilities of the component attribute levels. For the adaptive design, the most preferred profile includes each of the top attribute levels presented in the left panel of Figure G.9, with a final probability of being best of 0.338; the second-most preferred profile would include the same levels of the other attributes, but upping personal contribution limits to $100,000, with a final probability of being best of 0.213; the third-most preferred profile would similarly contain the same levels of the other attributes, but set personal contribution limits at $1 million, with a final probability of being best of 0.076. For the static design, the most preferred profile includes each of the top attribute levels presented in the right panel of Figure G.9, with a final probability of being best of 0.194; the second-most preferred profile would include the same levels of the other attributes, but requiring disclosures of all contributions, with a final probability of being best of 0.161; the third-most preferred profile would similarly contain the same levels for personal and corporate contributions, but would prohibit public funding and would require disclosures for contributions over $500, with a final probability of being best of 0.140. Both the adaptive and static designs offer a similar characterization of public preferences regarding campaign finance, but the adaptive design offers a more precise reading of the public's most preferred policies.

Substantively, what does the adaptive trial tell us about the structure of public preferences regarding campaign finance regulation? First, the constitutional amendment that most appeals to our adaptive sample is at odds with Supreme Court rulings. The Court in *McCutcheon v. Federal Election Commission*, 572 U.S. 185 (2014) struck down aggregate limits on how much individuals can donate to campaigns during an election cycle; by contrast, to our sample, the lower the personal limit, the better. Public preferences are also at odds with the current policy with respect to corporate contributions. Our sample would prohibit all corporate contributions, preferring an outright ban to a system that permits PAC contributions and, under *Citizens United v. Federal Election Commission*, 558 U.S. 310 (2010), allows unlimited electioneering communications but not direct contributions to parties or candidates. Second, our sample is also out of step with current federal law, which requires disclosures of contributions greater than $200. Our sample most prefers a system in which all contributions are disclosed. Finally, although our sample takes a "reformist" stance on contribution limits and disclosure, it remains reluctant to replace private contributions with public subsidies; its most preferred policy is one that prohibits public funding altogether. Overall, public opinion in the United States favors public policies often seen outside the U.S. whereby contributions are tightly restricted and public subsidies are minimal.

References

Agrawal, Shipra and Navin Goyal. 2017. "Near-optimal Regret Bounds for Thompson Sampling." Journal of the ACM 64(5):1–24.

Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins and Teppei Yamamoto. 2018. "The Number of Choice Tasks and Survey Satisficing in Conjoint Experiments." Political Analysis 26(1):112–119.

Bowden, Jack and Lorenzo Trippa. 2017. "Unbiased estimation for response adaptive clinical trials." Statistical methods in medical research 26(5):2376–2388.

Deshpande, Yash, Lester Mackey, Vasilis Syrgkanis and Matt Taddy. 2020. "Accurate inference for adaptive linear models." arXiv e-prints arXiv:1712.06695v5 [stat.ML].

Dimakopoulou, Maria, Zhengyuan Zhou, Susan Athey and Guido Imbens. 2017. "Estimation considerations in contextual bandits." arXiv e-prints arXiv:1711.07077v4 [stat.ML].

Dimmery, Drew. 2018. "Adaptive Experimental Design for Social Science." Working Paper. Accessed October 4, 2020, *https://dimmery.com/publication/adaptive-neyman/*

Filippi, Sarah, Olivier Cappe, Aurélien Garivier and Csaba Szepesvári. 2010. Parametric Bandits: The Generalized Linear Case. In Advances in Neural Information Processing Systems 23, ed. J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel and A. Culotta. Curran Associates, Inc. pp. 586–594.
Accessed October 4, 2020, *http://papers.nips.cc/paper/4166-parametric-bandits-the-generalized-linear-case*

Gao, Zijun, Yanjun Han, Zhimei Ren and Zhengqing Zhou. 2019. "Batched multi-armed bandits problem." arXiv e-prints arXiv:1904.01763v3 [stat.ML].

Hadad, Vitor, David A Hirshberg, Ruohan Zhan, Stefan Wager and Susan Athey. 2019. "Confidence Intervals for Policy Evaluation in Adaptive Experiments." arXiv e-prints arXiv:1911.02768v3 [stat.ML].

Hájek, Jaroslav. 1971. Comment on "An Essay on the Logical Foundations of Survey Sampling, Part One". In Foundations of Statistical Inference, ed. Vidyadhar P. Godambe and David A. Sprott. Vol. 236 Toronto: Holt, Rinehart, and Winston.

Horvitz, Daniel G and Donovan J Thompson. 1952. "A generalization of sampling without replacement from a finite universe." Journal of the American Statistical Association 47(260):663–685.

Howard, Steven R, Aaditya Ramdas, Jon McAuliffe and Jasjeet Sekhon. 2018. "Uniform, nonparametric, non-asymptotic confidence sequences." arXiv e-prints arXiv:1810.08240v5 [math.ST].

Jun, Kwang-Sung, Kevin Jamieson, Robert Nowak and Xiaojin Zhu. 2016. Top Arm Identification in Multi-Armed Bandits with Batch Arm Pulls. In Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, ed. Arthur Gretton and Christian C. Robert. Vol. 51 of Proceedings of Machine Learning Research Cadiz, Spain: PMLR pp. 139–148.
Accessed October 4, 2020, *http://proceedings.mlr.press/v51/jun16.html*

Perchet, Vianney, Philippe Rigollet, Sylvain Chassang, Erik Snowberg et al. 2016. "Batched bandit problems." The Annals of Statistics 44(2):660–681.

Robbins, Herbert. 1952. "Some aspects of the sequential design of experiments." Bulletin of the American Mathematical Society 58(5):527–535.

Russo, Daniel. 2016. Simple bayesian algorithms for best arm identification. In Proceedings of the 29th Annual Conference on Learning Theory. Vol. 49 pp. 1417–1418. Accessed October 4, 2020, *http://proceedings.mlr.press/v49/russo16.html*

Shahriari, Bobak, Kevin Swersky, Ziyu Wang, Ryan P Adams and Nando De Freitas. 2016. "Taking the Human Out of the Loop: A Review of Bayesian Optimization." Proceedings of the IEEE 104(1):148–175.

Wasserman, Larry. 2013. All of Statistics: A Concise Course in Statistical Inference. New York: Springer-Verlag.

Zhang, Kelly W, Lucas Janson and Susan A Murphy. 2020. "Inference for Batched Bandits." arXiv e-prints arXiv:2002.03217v2 [cs.LG].