Adaptive Experimental Design: Prospects and Applications in Political Science

Molly Offer-Westort,[*] Alexander Coppock,[†] and Donald P. Green[‡]

February 12, 2019

Abstract

Experimental researchers in political science frequently face the following inference problem: which of several treatment arms produces the greatest return, where returns may be expressed in terms of campaign donations, votes for a candidate, or some other support for a political cause? Multi-arm trials are typically conducted using a static design in which fixed proportions of study participants are allocated to each arm. However, a growing statistical literature suggests that adaptive experimental designs may be far more efficient in finding the most effective treatment arm. An important class of adaptive designs uses randomized probability matching to dynamically allocate subjects to treatment arms. We review the operating characteristics of randomized probability matching and suggest that it has many potential applications in political science. We discuss the design and analysis of original experiments using this approach and compare their efficiency to traditional static designs.

Word count: 8,791

[*]Molly Offer-Westort is a Ph.D. Candidate in the Departments of Political Science and Statistics and Data Science, Yale University

[†]Alexander Coppock is Assistant Professor of Political Science, Yale University.

[‡]Donald P. Green is the Burgess Professor of Political Science, Columbia University

Experimentation in the social sciences often boils down to a search for the intervention that maximizes some desired outcome. What pricing strategy maximizes demand for vaccines in low-income countries (Cook et al. 2009)? Which of the many ways of monitoring corruption among public officials minimizes the amount of missing public funds (Olken 2007)? What combination of personal attributes makes an applicant for naturalization most attractive to voters in the receiving country (Hainmueller and Hangartner 2013)? In many cases, this search dovetails with other academic objectives, such as discerning the causal mechanisms that make certain interventions especially effective (Ludwig et al. 2011).

Experiments that assess the relative effectiveness of competing interventions, be they policies or messages, often confront a fundamental problem: the list of interventions under consideration is so long that it is prohibitively costly and time-consuming to test the full range of treatment arms. Furthermore, even if money were no object, a prolonged search for the best alternative may impose excessive and unethical costs on human subjects and delay the implementation of interventions that would be superior to the status quo.

Adaptive trials represent a design-based attempt to increase the speed and efficiency with which multi-arm trials discern the best-performing intervention or interventions. In contrast to conventional static designs that allocate a fixed proportion of subjects to each arm throughout the trial, adaptive trials adjust the allocation as the trial unfolds, investing an ever-larger share of the subject pool in more promising treatment arms.[1] Adaptive trials are most likely to pick the true winner when the best arm is substantially better than its competitors; in such cases the design often declares a winner with much greater confidence than would have been possible under a static design. In substantive domains from advertising (Graepel et al. 2010; Li et al. 2010; Scott 2015) to biomedical research (Chow and Chang

---

[1]Adaptive trials encompass a broad class of designs that potentially evolve based on interim results. Here we focus exclusively on the allocation of subjects, but other adaptive design adjustments include changing the treatments or halting the trial entirely (Chin 2016).

2

2008; Sydes et al. 2012; Villar et al. 2015; Chin 2016), adaptive trials are regularly used to speed the search for the best-performing intervention.

That said, adaptive designs are no panacea. In situations where several treatment arms are equally effective (or nearly so), adaptive algorithms may equivocate, allocating more subjects to arms whose initial success was due only to sampling fluctuation. In the end, there is no guarantee that the researcher will discover the truly best intervention and no guarantee that the adaptive design will have allocated more subjects to the truly best option than the static design. Given this uneasy combination of upside potential and downside risk, the literature on adaptive designs abounds with proposals for allocating subjects in ways that guard against false positives and give early warning signals about futile searches among roughly equally effective (or ineffective) interventions.

The aim of this paper is to introduce political scientists to adaptive designs, highlighting the conditions under which they do or do not outperform conventional static designs. We begin by introducing the most commonly used adaptive design, known in the literature as Thompson sampling, or more generally, randomized probability matching (Thompson 1933, 1935). The features of this design are illustrated through simulation of multi-arm trials, some of which are more favorable to adaptive designs, while others expose their limitations. Next, we turn to empirical applications involving the wording of ballot measures, specifically, a pair of multi-arm trials. The first assembles actual ballot measures proposing changes in the minimum wage; we conduct an adaptive trial to discern which wording maximizes voter support. The second experiment uses the same design to maximize voter support for right-to-work proposals. These two examples illustrate the conditions under which adaptive designs work well. In the case of right-to-work proposals, adaptive design quickly identifies a clear winner with a high degree of statistical precision. Results are more ambiguous for minimum wage proposals, where several proposals seem equally promising. We then present a design varying multiple features of a ballot measure on campaign finance reform, where the objective

is to find the combination of features that maximizes voter support. When multiple factors are randomized, the number of unique treatment combinations may be so large that it is not feasible to test each combination independently. We conduct an experiment using a model-based adaptive conjoint design, and compare the results to those obtained from an experiment using a traditional static conjoint design.

Adaptive Trials

Adaptive trials are frequently positioned in the framework of the multi-armed bandit problem, first posed by Thompson (1933, 1935), where the experimenter is tasked with sequentially allocating finite resources across multiple treatment arms.[2] The objective is to maximize expected reward, framed here as the probability that the arm results in a success. The properties of the arms are not fully known at the outset of the experiment, but the experimenter may learn more about these properties over time by observing outcomes under different treatments. The typical tradeoff addressed in such settings is between exploration and exploitation. Experimenters would like to *explore* by obtaining information about the probability of success of each arm so that they can be confident in selecting the best arm or arms. They would also like to *exploit* the best performing arms by allocating large proportions of subjects to them. These two objectives are in tension with one another. On the one hand, too much exploration means wasting draws on under-performing arms. On the other hand, over-exploitation of early frontrunners risks ignoring potentially superior arms.

A heuristic approach to navigating the tradeoff between exploration and exploitation is Thompson sampling, which randomly allocates subjects to treatment arms according to

---

[2]For a history of the bandit problem and an overview of problem formulations and general approaches, see Berry and Fristedt (1985). For an updated overview of the general field of reinforcement learning, see Sutton and Barto (2018).

their probability of returning the highest reward under a Bayesian posterior.[3] When the researcher does not have much information about which arm is the best, the algorithm will facilitate exploration. As more information is gained, the best-performing arms are increasingly exploited.

While this approach is generalizable to continuous rewards, for ease of exposition we consider binary rewards, where each observation is either a success or a failure, $x \in \{0, 1\}$. Here K arms have an unknown probability of success $\theta_1, \ldots, \theta_K$, following their respective Bernoulli distributions, with likelihoods

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

An experimenter assigns some prior $f_{\Theta_k}(\theta_k)$ to the probability of success of each arm. When researchers are initially agnostic about the relative performance of the K arms, priors are distributed uniformly over the parameter space, i.e., Beta$(1, 1)$. In each period t, they assign treatments and observe $n_{k,t}$ responses for each arm k respectively, where $\sum_k n_{k,t}$ is fixed. 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. The distribution of each Θ_k given the data $X_k^{\{n_{k,t}\}}$ in time t is then

$$f_{\Theta_k|X_k^{\{n_{k,t}\}}}(\theta_k|x_k^{\{n_{k,t}\}}) \propto f_{X_k^{\{n_{k,t}\}}|\Theta_k}(x_k^{\{n_{k,t}\}}|\theta_k) f_{\Theta_k}(\theta_k).$$

---

[3]We will focus primarily on the Thompson sampling algorithm here, although there are many alternative approaches to the multi-armed bandit problem, such as (ϵ-)Greedy methods, which select the arm with the highest empirical mean (or with some probability ϵ of random selection), and the upper confidence bound (UCB) algorithm, which selects the arm with the highest upper bound on an uncertainty interval around its estimated value (Sutton and Barto 2018). The relative performance of each allocation rule depends on the yardstick used to measure success: statistical power, type-I error rates, and bias. Performance will also depend on the time-horizon of the trial (Villar et al. 2015).

The Beta distribution is a conjugate prior for Binomial likelihoods, and consequently the posterior follows a Beta distribution. The posterior α parameter is equal to one plus the total number of successes observed for that arm and the posterior β parameter is equal to one plus the total number of failures observed from that arm. (See appendix A.1.)

In each period t, treatment is randomly assigned according to the probability of arms being best i.e.,

$$P\left[\theta_k = \max\{\theta_1, \ldots, \theta_K\}|(x^{\{n_{1,t}\}}, \ldots, y_{K,t})\right],$$

and rewards are observed.[4] At the end of the period, the posterior is updated according to the successes and failures observed in that period, and the probability that each arm is best is re-calculated. In the subsequent period, treatment assignment continues according to the updated probabilities.

Thompson sampling can be adapted to allow for drift in parameter values over time (Gupta et al. 2011) or can account for reward probabilities that vary based on other variables that describe the context in which the action is taken (referred to as contextual bandits, see Agrawal and Goyal 2012). It can also be applied to more complex problems considered under reinforcement learning, where actions can affect future states and information about rewards may be delayed or sparse (Russo et al. 2017; Sutton and Barto 2018). In some applications, such as the ones considered here, adaptive trials end after a fixed period or when a pre-determined number of subjects have participated in the trial. In other applications, the trial stops when any arm achieves a pre-specified probability of being best; when used to establish statistical significance of effects, such stopping rules can run the risk of producing

---

[4]This value may be calculated analytically, as demonstrated in appendix A.2. In practice, we estimate the value through simulation, taking a series of random draws from the posterior probability distributions of all arms, and calculating the share of the series in which each arm had the highest draw, as implemented in the `bandit` package in `R` statistical software (Lotze and Loecher 2014).

a false discovery, as the trial may stop if the best-performing arm surpasses the target due to chance (Berman et al. 2018).

Here, the main objective is to maximize reward, not, as is common in the social sciences, to estimate average treatment effects. Indeed, if the "best" treatment arm has a much higher probability of success than the control arm, relatively few subjects will be assigned to the control arm, in which case estimates of the average treatment effect will typically have a larger standard error than under a static, balanced design. To improve power in such settings, Villar et al. (2015) propose a design that combines adaptive and static elements, where treatments are allocated adaptively, but in which a fixed portion of patients are allocated to the control group. Additionally, Nie et al. (2017) demonstrate that under certain common conditions, sample means from adaptive experiments are systematically negatively biased. (See appendix A.2.) Ex-post approaches to estimation to reduce bias in such settings, such as inverse propensity score weighting methods, can exhibit large variance (Nie et al. 2017). Consequently, if unbiased estimates of treatment effects are the primary objective of a study, a standard static design may be preferable. Alternatively, procedures to reduce bias may be integrated into the design stage of adaptive experiments.[5]

We consider an additional setting where arms are composed of factorial components, with each component contributing to the success rate of the arm. 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., Scott 2010; Shahriari et al. 2016; for more general cases, see Filippi et al. 2010). Modeling assumptions, such as the stipulation that the factors

---

[5]A static experiment may be conducted following, or alongside, an adaptive experiment to produce unbiased treatment effect estimates. Nie et al. (2017) propose a randomization algorithm founded on selective inference methods that accounts for selection effects and may have lower RMSE compared to data splitting designs with comparable sample sizes.

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.

Simulations Illustrating How Adaptive Designs Work

We illustrate properties of adaptive designs under several different simulated scenarios. In each case, there are nine arms, each with some true probability of success set according to $(\theta_1, \ldots, \theta_9)$. We set a uniform prior for all arms. As the probability of being best is equal for all of the arms prior to the start of the experiment, each arm is sampled with equal probability in the first period. We sample 100 observations for each of 10 periods, updating posterior probability of being best after each period, and assign treatment probabilities in the subsequent period accordingly. The choice of 10 periods is arbitrary, but anticipates the empirical examples presented below, which run for 10 days. In each setting, we also run a static trial with the same number of observations, but where treatment assignment probability does not change during the experiment.
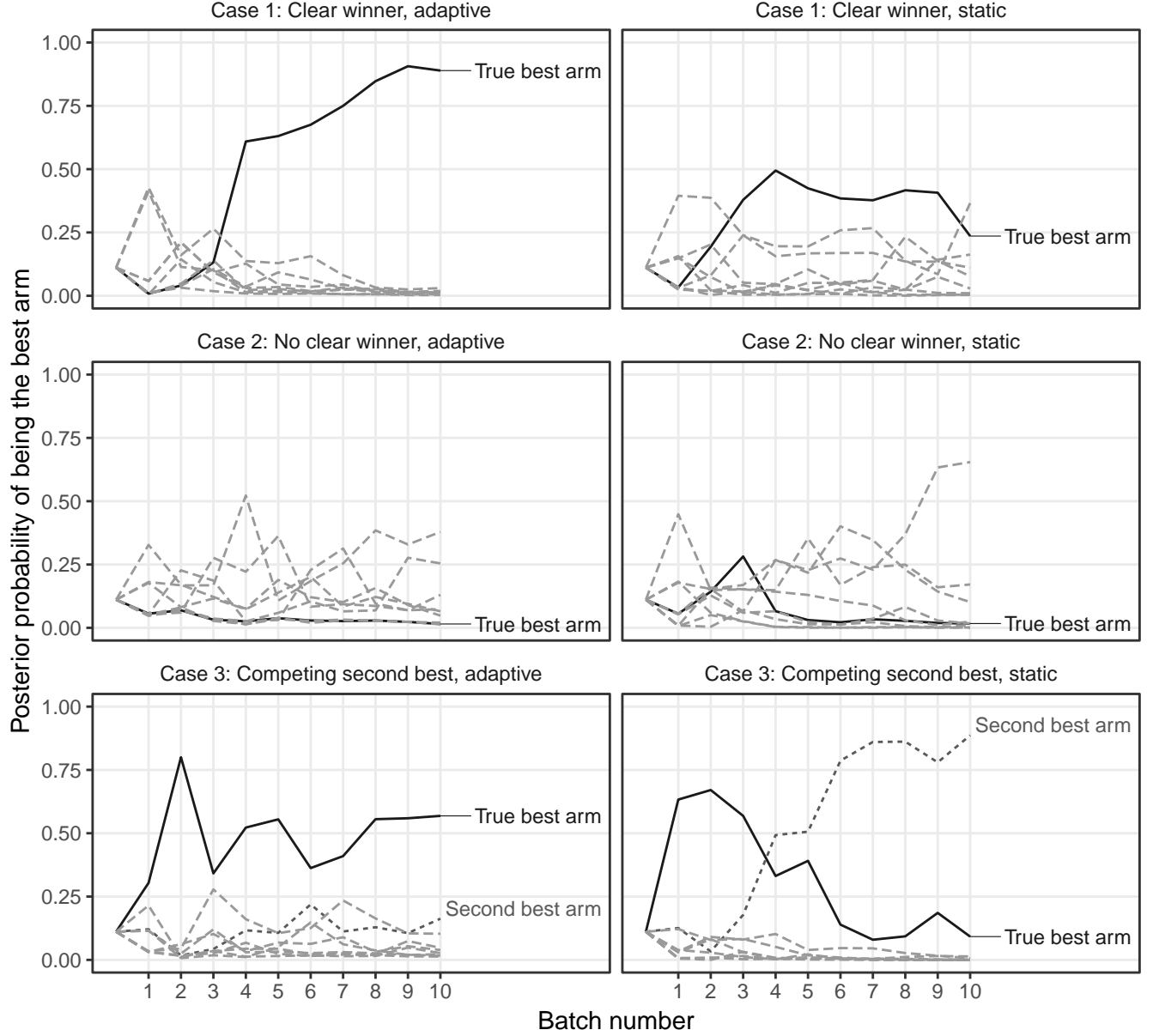
In the first case, one arm has a true 0.20 probability of success, and the remaining eight arms have a 0.10 probability of success. Within four periods, the true best arm (presented as a solid line) takes a clear lead in probability of being best in the adaptive design, shown in the top left panel of figure 1. By the end of the 10-period experiment, the true best arm is assigned a 0.89 probability of being best. In the top right panel, we plot probability of being best for each arm in a static trial. By the end of the experiment, we assign the true best arm 0.24 probability of being best in the static trial, but this is second to an alternative arm, which is assigned a 0.37 probability of being best. The "clear winner" scenario highlights the advantages of adaptive design over static design.

In the second case, the best arm has only a 0.11 true probability of success, compared to a 0.10 probability of success for the remaining 8 arms. In our simulation, during the 10-period experiment we do not correctly identify the true best arm in either the adaptive or static trials. Indeed, in our illustrative example of the adaptive design we assign the best arm only 0.02 probability of being best, whereas we assign an inferior arm 0.38 probability of being best, shown in the center left panel of figure 1. For the static experiment, shown in the center right panel, we assign the best arm 0.02 probability of being best, and an inferior arm a 0.65 probability of being best. The "no clear winner" scenario demonstrates a case where both static and adaptive designs fail to correctly identify the true best arm.

Finally, we consider a case where the best arm has a 0.20 true probability of success, a second-best arm has a 0.18 probability of success, and the remaining seven arms have 0.10 probability of success. In the adaptive design, we assign the best arm a 0.57 probability of being best, followed by the second best arm with a 0.16 probability of being best, shown in the bottom left panel of figure 1. In the static design, shown in the bottom right panel, the best and second-best arms trade places during the experiment, and by the end of the experiment we assign the second-best arm a 0.89 probability of being best, while we assign the true best arm only a 0.09 probability of being best.

To compare the performance of adaptive trials in the three scenarios more rigorously, we plot the distributions of posterior probability of being best for the true best arm across 10,000 simulations in figure 2. Means for the true best arm's probability of being best and the percent of simulations under which the true best arm was selected are presented in table 1. Across cases, the true best arm is selected as best more frequently under adaptive as compared to static designs. The posterior probability of being best is higher for the true best arm under adaptive designs for the first and third cases, but is higher under static designs in the second case.

9

Figure 1: Posterior Probabilities Over Time

The arm with the true highest probability of success is represented as a solid black line in all panels. In case 1, the best arm has a 0.20 probability of success, and the remaining eight arms have a 0.10 probability of success. In case 2, the best arm has a 0.11 probability of success, and the remaining eight arms have a 0.10 probability of success. In case 3, the best arm has a 0.20 probability of success, there is a second best arm with a 0.18 probability of success, and the remaining seven arms have 0.10 probability of success.

10

Figure 2: Distribution of Probabilities of Being the Best for True Best Arm after 10 Periods, Adaptive and Static Designs
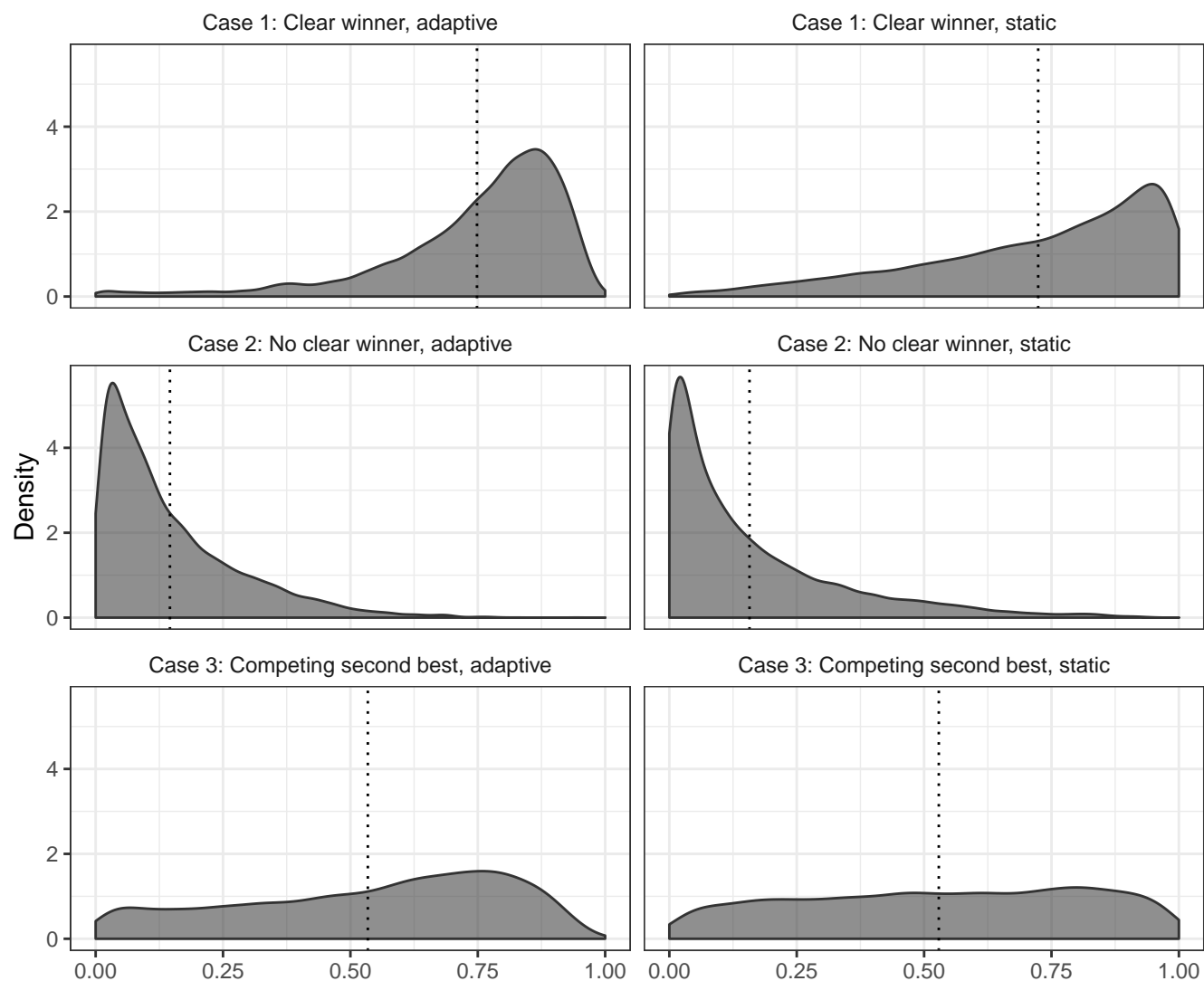
Each panel represents 10,000 simulations. Averages are represented as dotted vertical lines.

Table 1: Simulation Results

	Adaptive		Static	
	Probability best	Best arm selected	Probability best	Best arm selected
Case 1	0.968	0.748	0.921	0.724
Case 2	0.187	0.146	0.182	0.157
Case 3	0.708	0.534	0.634	0.529

Note: "Probability best" columns present the mean final posterior probability of being best assigned to the true best arm. "Best selected" columns present the portion of simulations under which the true best arm was selected.

Study One: Two Multi-Arm Adaptive Trials

Simulations provide useful intuition about the conditions under which adaptive designs are helpful, but empirical applications allow us to illustrate how adaptive designs are implemented and analyzed in settings relevant to political scientists. Our two empirical applications address the wording of ballot measures. For the first study, we recruited 1,000 subjects from Amazon's Mechanical Turk (MTurk). Convenience samples obtained on MTurk are far from representative of the national population but do provide a fertile testing ground for experimental studies. Recent meta-analyses have revealed a close correspondence of experimental estimates obtained on MTurk and probability samples (Mullinix et al. 2015; Coppock 2018; Coppock et al. 2018). Our study ran from June 21st, 2018 to June 30th, 2018, and we paid subjects $1 each for their participation.

Design

After answering a series of demographic questions, all subjects rated two ballot measures, one on minimum wage and one on right-to-work. We adapted the wording of the proposed measures from real proposals, making only small wording changes to facilitate consistency across arms. We implemented a composite adaptive/static design similar to the controlled Gittins approach recommended by Villar et al. (2015). For each type of ballot measure 90% of treatments were assigned according to Thompson sampling, and 10% were assigned under simple random assignment, with equal probability of assignment for each treatment.

The minimum wage treatments were drawn from ballot measures proposed in Colorado, Florida, Illinois, Nevada, and New Jersey. We generated two versions of each of these five proposals, varying whether the current value of the minimum wage was displayed, resulting

in ten unique minimum wage treatments.[6] The right-to-work treatments were adapted from ballot measures in Missouri, North Dakota, Oklahoma, and South Dakota. For each of these, we created versions that did or did not describe the ballot measure as a "constitutional amendment," resulting in eight unique right-to-work treatments. For both rating tasks, the outcome question asked was, "If this measure were on the ballot in your state, would you vote in favor or against?" The full text of all treatments is presented in table 2.

Our analysis procedure relies on an assumption that the true parameter values remain constant over the duration of our study. This assumption is justified by the qualitative expectation that day-to-day changes in the latent attractiveness of each measure are unlikely to be meaningful. An alternative (but much higher variance) analysis strategy would employ inverse probability weighting.

Results

We present two sets of results. The first, presented in figure 3, is the over time development of the posterior probability that each arm is optimal. The second, presented in figure 4, is a straightforward comparison of the average approval of each proposal.[7]

The minimum wage study yielded no clear winner. The winning arm, by a hair, was Proposal 3 (B, without current minimum wage), with a posterior probability of being best

---

[6]Minimum wage rates are presented in appendix table B.5. For states that do not have a state minimum wage, we imputed the federal minimum wage value.
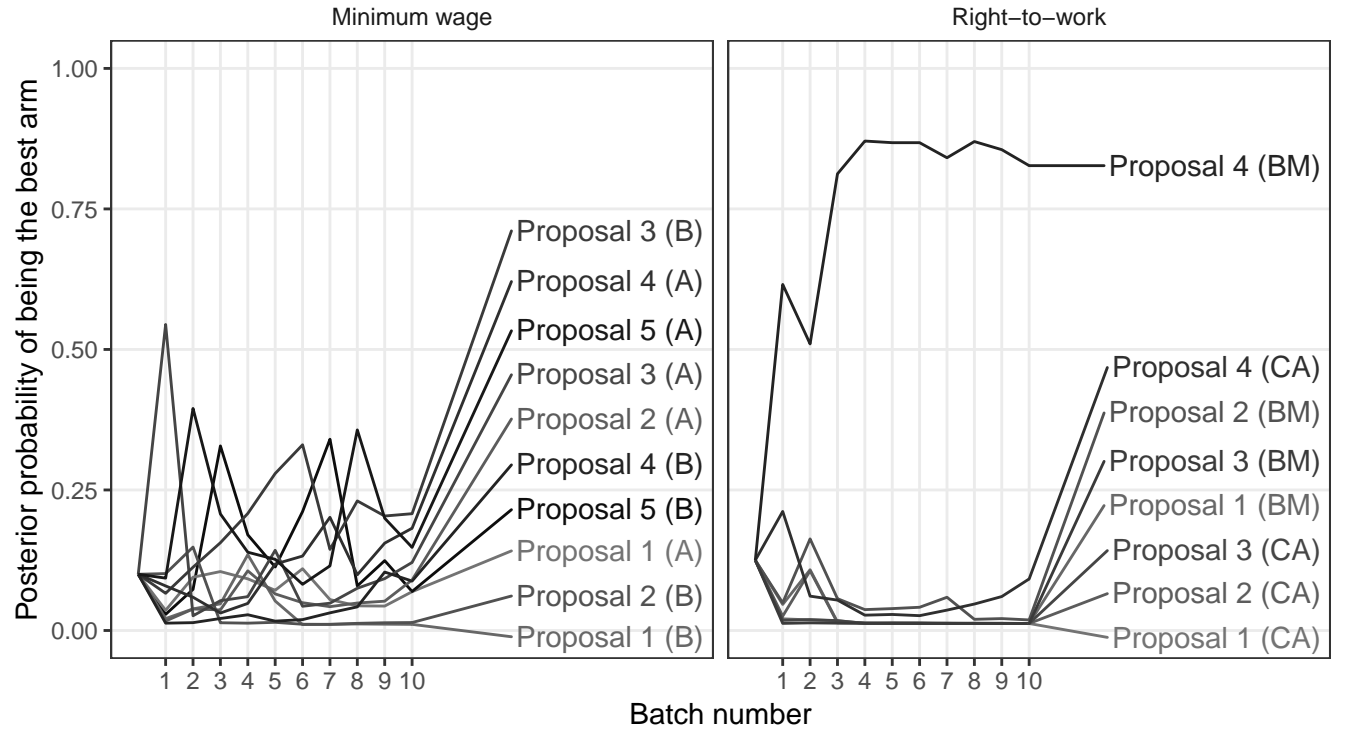
[7]A key feature of adaptive designs is that the probability of assignment to each condition varies over time. A default analytic approach when the probabilities of assignment are different for different units is to weight each observation by the inverse of the probability of assignment to the condition that it is in (see Gerber and Green 2012, chapter 4 for a textbook introduction to inverse probability weighting, or IPW). We present IPW estimates, which are similar to the unweighted results presented in the text, in appendix figure C.8.

Table 2: Treatments and Outcome Measures

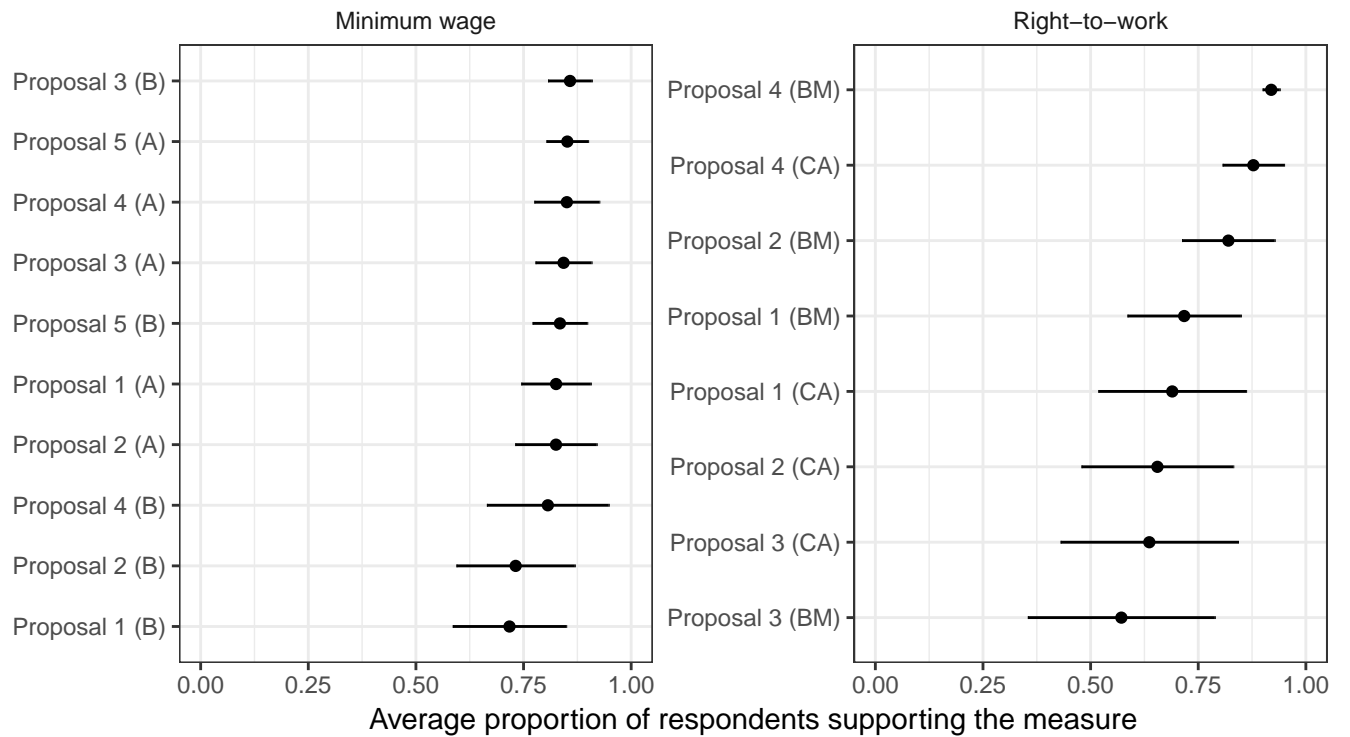|	Minimum Wage	Right to Work
Question Text	Imagine that the following ballot measure were up for a vote in your state. The measure would: [**ballot measure text**]. If this measure were on the ballot in your state, would you vote in favor or against? [I would vote in favor of this measure; I would vote against this measure]	Imagine that the following ballot measure were up for a vote in your state. The measure would [**amend the State Constitution to**]: [**ballot measure text**]. If this measure were on the ballot in your state, would you vote in favor or against? [I would vote in favor of this measure; I would vote against this measure]
Proposal 1	increase the minimum wage [**from {current}**] to {current + 1} per hour, adjusted annually for inflation, and provide that no more than $3.02 per hour in tip income may be used to offset the minimum wage of employees who regularly receive tips.	prohibit, as a condition of employment, forced membership in a labor organization (union) or forced payments of dues or fees, in full or pro-rata ("fair-share"), to a union. The measure will also make any activity which violates employees' rights provided by the bill illegal and ineffective and allow legal remedies for anyone injured as a result of another person violating or threatening to violate those employees' rights. The measure will not apply to union agreements entered into before the effective date of the measure, unless those agreements are amended or renewed after the effective date of the measure.
Proposal 2	raise the minimum wage [**from {current}**] to {current + 1} per hour effective September 30th, 2021. Each September 30th thereafter, minimum wage shall increase by $1.00 per hour until the minimum wage reaches {current + 5} per hour on September 30th, 2026. From that point forward, future minimum wage increases shall revert to being adjusted annually for inflation starting September 30th, 2027.	The right of persons to work may not be denied or abridged on account of membership or nonmembership in any labor union or labor organization, and all contracts in negation or abrogation of such rights are hereby declared to be invalid, void, and unenforceable.
Proposal 3	Shall the minimum wage for adults over the age of 18 be raised [**from {current}**] to {current + 1} per hour by January 1, 2019?	ban any new employment contract that requires employee to resign from or belong to a union, pay union dues, or make other payment to a union. Required contributions to charity or other third party instead of payments to union are also banned. Employees must authorize payroll deduction to unions. Violations of the section is a misdemeanor.
Proposal 4	raise the minimum wage [**from {current}**] to {current + 1} per hour worked if the employer provides health benefits, or {current + 2} per hour worked if the employer does not provide health benefits.	No person shall be deprived of life, liberty or property without due process of law. The right of persons to work shall not be denied or abridged on account of membership or nonmembership in any labor union, or labor organization.
Proposal 5	raise the State minimum wage rate [**from {current}**] to at least {current + 1} per hour, and require annual increases in that rate if there are annual increases in the cost of living.	

Boldface text indicates randomly varied elements.

Figure 3: Study One: Overtime Best Probabilities

Posterior probabilities updated after each day's data collection according to the algorithm described in footnote 4. "A" versions of the minimum wage proposals include the current minimum wage and "B" versions do not. "CA" versions of the right-to-work proposals are describes as "constitutional amendments" and "BM" versions are not.

Figure 4: Mean Vote Outcomes for Study One

Group means are unweighted. "A" versions of the minimum wage proposals include the current minimum wage and "B" versions do not. "CA" versions of the right-to-work proposals are describes as "constitutional amendments" and "BM" versions are not.

of 0.208 and a raw success rate of 0.858 over 183 trials. This arm was closely followed by Proposal 5 (A, with current minimum wage), Proposal 4 (A, with current minimum wage), and Proposal 3 (A, with current minimum wage). Out of 10 arms, only two had raw success rates under 0.8; with similarly high probabilities of success across several arms, the best arm was not easily distinguishable.

By contrast, the right-to-work experiment immediately produced a standout arm that proved to be very successful throughout the study. Proposal 4 (framed as a ballot measure) ended with an 0.827 posterior probability of being best, with a raw success rate of 0.920 over 721 trials. The second-best arm was also Proposal 4 (framed as a constitutional amendment) with a posterior probability of being best of 0.092 and a raw success rate of 0.878 over 82 trials. This example illustrates the fact that even a small (4.2 percentage point) difference in success rates can translate into a very large (73.5 percentage point) difference in the posterior probability of being best.

We can use these results to inform guesses about how our experiment would have fared if we had used a standard static design instead of the adaptive design. The static design would have sampled each of the 10 arms in the minimum wage experiment 100 times each and each of the 8 arms in the right-to-work experiment 125 times each. Treating observed success rates as the truth in simulations of the minimum wage experiment, we picked the best proposal 28% of the time in adaptive experiments and 24% of the time in static experiments. On average, we assigned this arm a posterior probability of being the best of 0.21 in both adaptive and static experiments. Conversely, in the right-to-work experiment we picked the best proposal 96% of the time in adaptive experiments and 86% of the time in static experiments. On average, we assigned this arm a posterior probability of being the best of 0.88 in adaptive experiments and 0.76 in static experiments.

Considering figure 4, we note that a feature of the adaptive design is that the proposals with the highest raw success rates also have the smallest standard errors, as these arms tend

to receive more samples than arms with lower success rates. This is appropriate when the performance of the best arm is considered a priority by the researcher and success rates of poorly-performing arms are not of particular interest. For the minimum wage experiment, standard errors around our estimate of probability of success for the best arm in a simulated static design would have been, on average, 89% as large as those under a comparable adaptive design. This reflects the relatively large standard errors in the adaptive trial when the best arm is under-sampled. For the right-to-work experiment, standard errors on the probability of success of the best arm in a static experiment would have been on average over 200% as large as those under a comparable adaptive design. The adaptive design appears to offer advantages in terms of the precision with which the best performing arm's success rate is evaluated, when there is a standout arm.

Study Two: An Adaptive Conjoint Trial

Study two extends the adaptive design beyond simple comparisons of unique treatments to conjoint designs, where multiple dimensions of treatments are varied.[8] 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 treatments.

In such cases, if we consider all possible combinations of factors, discovery of the best treatment arm may not be feasible under implementation constraints using the approach taken in the first study, where we updated our priors on each arm separately. Instead,

---

[8]Conjoint experiments were introduced to political science by Hainmueller et al. (2014). For an application of the conjoint design to the study of candidate partisanship on preferences over mayoral candidate attributes, see Kirkland and Coppock (2018).

we use a statistical model to pool information across conditions. To keep the number of parameters manageable, we impose an additive probit model, allowing us to then calculate the posterior probability that each arm is best.[9]

Here, 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.[10] Modeling could account for interactions through regularization (following, e.g., Imai and Ratkovic 2013, or imposing increasing penalties on higher-order interactions).

For this study, we recruited a convenience sample of 979 subjects from Lucid, collected as a target of 100 responses a day for each of 10 waves. As with convenience samples obtained on Mechanical Turk, samples from Lucid are not nationally representative; their suitability as a methodological testing ground for social science applications is discussed in Coppock and McClellan (2019). Our study ran from November 26, 2018 to December 7, 2018. We paid \$1 for each survey response; participants were compensated in money, points,

---

[9]Scott (2010) refers to this general approach as a *fractional factorial bandit*; we use the term *adaptive conjoint* to reflect that the design represents a conjoint experiment with attributes whose probability distributions change over time. This nomenclature also avoids confusion with use of the term "fractional factorial" to refer to subsetting of full factorial designs to facilitate estimation of targeted effects.

[10]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.

or other rewards depending on how they were enrolled in the Lucid subject pool.[11] The experimental design and analysis plan were pre-registered with EGAP. 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 4–12, 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. Subjects were paid $1 for their participation in each wave.

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 3, 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

---

[11]Due to constraints in survey implementation, 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. Because we assume that the latent attractiveness of each policy remains constant over time, these implementation issues have no effect on our analysis.

Table 3: 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

of main effects. Following the first wave of data collection, we then update this prior by modeling probability of success 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 probability of success. On the subsequent wave, sampling of each condition is conducted in proportion to this probability.[12] Probabilities are updated with additional data following each wave's sampling.[13]

---

[12]Posterior probability is calculated by simulation and is not exact. In a given 10,000 simulated draws, we may have an arm that does not have highest predicted probability of success in any of the observed draws. As a result, we may not sample that arm in a given wave. However, this does not mean that the probability of sampling this arm is zero, as, given a different set of draws, there is always some probability of that arm having the highest predicted probability of success. If the researcher is concerned with ensuring a minimal probability of sampling for each arm, they may follow a version of the composite approach implemented in the first study. Simulation is implemented with the `MCMCpack` package in `R` statistical software (Martin et al. 2011).

[13]This design depends on a parametric model and does not straightforwardly facilitate non-parametric strategies to estimate the average marginal component effects (AMCE) proposed by Hainmueller et al. (2014). This estimand marginalizes over the joint distribution of other attribute levels, and, in conventional conjoint settings, AMCEs may be estimated directly from linear regression. However, as both the marginal and joint distributions of attribute levels change throughout the duration of an adaptive study, AMCEs for different levels within an attribute estimated in this way would marginalize over different distributions, and would need to be re-weighted to the desired distribution.

Results

We compare the results from the adaptive conjoint experiment to those from a conventional conjoint experiment with identical treatments and outcome measures, but implemented under a static design. Figure 5 reports attribute effect estimates under the static design on the left side, and the adaptive design on the right side. Comparatively, the effect estimates under the adaptive design presented are generally consistent in order and magnitude with those under the static design, but with much larger standard errors. This is to be expected in such adaptive designs, as coefficients for factor levels are estimated as a difference from the baseline. The imbalance in sampling that is an intended consequence of the adaptive design will generally lead to larger standard errors when estimating differences in factor levels.

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 to four separate adaptive experiments, each running over the course of the study. Considering the left panel in figure 6, 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 6, 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 maintain the status quo for public funding and disclosures.

Figure 7 presents the posterior probability of being best for each condition as the joint probabilities of the component attribute levels. For the adaptive design, the most preferred

Figure 5: Study Two: Estimated Average Effects of Attributes, Static and Adaptive Designs
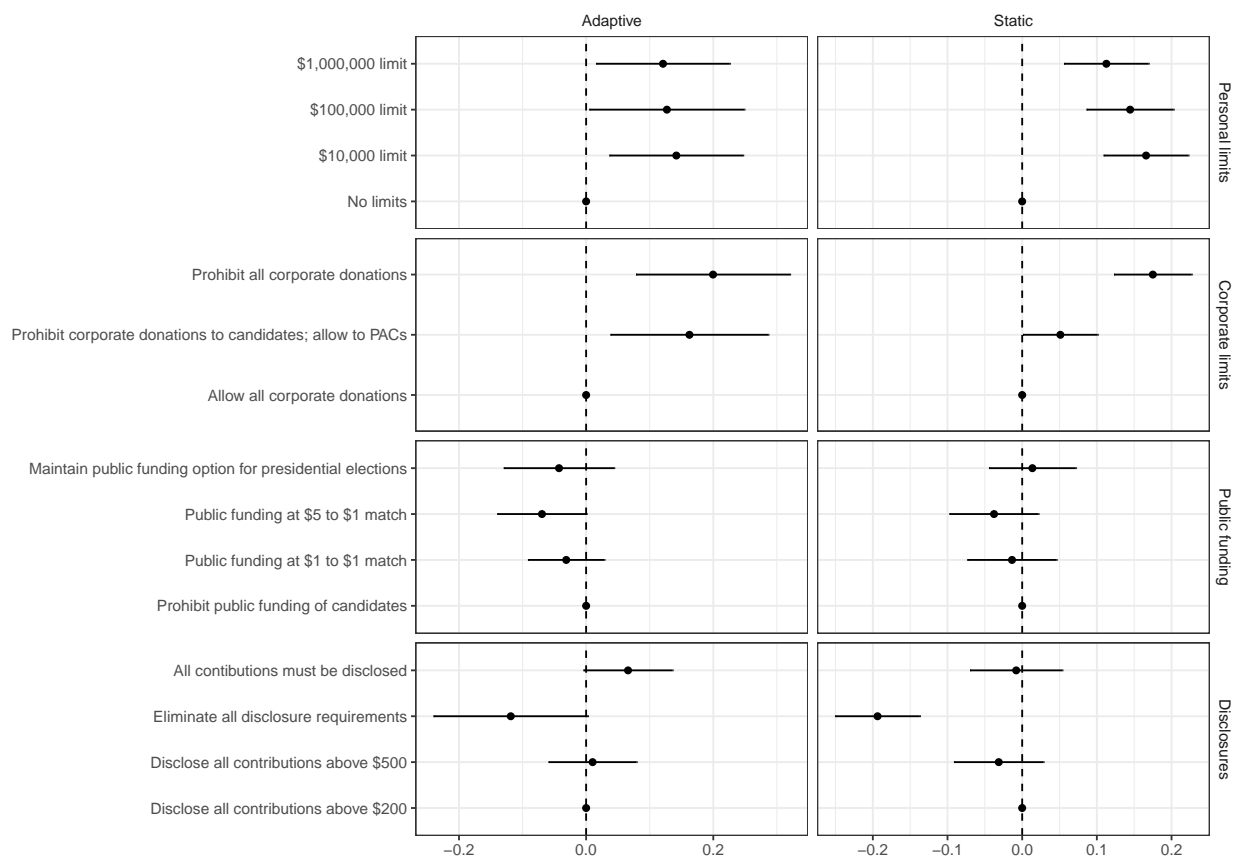
Estimates are coefficients of linear regression estimators, with standard errors clustered at the subject level. Estimates with continuous outcomes are presented in appendix figure C.9

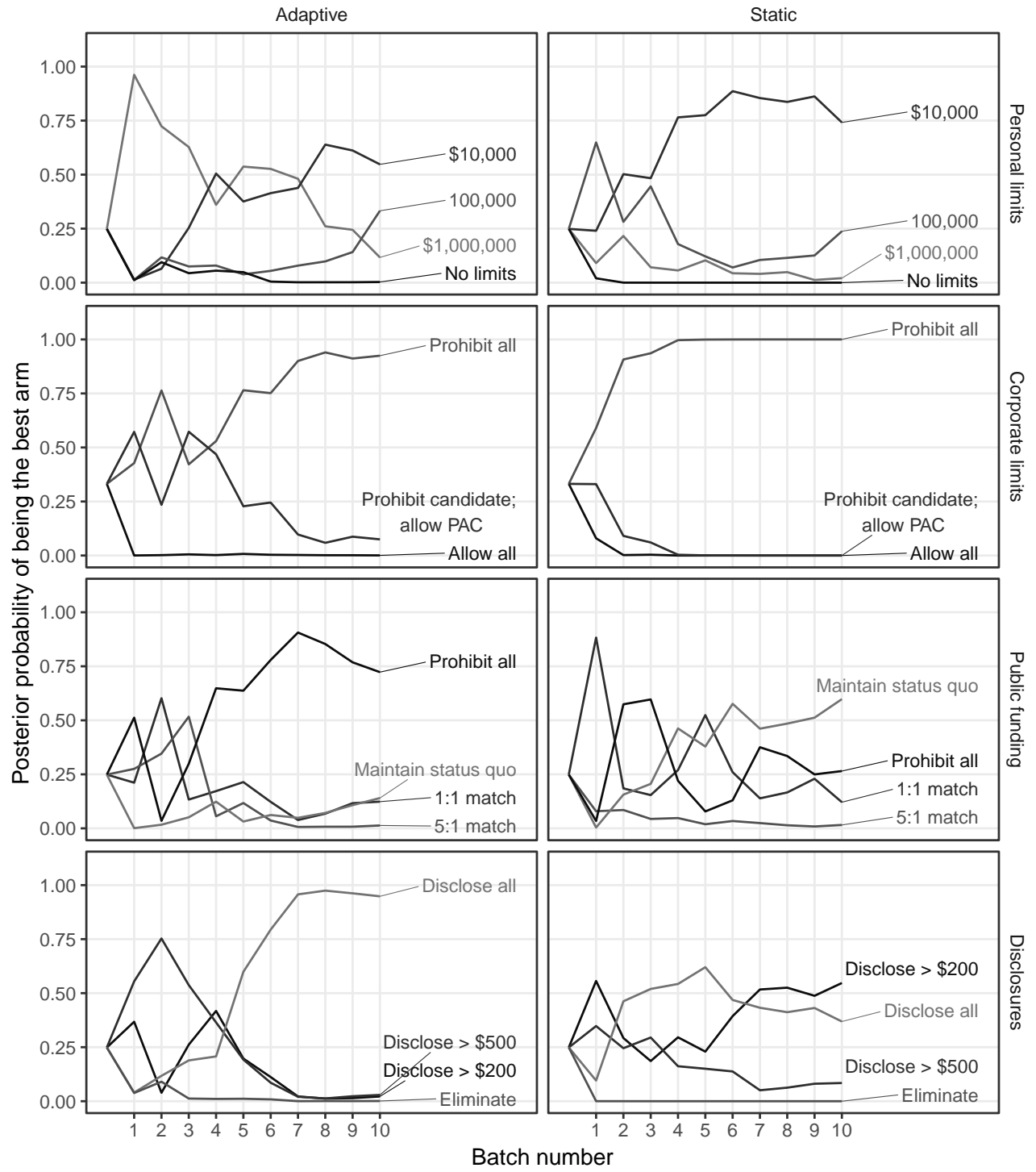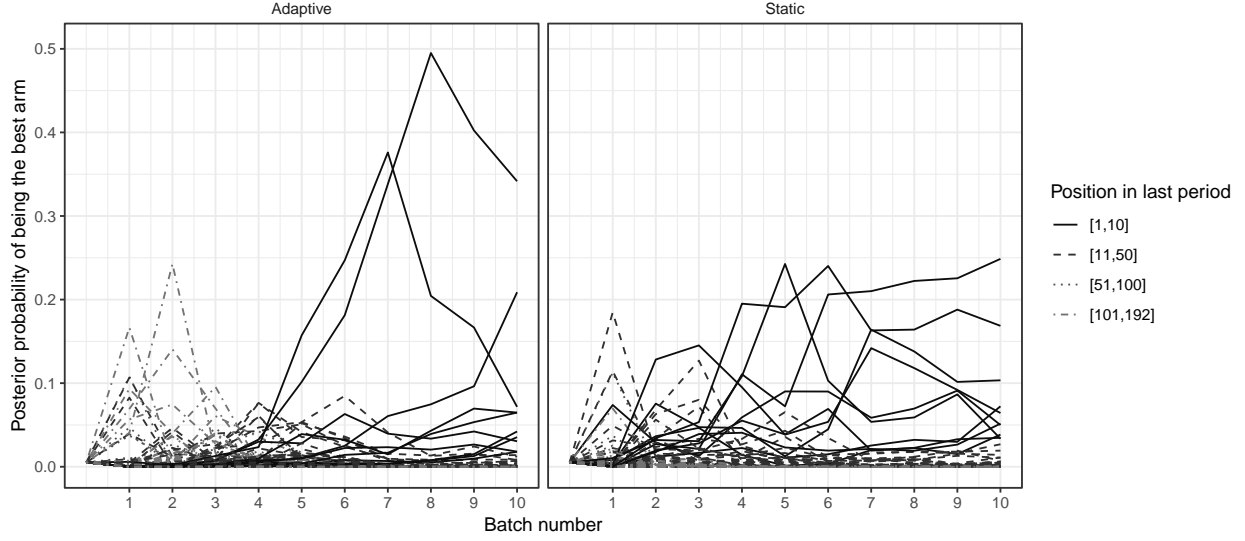Figure 6: Study Two: Overtime Posterior Probabilities by Attribute

Figure 7: Study Two: Overtime Posterior Probabilities, Joint

profile includes each of the top attribute levels presented in the left panel of figure 6, with a final posterior probability of being best of 0.342; 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 posterior probability of being best of 0.209; 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 posterior probability of being best of 0.072. For the static design, the most preferred profile includes each of the top attribute levels presented in the right panel of figure 6, with a final posterior probability of being best of 0.249; the second-most preferred profile would include the same levels of the other attributes, but requiring disclosures of all contributions, with a final posterior probability of being best of 0.169; 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 $200, with a final posterior probability of being best of 0.103. 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 the public 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 the public, the lower the personal limit, the better. Public preferences are also at odds with the current policy with respect to corporate contributions. The public 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, the public is also out of step with current federal law, which requires disclosures of contributions greater than $200. The public most prefers a system in which all contributions are disclosed. Finally, although the public 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.

Discussion

The growth and development of experimentation in the social sciences has led to increasing sophistication in the design of multi-arm trials. Although the adaptive allocation of subjects to treatment arms over time adds complexity to a trial's implementation and analysis, the payoff may be considerable. When one arm is truly superior to the others, an adaptive trial can locate the winning arm more reliably than a static design. Moreover, because the

adaptive trial allocates more sample to the winning arm, the experimenter learns more about the attributes of the winner at the conclusion of the study. Our simulations and the empirical example of right-to-work ballot measures illustrate just how valuable adaptive designs can be in the context of a truly superior arm. The level of public support for the winning right-to-work ballot measure was estimated with a standard error that was 42% as large as would have been the case under a static design.

The adaptive allocation of subjects, however, is of little value when no treatment arm truly stands above the others. In such cases, adaptive allocation follows clues that are the product of sampling variability rather than the true superiority of an arm. As the minimum wage application suggests, at best an adaptive design winnows out some inferior arms and reallocates the sample to obtain a somewhat more precise assessment of the winning arm's payoff. In this application, an adaptive design still outperformed the expected outcome from a static design in terms of the precision with which the winning arm's payoff was estimated, but the gains were far less dramatic than those of the right-to-work application. Researchers considering the use of adaptive sample allocation should therefore reflect on their prior beliefs about the effectiveness of the treatment arms they plan to investigate. The more variable the true effects, the more valuable an adaptive design is likely to be.

This point also holds for studies in which the aim is to compare treatment arms to an untreated control group. Consider a hybrid design in which a static allocation is made to an untreated control group throughout the trial, but sample is allocated adaptively to the treatment arms. The adaptive component of the design aims to locate the best performing treatment arm, but the static component ensures that the control group always receives ample subjects regardless of how it performs over the trial. When one treatment arm is truly superior, this design will allocate substantially more subjects to it and will therefore render a more precise estimate of the treatment effect vis-à-vis the control group. On the other hand, the gains may be negligible if the treatment arms are in fact similarly effective.

29

One important research frontier is the efficient allocation of sample in the context of factorial designs. To our knowledge, ours is the first paper to consider adaptive design in the context of conjoint experiments in the social sciences, where the research aim is to find the combination of traits with the highest payoff. Because the number of possible treatment arms is large relative to the number of subjects, adaptive design alone may be unable to isolate the best treatment combination with high probability over a fixed data collection schedule. In this case, adaptive design requires the assistance of modeling assumptions to reduce the set of promising treatment combinations. Further empirical tests are needed in order to assess whether additive models perform well in other substantive domains and whether the performance of more sophisticated models could be improved by restructuring the initial search phase to more reliably explore interactions among factors.

Another important research frontier is designing adaptive trials when the true underlying performance of the treatment arms is believed to be changing over time. The empirical applications discussed above sidestepped this concern by focusing on domains of public opinion (e.g., views on campaign finance regulation) that are unlikely to be affected by day-to-day events and by measuring opinion over a relatively short span of time. However, one can readily imagine other applications where the "payoffs" associated with various treatment arms change over time during a trial. For example, an adaptive trial designed to gauge the political campaign ad that attracts the most support for the advertising candidate might operate in a fast-moving environment in which a candidate's popularity erodes over time. An adaptive trial might start out with one front-running ad, allocate more sample to it, and then see the apparent performance of that ad deteriorate as the candidate's popularity drifts downward. Future research, building on the work of Granmo and Berg (2010) and Gupta et al. (2011), will need to grapple with the added complexity of dynamics in both the design and analysis of adaptive trials.

References

Agrawal, Shipra and Navin Goyal. 2012. "Thompson Sampling for Contextual Bandits with Linear Payoffs." arXiv e-prints arXiv:1209.3352.

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.

Berman, Ron, Leonid Pekelis, Aisling Scott and Christophe Van den Bulte. 2018. "p-Hacking and False Discovery in A/B Testing.".

Berry, Donald A and Bert Fristedt. 1985. *Bandit problems: Sequential Allocation of Experiments.* New York and London: Chapman and Hall.

Chin, Richard. 2016. *Adaptive and Flexible Clinical Trials.* Boca Raton: CRC Press.

Chow, Shein-Chung and Mark Chang. 2008. "Adaptive design methods in clinical trials–a review." *Orphanet Journal of Rare Diseases* 3(1):11.

Cook, Joseph, Marc Jeuland, Brian Maskery, Donald Lauria, Dipika Sur, John Clemens and Dale Whittington. 2009. "Using Private Demand Studies to Calculate Socially Optimal Vaccine Subsidies in Developing Countries." *Journal of Policy Analysis and Management* 28(1):6–28.

Coppock, Alexander. 2018. "Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach." *Political Science Research and Methods* pp. 1–16.

Coppock, Alexander and Oliver A. McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research & Politics* . Forthcoming.

Coppock, Alexander, Thomas J. Leeper and Kevin J. Mullinix. 2018. "Generalizability of heterogeneous treatment effect estimates across samples." *Proceedings of the National Academy of Sciences* 115(49):12441–12446.

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.

Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W.W. Norton.

Graepel, Thore, Joaquin Quiñonero Candela, Thomas Borchert and Ralf Herbrich. 2010. Web-Scale Bayesian Click-Through Rate Prediction for Sponsored Search Advertising in Microsoft's Bing Search Engine. In *Proceedings of the 27th International Conference on Machine Learning (ICML 2010)*, ed. J. Fürnkranz and T. Joachims. Omnipress pp. 13–20.
URL: *http://www.icml2010.org/papers/901.pdf*

Granmo, Ole-Christoffer and Stian Berg. 2010. Solving Non-stationary Bandit Problems by Random Sampling from Sibling Kalman Filters. In *23rd International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2010, Cordoba, Spain, June 1-4, 2010, Proceedings, Part III (IEA-AIE 2010)*, ed. Nicolás García-Pedrajas, Francisco Herrera, Colin Fyfe, José Manuel Benítez and Moonis Ali. Vol. 6098 of *Lecture Notes in Computer Science* Berlin, Heidelberg: Springer-Verlag pp. 199–208.
URL: *doi.org/10.1007/978-3-642-13033-5_21*

Gupta, Neha, Ole-Christoffer Granmo and Ashok Agrawala. 2011. Thompson Sampling for Dynamic Multi-armed Bandits. In *Proceedings of the 2011 10th International Conference*

on Machine Learning and Applications and Workshops - Volume 01. ICMLA '11 Washington, DC: IEEE Computer Society pp. 484–489.

URL: *doi.org/10.1109/ICMLA.2011.144*

Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22(1):1–30.

Hainmueller, Jens and Dominik Hangartner. 2013. "Who Gets a Swiss Passport? A Natural Experiment in Immigrant Discrimination." *American Political Science Review* 107(1):159–187.

Imai, Kosuke and Marc Ratkovic. 2013. "Estimating treatment effect heterogeneity in randomized program evaluation." *The Annals of Applied Statistics* 7(1):443–470.

Kirkland, Patricia A. and Alexander Coppock. 2018. "Candidate Choice Without Party Labels: New Insights from Conjoint Survey Experiments." *Political Behavior* 40(3):571–591.

Li, Lihong, Wei Chu, John Langford and Robert E. Schapire. 2010. "A Contextual-Bandit Approach to Personalized News Article Recommendation." arXiv e-prints arXiv:1003.0146.

Lotze, Thomas and Markus Loecher. 2014. *bandit: Functions for simple A/B split test and multi-armed bandit analysis.* R package version 0.5.0.

Ludwig, Jens, Jeffrey R Kling and Sendhil Mullainathan. 2011. "Mechanism Experiments and Policy Evaluations." *Journal of Economic Perspectives* 25(3):17–38.

Martin, Andrew, Kevin Quinn and Jong Hee Park. 2011. "MCMCpack: Markov Chain Monte Carlo in R." *Journal of Statistical Software, Articles* 42(9):1–21.

Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2:109–138.

Nie, Xinkun, Xiaoying Tian, Jonathan Taylor and James Zou. 2017. "Why Adaptively Collected Data Have Negative Bias and How to Correct for It." arXiv e-prints arXiv:1708.01977.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(2):200–249.

Russo, Daniel, Benjamin Van Roy, Abbas Kazerouni, Ian Osband and Zheng Wen. 2017. "A Tutorial on Thompson Sampling." arXiv e-prints arXiv:1707.02038.

Scott, Steven L. 2010. "A modern Bayesian look at the multi-armed bandit." *Applied Stochastic Models in Business and Industry* 26(6):639–658.

Scott, Steven L. 2015. "Multi-armed bandit experiments in the online service economy." *Applied Stochastic Models in Business and Industry* 31(1):37–45.

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.

Sutton, Richard S and Andrew G Barto. 2018. *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT press.

Sydes, Matthew R, Mahesh K. B. Parmar, Malcolm D. Mason, Noel W. Clarke, Claire Amos, John Anderson, Johann de Bono, David P. Dearnaley, John Dwyer, Charlene Green, Jordana Jovic, Alistair W. S. Ritchie, J. Martin Russell, Karen Sanders, George Thalmann

and Nicholas D. James. 2012. "Flexible trial design in practice-stopping arms for lack-of-benefit and adding research arms mid-trial in STAMPEDE: a multi-arm multi-stage randomized controlled trial." *Trials* 13(1):168.

Thompson, William R. 1933. "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples." *Biometrika* 25(3/4):285–294.

Thompson, William R. 1935. "On the theory of apportionment." *American Journal of Mathematics* 57(2):450–456.

Villar, Sofía S, Jack Bowden and James Wason. 2015. "Multi-armed Bandit Models for the Optimal Design of Clinical Trials: Benefits and Challenges." *Statistical Science* 30(2):199.

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

Supporting Information for Adaptive Experimental Design:

Prospects and Applications in Political Science

February 12, 2019

This document provides additional information referenced in the main paper.

A   Worked Example

A.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 $\text{Beta}(1,1)$, with density

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

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,

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

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.

A.2 Naive Estimation under Thompson Sampling

As an example, consider an experiment where we are comparing two treatment arms, with probability of success 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., $\text{Beta}(1,1)$. 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 $\text{Beta}(2,1)$, while the posterior for Θ_2 is $\text{Beta}(1,2)$.

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

$$
\begin{aligned}
P\left(\theta_1 \geq \theta_2 \mid x^{\{n_{1,t}\}}, x^{\{n_{2,t}\}}\right) &= \int_{\theta_1=0}^{1} \int_{\theta_2=0}^{\theta_1} f_{\Theta_2 \mid X^{\{n_{2,t}\}}}(\theta_2 \mid x^{\{n_{2,t}\}}) f_{\Theta_1 \mid X^{\{n_{1,t}\}}}(\theta_1 \mid x^{\{n_{1,t}\}}) d\theta_2 d\theta_1 \\
&= \int_{\theta_1=0}^{1} c_{1,t} \theta_1^{\alpha_{1,t}-1}(1-\theta_1)^{\beta_{1,t}-1} \int_{\theta_2=0}^{\theta_1} c_{2,t} \theta_2^{\alpha_{1,t}-1}(1-\theta_2)^{\beta_{1,t}-1} d\theta_2 d\theta_1
\end{aligned}
$$

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 $t = 1$,

$$= \frac{\Gamma(3)}{\Gamma(2)\Gamma(1)} \frac{\Gamma(3)}{\Gamma(1)\Gamma(2)} \int_{\theta_1=0}^{1} \theta_1{}^1 (1-\theta_1)^0 \int_{\theta_2=0}^{\theta_1} \theta_2{}^0 (1-\theta_2)^1 d\theta_2 d\theta_1$$

$$= 4 \int_{\theta_1=0}^{1} \theta_1 \left(\int_{\theta_2=0}^{\theta_1} (1-\theta_2) d\theta_2 \right) d\theta_1$$

$$= 4 \int_{\theta_1=0}^{1} \theta_1 \left(\theta_1 - \frac{\theta_1{}^2}{2} \right) d\theta_1$$

$$= \frac{5}{6}$$

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 A.4. A state is defined by the number of times we see each treatment and the number of successes we observe under each treatment,

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

While marginal sample means are systematically negatively biased, conditional on being selected as best, the sample mean is positively biased. The sample mean 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 A.4: Possible States after Two Periods

$n_{1,2}$	$n_{2,2}$	$|x_1^{\{n_{1,2}\}}|$	$|x_2^{\{n_{2,2}\}}|$	$P(\theta_1 \geq \theta_2 | s)$	$P(s)$
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	0.4	0.063
1	2	0	0	0.6	0.063

arm is, in expectation, only 0.458.

$$\mathrm{E}\left[\widehat{\theta}_k\right] = \mathrm{E}\left[\mathrm{E}\left[\widehat{\theta}_k | s\right]\right] = \sum_s \mathrm{E}\left[\widehat{\theta}_k | s\right] \cdot P(s)$$

$$\approx 0.458$$

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.

$$\mathrm{E}\left[\widehat{\theta}_{\max}\right] = \mathrm{E}\left[\mathrm{E}\left[\widehat{\theta}_{\max} | s\right]\right] = \sum_s \mathrm{E}\left[\widehat{\theta}_{\max} | s\right] \cdot P(s)$$

$$\approx 0.708$$

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.

B   Minimum Wage Rates

Table B.5: 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

C Additional Analyses

Figure C.8: Mean Vote Outcomes for Study One (IPW)
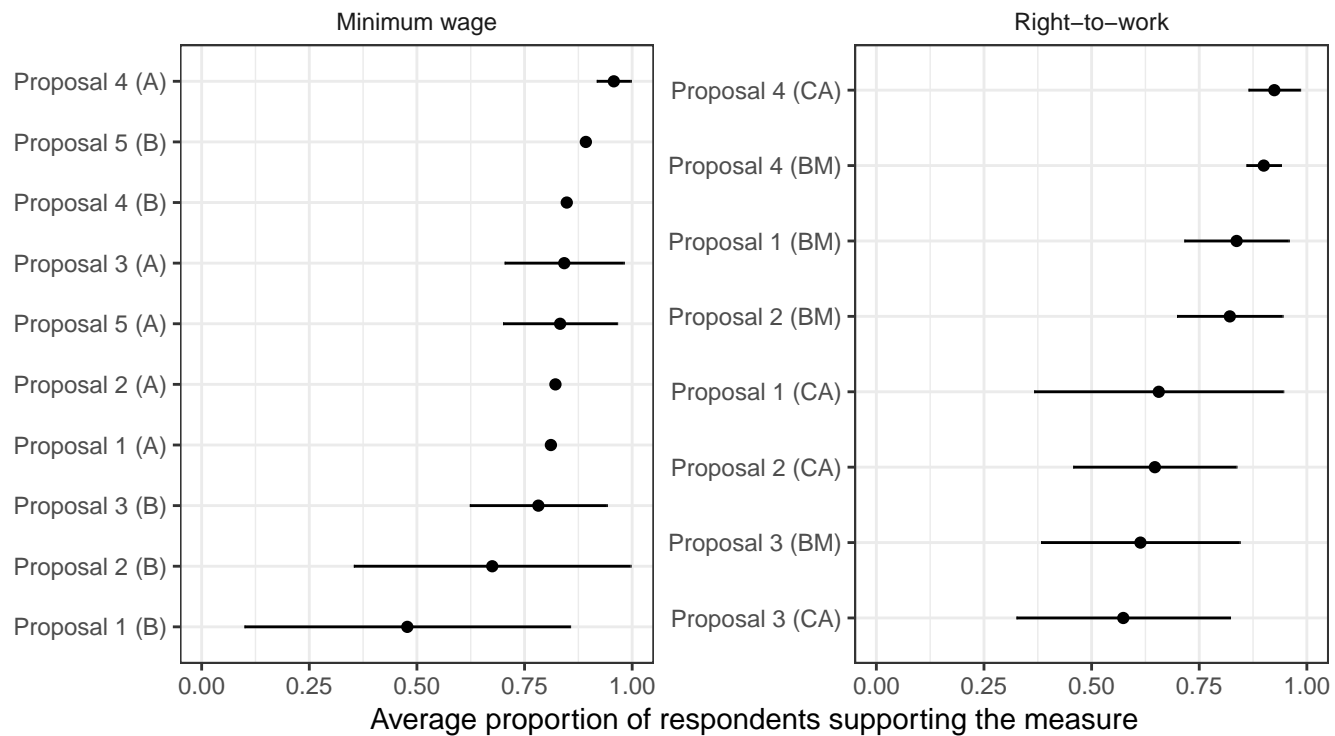
Group means are weighted by inverse probability weights. "A" versions of the minimum wage proposals include the current minimum wage and "B" versions do not. "CA" versions of the right-to-work proposals are describes as "constitutional amendments" and "BM" versions are not.

Figure C.9: Study Two: Effects of Attributes, Static and Adaptive Designs, Continuous Outcomes
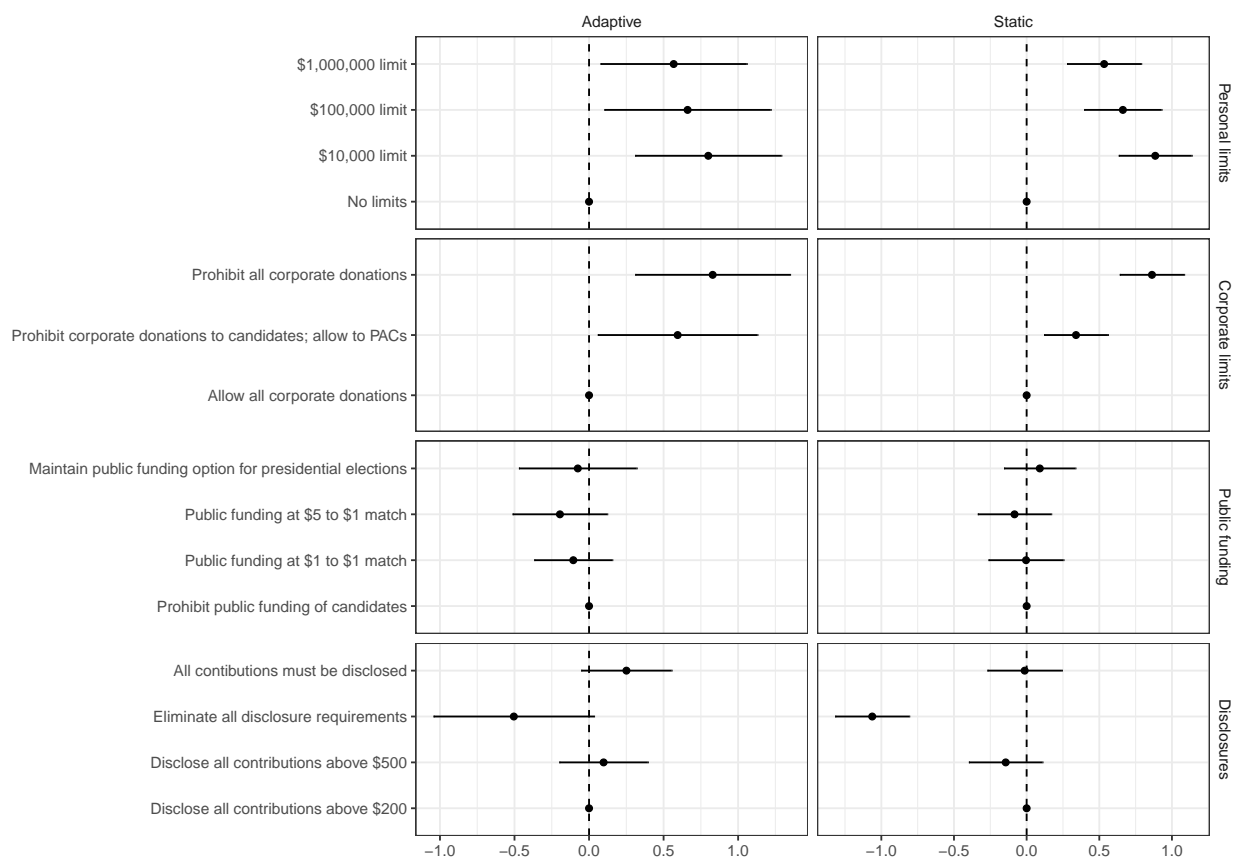
Estimates are coefficients of linear regression estimators, with standard errors clustered at the subject level.

D   Value Remaining in Simulation Studies

We consider an additional metric in our simulation studies: potential value remaining, or per-play regret, a measure of how much success rates might be improved by switching to another arm (Scott 2015).

Figure D.10: Value Remaining over Time