Truthful Incentives in Crowdsourcing Tasks using Regret Minimization Mechanisms

Adish Singla
ETH Zurich
Zurich Switzerland
adish.singla@inf.ethz.ch

Andreas Krause
ETH Zurich
Zurich Switzerland
krausea@ethz.ch

ABSTRACT

What price should be offered to a worker for a task in an online labor market? How can one enable workers to express the amount they desire to receive for the task completion? Designing optimal pricing policies and determining the right monetary incentives is central to maximizing requester's utility and workers' profits. Yet, current crowdsourcing platforms only offer a limited capability to the requester in designing the pricing policies and often rules of thumb are used to price tasks. This limitation could result in inefficient use of the requester's budget or workers becoming disinterested in the task.

In this paper, we address these questions and present mechanisms using the approach of regret minimization in online learning. We exploit a link between procurement auctions and multi-armed bandits to design mechanisms that are budget feasible, achieve near-optimal utility for the requester, are incentive compatible (truthful) for workers and make minimal assumptions about the distribution of workers' true costs. Our main contribution is a novel, no-regret posted price mechanism, BP-UCB, for budgeted procurement in stochastic online settings. We prove strong theoretical guarantees about our mechanism, and extensively evaluate it in simulations as well as on real data from the Mechanical Turk platform. Compared to the state of the art, our approach leads to a 180% increase in utility.

Categories and Subject Descriptors

K.4.4 [**Computers and Society**]: Electronic Commerce - Payment schemes; H.2.8 [**Database Management**]: Database applications - Data mining

General Terms

Algorithms, Economics, Experimentation, Human Factors, Theory

Keywords

Crowdsourcing, incentive compatible mechanisms, procurement auctions, posted prices, regret minimization, multi-armed bandits

1. INTRODUCTION

The growth of the Internet has created numerous opportunities for crowdsourcing tasks to online "workers". Specialized marketplaces for crowdsourcing have emerged, including Amazon's Mechanical Turk (henceforth MTurk [1]) and Click Worker[1], enabling "requesters" to post HITs (Human Intelligence Tasks), which can then be carried out by pools of workers available online and suitable for the task. Some of the tasks that are posted on these platforms include image annotation, rating the relevance of web pages for a query in search engines, translation of text or transcription of an audio recording. Similarly, in platforms like social networks, users can be compensated for participation in a viral marketing campaign. The requester generally has a limited budget for the task and needs to come up with a payment scheme for workers in order to maximize the utility derived from the task. For workers, the main goal is to maximize their individual profit by deciding which tasks to perform and at what price.

Monetary incentives in crowdsourcing tasks. One of the central components of these platforms is to design the right incentive structure and pricing policies for workers that maximize the benefits of both requester and the workers. Overpricing the tasks would result in inefficient use of the requester's budget, whereas underpricing could lead to task "starvation" because of unavailability of the workers willing to participate. In this light, how can one design optimal pricing policies? How can workers communicate and negotiate the price with requesters? How would the market behave if workers act strategically by misreporting their costs for their benefit? These are some of the questions that naturally come to mind while studying incentive structures for these online markets, yet they are not well understood.

Pricing models. Current crowdsourcing platforms offer limited capability to the requester in designing the pricing policies, mostly limiting them to a single fixed price ("*fixed price model*"). One way to set prices under such models is to estimate workers' costs via a market analysis and then compute an optimal fixed price which would maximize the utility. However, there are many difficulties in inferring this optimal fixed price, including the high cost of market surveys, the dynamic and online nature of the labor markets, inexperience of the requester and challenges in soliciting true costs from workers because of their self-interest. An alternate approach is to use tools of online procurement auctions where workers can bid on the price they are willing to receive

---

[1] http://www.clickworker.com/

and the requester's mechanism can decide on the allocation and prices to be paid to workers. In this "*bidding model*", mechanisms need to be truthful: it should be a dominant strategy for rational workers to bid their true cost. However, communicating these true costs to the requester may be challenging in real world settings. The worker may typically not trust the requester and understand the mechanism to reveal their true cost or the cost may not even be known to a worker and perhaps difficult to determine. Instead of soliciting the workers' costs, an often more natural setting is the "*posted price model*" where workers are offered a take-it-or-leave-it price offer. The mechanism interacts with each worker once in a sequential manner and adjusts the offered price from past responses of the workers.

1.1 Our Results

In this paper, we present a novel posted-price mechanism, BP-UCB, for online budgeted procurement, which is guaranteed to be budget feasible, achieve near-optimal utility for the requester, be incentive compatible (truthful) for workers and make minimal assumptions about the distribution of workers' true costs.

On the theoretical side, we present a novel mathematical analysis which exploits a link between procurement auctions and multi-armed bandits – a classical problem in online learning and experimental design – to prove regret bounds for the mechanism. BP-UCB builds on and extends existing mechanisms using multi-armed bandits for online auctions [25, 6] to procurement auctions under budget constraints. However, the mechanisms of [25, 6] are not directly applicable as they optimize a different objective, which leads to a substantially different mathematical analysis. Our analysis further yields insights into an explicit separation of the regret in terms of wasted budget through overpayment and rejected offers through underpayment. Additionally, our BP-UCB approach substantially improves upon the existing mechanisms for procurement auctions which are designed to achieve constant multiplicative approximation ratios [32, 8], which can lead to high additive regret.

We further carry out extensive experiments to compare the performance of BP-UCB with optimal benchmarks, as well as the state of the art mechanism of [8]. To the best of our knowledge, this is the first empirical study of posted-price mechanisms in procurement auctions. Apart from experimenting with simulated workers' cost distributions, we perform experiments using data gathered from an MTurk study to demonstrate the effectiveness of our approach on real world inputs. The results confirm the effectiveness and the practical applicability of using BP-UCB as posted price mechanism on crowdsourcing platforms.

2. RELATED WORK

Understanding incentives in crowdsourcing tasks. There has been growing interest in understanding the right incentives for workers in online labor markets. [21] present a model of workers and introduce methods to estimate workers' appropriate wages. The "hagglebot" of [22] negotiates payment rates for an image-labeling task with workers on MTurk. [28, 29] study other, non-monetary incentives that could improve the quality of workers' performance. [17] applies no-regret learning to better understand prediction markets and improve the results of information aggregation from crowds.

Multi-armed bandits (MAB) & regret minimization. The MAB problem is a natural formalism for studying settings where an agent repeatedly chooses among actions with uncertain rewards, and must trade exploration (gathering information about rewards) and exploitation (maximizing rewards obtained). A primary objective is to design no-regret algorithms which guarantee that the average regret approaches zero asymptotically over time when compared to the single best action in hindsight. MAB and regret minimization algorithms have been studied extensively and [13] gives a good overview. [3] introduces the UCB1 algorithm, which maintains an index (known as Upper Confidence Bound) on the actions and avoids explicit separation of exploration and exploitation by picking the action with the highest index. [24, 4, 10, 33] extend this approach to handle complex (possibly infinite) action spaces. Recently, budgeted variants of the MAB problem, where actions have different known costs, have been considered in [34, 36, 35]. [37] solves the crowdsourcing task whereby the goal is to learn workers' effectiveness as part of exploration. In our setting, in contrast, the costs are unknown and budget is utilized only in rounds when the offer is accepted by the worker – none of the standard approaches apply to this setting.

Learning in online auctions. Competitive online auctions were introduced in [26, 9]. These results were further extended and improved by using insights from regret minimization algorithms in [12, 11, 25]. [16] further extend the online posted price mechanisms for multi-parameter domains. [20, 18] study the auction problem with limited supply in the bidding model under stochastic arrival of the agents. [6] extends these results to the posted price model, by using insights from MAB problems. Our mathematical analysis builds on the results of [6]. However, in contrast, we consider dynamic pricing for procurement (reverse) auctions under a budget constraint. The results from [25, 6] are not applicable to this setting. In fact, straightforward application of the mechanisms of [25, 6] in our setting would simply offer the highest price, as that maximizes the utility of the action (acceptance of the price by the worker), though quickly exhaust the budget (and incur large regret).

Online procurement auctions. Mechanisms for procurement auctions have been extensively studied. Earlier work [2, 23, 14] concerns the frugality of mechanisms with the goal of procuring a feasible solution to a complex problem while minimizing the budget spent. In contrast, we are interested in studying truthful budget feasible mechanisms initiated recently in [30, 31]. Recent research addresses various models of budget constraints including the online knapsack secretary problem [7] and the weighted secretary problem [5]. However these are not directly applicable to truthful procurement mechanisms. [32] and [8] study a problem that is perhaps most similar to ours: they develop mechanisms for budgeted procurement in the stochastic setting for the bidding and posted price model respectively, and prove constant multiplicative bounds. In contrast, our mechanisms use the regret minimization framework, and we prove additive bounds on the regret. We note that mechanisms of constant multiplicative bounds could have arbitrarily poor performance in terms of additive regret. [32, 8] are also used as benchmarks for our experiments and our mechanism BP-UCB shows a substantial improvement over the state of the art mechanism of [8].

3. PROBLEM STATEMENT

We now formalize the problem addressed in this paper.

The requester and workers. There is a principal agent, the "requester", who posts the crowdsourcing task. A task is composed of atomic assignments, which can be performed by individual "workers". The requester has a budget $B > 0$ and a utility function over completed assignments. In this work, we assume that each assignment performed by a worker has unit value, thus the requester wishes to maximize the number of completed assignments subject to the budget constraint. There is a finite pool of workers, denoted by W. Each worker $w_i \in W$ is associated with a private cost $c_i \in \mathbb{R}_{\geq 0}$ for performing an assignment and $b_i \in \mathbb{R}_{\geq 0}$ is their bid or reported cost. We are interested in *truthful* mechanisms where it is a dominant strategy for worker w_i to report $b_i = c_i$. We assume that costs have known bounded support, i.e., $c_i \in [c_{\min}, c_{\max}]$ where $c_{\min}$ and $c_{\max}$ are the parameters of the problem, scaled such that $c_{\min} > 0$ and $c_{\max} \leq 1$. We note that the assumption of bounded costs naturally holds in online crowdsourcing platforms like MTurk, which generally enforce a publicly known minimal and maximal allowed payment for the assignments. We will keep the range of the costs fixed, and consider varying the budget. Also, the number of assignments per worker is normally set to one in MTurk. Furthermore, we assume that there are at least N workers where $N \geq \frac{B}{c_{\min}}$. We note that limiting the pool size to a much smaller number would lead to further constraints in our mechanism design in addition to the budget constraint and is beyond the scope of this work. However, having a very large, essentially infinite pool of workers would make the problem trivial as the mechanism can offer lowest possible prices without any overall loss of utility. We further discuss this issue in Section 6.

Online arrival of workers. We are interested in online settings where workers arrive one at a time. We focus on *stochastic* arrival of workers, where their costs are *i.i.d.* sampled from a distribution f (*i.i.d model*). We let $F : [c_{\min}, c_{\max}] \to [0, 1]$ denote the cumulative distribution function (CDF) of costs associated with the workers. Note that the stochastic arrival assumption may be violated in real online markets. This could be because of various factors, for example, if the workers' value of service increases over time and hence so does their cost. This non-stochastic setting is often called the *oblivious adversary model*. In our experiments, we empirically evaluate the robustness of our mechanisms in presence of such adversarial noise.

Optimal benchmarks. Consider an (unrealistic) offline mechanism with complete access to the pool of workers' true costs. The maximal utility in this setting can be achieved by sorting the workers by their increasing costs and offering each worker their true cost until the budget is exhausted. We denote this benchmark by OPT-VAR, i.e., the optimal variable price benchmark. An alternate benchmark of interest is a mechanism that is limited to offer a single fixed price to all workers, though this price is computed optimally assuming full knowledge of workers' costs. We denote this by OPT-FIX. Note that these benchmarks are offline, untruthful and assume full knowledge of the workers' costs. It seems natural to compare our online truthful mechanisms to the optimal truthful mechanism in offline settings. The utility of any single priced offline truthful mechanism is bounded by OPT-FIX and [19] shows that the performance of OPT-FIX is close to that of OPT-VAR. Further, recent results in pro-

curement auctions from [30, 8] show that OPT-FIX is only a factor of 2 away from OPT-VAR for modular as well as symmetric submodular functions and this is the best approximation that *any* truthful mechanism can achieve. Therefore we will compare our mechanisms against this benchmark. This optimal offline fixed price denoted by p^*, as illustrated in Figure 1, is given by:

$$p^* = \arg\max_p \min \left\{ F(p), \frac{B}{N \cdot p} \right\} \ s.t. \ \ p \in [c_{\min} \ldots c_{\max}].$$

Utility and Regret. For a fixed budget B, let $U(M, B)$ denote the expected utility of mechanism M and $U(p, B)$ denote the expected utility of fixed price p. In the regret minimization framework, we are interested in comparing the regret w.r.t. to the best single price p^* offered in hindsight. The expected regret of mechanism M is given by

$$R_M(B) = U(p^*, B) - U(M, B),$$

where $U(p^*, B) = \frac{B}{p^*}$. We write R instead of R_M where the mechanism M is clear from context. Using this notation, our goal is to design mechanisms M where the average regret of the mechanism approaches zero asymptotically, i.e.,

$$\lim_{B \to \infty} \frac{R(B)}{B} = 0.$$

4. MECHANISMS

We begin by describing the high level ideas behind our mechanisms. Then, we design BP-DGREEDY for the bidding model and subsequently extend it to arrive at our main mechanism, BP-UCB, for posted prices. In Section 4.4, we analyze the mechanisms and prove the regret bounds.

4.1 Methodology

Background on classical MAB. In the classical MAB setting [13, 3], there are K independent choices ("arms") associated with unknown reward distributions. A MAB algorithm operates in discrete timesteps (rounds) and pulls an arm in each round to get a stochastic reward associated with that arm. The algorithm needs to "explore" by experimenting with potentially suboptimal arms so as to learn about the optimal arm. Meanwhile, to maximize the reward, it has to "exploit" its learning by pulling the arm that appears best. The goal of the algorithm is to minimize the regret by quickly converging to the optimal arm.

Learning the cost curve and connection to MAB. The main challenge in deciding the payments in our problem is the unknown distribution of the workers' cost ("cost curve"). The mechanism interacts with workers sequentially in discrete timesteps denoted by t, offering a price p^t at each timestep to worker w^t and adjusting the estimates of the cost curve based on observed feedback. In order to cast our problem in the MAB framework, we discretize the prices by creating a set of K price "arms" using a multiplicative factor of $(1 + \alpha)$, where α is a parameter of the mechanism, similar to [12, 6], as illustrated in Figure 1. For these K arms, we maintain F_i^t as an estimate of the CDF of workers' costs for price p_i at time t. The mechanism stops execution when the budget or the pool of workers is exhausted. At each timestep, our mechanisms will pick the arm i^t based on some optimization criterion. Unfortunately, the presence of the budget constraint breaks the standard MAB algorithms: The optimal arm in terms of utility is the one corresponding to the maximal price, though it would quickly exhaust the
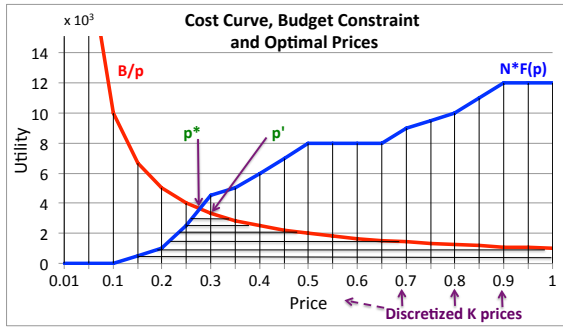
Figure 1: B/p represents the budget constraint, i.e. maximum utility that can be achieved given an infinite pool of workers. $N \cdot F(p)$ represents the utility with unlimited budget. The optimal price p^* lies at the intersection of these two curves. Discretized K prices are used by our mechanism, p' corresponds to the optimal price among these K prices.

available budget, leading to diminished utility. Further, we can exploit the fact that the price arms are correlated in our case: Acceptance at an offered price means acceptance for all more "expensive" arms and rejection at an offered price means rejection for all "cheaper" arms.

Additional notation. We need to introduce some more notation in order to describe our mechanisms. Let $i' \in \{1, \ldots, K\}$ denote the index of the optimal price among the K discretized prices, p' denote the corresponding price and F' the value of function $F_{i'} = F(p')$. As illustrated in Figure 1, i' is given by:

$$i' = \arg\max_i \ \min\left\{F_i, \frac{B}{N \cdot p_i}\right\} \quad \forall i \in [1 \ldots K].$$

We let B^t be the budget remaining at time t, N_i^t be the number of times p_i price has been offered, S_i^t be the indicator random variable indicating $b^t \leq p_i$ and T be the total number of rounds of execution of the mechanism (until the budget is exhausted or all the workers have been seen by the mechanism). Let us define $V_i = \min\left\{F_i, \frac{B}{N \cdot p_i}\right\}$, V_i^t be the estimate of V_i at round t. We further use $\Delta_i = V_{i'} - V_i$.

4.2 BP-DGREEDY for bidding model

We begin by designing a mechanism, BP-DGREEDY (deterministic greedy for budgeted procurement), for the bidding model (see Mechanism 1) in order to gain insights into the problem and develop mathematical foundations for the posted price model. The mechanism solicits the workers' bid b^t about service cost and then offers them a price p^t, based on past observations of workers' bids, which the worker can accept or reject. A natural approach towards ensuring truthful bids is to make the offered price *independent* of the bid of the current worker. Because of this truthfulness, we have $b^t = c^t$, which makes this mechanism resemble online learning with full information (refer [13, 27, 15]). This intuitively means that the mechanism gets to compute response feedback it would have received for any possible action.

We now discuss how to pick the arm i^t to make a price offer p^t independent of b^t. The intuition is simple: BP-DGREEDY just tracks the expected utility V_i^t of the arms based on the estimated F_i^t and offers the price corresponding to the best arm at time t. Based on the observed bid, it updates the estimates of F_i^t for all the arms by simply maintaining the average response of acceptance at arm i.

Mechanism 1: BP-DGREEDY
1 Parameters: B; N; $\alpha = (0, 1]$; $c_{\min}$; $c_{\max}$;
2 Initialize:
- **Prices**. $p_0 = c_{\min}$; $p_i = (1 + \alpha) \cdot p_{i-1}$; $p_K = c_{\max}$;
- **Variables**. *time*: $t = 0$; *budget* $B^t = B$; *utility*: $U = 0$;
- **Value estimates**. $F_i^t = 0$;

begin
3   **while** $B^t > c_{\min}$ & $t < N$ **do**
4 $i^t = \arg\max_i V_i^t \ s.t. \ p_i \leq B^t$;
      /* ties broken by picking lowest i */ ;
5   Offer price $p^t = p_{i^t}$ to worker w^t;
6   Observe bid b^t;
7 $\forall i$, Update $F_i^t = F_i^t + \frac{(S_i^t - F_i^t)}{(t+1)}$;
8   Set $U = U + S_{i^t}^t$; $B^{t+1} = B^t - p^t \cdot S_{i^t}^t$; $t = t + 1$;
   end
end
9 Output: U

4.3 BP-UCB for posted price model

Next, we present our main mechanism, BP-UCB (UCB for budgeted procurement, see Mechanism 2) for the posted price model, where we get to see acceptance or rejection feedback only for the price offered. This limited feedback leads to a natural exploration-exploitation tradeoff as in MAB problems. We tackle this problem by modifying BP-DGREEDY, whereby we maintain upper confidence bounds on F_i^t, denoted by $\widetilde{F_i^t}$, which are then used to optimistically estimate the values $\widetilde{V_i^t}$. The mechanism then picks the best arm based on $\widetilde{V_i^t}$ and offers a take-it-or-leave it price p^t. Based on the feedback from the worker, it updates the parameters only for the arm i^t. This approach is inspired from the classical UCB1 algorithm [3]. However, the budget constraints make the analysis of regret bounds non-trivial.

We further exploit the correlation between the arms, as a rejection response actually means rejection for all "cheaper" arms. Similar to [11], we use this correlation to further improve the execution performance of the mechanism by keeping an estimate of the lower bound of the cost curve's support and keeping all the arms (except one) below this estimate as inactive. This modification does not hurt the theoretical guarantees described in Section 4.4.

4.4 Performance Analysis

We now prove regret bounds for our mechanisms BP-DGREEDY and BP-UCB. We begin with the analysis of BP-DGREEDY and develop the mathematical tools that will be useful for the analysis of BP-UCB. A crucial challenge in dealing with the budget constraint lies in the fact that while higher prices are more "effective" since more workers would accept the offer, they would quickly exhaust the budget leading to reduced overall utility.

Components contributing to the regret. There are essentially three components contributing to the regret. The first is the discretization of the prices: Since the mechanism does not have access to the optimal price p^*, p' is the best price available. After accounting for the regret of "discretization", we can consider an alternative mechanism M' which has access to an additional arm corresponding to p^*. Considering M', the second component of the regret is attributed to pulling arms with prices $p_i < p^*$ as cheaper arms are less effective and result in more rejected offers. The third

---
Mechanism 2: BP-UCB
---
1 Parameters: B; N; $\alpha = (0, 1]$; $c_{\min}$; $c_{\max}$;
2 Initialize:
   - **Prices**. $p_0 = c_{\min}$; $p_i = (1 + \alpha)\cdot p_{i-1}$; $p_K = c_{\max}$;
   - **Variables**. *time*: $t = 0$; *budget* $B^t = B$; *utility*: $U = 0$;
   - **Value estimates**. $N_i^t = 0$; $F_i^t = 0$;

   begin
3 | while $B^t > c_{\min}$ & $t < N$ do
4 | | $\widetilde{F_i^t} = F_i^t + \sqrt{\frac{2\cdot \ln(t)}{N_i^t}}$;
5 | | $\widetilde{V_i^t} = \min\left\{\widetilde{F_i^t}, \frac{B}{N\cdot p_i}\right\}$;
6 | | $i^t = \arg\max_i \widetilde{V_i^t}$ s.t. $p_i \leq B^t$;
   | | /* ties broken by picking lowest i */ ;
7 | | Offer price $p^t = p_{i^t}$ to worker w^t;
8 | | Observe acceptance decision y^t;
9 | | Update $F_{i^t}^t = F_{i^t}^t + \frac{(y^t - F_{i^t}^t)}{(N_{i^t}^t + 1)}$; $N_{i^t}^t = N_{i^t}^t + 1$;
10 | | Update $U = U + y^t$; $B^{t+1} = B^t - p^t\cdot y^t$; $t = t + 1$;
   | **end**
   end
11 Output: U
---

component of the regret is attributed to pulling arms with prices $p_i > p^*$. Though these "expensive" arms are more effective than the price p^*, they overpay and quickly exhaust the budget. We formalize the above discussion in Lemma 1.

Lemma 1. *The expected regret $R_M(B)$ of any mechanism M can be expressed in terms of three components as follows:*

$$R_M(B) < \underbrace{\left(\frac{B}{p^*} - \frac{B}{p'}\right)}_{Discretization} + \underbrace{\sum_{i:p_i<p^*} \mathbb{E}\left[N_i^T\right]\cdot\left(F^* - F_i\right)}_{Rejected\ offers}$$

$$+ \underbrace{\sum_{i:p_i>p^*} \frac{\mathbb{E}\left[N_i^T\right]\cdot\left(p_i\cdot F_i - p^*\cdot F^*\right)}{p^*} + \frac{c_{\min}}{p^*}}_{Wasted\ budget\ through\ overpayment}$$

The proof is given in Appendix A.

Regret bounds for BP-DGREEDY. To obtain the desired regret bounds for mechanism BP-DGREEDY from Lemma 1, we need to bound N_i^T, as well as the regret of discretization. We bound the N_i^T for an execution of BP-DGREEDY using the Chernoff-Hoeffding concentration inequalities as in [3]. By exploiting the ordering of arms, we are able to separately provide bounds for N_i^T for arms with prices $p_i < p'$ and $p_i > p'$. For our analysis, we consider four separate cases based on whether $F_{i'}^t$ is less or greater than $\frac{B}{N\cdot p'}$ and based on relative ordering of p' compared to p^*. This insight crucially simplifies the analysis, enabling us to use the tools from the original UCB1 analysis in bounding each one of these four cases separately. Theorem 1 provides the desired regret bounds for BP-DGREEDY.

Theorem 1. *The expected regret of mechanism* BP-DGREEDY *is upper-bounded as follows:*

$$R_{\text{BP-DGREEDY}}(B) < \frac{\alpha\cdot B}{p^*} + \sum_{i:p_i<p^*} \frac{4\cdot\left(F^* - F_i\right)}{\Delta_i^2}$$

$$+ \sum_{i:p_i>p^*} \frac{\left(p_i\cdot F_i - p^*\cdot F^*\right)}{2\cdot\Delta_i^2\cdot p^*} + \frac{c_{\min}}{p^*}$$

The proof is given in Appendix A.2. Next, we prove the no-regret property of mechanism BP-DGREEDY by using an appropriate choice of the discretization factor α, similar to the choice made for the problem of online auctions with limited supply in [6].

Corollary 1. *The expected average regret of mechanism* BP-DGREEDY *w.r.t. the budget size B goes to zero asymptotically for appropriate choice of α.*

Proof. By setting α to $\mathcal{O}\left(\frac{\ln(B/c_{\min})}{B}\right)$, the expected average regret of BP-DGREEDY *w.r.t.* B in the limit $\lim_{B\to\infty}$ is given as:

$$\lim_{B\to\infty} \mathbb{E}\left[\frac{R_{\text{BP-DGREEDY}}(B))}{B}\right] = \lim_{B\to\infty} \frac{\mathcal{O}\left(\ln(B/c_{\min})\right)}{B} = 0. \;\square$$

Regret bounds for BP-UCB. We now extend the analysis of BP-DGREEDY to BP-UCB. Theorem 2 provides the desired regret bounds for BP-UCB.

Theorem 2. *The expected regret of mechanism* BP-UCB *is upper-bounded as follows:*

$$R_{\text{BP-UCB}}(B) < \frac{\alpha\cdot B}{p^*}$$

$$+ \sum_{i:p_i<p^*} \left(\frac{8\cdot\ln(B/c_{\min})}{\Delta_i^2} + \frac{\pi^2}{3} + 1\right)\cdot\left(F^* - F_i\right)$$

$$+ \sum_{i:p_i>p^*} \frac{\pi^2\cdot\left(p_i\cdot F_i - p^*\cdot F^*\right)}{6\cdot p^*} + \frac{c_{\min}}{p^*}$$

The proof is given in Appendix A.3. Corollary 2 proves the no-regret property of the mechanism.

Corollary 2. *The expected average of regret of mechanism* BP-UCB *w.r.t. budget size B goes to zero asymptotically for appropriate choice of α.*

Proof. The proof follows by using exactly the same arguments as in Corollary 1. $\square$

5. EXPERIMENTAL EVALUATION

In this section, we carry out extensive experiments to understand the practical performance of our mechanisms on simulated cost distributions, as well as on costs derived from an actual MTurk study. We begin by describing our benchmarks, metrics and experimental setup.

Benchmarks. We compare our mechanisms against the following benchmarks and state-of-art mechanisms:
- OPT-VAR and OPT-FIX: These are offline, untruthful mechanisms with full information of the workers' true costs as discussed in Section 3.
- MEAN: Another offline mechanism that operates under the bidding model. It offers a fixed price computed as the mean value of the workers' bids. This mechanism serves as a rule of thumb to determine fixed prices for tasks, as possibly used by inexperienced requesters.
- BS'11: We implemented the mechanism for the bidding model based on sampling bids from [32]. This mechanism assumes that workers' arrival order is stochastic *i.i.d.*
- PP'12: This is the online posted price mechanism from [8] designed for the stochastic setting. We found that the recommended parameters used for proving theoretical guarantees did not work in practice. We therefore
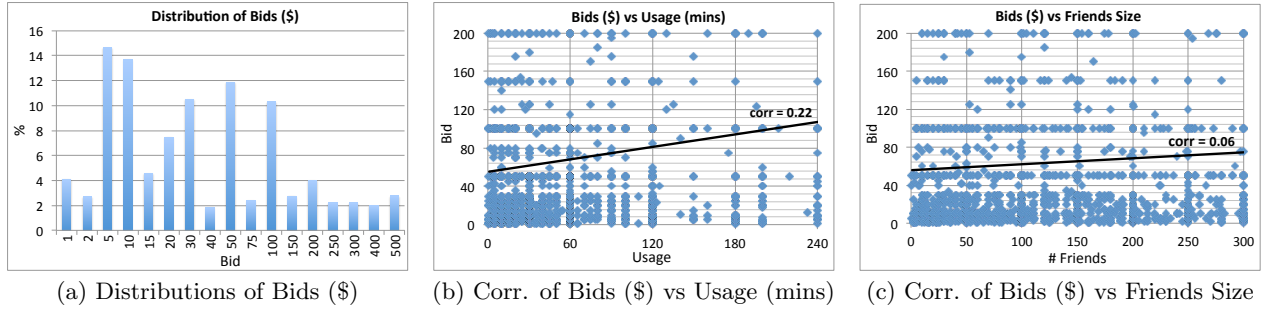
(a) Distributions of Bids ($)   (b) Corr. of Bids ($) vs Usage (mins)   (c) Corr. of Bids ($) vs Friends Size

Figure 2: **(a)** Distribution of workers' bids, Correlation with **(b)** usage time and with **(c)** number of friends online

manually tuned parameters to optimize the performance of this benchmark. Specifically, we ignored the parameter z in determining the price of the highest arm as $\frac{B}{z}$ and instead used $c_{\max}$ as a bound. Also, we used $a = 5$ instead of 4000, which would need an extremely large pool of workers for execution.

Metrics and experiments. The primary metric we track is the utility of the mechanism as we vary the budget B, setting $N = \frac{B}{c_{\min}}$. We also compute the average regret of the mechanism *w.r.t.* increasing budget to verify its no-regret property. To study the effect of the worker pool size, we also look into varying N for a fixed budget. To gain insight into execution of the mechanisms, we measure their rate of convergence by determining the unique price to which the mechanism converges in the end and measuring the number of times this price has been offered so far with increasing timesteps. Lastly, we evaluate the utility over time to understand the dynamics of how quickly the budget is exhausted.

Parameter choices. We used $c_{\min} = 0.01$ and $c_{\max} = 1$ based on the payment bounds typically seen on MTurk. The price discretization factor α is set to 0.2. We note that setting α to $\mathcal{O}\left(\frac{\ln(B/c_{\min})}{B}\right)$ guarantees asymptotic bounds of $\mathcal{O}\left(\ln(^B/_{c_{\min}})\right)$, however smaller values of α would increase the number of arms, leading to slower convergence.

Cost distributions. We considered cost distributions based on simulations as well as gathered from an actual MTurk study. We considered various simulated distributions for analyzing our algorithms, including uniform, normal, exponential and more complex ones including mixture of two uniform or two Gaussian distributions. Also, we considered various settings to simulate the arrival of workers, including ordering by ascending bids to simulate adversarial arrival. To simulate a more realistic non-stochastic setting, we considered groups of two distributions arriving one after another, ordered by their increasing means.

5.1 Mechanical Turk Study

The primary objective of this study was to get cost distributions from a realistic crowdsourcing task, as well as to understand the non-stochastic nature of real inputs that could arise because of the task's temporal nature.

Data Collection. We posted a HIT (Human Intelligent Task) on MTurk in form of a survey, where workers were told about an option to participate in a hypothetical advertisement system for a social networking site. In this hypothetical system, they can opt to use the top of their homepage for banner ads and obtain some monthly payment from the publishers. Workers were asked to bid on the monthly payment they would like to receive, in addition to providing informa-

| Feature	$July10$	$July11$	$July12$	all
# Active workers	641	833	867	867
# Friends (mean)	165	230	318	242
# Friends (median)	100	160	250	157
Usage mins (mean)	64	73	83	74
Usage mins (median)	30	50	60	50
Bids $ (mean)	55	66	90	71
Bids $ (median)	20	25	30	25

Table 1: Statistics of the data reported for different years

tion like years of being active on social networks and time spent there, approximate geographical location, number of friends and optional comments. They were asked to provide this information for different times including July 2012, July 2011 and July 2010. The goal of this survey was twofold. Firstly, we wanted to get the workers' cost distribution for a realistic scenario which fits our procurement auction task. Secondly, we wanted to understand whether the assumption of stochastic costs holds true in real world inputs. A total of 1200 workers participated in our HIT, which was online for one week, restricted to workers with more than 90% approval rate. Workers were paid a fixed amount for participation in the HIT. We did not restrict the workers to any geographical region. Additionally, we made a bonus payment to selected 20 individuals based on their insightful comments about factors affecting their payment choice.

Statistics. The workers represented more than 20 different countries with 44.5% from USA and 44.0% from India. In total 72.25% of the workers agreed to participate in the hypothetical online advertisement system, and we analyze the statistics from these workers below. Workers from USA have lower acceptance rate of participation (59.3%) compared to workers from India (86.1%), which shows interesting dependence on geographical factors in determining the pricing model of the workers. Table 1 shows the mean and median values of various features. The data shows an increase in social activity (in terms of friend count and service usage) as well as the bids reported by workers for their service. Figure 2 shows the distribution of bids collected as well as the correlation with usage time and number of friends. The data is skewed towards lower bids and is discretized because of the tendency of workers to bid at rounded numbers. In total 75.8% of the workers provided subjective feedback in the comments section about the pricing factors. Common themes reported by the workers for the pricing factors were the usage time, friend count and nature of the ads. The statistics related to friend size matches closely with those of publicly available numbers supporting the quality of the data obtained from the workers.
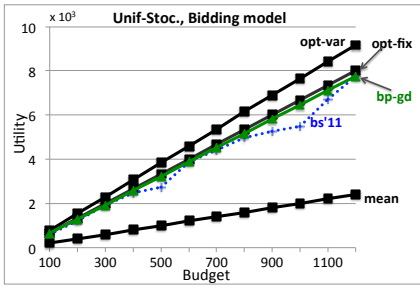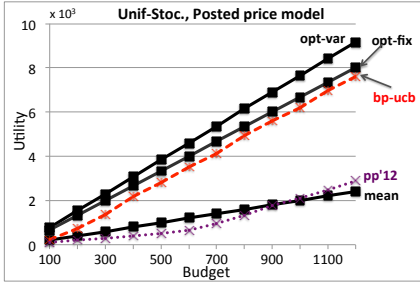
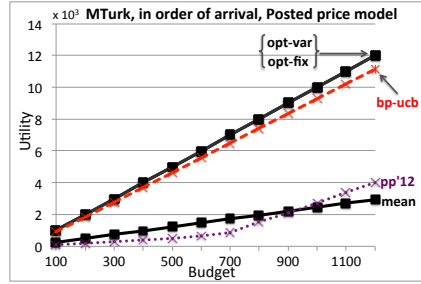(a) Uniform-Stoc., Bidding model
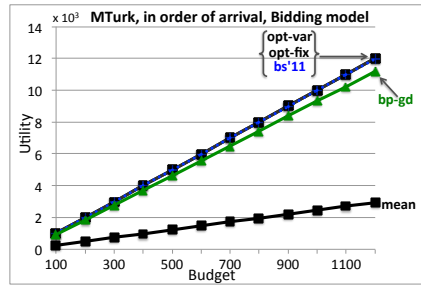
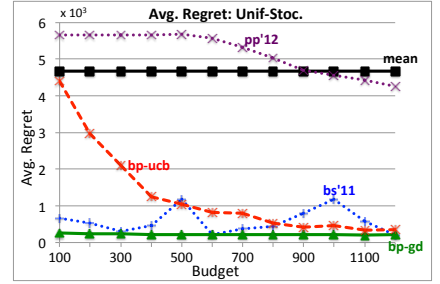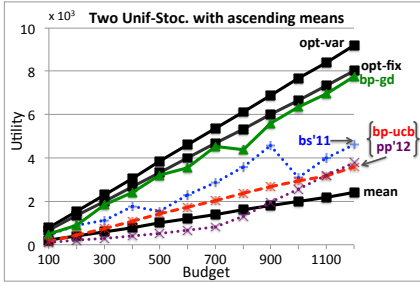(a) MTurk arrival, Bidding model
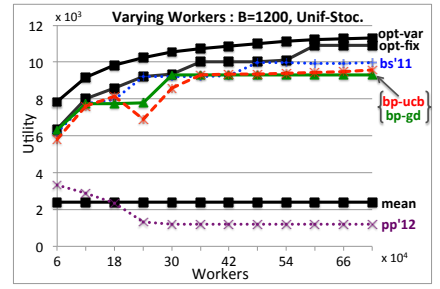
(a) Varying workers

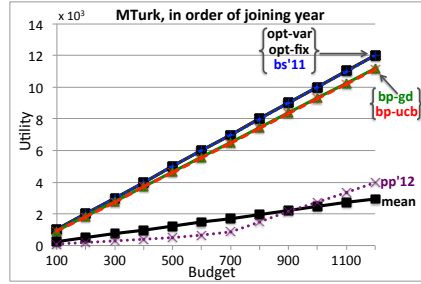(b) Uniform-Stoc., Posted price model
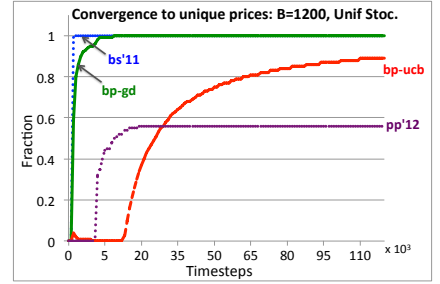

(b) MTurk arrival, Posted price model
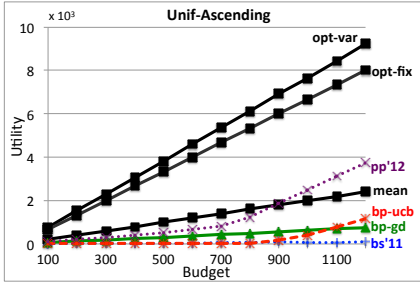

(b) Average regret

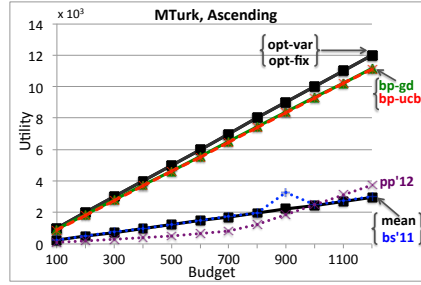(c) Two Uniform-Stoc. with asc. means

(c) MTurk, Arrival by year of joining
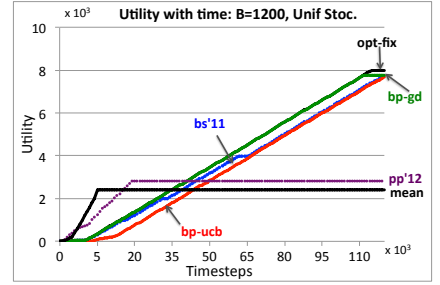

(c) Convergence

(d) Uniform, Ascending bids

(d) MTurk, Arrival by ascending bids

(d) Utility with time

Figure 3: 1st column   **Figure 4: 2nd column**   **Figure 5: 3rd column**

Figure 3 (1st column): Utility for uniform distribution in [0.1, 0.9], varying budget. In **(a)**, BP-UCB outperforms PP'12 by over 150% increase in utility for the stochastic settings. **(c)** uses two uniform distributions in [0.1, 0.5) and [0.5, 0.9].
Figure 4 (2nd column): Utility for MTurk distribution for bids in [10, 100], varying budget. In **(a)**, BP-UCB outperforms PP'12 by over 180% increase in utility. Also, BP-UCB and BP-DGREEDY are robust against all the online settings above.
Figure 5 (3rd column): Uniform distribution in [0.1, 0.9], stochastic settings. In **(b)**, no-regret properties of BP-UCB can be seen as the average regret diminishes with increase in budget. **(c)** shows better convergence rate of BP-UCB compared to PP'12. **(d)** shows that BP-UCB makes low offers in beginning, in contrast to PP'12 which quickly exhausts the budget.

5.2 Results

We now present and discuss the findings from our experiments. In the figures, we denote BP-DGREEDY as *bp-gd*.

Utility for simulated distributions. Figure 3 shows results for costs uniformly distributed in the range [0.1, 0.9], though the results are qualitatively similar for other distri-

butions and ranges. We consider the online setting with stochastic arrival of workers and also assess the robustness of the mechanisms when these assumptions are violated. In Figure 3(a), we can see that the mechanism BP-DGREEDY performs very close to OPT-FIX and slightly outperforms the state of the art mechanism BS'11 for the bidding model.

Somewhat surprisingly, as we can see in Figure 3(b), our mechanism BP-UCB for the posted price model performs as good as BP-DGreedy even though it operates under limited feedback. It clearly outperforms PP'12 by an over 150% increase in utility for all the budgets considered. Mean is much lower compared to both of our mechanisms, suggesting that rules of thumb prices may not be optimal.

We also simulate arrival of the workers in order of ascending bids, violating the stochastic *i.i.d.* assumptions. In Figure 3(d), we see that all the mechanisms perform quite poorly in this somewhat unrealistic case. Figure 3(c) shows results for a perhaps more meaningful non-stochastic setting where two groups with bids uniformly distributed in [0.0, 0.5) and [0.5, 0.9] respectively arrive after another. Therefore, a natural question is how robust the algorithms are w.r.t. more realistic cost distributions, as we analyze next.

Utility for MTurk distributions. Figure 4 shows results for cost distributions from the MTurk study. We considered bids ranging in [10$, 100$], scaled down by 100, although the results are qualitatively similar for other ranges. Note that scaling down the costs is equivalent to scaling up the budget. We sampled with replacement from the bids to generate the entire pool of workers. We considered various online settings to simulate the arrival order of workers: the actual order in which workers arrived on MTurk for completing the task; ordered by their usage time; ordered by number of friends and by the year of joining the social network. Here, we discuss the results for the arguably most natural orderings. For the bidding model, we can see in Figure 4(a) that both OPT-Fix and BS'11 coincide exactly with OPT-Var, in the case of workers arriving according to the actual MTurk ordering. We attribute this to the highly skewed nature of bids at low prices, as the optimal strategy for all these three mechanisms is to offer a single fixed price corresponding to the lowest bid. For the posted price model in Figure 4(b), BP-UCB clearly outperforms PP'12, increasing the utility by over 180% for all the budgets considered.

Figure 4(c) shows results where workers are ordered by the year in which they joined and Figure 4(d) shows the results where workers are ordered by their increasing bids on MTurk. Interestingly, BP-UCB and BP-DGreedy continue to perform well in both the settings, whereas BS'11 degrades in Figure 4(d) and PP'12 performs poorly in both.

Effect of varying N. Apart from varying the budget, it is interesting to compare the impact of workers' pool size on the mechanisms for a fixed budget. Note that the availability of more workers (larger N) shifts the optimal solution towards lower prices. Figure 5(a) shows the impact of varying N. As one would expect, our mechanisms BP-DGreedy and BP-UCB as well as BS'11 show an increase in utility exhibiting diminishing returns. Interestingly, PP'12 shows a decrease in utility as the number of workers increases.

Average regret. Figure 5(b) shows the average regret of the mechanisms with increasing budget. Note that the average regret of BP-UCB decreases at a much faster rate compared to that of PP'12.

Rate of convergence. Here, we look at the rate of convergence of the mechanisms in Figure 5(c), by computing the proportion of times the unique price, to which the mechanism converges in end, has been offered so far with increasing timesteps. BS'11 rapidly converges to the unique price, favorably compared to BP-DGreedy. We can see an initial phase of "exploration" for BP-UCB followed by "exploitation" as the mechanism converges. However, PP'12 stabilizes at 50% convergence as the Markov model used by the mechanism flips back and forth between the equilibrium prices.

Utility with timesteps. Lastly, we study how the mechanisms accrue utility over time (Figure 5(d)). BP-UCB offers very low prices in the initial phase of "exploration", followed by convergence to a unique price, after which the utility increases almost linearly. In contrast, PP'12 quickly exhausts the budget in the beginning by offering high prices, leading to overall reduced utility.

6. CONCLUSIONS AND FUTURE WORK

We designed mechanisms for online budgeted procurement using a regret minimization approach. We started with mechanism BP-DGreedy for the bidding model and then extended it to our main mechanism BP-UCB for the posted price model. These are the first provable no-regret mechanisms for online budgeted procurement. Apart from theoretical guarantees, we prove that they are empirically efficient compared to optimal benchmarks, and dramatically outperform the state of the art posted price mechanism. Our experiments on MTurk further supports the practical applicability of our mechanisms on crowdsourcing platforms. We believe that our results provide an important step towards developing practical, yet theoretically well-founded techniques for increasing the efficiency of crowdsourcing.

There are some natural extensions for future work. Here, we considered a simple additive utility function for the requester. It would be useful to extend our approach to more complex utility functions. Additionally, we assumed a homogeneous pool of workers, although it would be more practical to design mechanisms which can take into account skills and different utility values of the workers.

Our experiments on MTurk suggest that real world inputs may violate stochastic assumptions. While our mechanisms are robust against our study's cost distribution, one can force all the currently available mechanisms to perform poorly by carefully designing (unrealistic) cost distributions. It would be of interest to develop mechanisms that are more robust, and extend to the oblivious adversary model.

We used the knowledge of known bounded support and furthermore discretized the price space. Results from continuous arm bandits in [24, 4, 10, 33] can be applied here by making more realistic assumptions about cost distributions. This would enable learning the support as part of the mechanism itself and remove the regret from discretization.

In this work, we assumed a finite yet large pool of available workers. A perhaps more natural approach is to use time discounted rewards where a mechanism's goal would be *timely* completion of the task. Existing crowdsourcing platforms support only fixed price mechanisms and limited capabilities to design pricing policies. Our experiments show that simple mechanisms like Mean perform quite poorly, although inexperienced requesters may be tempted to use them as rule of thumb. It would be interesting to build applications and conduct studies where we can actually run our mechanisms in real time on crowdsourcing platforms.

APPENDIX

A. PROOFS

A.1 Components contributing to the regret

We begin by expressing the expected utility of a mechanism in terms of N_i^T in Lemma 2. We use I_i^t as indicator variable indicating that price p_i was offered at time t.

Lemma 2. *The expected utility of the mechanism M is given by $U(M, B) = \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot F_i$*

Proof. From definition of S_{it}^t, we have

$$U(M, B) = \mathbb{E}_T\left[\mathbb{E}\left[\sum_{t=1}^T S_{it}^t \mid T\right]\right] = \sum_{t=1}^T \sum_{i=1}^K \mathbb{E}_T\left[\mathbb{E}\left[I_i^t \cdot S_i^t \mid T\right]\right]$$

$$= \sum_{t=1}^T \sum_{i=1}^K \mathbb{E}_T\left[\mathbb{E}[I_i^t \mid T]\right] \cdot \mathbb{E}_T\left[\mathbb{E}[S_i^t]\right] \quad (1)$$

$$= \sum_{i=1}^K \mathbb{E}_T\left[\mathbb{E}\left[\sum_{t=1}^T I_i^t \mid T\right]\right] \cdot F_i = \sum_{i=1}^K \mathbb{E}_T\left[\mathbb{E}[N_i^T \mid T]\right] \cdot F_i \quad (2)$$

$$= \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot F_i$$

Unlike UCB1, T is a random variable here. Step 1 uses the fact that S_i^t only depends on the order of bids and is therefore independent of the mechanism (i.e., of I_i^t and T). Step 2 follows from definitions of N_i^T and F_i. □

Next, in Lemma 3, we provide a lower bound on the expected timesteps T in the execution of a mechanism, unlike in standard UCB1 where T is fixed. This ensures that the mechanism's regret coming from the component "Wasted budget through overpayment" in Lemma 1 is bounded.

Lemma 3. *The expected number of timesteps T in the execution of the mechanism has lower bound as follows:*

$$\mathbb{E}[T] > \frac{B}{p^* \cdot F^*} - \frac{c_{\min}}{p^* \cdot F^*} - \sum_{i:p_i > p^*} \frac{\mathbb{E}[N_i^T] \cdot (p_i \cdot F_i - p^* \cdot F^*)}{p^* \cdot F^*}$$

Proof. The algorithm terminates when either *i)* $t > N$ or *ii)* $B^t < c_{\min}$. In *i)*, $T = N$ and hence bounds in the equation hold trivially since $N \geq \frac{B}{c_{\min}} > \frac{B}{p^*} > \frac{B}{p^* \cdot F^*}$. We will prove the bounds for *ii)* below by bounding the sum of the prices offered and accepted by workers $\sum_{t=1}^T S_{it}^t \cdot p^t$.

$$B - c_{\min} < \mathbb{E}_T\left[\mathbb{E}\left[\sum_{t=1}^T S_{it}^t \cdot p^t \mid T\right]\right] = \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot F_i \cdot p_i \quad (1)$$

$$= \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot (F_i \cdot p_i - F^* \cdot p^*) + \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot F^* \cdot p^*$$

$$\mathbb{E}[T] \cdot F^* \cdot p^* > B - c_{\min} - \sum_{i:p_i < p^*} \mathbb{E}[N_i^T] \cdot (F_i \cdot p_i - F^* \cdot p^*)$$
$$- \sum_{i:p_i > p^*} \mathbb{E}[N_i^T] \cdot (F_i \cdot p_i - F^* \cdot p^*) \quad (2)$$

Step 1 follows by using the same arguments as in Lemma 2. Step 2 replaces $\sum_i^K N_i^T$ by T and we get the desired bounds by using the fact that $(F_i \cdot p_i - F^* \cdot p^*) < 0$ for $p_i < p^*$. □

We now prove Lemma 1 by using the above results.

Proof of Lemma 1. Consider an alternate mechanism M' which has access to an additional arm corresponding to price p^*. We first analyze the regret of M' below.

$$R_{M'}(B) = U(p^*, B) - U(M', B) = \frac{B}{p^*} - \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot F_i$$

$$= \frac{B}{p^*} + \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot (F^* - F_i) - \sum_{i=1}^K \mathbb{E}[N_i^T] \cdot F^*$$

$$= \frac{B}{p^*} - \mathbb{E}[T] \cdot F^* + \sum_{i:p_i < p^*} \mathbb{E}[N_i^T] \cdot (F^* - F_i)$$
$$+ \sum_{i:p_i > p^*} \mathbb{E}[N_i^T] \cdot (F^* - F_i) \quad (1)$$

$$\leq \frac{B}{p^*} - \mathbb{E}[T] \cdot F^* + \sum_{i:p_i < p^*} \mathbb{E}[N_i^T] \cdot (F^* - F_i) \quad (2)$$

$$< \frac{c_{\min}}{p^*} + \sum_{i:p_i < p^*} \mathbb{E}[N_i^T] \cdot (F^* - F_i)$$
$$+ \sum_{i:p_i > p^*} \frac{\mathbb{E}[N_i^T] \cdot (p_i \cdot F_i - p^* \cdot F^*)}{p^*} \quad (3)$$

Step 1 replaces $\sum_i^K N_i^T$ by T and step 2 uses the fact that $(F^* - F_i) < 0$ for $p_i > p^*$. And, step 3 follows from the results of Lemma 3.

The regret of mechanism M is given by $R_M(B) = R_{M'}(B) + R_{p'}(B)$ where $R_{p'}(B) = \left(\frac{B}{p^*} - \frac{B}{p'}\right)$. Using the value of $R_{M'}$ from above completes the proof. □

A.2 Regret bounds for BP-DGREEDY

In Lemma 4, we show that discretization to a power of $(1 + \alpha)$ results in loss of utility by at most a factor of $(1 + \alpha)$, similar to the online auctions as shown in [12, 6].

Lemma 4. *The regret component "discretization" in Lemma 1 is upper-bounded by $R_{p'}(B) \leq \frac{\alpha \cdot B}{p^*}$.*

Proof. Consider price p^h given by:

$$p^h = \inf_{p_i} \{\forall i \in [1 \ldots K] \ s.t. \ p_i \geq p^*\}$$

By the design of discretization, $p^h < (1 + \alpha) \cdot p^*$. Now, consider that price p^h is offered to every worker instead of p'.

$$R_{p'}(B) = U(p^*, B) - U(p', B) \leq \frac{B}{p^*} - \frac{B}{p^h} \leq \frac{\alpha \cdot B}{p^*}$$

We use the fact that for $p^h \geq p^*$, $U(p^h; B) = \frac{B}{p^h}$. □

Lemma 5. *$\forall i \ s.t. \ p_i < p'$, expected number of times a suboptimal arm i is played is upper-bounded by $\mathbb{E}[N_i^T] \leq \frac{4}{\Delta_i^2}$.*

Proof. A suboptimal arm i is picked at time t when $V_i^t \leq V_{i'}^t$ where $V_j^t = \min\left\{F_i^t, \frac{B}{N \cdot p_j}\right\}$. We consider the following cases:

Case a) $p' \leq p^*$ and $F_{i'}^t \leq \frac{B}{N \cdot p'}$:

$$F_i^t \geq F_{i'}^t$$

$$\left(F_i^t - F_i - \frac{F_{i'} - F_i}{2}\right) + \left(F' - F_{i'}^t - \frac{F_{i'} - F_i}{2}\right) \geq 0$$

$$\left((F_i^t - F_i) - \frac{\Delta_i}{2}\right) + \left((F' - F_{i'}^t) - \frac{\Delta_i}{2}\right) \geq 0 \quad (1)$$

Case b) $p' \leq p^*$ and $F_{i'}^t > \frac{B}{N \cdot p'}$:

$$F_i^t \geq \frac{B}{N \cdot p'} \geq F' \implies F_i^t - F_i \geq \Delta_i \quad (2)$$

Case c) $p' > p^*$ and $F_{i'}^t \le \frac{B}{N \cdot p'}$:

$$F_i^t \ge F_{i'}^t$$

$$\left(F_i^t - F_i - \frac{\frac{B}{N \cdot p'} - F_i}{2}\right) + \left(\frac{B}{N \cdot p'} - F_{i'}^t - \frac{\frac{B}{N \cdot p'} - F_i}{2}\right) \ge 0$$

$$\left((F_i^t - F_i) - \frac{\Delta_i}{2}\right) + \left((F' - F_{i'}^t) - \frac{\Delta_i}{2}\right) \ge 0 \qquad (3)$$

Case d) $p' > p^*$ and $F_{i'}^t > \frac{B}{N \cdot p'}$:

$$F_i^t \ge \frac{B}{N \cdot p'} = \frac{B}{N \cdot p'} + F_i - F_i \implies F_i^t - F_i \ge \Delta_i \qquad (4)$$

Using Chernoff-Hoeffding inequality and the fact that $\mathbb{P}(A + B \ge 0) \le \mathbb{P}(A \ge 0) + \mathbb{P}(B \ge 0)$, we bound step 1 and 3 as:

$$\mathbb{P}\left(\left(F_i^t - F_i - \frac{\Delta_i}{2}\right) + \left(F' - F_{i'}^t - \frac{\Delta_i}{2}\right) \ge 0\right) \le 2 \cdot e^{-\frac{\Delta_i^2}{2} \cdot t}$$

For step 2 and 4, we have the following bounds:

$$\mathbb{P}\left(F_i^t - F_i \ge \Delta_i\right) \le e^{-2 \cdot \Delta_i^2 \cdot t} \le 2 \cdot e^{-\frac{\Delta_i^2}{2} \cdot t}$$

Combining the bounds for above cases, we have:

$$\mathbb{E}\left[N_i^T\right] \le \sum_{t=1}^{T} 2 \cdot e^{-\frac{\Delta_i^2}{2} \cdot t} \le \sum_{t=1}^{\infty} 2 \cdot e^{-\frac{\Delta_i^2}{2} \cdot t} = \frac{4}{\Delta_i^2} \quad \square$$

Lemma 6. $\forall i$ s.t. $p_i > p'$, expected number of times a suboptimal arm i is played is upper-bounded by $\mathbb{E}\left[N_i^T\right] \le \frac{1}{2 \cdot \Delta_i^2}$.

Proof. A suboptimal arm i is picked at time t when $V_i^t > V_{i'}^t$. We consider the following cases and the conditions that need to hold for picking arm i.

Case a) $p' \le p^*$:

$$F_{i'}^t < \frac{B}{N \cdot p_i} = F' - \Delta_i \implies F' - F_{i'}^t > \Delta_i \qquad (1)$$

Case b) $p' > p^*$:

$$F_{i'}^t < \frac{B}{N \cdot p_i} = \frac{B}{N \cdot p'} - \Delta_i \implies F' - F_{i'}^t > \Delta_i \qquad (2)$$

Using Chernoff-Hoeffding inequality, step 1 and step 2 are bounded by $e^{-2 \cdot \Delta_i^2 \cdot t}$. Combining the bounds for above cases, we have:

$$\mathbb{E}\left[N_i^T\right] \le \sum_{t=1}^{T} e^{-2 \cdot \Delta_i^2 \cdot t} \le \sum_{t=1}^{\infty} e^{-2 \cdot \Delta_i^2 \cdot t} = \frac{1}{2 \cdot \Delta_i^2} \qquad \square$$

Proof of Theorem 1. The proof directly follows from the bounds of N_i from Lemmas 5 and 6. By putting in these bounds in Lemma 1 along with the regret of discretization from Lemma 4, we get the desired results. $\square$

A.3 Regret bounds for BP-UCB

Lemma 7. $\forall i$ s.t. $p_i < p'$, expected number of times a suboptimal arm i is played is upper-bounded by

$$\mathbb{E}\left[N_i^T\right] \le \frac{8 \cdot \ln(B/c_{\min})}{\Delta_i^2} + \frac{\pi^2}{3} + 1.$$

Proof. A suboptimal arm i is picked at time t when $\widetilde{V_i^t} \le \widetilde{V_{i'}^t}$ where $\widetilde{V_j^t} = \min\left\{F_i^t + \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}}, \frac{B}{N \cdot p_j}\right\}$. We consider the following cases:

Case a) $p' \le p^*$ and $F_{i'}^t + \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}} \le \frac{B}{N \cdot p'}$:

$$F_i^t + \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}} \ge F_{i'}^t + \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}}$$

$$\left((F_i^t - F_i) - \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}}\right) + \left((F' - F_{i'}^t) - \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}}\right)$$

$$+ \left(2 \cdot \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}} - \Delta_i\right) \ge 0 \quad (1)$$

Case b) $p' \le p^*$ and $F_{i'}^t + \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}} > \frac{B}{N \cdot p'}$:

$$F_i^t + \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}} \ge \frac{B}{N \cdot p'} \ge F'$$

$$\left((F_i^t - F_i) - \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}}\right) + \left(2 \cdot \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}} - \Delta_i\right) \ge 0 \quad (2)$$

Case c) $p' > p^*$ and $F_{i'}^t + \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}} \le \frac{B}{N \cdot p'}$:

This case is analogous to **(1)** in **Case (a)** with some algebraic manipulations.

Case d) $p' > p^*$ and $F_{i'}^t + \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}} > \frac{B}{N \cdot p'}$:

This case is analogous to **(2)** in **Case (b)** with some algebraic manipulations.

Using the Chernoff-Hoeffding inequality, step 1 and step 2 are bounded by $2 \cdot t^{-4}$ and t^{-4} respectively. Once suboptimal arm i has been played sufficient number of times, given by $N_i^t \ge \left\lceil \frac{8 \cdot \ln(B/c_{\min})}{\Delta_i^2} \right\rceil$, we have:

$$\mathbb{P}\left(2 \cdot \sqrt{\frac{2 \cdot \ln(t)}{N_i^t}} - \Delta_i > 0\right) = 0$$

Combining these together, we have the following:

$$\mathbb{E}\left[N_i^T\right] \le \left\lceil \frac{8 \cdot \ln(B/c_{\min})}{\Delta_i^2} \right\rceil + \sum_{t=1}^{T} 2 \cdot t^{-4}$$

$$\le \left\lceil \frac{8 \cdot \ln(B/c_{\min})}{\Delta_i^2} \right\rceil + \sum_{t=1}^{\infty} 2 \cdot t^{-4} < \frac{8 \cdot \ln(B/c_{\min})}{\Delta_i^2} + \frac{\pi^2}{3} + 1 \square$$

Lemma 8. $\forall i$ s.t. $p_i > p'$, expected number of times a suboptimal arm i is played is upper-bounded by $\mathbb{E}\left[N_i^T\right] \le \frac{\pi^2}{6}$.

Proof. A suboptimal arm i is picked at time t when $\widetilde{V_i^t} > \widetilde{V_{i'}^t}$. Irrespective of the relation between p' and p^*, the following must hold true:

$$F_{i'}^t + \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}} < \frac{B}{N \cdot p_i} < F' \implies F' - F_{i'}^t > \sqrt{\frac{2 \cdot \ln(t)}{N_{i'}^t}}$$

Using Chernoff-Hoeffding inequality, above case is bounded by t^{-4}. We have:

$$\mathbb{E}\left[N_i^T\right] \le \sum_{t=1}^{T} t^{-4} \le \sum_{t=1}^{\infty} t^{-4} < \frac{\pi^2}{6} \qquad \square$$

Proof of Theorem 2. The proof follows by using exactly the same arguments as in Theorem 1.

$\square$

B. REFERENCES

[1] Mechanical turk platform. `https://www.mturk.com/`.

[2] A. Archer and E. Tardos. Frugal path mechanisms. In *SODA*, pages 991–999, 2002.

[3] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256, 2002.

[4] P. Auer, R. Ortner, and C. Szepesvari. Improved rates for the stochastic continuum-armed bandit problem. In *COLT*, pages 454–468, 2007.

[5] M. Babaioff, M. Dinitz, A. Gupta, N. Immorlica, and K. Tal-war. Secretary problems: weights and discounts. In *SODA*, pages 1245–1254, 2009.

[6] M. Babaioff, S. Dughmi, R. Kleinberg, and A. Slivkins. Dynamic pricing with limited supply. In *EC*, 2012.

[7] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg. A knapsack secretary problem with applications. In *APPROX-RANDOM*, pages 16–28, 2007.

[8] A. Badanidiyuru, R. Kleinberg, and Y. Singer. Learning on a budget: Posted price mechanisms for online procurement. In *EC*, 2012.

[9] Z. Bar-yossef, K. Hildrum, and F. Wu. Incentive-compatible online auctions for digital goods. In *SODA*, pages 964–970, 2002.

[10] O. Besbes and A. Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research*, 2009.

[11] A. Blum and J. Hartline. Near-optimal online auctions. In *SODA*, 2005.

[12] A. Blum, V. Kumar, A. Rudra, and F. Wu. Online learning in online auctions. In *SODA*, 2003.

[13] A. Blum and Y. Monsour. *Learning, Regret minimization, and Equilibria in Algorithmic Game Theory*. Cambridge University Press, 2007.

[14] M. C. Cary, A. D. Flaxman, J. D. Hartline, and A. R. Karlin. Auctions for structured procurement. In *SODA*, pages 304–313, 2008.

[15] N. Cesa-Bianchi, Y. Freund, D. P. Helmbold, D. Haussler, R. Schapire, , and M. Warmuth. How to use expert advice. *Journal of the ACM*, 44(2):427–485, 1997.

[16] S. Chawla, J. D. Hartline, D. L. Malec, and B. Sivan. Multi-parameter mechanism design and sequential posted pricing. In *EC*, 2010.

[17] Y. Chen and J. W. Vaughan. A new understanding of prediction markets via no-regret learning. In *STOC*, 2010.

[18] N. Devanur and J. Hartline. Limited and online supply and the bayesian foundations of prior-free mechanism design. In *EC*, 2009.

[19] A. Goldberg, J. Hartline, and A. Wright. Competitive auctions and digital goods. In *SODA*, pages 735–744, 2001.

[20] M. T. Hajiaghayi, R. D. Kleinberg, , and D. C. Parkes. Adaptive limited-supply online auctions. In *EC*, 2004.

[21] J. J. Horton and L. B. Chilton. The labor economics of paid crowdsourcing. In *EC*, pages 209–218, 2010.

[22] J. J. Horton and R. J. Zeckhauser. Algorithmic wage negotiations: Applications to paid crowsourcing. In *CrowdConf*, 2010.

[23] R. Karlin, D. Kempe, and T. Tamir. Beyond vcg: Frugality of truthful mechanisms. In *FOCS*, 2005.

[24] R. Kleinberg. Nearly tight bounds for the continuum-armed bandit problem. In *NIPS*, 2004.

[25] R. D. .Kleinberg and F. T. Leighton. The value of knowing a demand curve:bounds on regret for online posted-price auctions. In *FOCS*, 2003.

[26] R. Lavi and N. Nisan. Competitive analysis of incentive compatible on-line auctions. In *EC*, pages 233–241, 2000.

[27] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Info and Computation*, 70(2):212–261, 1994.

[28] W. Mason and D. J. Watts. Financial incentives and the performance of crowds. In *SIGKDD Explor. Newsl.*, pages 100–108, 2010.

[29] J. J. Shaw, A. D.; Horton and D. L. Chen. Designing incentives for inexpert human raters. In *CSCW*, pages 275–284, 2011.

[30] Y. Singer. Budget feasible mechanisms. In *FOCS*, pages 765–774, 2010.

[31] Y. Singer. How to win friends and influence people, truthfully: Influence maximization mechanisms for social networks. In *WSDM*, 2011.

[32] Y. Singer and M. Mittal. Pricing tasks in online labor markets. In *Workshop on Human Computation*. ACM, 2011.

[33] N. Srinivas, A. Krause, S. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *ICML*, 2010.

[34] Tran-Thanh, L., C. A., J. E. M. de Cote, A. Rogers, and N. R. Jennings. Epsilon-first policies for budget-limited multi-armed bandits. In *AAAI*, pages 1211–1216, 2010.

[35] L. Tran-Thanh, A. Chapman, A. Rogers, , and N. R. Jennings. Knapsack based optimal policies for budget-limited multi-armed bandits. `http://arxiv.org/abs/1204.1909`, 2013. Technical report.

[36] L. Tran-Thanh, A. Chapman, A. Rogers, and N. R. Jennings. Knapsack based optimal policies for budget-limited multi-armed bandits. In *AAAI*, 2012.

[37] L. Tran-Thanh, S. Stein, A. Rogers, and N. R. Jennings. Efficient crowdsourcing of unknown experts using multi-armed bandits. In *AAAI*, 2012.