A Personalized BDM Mechanism for Efficient Market Intervention Experiments

IMANOL ARRIETA-IBARRA, Stanford University, USA
JOHAN UGANDER, Stanford University, USA

The BDM mechanism, introduced by Becker, DeGroot, and Marschack in the 1960's, employs a second-price auction against a random bidder to elicit the willingness to pay of a consumer. The BDM mechanism has been recently used as a treatment assignment mechanism in order to estimate the treatment effects of policy interventions while simultaneously measuring the demand for the intervention. In this work, we develop a personalized extension of the classic BDM mechanism, using modern machine learning algorithms to predict an individual's willingness to pay and personalize the "random bidder" based on covariates associated with each individual. We show through a mock experiment on Amazon Mechanical Turk that our personalized BDM mechanism results in a lower cost for the experimenter, provides better balance over covariates that are correlated with both the outcome and willingness to pay, and eliminates biases induced by ad-hoc boundaries in the classic BDM algorithm. We expect our mechanism to be of use for policy evaluation and market intervention experiments, in particular in development economics. Personalization can provide more efficient resource allocation when running experiments while maintaining statistical correctness.

CCS Concepts: • **Computing methodologies** → **Machine learning**; • **Applied computing** → **Economics**; Psychology;

Additional Key Words and Phrases: personalization; BDM; second price auctions; causal inference

1 INTRODUCTION

Measuring the success of a market intervention, where a new product or service is introduced to an existing market, requires evaluating the convolved causal effects of both the product itself and the effects of the price at which the product is introduced, while also estimating the market demand for the product at different prices. The BDM mechanism [9] provides an elegant method for evaluating market interventions, and it has been recently employed in field experiments by development economists to simultaneously evaluate causal treatment effects as well as market demand [10].

Imagine that we want to estimate the average treatment effect of a water filter on a household's health. A simple way to estimate the effect is to run a randomized controlled trial—randomly give away water filters to some households and not to others—and measure the treatment effect by comparing the health of the households. However, this experiment would produce no information about the demand for the product, since we'd be giving the product away. It would also prevent us from studying the interaction effect of the price on the water filter's effectiveness, where previous

The authors would like to thank Anup Malani for discussions that inspired this work. This work supported in part by NSF grant IIS-1657104 and a Facebook Faculty Research Grant.

Authors' addresses: Imanol Arrieta-Ibarra, Stanford University, Management Science and Engineering, Stanford, CA, 94305, USA; Johan Ugander, Stanford University, Management Science and Engineering, Stanford, CA, 94305, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2018 Association for Computing Machinery.
ACM EC'18, June 18–22, 2018, Ithaca, NY, USA. ACM ISBN 978-1-4503-5829-3/18/06…$15.00
https://doi.org/10.1145/3219166.3219220

field experiments have shown that price paid for a product can have a significant effect on its adoption [6]. As another approach, we could propose a (random) price to the treatment households where we randomly offer the product, applying the so-called "take it or leave it" mechanism. But this approach would tell us relatively little about the willingness of the household to pay for the product at other prices, and the households that purchase the product would skew towards wealthier homes, creating a difference between the treatment and control population that would be hard to discern.

The BDM mechanism, meanwhile, gathers the willingness to pay of the households in an incentive compatible way, allowing us to derive demand curves for the product while also making it possible to exactly compute the household's probability of treatment when bidding against the BDM mechanism's random bidder. The mechanism can therefore be viewed as performing a randomized controlled trial with heterogeneous but known probabilities of assignment to treatment or control (known once the willingness to pay has been elicited). In addition to the average treatment effect (ATE), we can use BDM to estimate the conditional average treatment effect (CATE) conditional on willingness to pay which is key to making policy decisions about an intervention.

In this work we observe that the BDM mechanism can be improved upon using personalization to provide less volatile causal effect estimates at a lower expected cost for researchers, all while maintaining the incentive compatibility of traditional BDM. In particular, we can use personalization to reduce (1) the variance of effect estimators, (2) the budget regret of running the experiment, all while (3) maintaining incentive compatibility. The variance is reduced by balancing the treatment probabilities so individuals are more equally likely to land it both treatment and control based on what we know about them a priori. The budget regret is reduced by essentially avoiding situations where we would be giving away the product for much less than what the person is willing to pay for it, again based on what we know about them a priori. In order to perform our personalization we employ standard tools from machine learning, taking care to maintain incentive compatibility.

We evaluate our personalized BDM (PBDM) mechanism under different simulation specifications and run a small field experiment on Amazon Mechanical Turk as a demonstration. The experiment consists of a data-labelling task where we simultaneously evaluate both the *causal effect of* and *demand against* time constraints during the task. How much better or worse are subjects at labelling when there are tight time constraints? How much would they be willing to pay to remove time constraints? The results of this field experiment suggest that a personalized BDM mechanism can provide large efficiency gains during market intervention experiments, where the personalization can also enable statistically efficient estimation of heterogeneous effects that would otherwise be infeasible to estimate reliably.

2 BACKGROUND

The BDM mechanism has been previously utilized as an assignment instrument in [10], which compared it against a "take it or leave it" mechanism for the provision of water filters in Ghana. The BDM mechanism is particularly convenient in such settings because of its double randomization, both of treatment and of prices paid, whereby it is possible to estimate both price effects as well as treatment effects conditioned on willingness to pay. Using BDM as an assignment tool thus improved over [27], which used a two-step intervention in order to achieve this double randomization. Furthermore, eliciting user's willingness to pay allows for the estimation of demand curves for the offered product or service. This capability can be of great use in circumstances where a clean randomized control trial can't be performed, or where there is both the need to estimate the demand for a product as well as its potential benefits.

The effectiveness of the BDM mechanism as an elicitation mechanism has been thoroughly tested in auction theory: for lotteries [19, 26, 40], applied to environmental commodities [12], and in comparison to "take it or leave it" mechanisms and Vickrey auctions and against other individuals

[10, 34]. However, many of these studies report some practical problems with BDM. One problem that we focus on is that, in practice, setting and announcing the bounds for the bidding distribution of the BDM mechanism has been found to introduce bias in individual bids in empirical settings [11], meaning that the mechanism is not "behaviorally incentive compatible." One consequence of our personalized mechanism is that it solves this issue by not requiring the specification of these bounds, which makes the BDM mechanism distribution independent [31]. For our personalized mechanism, the proposal distribution is learned as users make bids.

This work is aligned with a larger effort to design adaptive experiments that optimize resources. Similar endeavors include adaptive designs to elicit time preference from subjects [23] and the adaptation of experiments to subpopulations to increase the precision of estimators [4]. Most broadly this research question sits between machine learning and economics, where other recent work [1, 7, 8, 21, 28, 39] has started to bridge the gap between the two fields. Our proposed methodology exemplifies a clear case where the statistical methods for policy evaluations can be improved by the powerful predictive capacities of modern artificial intelligence and machine learning techniques.

3 ESTIMATING CAUSAL EFFECTS

The idea of estimating causal effects was first formalized into the *potential outcomes* framework by Neyman [33], but it was Fisher [16], some years later, who commented on the importance of randomization in order to be able to make causal claims. The field has since then been expanded to include controlled as well as observational studies through work by Rubin [36], who was the first to formalize how causal statements could still be made even when researchers were not in total control of the assignment mechanism. See also the recent textbook by Imbens and Rubin [24]. We'll begin this section with a brief review of causal inference in the potential outcomes framework, presenting four different estimators commonly used for estimating average treatment effects [5].

The potential outcomes framework assumes that users present different outcomes depending on whether they are assigned to a treatment condition or a control condition. Let T_i represent whether the observation unit i is assigned to treatment ($T_i = 1$) or control ($T_i = 0$), for example if a water filter was given to a household (with possibly impacting health) or if an individual received paid legal counsel (possibly impacting court trial outcomes). Let $Y_i(T_i)$ be the outcome of interest for unit i, e.g. the number of times members of the household became sick during a year or the outcome of a trial. We use $Y_i(1)$ and $Y_i(0)$ to denote the potential outcomes when the unit is assigned to treatment or control, respectively. We make the standard Stable Unit Treatment Value Assumption (SUTVA) that the potential outcomes for one unit is not affected by the treatments of other units, and the potential outcomes are fixed in that there are no different forms or versions of each treatment level. Under these conditions we denote the Average Treatment Effect (ATE) as $\tau = N^{-1} \sum_{i=1}^{N} Y_i(1) - Y_i(0)$, taken over the whole population of N units. We are interested in this average, but for each individual we can only observe one of $Y_i(0)$ or $Y_i(1)$.

One way to estimate the ATE τ is to simply take an average of the given quantity from a sample of observed values. The most commonly used estimator of the ATE is the difference in means (DM) estimator, where we let N_C be the number of units assigned to the control group while N_T are assigned to the treatment group:

$$\hat{\tau}_{DM} = \frac{1}{N_T} \sum_{i=1}^{N} Y_i(1) T_i - \frac{1}{N_C} \sum_{i=0}^{N} Y_i(0)(1 - T_i). \tag{1}$$

In order for this estimator to be unbiased we need to make three more assumptions: that each unit has positive probability of being treated or controlled, that the treatment of each unit is independent

from the treatment of the rest, and that conditional on observables, the assignment to treatment or control is random [24].

If the probability of assignment to treatment depends on observables, the difference in means estimator would be biased, but this bias can be corrected by inverse probability weighting. Let $e(X_i)$ be the *propensity score* or the probability of unit i being treated given a vector of covariates X_i, i.e. $e(X_i) = P(Ti = 1|X_i)$. Generally a great deal of attention is given to conditions under which these propensities can be estimated. However, in this work we are in control of the assignment mechanism and we'll be dealing with known true probabilities of assignment. We still call them propensities by convention.

Given these propensity scores, the Horvitz-Thompson estimator (HT) [22] is an unbiased estimator that employs inverse probability weighting:

$$\hat{\tau}_{HT} = \frac{1}{N}\left[\sum_{i=1}^{N} \frac{1}{e(X_i)} Y_i(1)T_i - \sum_{i=1}^{N} \frac{1}{1-e(X_i)} Y_i(0)(1-T_i)\right]. \tag{2}$$

Although the HT estimator is unbiased, it tends to be very volatile in practice. The reason is that the probability of being treated for some units may be very close to zero or one, making for very high variance. The Hajek estimator [20] is a refinement of the HT estimator where instead of dividing by N, one divides by the sum of the propensity weights (the expected value of which is N). This causes the normalizing denominator to move with the same magnitude as the weights; for this reason the Hajek estimator is sometimes called the self-normalizing estimator [38]. The Hajek estimator is defined as:

$$\hat{\tau}_{\text{Hajek}} = \frac{\sum_{i=1}^{N} \frac{1}{e(X_i)} Y_i(1)T_i}{\sum_{i=1}^{N} \frac{1}{e(X_i)} T_i} - \frac{\sum_{i=1}^{N} \frac{1}{1-e(X_i)} Y_i(0)(1-T_i)}{\sum_{i=1}^{N} \frac{1}{(1-e(X_i))} (1-T_i)}. \tag{3}$$

A final family of estimators we consider is that of *blocking estimators* that employ stratification. These estimators recognize that, for populations with close propensity scores, we have an almost random assignment to treatment or control. We can then estimate the ATE using post-stratification techniques assuming some smoothness in the effects as follows:

(1) Choose M points $q_1, ..., q_M$ from the domain of $e(X_i)$.
(2) Given an $\epsilon > 0$, for each point q_j select all units i whose propensity score is close to this point, $\{X_i : |e(X_i) - q_j| < \epsilon\}$.
(3) For each $j = 1, \ldots, M$, compute the Horvitz-Thompson estimator for those units, $\hat{\tau}^j_{HT}$.
(4) Compute $\hat{\tau}_{\text{block}} = \sum_{j=1}^{M} \frac{N_j}{N} \hat{\tau}^j_{HT}$ as the post-stratified weighted average of the HT estimators, where $N_j = |\{X_i : |e(X_i) - q_j| < \epsilon\}|$ is the number of units near q_j.

For this estimator it is often recommended to do some trimming beforehand to improve its variance; for indications about how to perform this trimming without falling into involuntary p-hacking, see [24].

Beyond average treatment effects, there has been a growing interest in the estimation of so-called heterogeneous treatment effects in contexts where treatment effects are thought to differ widely for different subpopulations [7]. In these contexts, one aims to estimate the conditional average treatment effect (CATE), $\tau(x) = \frac{1}{|\{i:x_i=x\}|} \sum_{i \in \{i:x_i=x\}} Y_i(1) - Y_i(0)$. We estimate the CATE conditional on willingness to pay, making it possible to investigate how the treatment effect varies with revealed wealth.

4 A PERSONALIZED BDM MECHANISM

In this section we propose a way of personalizing the BDM mechanism that we call simply personalized BDM (or PBDM for short). We first show how BDM achieves incentive compatibility, but can produce high variance estimators with high budget regret. By contrast a simple "randomized controlled trial mechanism" (randomization without any willingness to pay component) would minimize the variance of the treatment effect estimators, but at the same time it would only be weakly incentive compatible and it would also achieve the upper bound of budget regret. Personalized BDM can be incentive compatible, produce estimators with lower expected variance, and achieves lower budget regret than both the BDM and RCT mechanisms. As a concrete summary, we will discuss how personalization can deliver a mechanism that is superior to BDM in three specific aspects:

(1) PBDM results in lower variance than BDM for the HT and Hajek estimators.
(2) PBDM has a smaller budget regret than BDM (the ex-post cost from treated units is smaller).
(3) Under PBDM, elicited valuations are distribution independent.

It is important to note that the types of interventions that have been analyzed in this context, and the ones we experiment with, involve giving out subsidies, e.g. water filters to households or legal representation to accused defendants, that would normally cost C on the open market. Then, because of the nature of the problem, we are interested only in units with a willingness to pay $W < C$. However, not all market interventions have an easily knowable upper bound for the population of interest, and indeed that is a difficulty highlighted by the work by Bohm et al. [11]. And even knowing C, a researcher could run an inefficient implementation of the BDM mechanism. Imagine, for example, that a vaccine costs C dollars to produce, but the maximum people are willing to pay for it is $\frac{C}{10}$. Every treated user would be very expensive to subsidize, but we won't know if the medical benefits of the vaccine outweigh the cost of the subsidy without a controlled experiment.

Let Φ be the randomized price offered by BDM's phantom bidder. At the heart of our personalized mechanism is a choice of conditional price distribution $\Phi|X_i$, where X_i are observable characteristics of i that can not be manipulated. Given a choice of $\Phi|X_i$, the mechanism is carried out in the following manner:

(1) Offer a product with cost C to a subject with X_i observable characteristics,
(2) Draw a price ϕ from $\Phi|X_i$ without showing it to the subject,
(3) Ask the subject to report her willingness to pay w_i, which we assume is drawn from $W|X_i$,
(4) If $\phi < w$ the user gets the product and pays ϕ. Otherwise there is no exchange.

How should we design $\Phi|X_i$? It is clear that we want to make our choice using estimated properties of $W|X_i$, the willingness to pay of previous subjects conditional on their covariates. The problem is made easier by assuming that $W|X_i$ has support in $[0, C]$ for all X_i for a fixed outside cost C, as discussed above. We will now describe our choice of $\Phi|X_i$ before elaborating on why this choice embodies favorable tradeoffs between the three objectives at the start of this section.

The conditional price distribution $\Phi|X_i$ we propose consists of two point masses of probability $\epsilon/2$ at the lower and upper bounds of the price range (0 and C) combined with a uniform distribution over a range $[a, b]$, where a and b are quantiles of $W|X_i$ given by $b = F^{-1}_{W|X_i}(1-\frac{\delta}{2})$ and $a = F^{-1}_{W|X_i}(\frac{\delta}{2})$. The cumulative distribution function $F_{\Phi|X_i}(w; a, b, \epsilon)$ is then:

$$F_{\Phi|X_i}(w) = \frac{\epsilon}{2} + (1-\epsilon)\frac{w\mathbf{I}\{a \le x \le b\}}{b-a} + \frac{\epsilon}{2}\mathbf{I}\{w \ge C\}. \tag{4}$$

In practice a and b must be estimated from observed willingness to pay data, furnishing estimates $\hat{a}$ and $\hat{b}$. We leave the estimation of ϵ for future work, and choose to set it before the experiment begins based on simulations. For different specifications of the data generating process, we find

that $\epsilon = 0.05$ is a conservative value that balances cost and variance while maintaining incentive compatibility[1].

We propose the above family of distributions for two reasons. First, we want a mechanism that in the absence of any signal from the covariates reverts to BDM, which the mechanism clearly does (when $\epsilon = 0$). Second, if $\epsilon = 0$ there may be cases where the probability of assignment will either be zero or one (if someone bids outside the range of $[a, b]$), so the point masses of $\epsilon/2$ serve to bound the variance of our estimates (which is unbounded if any of the treatment probabilities reach 0 or 1).

Given that we estimate $\hat{F}_{W|X_i}$ for user i from the dataset $\{(x_1, w_1), ..., (x_{i-1}, w_{i-1})\}$, the probability of being treated depends heavily on the history of users that have come before user i. This means that, contrary to BDM (where conditional on W the assignment to treatment is random), for PBDM one needs to condition on all the previous history. However all the previous history is summarized by the choice of $F_{\Phi|X_i}$ so that

$$P(T_i = 1|\{(x_1, w_1), ..., (x_{i-1}, w_{i-1})\}) = F_{\Phi|X_i}(w_i)$$

for the realized w_i. In order to compute the HT or Hajek estimators of the average treatment effect, it suffices to save $F_{\Phi|X_i}(w_i)$ at the point when it was computed.

There are three immediate benefits from this setup. First as [31] conclude, not making the exact distribution explicit to users prevents distribution-dependence bids. Second, in the case where there is little predictive power of X_i on W, the mechanism reverts to a well-centered BDM. Third, the point masses at the boundaries of this distribution help bound the variance of the ATE estimators. The following sub-sections show how the objectives specified in the introduction (low variance, low budget regret, and incentive compatibility) help inform the choice of this distribution for our mechanism. We detail how PBDM is superior to BDM in terms of lower variance, lower budget regret, and superior to a pure RCT in terms of stronger incentive-compatibility and lower budget regret. An RCT that gives away water filters can be thought of as randomizing people to either get a price of 0 (treatment) or C (control) since they could possibly still buy it on the open market.

4.1 HT estimator variance

One can use the BDM mechanism to compute the Average Treatment Effect (ATE) of a particular intervention [10] as explained in the previous section, utilizing the probabilities of treatment and control as "propensities" (conditional on the willingness to pay they are true probabilities, not estimates). Post-stratification can be used for populations that are unbalanced in terms of their willingness to pay. However, since the probability of treatment depends fully on W, these estimates may be highly volatile. Equation (5) makes evident how the variance of the HT estimator can grow to infinity if the probability of treatment is close to 0 or 1. The variance of the Horviz-Thompson estimator is [5]:

$$Var(\hat{\tau}_{HT}) = \frac{1}{N^2}\left(\sum_{i=1}^{N} \frac{1 - e(X_i)}{e(X_i)} Y_i(1)^2 + \frac{e(X_i)}{1 - e(X_i)} Y_i(0)^2\right). \tag{5}$$

In the event of unavoidable extreme propensity scores close to 0 or 1, one could restrict the population of interest (change the estimand) to a subset of the propensity scores through trimming [24], though this can quickly turn into an invitation for "p-hacking" if not preconceived as part of the experimental design. Therefore we would like to have better balance of treated and control individuals for every willingness to pay w. This achieves two goals. First, both the ATE and CATE estimates would have lower variance. Second, by reducing the "profit" of the treated consumers (the

[1]This is consistent with notions of trimming specified in [24].

difference between their willingness to pay and the price they paid), researchers would have less regret about their subsidy of the product as will be explained in more detail in the next subsection.

The Hajek estimator's variance can be related to the HT variance. The standard approach here is to linearize the Hajek estimator around the HT estimator and then take the variance and truncate higher order terms. We thus concentrate on the HT estimator for comparing variance across mechanisms; the results presented in this section apply for the Hajek estimator asymptotically[2].

How should we design the conditional price distribution, $\Phi|X$, to balance the treatment and control? The variance in equation (5) can be modified by replacing the propensities with $F_{\Phi|X}$, the (true) probability of being treated by our personalized mechanism:

$$Var(\hat{\tau}_{HT}) = \frac{1}{N^2}\left(\sum_{i=1}^{N} \frac{1 - F_{\Phi|X_i}(W_i)}{F_{\Phi|X_i}(W_i)} Y_i(1)^2 + \frac{F_{\Phi|X_i}(W_i)}{1 - F_{\Phi|X_i}(W_i)} Y_i(0)^2\right). \tag{6}$$

Under a (rather strong) null hypothesis of no individual treatment effect (Fisher's null, see [15]) where $Y_i(1) - Y_i(0) = 0\ \forall i$, meaning $Y_i(1) = Y_i(0) = \alpha_i$, we can write the variance as:

$$Var(\hat{\tau}_{HT}) = \frac{1}{N^2}\sum_{i=1}^{N} \alpha_i \left(\frac{1 - F_{\Phi|X_i}(W_i)}{F_{\Phi|X_i}(W_i)} + \frac{F_{\Phi|X_i}(W_i)}{1 - F_{\Phi|X_i}(W_i)}\right). \tag{7}$$

This expression is minimized when $F_{\Phi|X_i}(W_i) = \frac{1}{2}$ for all W_i , for any realization of $w_1, ..., w_N$. If we are only interested in minimizing $Var(\hat{\tau}_{HT})$ under the null, there is a simple solution: we can just place half of the probability mass of Φ on 0 and the other half on the upper bound for W_i. This provides an intuition for why randomized control trials are the gold standard for estimating causal effects. Yet, our objective is not only the estimation of an ATE: we are also interested in estimating the demand for the product and reducing the regret over our spent budget. Deploying an RCT actually achieves the upper bound for the budget regret (for any subject treated, they would pay 0 and we would lose the full value of the product, C). With respect to estimating the demand, RCTs provide no incentive to users for revealing their true willingness to pay, making the mechanism only weakly incentive compatible.

Guaranteeing minimum variance for any realization of W is a very strong condition. We instead focus on the expected variance, conditional on the observed characteristics X_i, with the expectation being taken over realizations of $W_1, ..., W_N$. Using a Taylor expansion of Equation (6) around the mean and taking a first degree approximation, we then are interested in minimizing:

$$E[Var(\hat{\tau}_{HT})|X_1, ...X_N] \approx \frac{1}{N^2}\sum_{i=1}^{N}\left(\frac{1 - F_{\Phi|X_i}(E[W_i|X_i])}{F_{\Phi|X_i}(E[W_i|X_i])} Y_i(1)^2 + \frac{F_{\Phi|X_i}(E[W_i|X_i])}{1 - F_{\Phi|X_i}(E[W_i|X_i])} Y_i(0)^2\right). \tag{8}$$

Again assuming the Fisher null of no individual effect, this expected variance is minimized whenever $F_{\Phi|X_i}(E[W_i|X_i]) = \frac{1}{2}$, meaning that the median from the price distribution thus needs to be equal to the conditional expectation of the willingness to pay distribution.

One of the biggest advantages of PBDM over BDM is that it "smooths" the probability of a unit to be treated or controlled, depending on the distribution of $W|X_i$. This is, if the distribution of $W|X_i$ is skewed to the right, where users would be rarely treated, PBDM would "focus" in this part of the distribution and would increase the probability of treatment. On the other hand if $W|X_i$ is skewed to the left then PBDM could diminish the probability of "high spenders" to be treated. Our objective however is clearly underspecified in that there is an infinite number of distributions that achieve this goal. We can therefore target additional goals within this space of distributions.

[2]See [37], pages 172–176.

In particular, we would like for the variance of our estimator to be bounded. In order to achieve this, it is enough to require for

$$P(F_{\Phi|X_i}(W_i) > 1 - \frac{\epsilon}{2}) = 0, \tag{9}$$

$$P(F_{\Phi|X_i}(W_i) < \frac{\epsilon}{2}) = 0. \tag{10}$$

These conditions can be achieved by placing a probability mass of $\frac{\epsilon}{2}$ at both 0 and C as PBDM does.

If we assume that $W|X_i$ is symmetric, this price distribution achieves the minimum expected conditional variance while at the same time bounding the probability of it escalating to infinity. It is important to note that BDM also achieves these two principles, but only in the case where the lower and upper bounds are chosen appropriately and W is independent from X_i. For any other case BDM would provide a higher variance in expectation than PBDM.

4.2 Budget Regret

If we knew in advance a user's willingness to pay, the optimal pricing mechanism would be to randomize the price uniformly between $w_i - \epsilon$ and $w_i + \epsilon$ for some small value of ϵ. This would mean that every user effectively pays their willingness to pay for the product or service and would result in the estimator with minimal variance. However, due to the uncertainty in W we incur some level of regret every time we assign someone to treatment.

We define *budget regret* as:

$$BR(X, W) = E_{\Phi}\left[(W - \Phi)|\Phi < W, X, W\right], \tag{11}$$

and *expected budget regret* as:

$$br(F_{\Phi}) = E_{X,W}[BR(X, W)]. \tag{12}$$

Expected budget regret is the expected subsidy paid when assigning a unit to treatment. For any realization of W, an RCT price mechanism achieves the upper bound for $br(\Phi)$ among all possible price distributions: when $\mathbf{I}\{W_i > \Phi_i\} = 1$ it must be that $\Phi_i = 0$.

Meanwhile, for BDM we have that the budget regret is equal to:

$$br_{BDM}(F_{\Phi}) = E\left[\frac{W}{2}\right]. \tag{13}$$

For PBDM, if $\hat{b} = F^{-1}_{W|X_i}(1 - \frac{\delta}{2})$ and $\hat{a} = F^{-1}_{W|X_i}(\frac{\delta}{2})$ and $\epsilon = 0$ (assume an asymptotic setting where $\hat{a}$, $\hat{b}$ have converged on a, b):

$$br_{PBDM}(F_{\Phi}) = E\left[\frac{W - \hat{a}}{2}\right]. \tag{14}$$

These budget regret expressions mean that less variance in $W|X$ translates to less regret on average.

4.3 Balancing Cost and Variance

The choice of ϵ and δ for the PBDM mechanism is equivalent to choosing a trade-off between experiment cost and estimator variance, where larger values of ϵ or smaller values of δ correspond to lower variance and higher costs. On the other hand, a good reason to want to decrease the cost of treating an individual is to be able to treat more people and thereby have lower variance for the experiment at a given budget (or for the same variance, have lower costs). We propose a way to summarize this trade-off under a single minimization problem. We'll assume that sending people to control is costless and that there is no information in X about W. Let the expected number of treated units under no budget constraints be N_u and the expected number of treated units under

budget constraint be N_b, taking the expectation over both willingness to pay and the random draws. It is evident that $N_u \geq N_b$. Let N be the maximum number of units that can be treated.

We can write N_u as:

$$N_u = \sum_{i=1}^{N} \mathbf{I}\{W_i > \Phi_i\} \approx Np(\epsilon, \delta), \tag{15}$$

where we define $p(\epsilon, \delta) = P(W > \Phi)$. Meanwhile in the setting with a budget B we want to approximately spend the full budget:

$$B \approx \sum_{i=1}^{N} (C_i - P_i)\mathbf{I}\{W_i > P_i\} \approx N_b c(\epsilon, \delta) \tag{16}$$

where $c(\epsilon, \delta) = E\left[(C - P)|W > P\right]$.

We then have that

$$N_u \geq N_b \Rightarrow Np(\epsilon, \delta) \gtrapprox \frac{B}{c(\epsilon, \delta)} \Rightarrow N \gtrapprox \frac{B}{p(\epsilon, \delta)c(\epsilon, \delta)}. \tag{17}$$

From Equation (7) we get that the variance for this simplified case is

$$Var(\hat{\tau}_{HT}|W) \propto \frac{1}{N^2}\left(\sum_{i=1}^{N} \frac{p(\epsilon, \delta, W_i)}{1 - p(\epsilon, \delta, W_i)} + \frac{1 - p(\epsilon, \delta, W_i)}{p(\epsilon, \delta, W_i)}\right) \tag{18}$$

$$\lessapprox \left(\frac{p(\epsilon, \delta)c(\epsilon, \delta)}{B}\right)^2 \left(\sum_{i=1}^{N} \frac{p(\epsilon, \delta, W)}{1 - p(\epsilon, \delta, W)} + \frac{1 - p(\epsilon, \delta, W)}{p(\epsilon, \delta, W)}\right). \tag{19}$$

For a given budget B we would like to minimize $E[Var(\hat{\tau}_{HT}|W)]$. For the purposes of this work, we chose to fix ϵ to preserve incentive compatibility (otherwise, the variance would be minimized with $\epsilon = 1$, corresponding to an RCT). We ran simulations to select δ by minimizing the expectation of Equation (19) over our prior belief of the distribution of W_i. We chose δ after some robustness checks prior to running the experiment. In practice, δ could be personalized by being calculated at every point in time after having better estimates of the distribution of W. We believe there is ample room to fine tune these parameters but leave this as future work.

4.4 Incentive Compatibility

Finally, for BDM and PBDM if a user is an expected-utility maximizer she gains nothing by providing a wrong valuation of the product offered. Let's say that the user has a valuation of v_i for the product. If she were to report $v_i - \epsilon$ as her valuation for all those prices between $v_i - \epsilon$ and v_i she would lose the product, which is still more valuable than the price she would have to pay. If she were to offer $v_i + \epsilon$ she would have to pay more than her valuation with ϵ probability.

For an RCT mechanism the user is indifferent between reporting her true valuation or not. This is because she would always be offered 0 or C as the price. This could happen in PBDM if the user's willingness to pay is between 0 and $F^{-1}_{W|X_i}(\frac{\epsilon}{2})$, but this would only happen with small probability. Even if users know PBDM is being used, they don't know their order of arrival. As a result they would be uncertain about the distribution from which prices are being sampled. And as Mazar et al. [31] show, as long as users don't know the underlying distribution, their offers will be distribution-independent.

4.5 Why not use the predicted median as the price?

We have thus far described a personalized mechanism that outperforms BDM in all of the described dimensions. One may ask however why not just predict the median $W_i|X_i$ and use this as the offered price. This was actually the original mechanism we envisioned for this problem, running a simple median regression on $W_i|X_i$. If the covariates are highly predictive of W_i we would achieve a 50-50 split and achieve close to minimum budget regret. However, we would not be able to assume that W_i is fixed (but previously unknown) for every user, since doing so would make the probability of being assigned to treatment *conditional on* W_i either 0 or 1, losing randomness in the conditional assignment and breaking the assumption that treatment needs to be assigned at random in order to make causal interpretations. Even if we assumed that the randomness comes from some probabilistic process in how people elicit their willingness to pay, one could have an unobserved variable correlated with both the willingness to pay and the outcome which would complicate causal interpretations.

5 EMAIL CLASSIFICATION MOCK EXPERIMENT

In order to test our method we designed a Mechanical Turk task where users performed email spam classification [17]. After performing several rounds of classification, the users were offered the opportunity to spend part of their wages in order to forgo a time restriction on their work. The time restriction made the task significantly more difficult, and they experienced the task both with and without the time restriction before being offered the opportunity to remove it. The offer to pay to remove the restriction was formulated as a willingness to pay experiment. Users were randomly assigned to either the standard BDM mechanism or our PBDM mechanism. The objective in this mock experiment was thus to estimate both the demand for time freedom (the willingness to pay for it) and the causal effect of being time-restricted on performance (better or worse spam classification).

We conducted our experiment on Amazon Mechanical Turk for its convenience and simplicity for running experiments. Mechanical Turk compares favorably to traditional ways of conducting behavioral experiments [13]; see [30, 35] for discussions of best practices for behavioral research using Mechanical Turk.

We collected a range of covariates as the basis for our personalization under the PBDM mechanism. First, our task was hosted on a server that allowed us to gather browser-level features of the web users. Second, we included two surveys at the onset of the task: a demographic survey in order to compare our population with that of Mechanical Turk [25] and a risk survey, adapted from [18]. For computing the risk profile of users from this survey we used the index for general risk aversion as in [29]. Lastly, we also collected additional covariates from the behavior of users during the early tutorial phase of our task.

The users were not made aware that their browser configuration or survey responses impacted the phantom bidder of the PBDM mechanism. We had no true way of verifying the validity of the demographic survey responses or the risk survey responses, and it is important to be clear that these responses ultimately impacted the price at which these users were offered the time freedom. As a result, in settings where the subject may realize the connection between their responses to early questions and the price offered in a later stage of an experiment, it may be appropriate to only use survey questions for which answers can be validated. Examples of questions from development economics field experiments used in similar contexts include "How many members are there in your household?" or "How many rooms do you have in your home?" [2]. Note, however, that even if the users are able to impact their offer distribution in their favor, it is still in their interest to bid their true willingness to pay, regardless of how much they've been able to manipulate the

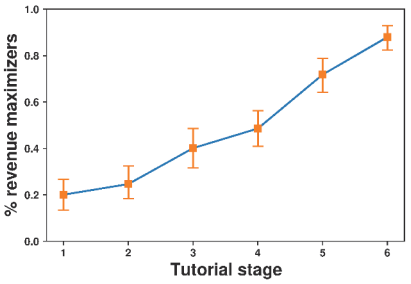

Fig. 1. As the tutorial progresses, more participants bid the revenue-maximizing number of credits.

mechanisms offer distribution (or not). It may, however, impact the price they pay. Thus if the main effect of interest is the effect of the price, it is critical to only use covariates that participants cannot strategically manipulate.

Following the initial demographic surveys, users were subject to three phases of labelling tasks designed to demonstrate and evaluate our task and mechanism. First, users were shown a selection of six emails to classify into spam or non-spam, with no time restriction. The proportion of spam/non-spam was always kept balanced (three spam, three non-spam), and the emails were showed in random order. Users were instructed that for each email classified correctly, they would earn 10 credits, otherwise they would lose 10 credits. Credits were converted to dollars at a rate of 0.5 cents per credit. These first six emails were then followed by another six emails, again balanced (three spam, three non-spam), but for these emails there was a time restriction of t seconds per email (we varied t between tasks), meaning that if the user did not answer within t seconds their answer was marked as wrong. If the user accumulated a net loss during these phases it did not carry on to later phases.

In the second phase, users were introduced to the idea of the BDM mechanism. For 6 rounds we offered users a chance to earn k credits without any work, where k was a random integer from 0 to 20. Users were instructed to make a bid stating the most they were willing to pay for the k credits and they were told that they would be in direct competition against an artificial intelligence. If the AI's offer was smaller than their bid, we gave them k credits and charged them the AI's price. The optimal bid under the assumption of expected utility maximization should have been k if the user fully understood the mechanism. We thus used this phase in order to evaluate users' understanding of the mechanism.

5.1 Understanding the mechanism

Figure 1 shows to what extent users were understanding the mechanism and behaving as rational agents. For the tutorial stages we measure the difference between the credits offered and a worker's willingness to pay. If users understand the mechanism, this number should be either 0 or 1, given that we are working with a discrete space. If offered k credits the user should be indifferent between bidding k or $k-1$; in both cases if the random draw is k they get nothing and any other draw gives them back the same amount of credits. In order to evaluate the last phase, we asked participants if they regretted their decision once our random offer was shown to them. Regret here does not necessarily mean that participants didn't understand the mechanism. It is not clear to what extent people are conscious of their willingness to pay, so when a price is shown it may work as an anchor on valuation. In total 88% of participants seemed to understand the mechanism (by bidding rationally) by the end of the tutorial. Meanwhile in a post-experiment survey 33% said they regretted their bid in the final (non-tutorial) round.

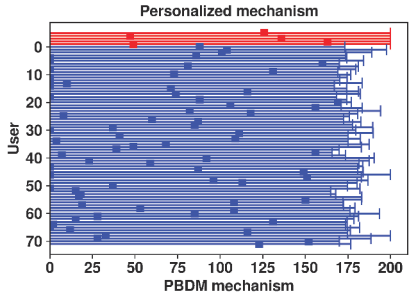

Fig. 2. For each user treated with the PBDM mechanism, intervals corresponding to the range of $\Phi|X_i$ with points correspond to the offer made by the mechanism. The first 5 users encounter the non-personalized BDM mechanism (and are counted towards that treatment arm). User 0 and onward encountered the PBDM mechanism, trained on the previous users. Probability mass below 0 and above 200 is placed at the respective boundaries (no user could have more than 200 credits). It is hard to visualize, but many of these intervals have significant probability mass at zero, which is their main difference compared to the BDM bidder distribution. Since ridge regression assumes normality and our intervals are very wide, we get that 10% of the normal intervals have lower bounds below zero.

The final phase begins with a final encounter with the willingness to pay mechanism where users were instructed to make a bid to remove the timer from a long round of spam labeling. If the AI's offer was below their bid, they will pay the AI's offer and not be subject to a timer for the final tasks. Otherwise they would be time-restricted for these tasks (though not have to pay anything). They were not told if they were in the BDM condition or the PBDM condition.

As mentioned before, we used the information from a brief demographic survey and risk profile survey executed before the first phase, the browser information, and behavioral measurements from the tutorials to inform the personalized mechanism for the final bidding. Since the mechanism is incentive compatible no matter the distribution from which we sample [9], we can easily evaluate how other offer distributions would have fared. Thus, we only needed to run one mock experiment and can then test different ways of predicting willingness to pay.

Our experiment assigned 142 users to either BDM or PBDM at random. For the personalization, we used Bayesian ridge regression (i.e. L_2-regularized linear regression) as our model for predicting willingness to pay, discussed in further detail below. In tests on synthetic data we saw little difference between this model and other models such as Random Forests, Lasso, or Gradient Boosting Decision Trees. PBDM can also be implemented without any user covariates with the goal of just estimating the mean and variance of the distribution for W. Even in this case a researcher would find improvements since the intervals over which samples are taken would be positioned optimally (but "personalized" only to the population, not individually personalized).

5.2 PBDM for this mock experiment

In order to specify the PBDM mechanism for our mock experiment we only need to determine how to compute $F_{\Phi|X_i}(w)$ as in Equation 4, which means estimating $a = F^{-1}_{W|X_i}(\frac{\epsilon}{2})$ and $b = F^{-1}_{W|X_i}(1-\frac{\epsilon}{2})$. Because of the small size of the participant pool for our mock experiment, we decided to restrict ourselves to linear predictors. In particular we chose a simple ridge regression with a prior on the size of the parameters. The Bayesian interpretation of ridge regression is key, allowing us to interpret a posterior on $W|X_i$ from the model. We used STAN [14] for all computations.

Under the reasonable assumption that willingness to pay responses were not influenced by the choice of the prediction algorithm, we can use their covariates and responses to test other

	BDM	PBDM
Number of participants	71	71
Percentage treated	31%	52%
Total cost (credits)	5469	6146
Cost per person treated (credits)	237	161
Average budget regret	65	45
HT ATE	-0.63 (-38,36)	12.09 (2.09,22.09)
Hajek ATE	2.23 (-4.55,9.03)	4.26 (2.78,5.73)
Block ATE	5.19 (3.56,6.81)	3.8 (1.91,5.70)

Table 1. PBDM improves on BDM in terms of cost per person, budget regret, and variance of the average treatment effect estimators.

prediction algorithms such as Gradient Boosting Decision Trees, Adaboost, and Lasso post-fact[3]. We observe that none of these algorithms did a good job of predicting the personalized willingness to pay from the covariates we gathered for our mock application, with all the models reverting to the population mean to predict W. This highlights an important basic benefit of PBDM: for all BDM experiments there is a basic population-level personalization problem of what the population of responses will look like. A PBDM mechanism that simply learns the population of bids with reasonable fidelity therefore delivers significant improvements over BDM, as our results show. Figure 2 shows the estimated intervals as well as the PBDM phantom bid for every user.

6 RESULTS

We use the email classification task in the previous section to evaluate our PBDM mechanism in contrast with an ordinary BDM mechanism. Our evaluation is five fold: we examine the cost of running our experiment, the cost per person treated, the variance of the average treatment effect (ATE) estimators, the CATE conditional on W, and the expected budget regret. Table 1 summarizes these differences. For costs, the credits we pay to user i is 10 times the difference between the number of correct answers, Y_i, and the number of incorrect answers (out of 20), $20 - Y_i$, minus the amount paid in case of treatment: $10(Y_i - (20 - Y_i)) - \Phi_i \mathbb{I}\{W_i > \Phi_i\}$.

For the variance of the average treatment effect estimators, we focus on the variance of the Horvitz-Thompson estimator and the consistent estimator of this variance [4]:

$$\widehat{Var}(\hat{\tau}_{HT}) = \frac{1}{N^2}\left(\sum_{i=1}^{N} T_i(1-e(X_i))\left[\frac{Y_i(1)}{e(X_i)}\right]^2 + (1-T_i)(e(X_i))\left[\frac{Y_i(0)}{1-e(X_i)}\right]^2\right). \tag{20}$$

Here $e(X_i)$ is the probability of assignment to treatment, which we can compute exactly given the known distribution of $\Phi_i|X_i$ at every stage of the experiment. As detailed in Section 4.1, we can obtain Hajek variances estimates from HT variance estimates by linearization and Taylor expansion to obtain $\widehat{Var}(\hat{\tau}_{\text{Hajek}})$. We use this estimate $\widehat{Var}(\hat{\tau}_{\text{Hajek}})$ to construct conservative Wald-type confidence intervals for the ATE, $\hat{\tau}_{\text{Hajek}} \pm z_{1-\alpha}\sqrt{\widehat{Var}(\hat{\tau}_{\text{Hajek}})}$. Furthermore, we build bootstrap confidence intervals for $\widehat{Var}(\hat{\tau}_{\text{Hajek}})$ under Fisher's null (no individual treatment effect [15]). We

[3]In practice, researchers can run a variety of prediction algorithms at runtime (rather than post-fact) and make predictions with the one with the best performance under a pre-specified metric.

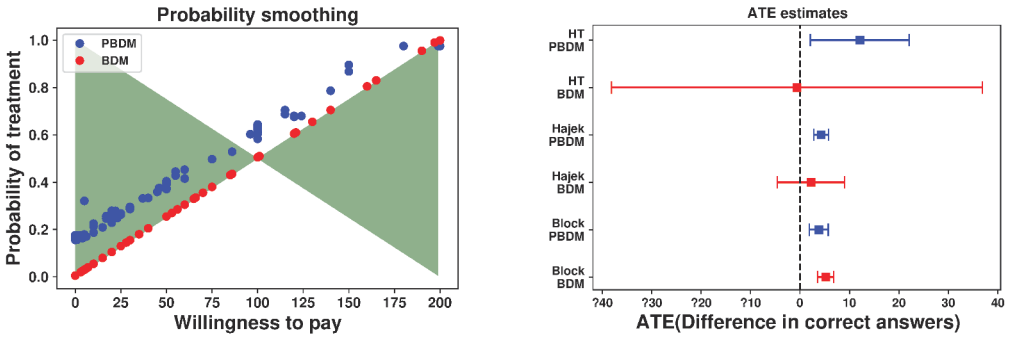

Fig. 3. (a) The probability of treatment as a function of willingness to pay for the BDM and PBDM mechanisms in our email labelling experiment. The green areas correspond to reduced contributions to the estimator variance. PBDM smoothes probabilities for most users. (b) The ATE estimates computed for both BDM and PBDM experiments for the HT, Hajek, and block estimators, with bootstrapped [5%, 95%] confidence intervals shown. For each of the three estimators, PBDM provides tighter confidence intervals than BDM.

require this null assumption because otherwise the probability of treatment is dependent on the order of arrival, and we only observe one of the two potential outcomes.

For the CATE we proceeded as in [10], computing a non-parametric kernel regression with the willingness to pay as the dependent variable of interest. We compute confidence intervals based on bootstrapping the residuals as in [32].

Finally, we estimate expected budget regret as:

$$\widehat{br} = \frac{\sum_{i=1}^{N} (W_i - \Phi_i)\mathbb{I}\{\Phi_i < W_i\}}{\sum_{i=1}^{N} \mathbb{I}\{\Phi_i < W_i\}}. \tag{21}$$

Table 1 shows a summary comparing PBDM to BDM under these varied metrics. The percentage of people treated is more balanced in PBDM compared to BDM, as can be seen in more detail in Figure 3a, helping reduce the variance of the estimators under Fisher's null.

In terms of costs, from a first glance at Table 1, it may seem that BDM is cheaper than PBDM. However, the cost per person treated is much lower for PBDM. In Figure 3b we how, for the same number of units in the experiment, PBDM achieves tighter confidence intervals than BDM.

As Figure 4 shows, at every budget level, the variance of the Hajek estimator is lower for PBDM than for BDM (we achieve the same results for the other two estimators). Finally Figure 5 shows that there is no difference between the two methods in terms of computing a demand curve (as there shouldn't be).

For our mock experiment of email labeling, these improvements were an example of a worst case scenario for the mechanism, in the sense that there was little individual personalization to learn. We were not able to predict the willingness to pay based on any of the covariates, and using the population mean and variance was the resulting strategy. As such, the main improvement of PBDM against BDM in this case was the correct centering of the distribution allowing us to increase the number of individuals treated at every level of willingness to pay. As Figure 6a shows, none of the covariates had a significant effect on W.

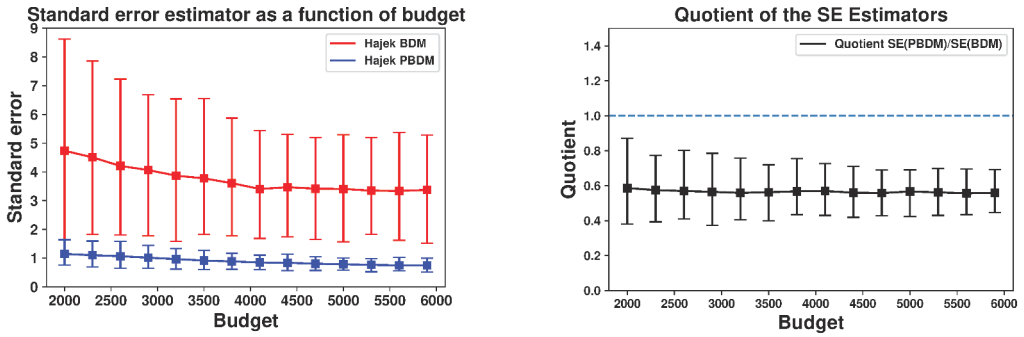

Fig. 4. (Left) Bootstrapped confidence intervals for the standard error of the Hajek ATE estimators under BDM and PBDM. The plot shows how, at each budget level, the standard errors computed for PBDM are lower than those for BDM. (Right) Bootstrapped confidence intervals for the quotient between the standard error of the Hajek ATE estimator for PBDM and BDM. The standard error under PBDM is estimated to be around 25% of the estimator under BDM.

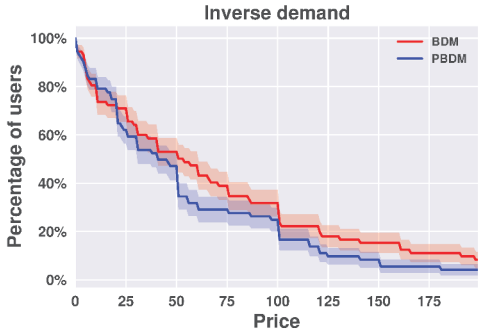

Fig. 5. The demand for not having a time limit in the email classification mock experiment, computed using the elicited willingness to pay of users under the two mechanisms, priced in units of credit.

Figure Figure 6b shows the CATE, the ATE conditional on willingness to pay. The wide confidence bands here show that the ATE did not vary significantly among people with different willingness to pay. Based on these estimates we believe that removing time restrictions on classification tasks such as ours would not have a heterogeneous effect across users of different willingness to pay levels. Our results here are all in terms of a mock experiment, but the real importance of this analysis would come by e.g. analyzing the CATE of water filters on health conditional on willingness to pay.

7 DISCUSSION AND EXTENSIONS

Overall our experiment shows how personalized BDM (PBDM) can improve on classic BDM in terms of budget regret, average cost, and estimator variance all at once. PBDM makes it possible to forego any assumptions about the bounds of the bid interval (except for the natural bounds that constrain the population of interest), which eliminates potential biases observed when telling users that interval. Under PBDM, the interval is quickly learned from data. While we found little to no signal in the covariates about a user's willingness to pay in our specific experiment, the

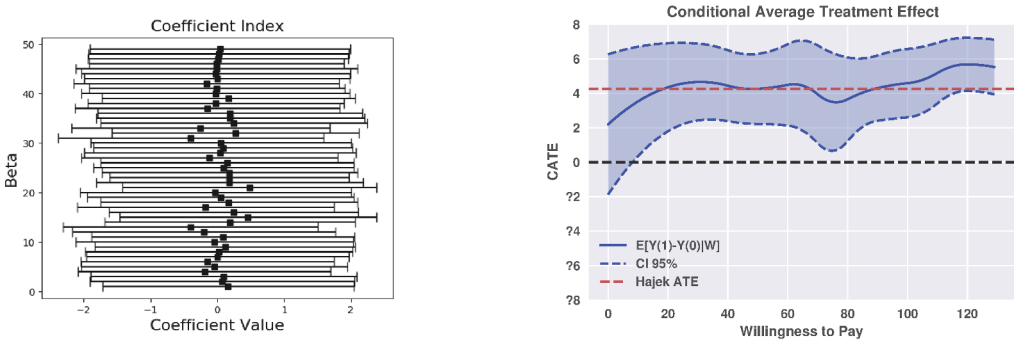

Fig. 6. (a) The coefficients corresponding to the predictors of W, with confidence intervals. We found no effect on any of these predictors in our mock experiment (and post-hoc found no effects with several other prediction models as well). (b) The Hajek estimator finds a significant average treatment effect (see Table 1), but no significant heterogeneous treatment effects as a function of W. The non-uniform confidence bands reveal that most bids were on the lower end of the interval.

personalization still served to effectively find the population distribution. We did not observe any evidence of conditional treatment effects, though it's not clear that we should have expected any: in our mock experiment a conditional treatment effect would imply that individuals with a greater willingness to pay were more impacted by the timing restriction than individuals with a smaller willingness to pay. It's not clear that we would expect people to be good judges of the value, to them, of the lifted time restriction.

In a more general note there are two considerations that we did not utilize in our current work but that we believe may be key to implementing these mechanisms in contexts where we can't assume unconfoundedness. There are different reasons why unconfoundedness may be violated. First, it may be that there are omitted variables correlated with outcomes and/or with the treatment. For example, wealthy people may have a higher willingness to pay for some health service but receive smaller benefits from it. It may also be that some users don't comply in certain settings once they have been assigned to treatment. Compliance is a fairly common problem when dealing with this type of mechanisms since it is often difficult to force people to pay after they win a bid. To address these confoundedness issues, one can use the instrumental variable framework as in [3]. The original BDM mechanism offers an external source of randomness that can be used as an instrument in these settings, as noted by [10]. For our mock experiment we didn't concern ourselves with this problem since we had perfect compliance and no dropouts. However in practice, and in particular in development settings, one could see how this would be an important issue to address.

8 CONCLUSIONS

In market intervention experiments the classic BDM mechanism has recently been shown to facilitate the simultaneous estimation of treatment effects (of an intervention) and demand (for the intervention). We show that both the statistical efficiency and cost of running such an experiment using BDM can be significantly improved by personalizing the mechanism. As a basic matter, we find that our personalized BDM (PBDM) mechanism can learn the overall distribution of the population of willingnesses to pay, W, which reduces both the cost of the experiment and variance

of the estimate. As a more advanced matter, we designed our PBDM mechanism to be able to tailor the mechanism to the known covariates of each individual, learning $W|X_i$. That said, in our practical demonstration using a spam labelling task, we found that no covariates that we collected were able to meaningfully predict individual willingnesses to pay.

In terms of personalization, we believe there are further improvements to be made. Recent work in the estimation of confidence intervals for machine learning techniques (see [7] would allow more sophisticated algorithms to be used as predictors of W's distribution. Furthermore, a more sophisticated characterization of the phantom bidder distribution could yield better results in contexts where the distributional properties of the demand may be known a priori, as in well established markets.

Even with the simple representation of the phantom bidder used in this work we have high confidence that a personalized BDM mechanism can result in substantial savings in experimentation costs as well as improvements in statistical efficiency for a broad class of willingness to pay experiments.

REFERENCES

[1] Alekh Agarwal, Sarah Bird, Markus Cozowicz, Luong Hoang, John Langford, Stephen Lee, Jiaji Li, Dan Melamed, Gal Oshri, Oswaldo Ribas, Siddhartha Sen, and Alex Slivkins. 2016. Making Contextual Decisions with Low Technical Debt. *arXiv preprint arXiv:1606.03966* (2016).

[2] Vivi Alatas, Ririn Purnamasari, Matthew Wai-Poi, Abhijit Banerjee, Benjamin A Olken, and Rema Hanna. 2016. Self-targeting: Evidence from a field experiment in Indonesia. *Journal of Political Economy* 124, 2 (2016), 371–427.

[3] Joshua D Angrist and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

[4] Peter M Aronow. 2016. Data-Adaptive Causal Effects and Superefficiency. *Journal of Causal Inference* 4, 2 (2016).

[5] Peter M Aronow and Cyrus Samii. 2017. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics* 11, 4 (2017), 1912–1947.

[6] Nava Ashraf, James Berry, and Jesse M Shapiro. 2010. Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review* 100, 5 (2010), 2383–2413.

[7] Susan Athey and Guido Imbens. 2016. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences* 113, 27 (2016), 7353–7360.

[8] Hamsa Bastani, Mohsen Bayati, and Khashayar Khosravi. 2017. Exploiting the Natural Exploration In Contextual Bandits. *arXiv preprint arXiv:1704.09011* (2017).

[9] Gordon M Becker, Morris H DeGroot, and Jacob Marschak. 1964. Measuring utility by a single-response sequential method. *Behavioral science* 9, 3 (1964), 226–232.

[10] James Berry, Greg Fischer, and Raymond P Guiteras. 2015. Eliciting and utilizing willingness to pay: evidence from field trials in Northern Ghana. (2015).

[11] Peter Bohm, Johan Lindén, and Joakim Sonnegård. 1997. Eliciting reservation prices: Becker–DeGroot–Marschak mechanisms vs. markets. *The Economic Journal* 107, 443 (1997), 1079–1089.

[12] Rebecca R Boyce, Thomas C Brown, Gary H McClelland, George L Peterson, and William D Schulze. 1992. An experimental examination of intrinsic values as a source of the WTA-WTP disparity. *The American Economic Review* 82, 5 (1992), 1366–1373.

[13] Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. 2011. Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on psychological science* 6, 1 (2011), 3–5.

[14] Bob Carpenter, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2016. Stan: A probabilistic programming language. *Journal of Statistical Software* 20, 2 (2016), 1–37.

[15] Peng Ding. 2017. A paradox from randomization-based causal inference. *Statistical science* 32, 3 (2017), 331–345.

[16] Ronald A Fisher. 1937. *The design of experiments.* Oliver And Boyd; Edinburgh; London.

[17] Daniel G Goldstein, Siddharth Suri, R Preston McAfee, Matthew Ekstrand-Abueg, and Fernando Diaz. 2014. The economic and cognitive costs of annoying display advertisements. *Journal of Marketing Research* 51, 6 (2014), 742–752.

[18] Ramu Govindasamy and John Italia. 1999. Predicting willingness-to-pay a premium for organically grown fresh produce. *Journal of Food Distribution Research* 30 (1999), 44–53.

[19] David M Grether and Charles R Plott. 1979. Economic theory of choice and the preference reversal phenomenon. *The American Economic Review* 69, 4 (1979), 623–638.

[20] Jaroslav Hájek. 1971. Comment on "An essay on the logical foundations of survey sampling, part one". *The foundations of survey sampling* 236 (1971).
[21] Jason Hartford, Greg Lewis, Kevin Leyton-Brown, and Matt Taddy. 2016. Counterfactual Prediction with Deep Instrumental Variables Networks. *arXiv preprint arXiv:1612.09596* (2016).
[22] Daniel G Horvitz and Donovan J Thompson. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association* 47, 260 (1952), 663–685.
[23] Taisuke Imai and Colin F Camerer. 2016. Estimating Time Preferences from Budget Set Choices Using Optimal Adaptive Design. *Working Paper* (2016).
[24] Guido W Imbens and Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
[25] Panagiotis G Ipeirotis. 2010. Demographics of mechanical turk. (2010).
[26] Steven J Kachelmeier and Mohamed Shehata. 1992. Examining risk preferences under high monetary incentives: Experimental evidence from the People's Republic of China. *The American Economic Review* (1992), 1120–1141.
[27] Dean Karlan and Jonathan Zinman. 2009. Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica* 77, 6 (2009), 1993–2008.
[28] Lihong Li, Wei Chu, John Langford, and Robert E Schapire. 2010. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*. ACM, 661–670.
[29] Carter A Mandrik and Yeqing Bao. 2005. Exploring the concept and measurement of general risk aversion. *NA-Advances in Consumer Research Volume 32* (2005).
[30] Winter Mason and Siddharth Suri. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods* 44, 1 (2012), 1–23.
[31] Nina Mazar, Botond Koszegi, and Dan Ariely. 2014. True context-dependent preferences? The causes of market-dependent valuations. *Journal of Behavioral Decision Making* 27, 3 (2014), 200–208.
[32] Timothy L McMurry and Dimitris N Politis. 2008. Bootstrap confidence intervals in nonparametric regression with built-in bias correction. *Statistics & Probability Letters* 78, 15 (2008), 2463–2469.
[33] Jerzey Neyman. 1923. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. Edited and translated by DM Dabrowska and TP Speed. *Statist. Sci.* 5 (1923), 465–72.
[34] Charles Noussair, Stephane Robin, and Bernard Ruffieux. 2004. Revealing consumers' willingness-to-pay: A comparison of the BDM mechanism and the Vickrey auction. *Journal of economic psychology* 25, 6 (2004), 725–741.
[35] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. 2010. Running experiments on amazon mechanical turk. (2010).
[36] Donald B Rubin. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology* 66, 5 (1974), 688.
[37] Carl-Erik Särndal, Bengt Swensson, and Jan Wretman. 1992. *Model assisted survey sampling*. Springer Science & Business Media.
[38] Adith Swaminathan and Thorsten Joachims. 2015. The self-normalized estimator for counterfactual learning. In *Advances in Neural Information Processing Systems*. 3231–3239.
[39] Stefan Wager and Susan Athey. 2017. Estimation and inference of heterogeneous treatment effects using random forests. *J. Amer. Statist. Assoc.* just-accepted (2017).
[40] Nathaniel T Wilcox. 1993. On a lottery pricing anomaly: time tells the tale. *Journal of Risk and Uncertainty* 7, 3 (1993), 311–324.