Dynamic Procurement Auction with Abandonment

Ali Khodabakhsh* Evdokia Nikolova† Emmanouil Pountourakis‡ Jimmy Horn§

Abstract

We study a dynamic model of procurement auctions in which the agents (sellers) will abandon the auction if their utility does not satisfy their private target, in any given round. We call this "abandonment" and analyze its consequences on the overall cost to the mechanism designer (buyer), as it reduces competition in future rounds of the auction and drives up the price. We show that in order to maintain competition and minimize the overall cost, the mechanism designer has to adopt an inefficient (per-round) allocation, namely to assign the demand to multiple agents in a single round. We focus on threshold mechanisms as a simple way to achieve ex-post incentive compatibility, akin to reserves in revenue-maximizing forward auctions. We then consider the optimization problem of finding the optimal thresholds. We show that even though our objective function does not have the optimal substructure property in general, if the underlying distributions satisfy some regularity properties, the global optimal solution lies within a region where the optimal thresholds are monotone and can be calculated with a greedy approach, or even more simply in a parallel fashion.

*Department of Electrical & Computer Engineering, University of Texas at Austin, ali.kh@utexas.edu.
†Department of Electrical & Computer Engineering, University of Texas at Austin, nikolova@austin.utexas.edu.
‡College of Computing & Informatics, Drexel University, manolis@drexel.edu.
§Horn Wind, LLC, horn.wind@yahoo.com.

1 Introduction

The wide applicability of auctions in real life, from the simple traditional sealed-bid and ascending/descending price auctions, to the modern sponsored search and eBay auctions, to government-run auctions for spectrum and carbon emissions, has inspired the development of a rich theory of auctions and mechanism design. The more prevalent auction design focuses on the so called 'regular' auctions, where the bidders are buyers wishing to buy an item from the mechanism designer (seller), who tries to maximize her revenue (see, e.g. [15]). Less prevalent are 'reverse' or 'procurement' auctions where the bidders are sellers and the mechanism designer is a buyer wanting to minimize cost.

A principal example of procurement auctions is public procurement—the process by which governments purchase goods, services and construction—which comprises a significant fraction, 10-20%, of a country's GDP [5]. Some of the more complex procurement auctions include the above mentioned spectrum and carbon emissions auctions, as well as the procurement of energy, a key motivation of this work.

A notable feature of the above examples that makes the corresponding auction processes especially complex in reality is their repeated nature with interdependencies among the different rounds. Thus, while the large majority of literature on auction and mechanism design focuses on *static* mechanisms that optimize the designer goals with a single round in mind, there has been a recent rise in the study of *dynamic mechanism design* which attempts to model and analyze mechanisms across time [1]. In the context of procurement, different strands of literature investigate different types of interdependencies, such as caused by a capacity constraint [21, 22], a switching cost from one service provider to another [9, 17], a backlog cost in dynamic inventory control models [18, 23], learning through experience [16, 20] and piecewise procurement where the subprojects of a large project have to be procured in a predetermined order [2, 24].

Yet very simple and basic models for dynamic procurement remain unexplored that provide fertile ground for theory exploration and progress. We propose one such model which takes the most basic reverse auction of multiple sellers needing to provide a unit of divisible good or service over repeated rounds, with the condition that a seller must make at a minimum her overhead cost in order to remain present in future rounds of the auction. This provides a coupling or interdependency of the different rounds of the auction that precludes existing mechanisms from applying and calls for new tools in mechanism design.

Our motivation for this model comes from the process for energy procurement called "economic dispatch": Electricity generation is currently managed by Independent System Operators (ISO) in a myopic way (day by day). Each generator submits a supply curve, namely one or more bids of how much it is able to generate at what unit cost, for the following day. The ISO then allocates generation, based on demand (and subject to any system constraints), so as to minimize the total generation cost. Economic dispatch is thus effectively a generalized version of the standard procurement auction.

In a lot of US markets, wind is typically the least expensive form of generation, thus it is favored by the current selection mechanism over conventional generation (nuclear, coal and gas). Coal, as the least competitive conventional generation, is gradually being driven out of business due to underuse. Wind though has higher variability and uncertainty, and requires increased use of expensive back-up generation, while conventionals are reliable and do not need to be backed up. Ultimately, this is pushing the system to the two extremes of cheap, variable renewables and expensive, back-up generation. As a result, this short-term cost-minimization approach yields a higher long-term cost and compromises system reliability [14].

In reality, a less competitive generator whose economic viability is threatened might be "saved" by the ISO if it is considered critical to system reliability, by entering a side contract with the ISO that guarantees it sufficient allocation and payment to help it remain viable. Such contracts are currently done behind closed doors in an ad hoc way, including the ISO's decision which generators it considers critical.

To improve system efficiency and transparency, our model here makes a first step toward providing a framework for systematic allocation and payments that minimize cost over multiple periods. Specifically, two issues stand out from the brief background on energy procurement above. One issue is the need to capture the agents' overhead costs necessary to stay in business as a model feature to make transparent the process of identifying and saving a needed agent. We call the phenomenon of permanently leaving the auction due to not having met the overhead cost in a given round as "abandonment". The second, related issue, is the tension of cost vs competition, or short- vs long-term outlook, namely that being optimal in the current round might be suboptimal from a long-term perspective. That is because cost minimization in a given round might result in fewer agent allocations and thus reduced competition in future rounds, which would lead to a higher cost in the future. We discuss these two issues in more detail in the context of our model and results below.

Modeling choices and assumptions. Our goal is to frame the above real-life situation as a simplified auction theory model that abstracts away many engineering components, which are important but not central to the core mechanism design challenges. What are the minimal features our model can be stripped down to, that make it as simple as possible yet expressive of the two above-mentioned issues of (i) abandonment and (ii) tradeoff of cost and competition?

We focus on a two-round model with n symmetric agents (sellers), each of whom can meet the entire demand of 1 unit of divisible good/service per round, and each of whom submits a bid for her overhead cost, namely the amount she needs to make this round to "be saved" and remain in the next round of the auction. The overhead costs are private values, independent and identically distributed according to some known distribution F, across agents and across rounds.

To keep the model tractable, we assume that a per unit production cost that sellers incur for providing the good/service is known and constant (which turns out mathematically equivalent to it being zero). For example, in the energy application above, the cost of producing energy can easily be estimated by the technology; however, the overhead costs of generators (such as financing, labor costs, maintenance, etc.) are private information.

In a given round, the auctioneer or mechanism designer collects the bids and decides on the allocations and payments which in turn determine which agents are going to be saved for the following round. We will argue later that in the last day, the mechanism designer is going to allocate the entire demand to a single agent (since there is no need to maintain the competition anymore). Further, to more succinctly capture the challenges that abandonment and competition issues present, we assume that even in that final round, the mechanism should satisfy the overhead cost constraint of that single agent: this is also equivalent to removing this assumption and having one extra round, namely a 3-round auction setting.

Abandonment. In both forward and reverse auctions, when the auction is repeated over several iterations, it has been noted that the agents may leave the platform. The typical assumptions used in the literature are of dynamic arrival and departure that are exogenous and are not related to the outcome of the auction [11, 12]. In a regular auction the agent may prefer to change her auction platform if she is not receiving enough utility. Similarly, in a reverse or procurement auction, for example in the energy sector, if the generators do not meet their overhead costs they are forced to close down.

Two natural modeling choices for the utility function of an agent stand out to capture the abandonment: the utility for being allocated zero or, more generally, for being paid less than one's overhead cost could be modeled as zero or as negative infinity (or, equivalently, a large negative constant). The first choice may appear more natural on the surface but it fails to align the incentives with the phenomenon of abandonment—specifically, it fails to represent the negative repercussions of a bankruptcy in reality, which is what we are trying to model with agents abandoning the auction. Indeed, if an agent ever goes out of business, the agent should not be incentivized to stay in the auction. Furthermore, zero utility for zero allocation is inaccurate in the energy context where power plants continue having overhead expenses (such as employee salaries and power plant maintenance) even if they are not allocated and not producing in a given time period, so effectively a zero or even insufficient positive allocation implies losses which are what ultimately drives plants to retire. We thus opt for the negative infinity model, which also emphasizes the "finality" of an agent's participation in the auction if she is not allocated or has not met her overhead cost in a given round.

We remark that with this modeling choice, the utility function will not satisfy individual rationality, in that participating in the auction may have lower expected utility than not participating. Again, this is consistent with the energy and likely a number of other applications where starting a business such as building and operating a power plant entails risk and is not guaranteed to break even. We note the relation of our utility function choice to regular auctions where a buyer has a budget and receives a utility of negative infinity for exceeding it (e.g., [6, 10]). Indeed, we can view the overhead cost that needs to be met each period as a reverse budget where, once the budget is exceeded, or in our case the reverse budget is not met, the agent is forced to abandon the auction.

Competition vs cost. In a one-shot environment it is well understood that competition lowers cost and a monopoly increases it. Over multiple rounds, however, the connection of cost and competition is not as straightforward. Already in our two-round setting, we can see that in order to meet the demand while satisfying agent overhead costs, it is cheapest for the mechanism designer to pick the single lowest bid agent and allocate the entire demand to her—cheapest in the first round, that is. This can be implemented by running a Vickrey-Clarke-Groves (VCG) mechanism where we ask each agent for their overhead cost. However, due to abandonment, this would result in all but one agent abandoning the auction after the first round, and would lead to a monopoly situation in the second round whereby the designer needs to pay the maximum amount to the single remaining agent—paying the upper bound of the agent's bid distribution. Allocating to multiple agents in the first round would be a lot more costly in that round: is the higher cost worth the savings that would result from increased competition in round 2? Indeed this question highlights one of the central design challenges for our mechanism discussed in Section 4, and a key technical challenge in the resulting multi-variate optimization problem discussed in Section 5.

Bid-sensitive vs bid-oblivious mechanisms. Following the above discussion, there is a trade-off between the cost we incur on day 1, and the competition that exists on day 2. We thus expect that there is an optimal point in the trade-off between cost and competition. In other words, there should be an optimal number of agents that balances the cost required to save those agents today, and the expected cost we incur tomorrow given the competition among this surviving number of agents. Let k denote this optimal number of agents. Now we can generalize the well known truthful second-price auction as follows: after the agents submit their bids, the mechanism assigns the demand to the cheapest k agents and all those k agents receive a payment equal to the $(k+1)$st bid. Note that the higher k is, the higher the cost would be in the current round, but there will be more competition (hence less cost) in future round(s), and vice versa. Hence, the optimal k can be calculated based on the distribution F and the number of rounds left.

In fact, we can do even better than the mechanism described above, which we call a "bid-oblivious" mechanism. In a bid-oblivious mechanism, we commit to save k agents, independently of their bids. Instead, we could let the bids determine the number of agents to be saved, via using appropriate thresholds. This is similar to revenue maximization in regular auctions, in which by setting a reserve price we ensure that the item is not sold at a cheap price. If all the bids happen to be less than the reserve price, the item would not be sold. Similarly here, we can save k agents only if the lowest k bids are below a certain threshold. In this way, we can dynamically decide on the number of agents to be saved, and achieve a lower cost. We call this a "bid-sensitive" mechanism, as the bids affect the number of agents that get allocated the service. We provide an example on how bid-sensitive mechanisms can achieve a lower cost compared to bid-oblivious mechanisms in Appendix A.

Preview of our results and challenges. Our goal is to design the optimal bid-sensitive mechanisms, i.e., to find the optimal thresholds for allocating the service to various number of agents at every round. Specifically, for our two-round auction, it suffices to set the corresponding thresholds for round one, as the final round has a trivial optimal mechanism. We denote by t_i the threshold for saving i agents in round one. Our main result is to show that the global optimization for the thresholds $t_2, t_3, ...$ can be done in a greedy fashion, even though the objective function does not have the optimal substructure property that usually leads to optimality of greedy approach (see Theorem 3).

To illustrate this, consider and example with three agents in which the overhead costs are drawn from a uniform $[0, 1]$ distribution. Let $C(t_2, t_3)$ be the expected cost of both rounds, given that we set thresholds t_2, t_3 for round one (note that $t_1 = 1$, since we always have to save at least one agent). We show in Example 2 in Section 4 that the optimal thresholds in this setting are $(t_2^*, t_3^*) = (\frac{1}{6}, \frac{1}{12})$. Yet this simple example reveals a few interesting observations:

1. If the underlying distribution satisfies certain properties, the optimal thresholds are monotone, i.e., $t_2^* \geq t_3^* \geq t_4^* \geq \cdots$ (see Lemma 1 and Theorem 3 in Section 5). Even though this may seem intuitive, we show in Example 5 in Appendix D that this is not always true for general distributions.

2. We can find the optimal threshold mechanism with a greedy algorithm, meaning that we can start by optimizing over t_2 (while setting $t_3, t_4, ...$ to any arbitrary values). This gives the optimal value for t_2. Then, using the optimal value of t_2, we can optimize over t_3. However, if we do this process in any other order, it does not result in the optimal threshold values. We illustrate this in Figure 1, where the black curve shows the optimal value of t_3 for any arbitrary value of t_2. Specifically, the figure shows that if we first optimize over t_3, we get a value different than $1/12$ (unless t_2 is set to a high enough value).

3. Figure 1 also shows the derivative of the objective function with respect to t_3 for any pair (t_2, t_3). We show in Section 5 that as long as $t_2 > t_3$ (under the diagonal line), the sign of the derivative is independent of the actual value of t_2. This may suggest that as long as $t_2 > \frac{1}{12}$, the optimal t_3 should be $1/12$. However, this is not the case. In fact, for $t_2 = \frac{1}{12} + \epsilon$ we see that the optimal t_3 is even higher that t_2. The situation becomes even more chaotic if we add one more agent, as the optimal value of t_3 could depend on t_4 as well. However, we show that fortunately none of these complicating behaviors happen, once we set the earlier thresholds

to their optimal values. More precisely, once we are optimizing over t_i after setting earlier thresholds to their optimal values, the resulting value we get for t_i would be smaller than all previous values, and it will not depend on the remaining thresholds.

4. Another point worth noting from this example is that the optimal values of t_2 and t_3 would remain the same if we added one or more agents to the initial pool of participants. This is not specific to the uniform distribution, as we prove it more generally in Lemma 1. We discuss the similarity of this independence to regular auctions next.

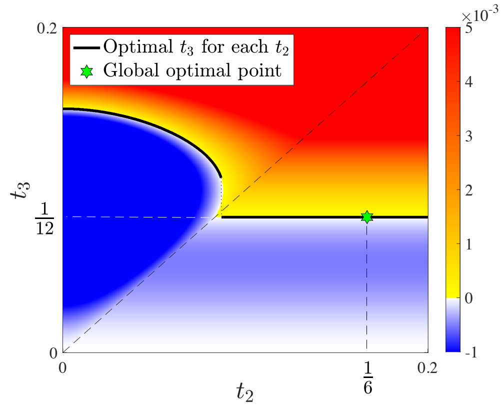


Figure 1: Derivative of the cost with respect to the threshold for saving 3 agents ($\partial C/\partial t_3$). The blue area represents the points (t_2, t_3) for which this derivative is negative, and the red/yellow area represents the positive region. The green star represents the global optimal solution of $(t_2^*, t_3^*) = (\frac{1}{6}, \frac{1}{12})$.

Similarities to Myersonian approach. Our main result is the characterization of optimal thresholds for our procurement auction setting and an algorithm to efficiently compute them. One important corollary of our analysis is that the optimal threshold mechanism is independent of the initial number of agents participating in the auction. More precisely, it turns out that the optimal threshold for saving k agents is the value where the marginal contribution to the cost of the current day (virtual cost w.r.t. F) is equal to the savings of having the k-th agent present in the future rounds. Note that the virtual cost function $(x + \frac{F(x)}{f(x)})$ is solely determined by the underlying distribution function F. Also, the marginal gain of having the k-th agent in the future round(s) does not depend on the pool of agents we start with. This is very similar to Myerson's result for revenue maximization in regular auctions: for n symmetric buyers, the optimal auction is a second-price auction with a reserve price. The reserve price is obtained by setting the virtual valuation (defined as $v - \frac{1-F(v)}{f(v)}$) to zero, and is again independent of the number of agents (n). This similarity is surprising, since in procurement auctions we have to always meet the demand and we cannot use price as a tool to trade off utility across different types. However, using thresholds for saving a different number of agents, we are able to trade off utility across different types and different rounds.

Another similarity is in how we achieve optimality. In revenue maximization, the reserve price is equivalent to *supply reduction*, meaning that depending on the bids, the seller has the right to not sell the item. Note that this means that the optimal auction is not efficient, as the seller will sometimes withhold the object even though the highest bidder has a strictly positive value. For our cost minimization problem, we achieve optimality via what is effectively a *demand increase*. The efficient outcome for each stage is to assign the service to the agent with the lowest overhead cost since every agent can satisfy the demand. However, the mechanism may allocate the production of the service to multiple agents in the hope of decreasing the future costs.[1]

Summary of results. Our results can be summarized as follows:

(a) We model a two-round dynamic procurement auction with abandonment, where the agents leave the auction if they do not meet their overhead costs in a given round. We focus on threshold mechanisms, as they are widely used in practice, and show that they are ex-post incentive compatible for our dynamic auction model. The thresholds are similar to setting reserves for revenue maximization in regular auctions.

(b) Next, we study the optimization problem for finding the optimal set of thresholds. We show that if the distribution F for overhead costs is regular (as defined later), the optimal thresholds are independent of the number of agents participating in the auction. In other words, we do not need to know the number of agents to determine the optimal set of thresholds.

(c) We prove that if the underlying distribution F satisfies certain properties, the optimal thresholds will be monotone, meaning that the optimal threshold for saving i agents is lower than the optimal threshold for saving j agents for any $i > j$. Moreover, we show that this monotonicity helps divide the optimization problem into n separate problems, which ultimately leads to an efficient algorithm to calculate the optimal thresholds in parallel.

2 Related Work

Single-parameter mechanism design has been extensively studied in theoretical computer science over the last decade and lead to several interesting results in the intersection of approximation and mechanism design (e.g. [13] and references therein). Over the last few years there has been an increased interest in dynamic mechanism design and specifically, revenue maximization in repeated auctions [3, 19]. The challenge in this line of work has been that depending on the assumptions about when the agents obtain their information, these models become multi-dimensional, leading to a notoriously hard problem in mechanism design (see [4] for a survey).

For example, Ashlagi *et al.* [3] study incentive compatible mechanisms for revenue maximization. In contrast to prior economic literature they require that the mechanism is strongly individually rational, namely the utility of each agent should be non-negative at any stage of the game. One interpretation of strong individual rationality in the context of a dynamic auction is that agents would abandon the service if they ever receive negative utility. Our model of abandonment in a procurement auction setting can be thought of as a relaxation of individual rationality, where each agent expects to achieve a specific level of utility and if she does not meet her target then she abandons the platform.

Different models of dynamic procurement auctions have been studied in the past. The common aspect between these different models is an intertemporal dependency, either on the procurer/buyer side or the suppliers/bidders, that ties the outcomes of the individual auctions. Examples of such dependencies include:

Capacity constraint: When the bidders are capacity-constrained, their costs increase if they win the current auction (due to higher future capacity utilization). Therefore, capacity-constrained

[1]Our solution keeps the total production the same, splits the allocation of the service equally and increases the payment rate accordingly. This is equivalent to increasing the total demand without inflating the payment rate.

firms face an intertemporal trade-off in sequential auctions: higher profits in the current period lead to lower profits in future periods. This model has been studied over both a finite [21] and an infinite horizon [22].

Switching cost: When a procurer buys goods from competing suppliers repeatedly over time, she may incur an additional switching cost each time she switches from one supplier to another. These costs arise because the buyer must acquire skill at using a new supplier's product, and affect the competition between the incumbent supplier and his rivals [9, 17].

Backlog/holding cost: In dynamic inventory control models, the procurer becomes a retailer who has to repeatedly run a procurement auction among a number of potential suppliers before observing the actual demand. At the end of each period, any unsatisfied demand will be backlogged with a backlog cost and any unsold inventory will be carried over to the next period with a holding cost [18, 23].

Learning through experience: In many industries learning by doing or learning through production experience enables suppliers to provide better service at lower costs. Lewis and Yildirim [16] consider such model in which the cost of each supplier at each round consists of a (public) intrinsic cost of production, which decreases every time that producer supplies the procurer, and a (private) transitory cost drawn according to a prior distribution. They study how buyer optimally manages dynamic competition among rival suppliers to exploit learning economies.

Piecewise procurement: Sometimes sequential procurement auctions belong to a large-scale project whose subprojects have to be procured in a predetermined order. The project yields its full value once it is completed. The question is then how the procurer optimally designs a procurement auction for each subproject, especially when she cannot write long-term contracts [2, 24].

In comparison to these previous models, we introduce the notion of **abandonment** to the procurement auction, meaning that the suppliers may leave the auction if their received payments do not cover their internal costs. Under this model, it is no longer true that repeating a single-round-optimal auction will lead to assigning the demand to the best set of agents at the best price [14]. To the best of our knowledge, this fundamental model has not been studied in the literature.

3 Preliminaries

There are 2 periods and a set of agents N, where $|N| = n$. Each period the mechanism designer wants to allocate a unit of production to a subset of agents. In period $j = 1, 2$, each agent i is characterized by her overhead cost M_i^j and her production cost c_i^j. We assume that the overhead costs are private and independently identically distributed according to a distribution F (independent across both agents and rounds). We will assume that F is a continuous distribution supported on $[0, 1]$.

Let x_i^j be the production percentage allocated to agent i in round j and p^j the anonymous payment rate for round j. The utility of agent i in round j is given by:

$$u_i^j(M_1^j, M_2^j, \cdots, M_n^j) = \begin{cases} x_i^j(p^j - c_i^j) & x_i^j(p^j - c_i^j) \geq M_i^j, \\ -\infty, & x_i^j(p^j - c_i^j) < M_i^j. \end{cases} \tag{1}$$

Agent i seeks to maximize her aggregate utility $u_i^1 + u_i^2$ from both rounds. The utility function is capturing the fact that if an agent does not meet her overhead cost M_i^j in round j, she goes out of business and loses everything she gained today. In addition we assume that if an agent receives $-\infty$ utility she will abandon the auction.

We further focus on the case where the individual production costs c_i^j are known to the designer and homogeneous across the agents. For simplicity all our results will assume $c_i^j = 0$ for all i and j but, as we show in Section 6, this can be generalized if they are the same for all agents in a particular round but not necessarily 0, and can vary across rounds. Hence, without loss of generality, the utility of agent i becomes:

$$u_i^j(M_1^j, M_2^j, \cdots, M_n^j) = \begin{cases} x_i^j \cdot p^j & x_i^j \cdot p^j \geq M_i^j, \\ -\infty, & x_i^j \cdot p^j < M_i^j. \end{cases} \tag{2}$$

The mechanism designer does not know the overhead costs, M_i^j, which are all identically and independently distributed according to a distribution F, i.e., $M_i^j \sim F$ independent across rounds $j = 1, 2$ and across agents $i = 1, \dots, n$.

Mechanism. Each agent reports her current overhead cost M_i^j to the designer during round j and the designer decides on the allocation $x_i^j(M_1^j, \cdots, M_n^j)$ for all $i \in N$ and the anonymous payment rate $p^j(M_1^j, \cdots, M_n^j)$. We seek to design a mechanism that minimizes the expected total cost of the outcome

$$\mathbb{E}_{M_i^j \sim F}\left[p^1(M_1^1, \dots, M_n^1) + p^2(\hat{M}_1^2, \dots, \hat{M}_n^2)\right],$$

where $\hat{M}_i^2 = M_i^2$ if $x_i^1 \cdot p^1 \geq M_i^1$ and $\hat{M}_i^2 = \infty$ otherwise.

Truthfulness. There are several generalizations of truthfulness once we depart from the standard single-shot environment. Ex-post incentive compatibility requires that agents want to report truthfully their overhead costs if this maximizes their aggregate utility even if they have access to the realization of their overhead costs in advance. For example, in our setting with two rounds, agent i should not have an incentive to report a different value than M_i^1 in round 1 despite knowing the value M_i^2. Periodic ex-post incentive compatibility relaxes this condition to agents having access to the history of the game and having only distributional assumptions for their future overhead costs. Nevertheless, the simple class of threshold mechanisms that we analyze in this paper satisfies the stronger notion of ex-post incentive compatibility. Each round j is characterized by a choice of n different thresholds $(t_1^j, t_2^j, \dots, t_n^j)$, where t_i^j represents the maximum amount that the mechanism is willing to pay to save the i-th agent in round j. This is more precisely described in the following definition.

Definition 1. *A single threshold mechanism using thresholds $t_1, \dots, t_n \in [0, 1]$ is defined as follows: Assume $M_1 < M_2 < \cdots < M_n$ and let us define the predicate $T_k(M_1, \dots, M_n) = 1$ if and only if $M_k \leq t_k$, in other words the k^{th} smallest value is less than the k^{th} threshold.*[2] *Let k be the highest index such that $T_k = 1$. Then the mechanism allocation is:*

$$x_i = \begin{cases} 1/k & \text{if } i \leq k \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

and the payment to agent i is $x_i \cdot p$, where p is the total mechanism payment (also the per unit cost of providing the demand) defined as $p = k \cdot \min\{t_k, M_{k+1}\}$. In other words, the cheapest k agents equally provide the service, while each receiving a payment of $\min\{t_k, M_{k+1}\}$.

[2] In the case of ties we need to slightly adjust the description of the mechanism. For the sake of clarity we present the more general version of the mechanism in Appendix B and prove that it is truthful. Since we assume continuous distributions, we can assume no ties for optimizing our objective, without loss of generality.

The mechanism uses the thresholds to determine the number of agents it wishes to allocate the service to. Note that allocating the service to more than one agent is inefficient. Allocating to multiple agents and respecting their overhead costs means that for every agent such that $x_i > 0$ it must be that the agent payment is at least her overhead cost, $x_i \cdot p \geq M_i$.

Proposition 1. *Any threshold mechanism is truthful*[3] *in the corresponding single-shot game and each agent that has non-zero allocation has non-negative utility.*

Proof. If an agent i is not allocated the service, she receives utility of $-\infty$. Bidding a lower overhead cost may result in her being allocated some part of the demand. There are two scenarios in which this may happen: (1) If there exists some k such that T_k is the highest true predicate both before and after agent i lowered her bid. In this case, it must be that her lower bid is less than or equal to $M_k < M_i$. This results in a payment equal to M_k, which makes her utility $-\infty$ again. (2) If T_k is not the highest true predicate after agent i lowers her bid. Assume that the new highest predicate satisfied is T_w for some $w > k$. Since T_w was not true before, it must be that the threshold t_w is now the critical value, therefore each agent receives a payment equal to t_w. But since T_w was false before, we know that $t_w < M_i$, meaning that agent i will receive $-\infty$ utility.

If agent i is allocated the service, notice that her payment is independent of her actual overhead cost. Reporting a lower overhead cost does not change her allocation nor payment. Similarly, if she reports a higher amount, she will receive the same payment, as long as she is still being allocated the service. If her increase makes her not being allocated, then her utility becomes $-\infty$. In neither case is deviating from reporting the true overhead cost profitable. □

We now define a threshold mechanism for a two-round game.

Definition 2. *A threshold mechanism for a two round game is characterized by two sets of thresholds* $\mathbf{t^1} = (t_1^1, \ldots, t_n^1)$ *and* $\mathbf{t^2} = (t_1^2, \ldots, t_n^2)$*. For any round j, we allocate the demand to at least i agents, if there are i bids below t_i^j.*

While technically the threshold mechanism defined in the second round could depend on the number of surviving agents, the optimal mechanism in the last round is oblivious to this fact; it will always allocate the service to a single agent and offer her a payment equal to the second lowest bid or the top of the distributional support if only one agent has survived. Since the mechanism is only feasible if it always allocates the entire demand, we need to have that $t_1^j = 1$ (the upper bound of the support of F) and therefore we will be omitting t_1 from now on.

Proposition 2. *A threshold mechanism for a dynamic game is ex-post incentive compatible.*

Proof. It is easy to see that for $j = 2$ (the last round), truthfulness of the threshold mechanism in the single-shot version implies that reporting the truth in the last round is optimal for each agent. For $j = 1$, we have to argue that deviating from the truth does not increase the aggregate utility for the agent. Since the mechanism is independent of the outcome of round 1, the only way that the reported overhead cost in round 1 affects the second round is if the agent is not allocated in the first round, hence has to abandon the auction. Instead, the agent could misreport a smaller overhead cost in order to ensure some allocation in round 1 so as to be considered in round 2. But in this case the aggregate utility of this agent remains $-\infty$, hence she cannot benefit from the deviation. □

[3] Assuming no ties.

As mentioned earlier, our objective in designing a threshold mechanism is to minimize the total payment of our allocation. In other words, we seek a mechanism (x, p) with thresholds $(\mathbf{t^1}, \mathbf{t^2})$ such that it minimizes the total payment. The optimal mechanism for the second round is independent of what happens during the first round and there is no reason to allocate the production of the service to more than one agent.

Proposition 3. *The optimal threshold mechanism for the second round of a two-round auction is always equal to* $t_2^2 = t_3^2 = \cdots = t_n^2 = 0$.

Proof. Setting $t_2^2 > 0$ means that with some probability two agents will be allocated the service resulting in a payment more than the second lowest bid. Note that allocating the service to the second lowest agent does not result in any benefit in the future (since this is the last round). On the contrary, setting all thresholds for round two to 0 ensures that we allocate the service to the agent with the lowest bid, and the payment would be equal to the second lowest bid. Similarly, setting any t_i^2 to a non-zero value is a sub-optimal choice. Therefore, the optimal mechanism in round 2 is to set all thresholds t_i^2 to zero for $i \geq 2$. (As always we have $t_1^2 = 1$.) □

Note that when the threshold mechanism allocates to an agent, it ensures that the payment she receive is at least her reported overhead cost so she will not abandon the auction. This is not necessarily needed for the second round according to the definition of our objective. If we allow the mechanism to allocate to an agent and not respect her overhead cost, then we could simply add an additional round. In that case, any feasible mechanism must ensure that one agent survives to the third round; therefore, the payments should satisfy her overhead cost in the second round as well. Thus our analysis exactly captures this case when we only focus on the first two rounds.

The main result of our paper is to characterize the first round optimal threshold mechanism for dynamic procurement. It is important to note the connection of our problem to revenue maximization where effectively we use a similar analysis in terms of virtual costs. An alternative way to interpret our mechanism is that it implements a form of supply increase to reduce the aggregate cost of the mechanism. Our results hold for natural assumptions on the distribution of the overhead cost defined below.

Definition 3 (Regularity [7])**.** *We say that a probability distribution f (with cumulative distribution function F) supported on* $[0, 1]$ *is regular if its virtual cost function defined as* $x + \frac{F(x)}{f(x)}$ *is monotone increasing.*

Definition 4 (Order statistics [8])**.** *Let $X_1, ..., X_n$ be a random sample of size n (independent) from a distribution F and $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ be the order statistics obtained by arranging X_i's in non-decreasing order. We denote by $\mu_{r:n}$ the expectation of the r^{th} order statistic, i.e.:*

$$\mu_{r:n} = \mathbb{E}[X_{r:n}]$$

Definition 5 (Diminishing returns of order statistics)**.** *We say that the r^{th} order statistic of a distribution F has the diminishing returns property if*

$$\mu_{r:n-1} - \mu_{r:n} \geq \mu_{r:n} - \mu_{r:n+1}, \qquad \forall n > r.$$

Our main theorem is stated below. A surprising property we find is that the mechanism does not need to know the initial number of agents that participate in any round of the auction.

Theorem 1. *If distribution F satisfies the regularity condition (Definition 3), and its second order statistic has the diminishing returns property (Definition 5), then the optimal threshold mechanism can be found in polynomial time.*

The remainder of our paper is organized as follows. In Section 4 we define the canonical threshold mechanism and provide a few examples. In Section 5 we present the main theorem of our paper proving the optimality of the canonical threshold mechanism. In Section 6 we discuss generalizations of the model where our results still hold. Finally, in Section 7 we discuss significant departures from our setting via breaking various types of homogeneity and symmetry and propose future directions.

4 Mechanism

Our objective in designing a threshold mechanism is to minimize the payment of our allocation. We will use $C_n(t_2, ..., t_n)$ to denote the aggregate cost of two rounds given a specific set of thresholds $(t_2, \dots, t_n)$ for the first round, where n is the number of agents in round 1.

Example 1 (2 agents, 2 rounds)**.** To develop intuition for the general problem, we first start with a simple example. Suppose we have 2 agents in round 1. We want to answer the following questions: "When is it beneficial to keep both agents alive for the second round? Is there a more efficient way than committing to save a particular number of agents a-priori?"

Since we have only two agents and one is always picked, we only need one threshold denoted by $t \in [0, 1]$ to determine whether or not picking the second agent is beneficial. If both bids are below t, we pick both agents and pay them t each; otherwise, we pick the cheaper agent and the payment would be equal to the higher bid. To find the optimal t, we have to calculate the expected cost as a function of t. We assume the overhead costs are drawn according to a uniform distribution supported on $[0, 1]$, i.e., $M_1, M_2 \sim U[0, 1]$. Three different cases can happen regarding the bids (in round one) as shown in Table 1.

For example, in the first case when both bids are below t, we decide to keep both agents alive; therefore, we have to pay t to each of them. There is also a cost of $\mu_{2:2} = 2/3$ which is the expected cost of the next round, given that both agents will be available and we have to pay the highest bid (out of those two uniform $[0, 1]$ bids) to the lowest bidder. On the other hand, when we keep one agent alive in cases 2 and 3, that single agent faces no competition in the next round and will bid $M_i = 1$ (which we will have to accept). Note that when we save only one agent, the payment is determined by the second bid, since we always assume $t_1 = 1$.

Multiplying the costs by their corresponding probabilities and adding up cases we get:

$$\mathbb{E}[cost] = t^2 \left(2t + \frac{2}{3}\right) + 2t(1-t)\left(\frac{1}{2} + \frac{t}{2} + 1\right) + (1-t)^2 \left(\frac{2}{3} + \frac{t}{3} + 1\right).$$

Optimizing over the threshold value t, we get $t^* = \frac{1}{6}$, which evaluates to a cost of $\frac{539}{324} \approx 1.6636$.

Example 2 (3 agents, 2 rounds)**.** In this example we want to answer the following questions: "When is it beneficial to keep all three agents alive for the second round? Also, how does the threshold for picking 2 agents change compared to the previous example with only 2 agents?"

Let t and t' be the thresholds for picking 2 and 3 agents in round one, respectively. Doing the same calculations as in the previous example, we get the following expected cost function: (we assume that $t' \leq t$, but we can also calculate the expected cost for the case of $t' > t$ and check that

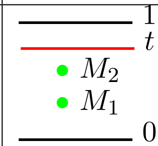
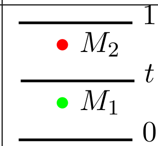
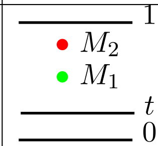
#	Case	Probability	Expected Cost	Picture
1	$M_1, M_2 \le t$	t^2	$2t + \frac{2}{3}$	
2	$M_1 \le t < M_2$ or $M_2 \le t < M_1$	$2t(1-t)$	$t + \frac{1}{2}(1-t) + 1$	
3	$M_1, M_2 > t$	$(1-t)^2$	$t + \frac{2}{3}(1-t) + 1$	

Table 1: Different scenarios of the bids of two agents (Example 1) with respect to the threshold t. Here green represents which agents survive to the next round, and red represents the payment to each of those green agents.

the optimal solution is indeed in the $t' \le t$ region.)

$$\begin{aligned}\mathbb{E}[cost] =& t'^3\left(3t' + \frac{1}{2}\right) + 3t'^2(t-t')\left(2(t' + \frac{t-t'}{2}) + \frac{2}{3}\right) + 3t'(t-t')^2\left(2(t' + \frac{2(t-t')}{3}) + \frac{2}{3}\right) \\ &+ (t-t')^3\left(2(t' + \frac{3(t-t')}{4}) + \frac{2}{3}\right) + 3t^2(1-t)\left(2t + \frac{2}{3}\right) \\ &+ 3t(1-t)^2\left(\frac{1+2t}{3} + 1\right) + (1-t)^3\left(\frac{1+t}{2} + 1\right). \end{aligned} \tag{4}$$

Optimizing over threshold values t, t' we get $(t, t') = (\frac{1}{6}, \frac{1}{12})$ which evaluates to a cost of 1.49149.

Observe that in the previous 2 examples, the threshold of saving 2 agents was $\frac{1}{6}$, regardless of whether we started with 2 or 3 agents. In addition, the expected cost (4) has the property that for any t', the optimal value of t is 1/6. Also, for any value of $t \ge \frac{1}{6}$, the cost is minimized at $t' = \frac{1}{12}$. These observations lead us to the idea that the optimal thresholds can be calculated individually, through the following notion of canonical thresholds.

Definition 6 (Canonical thresholds)**.** *The canonical threshold for saving i agents, denoted by $\hat{t}_i$, is the optimal value for t_i when all previous thresholds are set to one, and all remaining thresholds are set to zero. More precisely,*

$$\begin{aligned}\hat{t}_i \equiv \quad & \underset{t_i}{argmin} \quad C_n(t_2, ..., t_n) \\ & s.t. \qquad t_2 = \cdots = t_{i-1} = 1, \\ & \qquad\qquad t_{i+1} = \cdots = t_n = 0.\end{aligned} \tag{5}$$

In Section 5 we show that the canonical thresholds defined above are indeed optimal thresholds for minimizing the objective function $C_n(t_2, ..., t_n)$. To prepare the ground for this result, we first establish some properties of our objective function. In particular, in Theorem 2, we calculate the partial derivative of the objective function with respect to any threshold t_i. For this theorem, we have to define the following notation.

Notation. Recall that we defined the predicate $T_k(M_1, ..., M_n) = 1$ if there are at least k bids below t_k. We define the vector $\mathbf{M} = (M_1, ..., M_n)$ to be the vector of all private values and we write $T_k(\mathbf{M}) = 1$, or for short $T_k = 1$, if the k^{th} predicate is satisfied. Otherwise, we write $T_k = 0$.

Given that the first i bids are below t_i (hence predicate i is satisfied), we define $P_{i,n}$ as the probability that the remaining bids are above t_i so as not to satisfy any higher predicate $(T_{i+1}, ..., T_n)$. More precisely, we define

$$P_{i,n} = \Pr\left[M_{i+1}, ..., M_n > t_i, T_{i+1} = ... = T_n = 0 \mid M_1, ..., M_i \le t_i\right], \tag{6}$$

where we assume $P_{n,n} = 1$ (since higher bids/thresholds do not exist for this case). An important property that we use in our proofs is that by this definition, $P_{i,n}$ is independent of all lower thresholds $(t_2, ..., t_{i-1})$.

Finally, given a vector of all private values $\mathbf{M}$, we define $g(\mathbf{M}, t_2, ..., t_n)$ to be the total cost of the mechanism using thresholds $t_2, ..., t_n$. This total cost consists of a deterministic cost for the current round (since the bids are given by $\mathbf{M}$) and an expected cost for the future round(s). With our earlier notation, $C_n(t_2, ..., t_n) = \mathbb{E}_{\mathbf{M}}[g(\mathbf{M}, t_2, ..., t_n)]$. We are now ready to calculate the partial derivative of the objective function.

Theorem 2. *The derivative of the cost with respect to any threshold t_i is given by:*

$$\frac{\partial C_n(t_2, ..., t_n)}{\partial t_i} = i\binom{n}{i} P_{i,n} \times \tag{7}$$

$$\left[F(t_i)^i + F(t_i)^{i-1} f(t_i) \mathbb{E}_{\mathbf{M}}\left[i \times t_i + \mu_{2:i} - g(\mathbf{M}, t_2, ..., t_n) \mid T_i = ... = T_n = 0, t_i < M_{i:n} < t_i + \epsilon\right]\right],$$

where $\epsilon \to 0$.

Before proving (7), let us provide some intuition on different parts of this expression. Roughly speaking, the cost $C_n(t_2, ..., t_n)$ is determined by the "active" threshold, which corresponds to the highest predicate that is satisfied. As long as we do not change the active threshold, perturbing the remaining thresholds should not change the cost, therefore the derivative should be zero with respect to them. When we think of the derivative with respect to a particular t_i, we want to know how much the cost would increase/decrease if we change t_i to $t_i + \epsilon$. There are two scenarios where this perturbation changes the cost:

1. The first scenario is when there are exactly i agents below t_i. This corresponds to $\binom{n}{i} F(t_i)^i$ in (7). We also want the remaining bids to be above t_i in a way that higher predicates are not satisfied (so that t_i is active), which is captured by $P_{i,n}$. Finally in this case, when we add ϵ to t_i, all those i agents receive ϵ more payment, which corresponds to the multiplicative term i in (7).

2. The second scenario is when t_i becomes active after we add ϵ to it. This requires that the i^{th} bid is between t_i and $t_i + \epsilon$, which is why we get $i\binom{n}{i} F(t_i)^{i-1} f(t_i)$. We again need the remaining bids to be above $t_i + \epsilon$ and to not satisfy any higher predicate ($P_{i,n}$). The change in the cost is more complicated in this scenario. We know that at $t_i + \epsilon$ we are going to save i agents and therefore the cost would be roughly $i \times t_i$ for this round, and $\mu_{2:i}$ for the next round. However, it is not clear how many agents we were saving at t_i. That is why we have the expectation of the cost with negative sign, while the expectation is conditioned to this particular scenario in which there are exactly $i - 1$ agents with bids below t_i.

Proof. Using linearity of expectation and law of total probability we have:

$$\frac{\partial C_n(t_2,...,t_n)}{\partial t_i} = \frac{\partial}{\partial t_i}\mathbb{E}_{\mathbf{M}}\left[g(\mathbf{M},t_2,...,t_n)\right] = \mathbb{E}_{\mathbf{M}}\left[\frac{\partial g(\mathbf{M},t_2,...,t_n)}{\partial t_i}\right] \tag{8}$$

$$= \sum_{j\geq i}^{n} \Pr(T_j=1,T_{j+1}=...=T_n=0)\mathbb{E}_{\mathbf{M}}\left[\frac{\partial g(\mathbf{M},t_2,...,t_n)}{\partial t_i} \mid T_j=1,T_{j+1}=...=T_n=0\right] \tag{9}$$

$$+\Pr(T_i=...=T_n=0)\mathbb{E}_{\mathbf{M}}\left[\frac{\partial g(\mathbf{M},t_2,...,t_n)}{\partial t_i} \mid T_i=...=T_n=0\right] \tag{10}$$

Note that for $j \geq i+1$, the derivative (and therefore the conditional expectation) is zero because we save at least $i+1$ agents and the payment is independent of t_i. For $j=i$, if the $(i+1)^{\text{th}}$ bid (i.e., $M_{i+1:n}$) is more than t_i, then we have $g(\mathbf{M},t_2,...,t_n) = i \times t_i + \mu_{2:i}$, and hence

$$\mathbb{E}_{\mathbf{M}}\left[\frac{\partial g(\mathbf{M},t_2,...,t_n)}{\partial t_i} \mid T_i=1,T_{i+1}=...=T_n=0, M_{i+1:n}>t_i\right] = i.$$

In addition, the corresponding probability of such an event is

$$\Pr(T_i=1,T_{i+1}=...=T_n=0,M_{i+1:n}>t_i) \tag{11}$$

$$= \binom{n}{i}F(t_i)^i \Pr\left[M_{i+1},...,M_n>t_i, T_{i+1}=...=T_n=0 \mid M_1,...,M_i \leq t_i\right] = \binom{n}{i}F(t_i)^i P_{i,n} \tag{12}$$

For the case of $T_i=...=T_n=0$, we save at most $i-1$ agents and the derivative of the cost with respect to t_i is zero, unless there is a bid in $(t_i,t_i+\epsilon)$ such that increasing t_i by ϵ makes $T_i=1$. In this special case we use the fact that:

$$\frac{\partial g(\mathbf{M},t_2,...,t_n)}{\partial t_i} = \lim_{\epsilon\to 0}\frac{g(\mathbf{M},t_2,...,t_i+\epsilon,...,t_n) - g(\mathbf{M},t_2,...,t_i,...,t_n)}{\epsilon}.$$

Note that this derivative would be infinity as we have a sudden increase in g, however the probability of this event is proportional to ϵ which makes a finite product. In particular, we have:

$$\Pr(T_i=...=T_n=0,t_i<M_{i:n}<t_i+\epsilon) \tag{13}$$

$$= n\binom{n-1}{i-1}F(t_i)^{i-1}\epsilon f(t_i)\Pr\left[M_{i+1},...,M_n>t_i+\epsilon, T_{i+1}=...=T_n=0 \mid M_1,...,M_i\leq t_i+\epsilon\right]$$

$$= n\binom{n-1}{i-1}F(t_i)^{i-1}\epsilon f(t_i)P_{i,n}$$

In this case, $g(\mathbf{M},t_2,...,t_i+\epsilon,...,t_n) = i\times(t_i+\epsilon)+\mu_{2:i}$. Even though the cost $g(\mathbf{M},t_2,...,t_i,...,t_n)$ depends on the thresholds $t_2,...,t_i$ and the realization of the bids, its expectation is independent of the number of agents n, since we have already conditioned on the fact that exactly $i-1$ bids are below t_i (and the assignment and payments only depend on those $i-1$ bids). Putting the results back into the original summation (10), we get:

$$\frac{\partial C_n(t_2,...,t_n)}{\partial t_i} = i\times\binom{n}{i}F(t_i)^i P_{i,n} + \tag{14}$$

$$n\binom{n-1}{i-1}F(t_i)^{i-1}f(t_i)P_{i,n}\mathbb{E}_{\mathbf{M}}\left[i\times t_i+\mu_{2:i}-g(\mathbf{M},t_2,...,t_n) \mid T_i=...=T_n=0,t_i<M_{i:n}<t_i+\epsilon\right]$$

Using the identity $i\binom{n}{i} = n\binom{n-1}{i-1}$ we get:

$$\frac{\partial C_n(t_2, ..., t_n)}{\partial t_i} = i\binom{n}{i} P_{i,n} \times \tag{15}$$
$$\Big[F(t_i)^i + F(t_i)^{i-1} f(t_i) \mathbb{E}_{\mathbf{M}}\big[i \times t_i + \mu_{2:i} - g(\mathbf{M}, t_2, ..., t_n) \mid T_i = ... = T_n = 0, t_i < M_{i:n} < t_i + \epsilon\big]\Big],$$

which completes the proof. □

5 Optimality of Canonical Thresholds

In this section, we study the optimality of the canonical thresholds. We first begin by showing that canonical thresholds form a monotone decreasing sequence. While this property seems intuitive, it is not necessarily true if the underlying distribution F does not satisfy our two assumptions of regularity and diminishing returns property of the second-order statistic, discussed in Section 3. See Appendix D for an example of an irregular distribution F, for which the optimal thresholds are non-monotone.

Lemma 1. *For a regular distribution F that its second order statistic has the diminishing returns property (Definition 5), the canonical thresholds are monotone non-increasing and independent of the number of agents n.*

Proof. By definition, $\hat{t}_i$ is the optimal value for t_i when $t_2 = \cdots = t_{i-1} = 1$ and $t_{i+1} = \cdots = t_n = 0$. To find the optimal t_i, we start from the general expression (7) for the derivative and show that it simplifies as follows whenever $t_i \le t_{i-1}$ (which is true here since $t_{i-1} = 1$). When the i^{th} bid is between t_i and $t_i + \epsilon$, $T_i = ... = T_n = 0$, and $t_i \le t_{i-1}$, we would save $i-1$ agents and therefore $g(\mathbf{M}, t_2, ..., t_n) = (i-1)t_i + \mu_{2:i-1}$. This simplifies equation (7) to:[4]

$$\begin{aligned}
\frac{\partial C_n(t_2, ..., t_n)}{\partial t_i} &= i\binom{n}{i} P_{i,n} \times \Big[F(t_i)^i + F(t_i)^{i-1} f(t_i)\big[t_i + \mu_{2:i} - \mu_{2:i-1}\big]\Big] \\
&= i\binom{n}{i} P_{i,n} F(t_i)^{i-1} f(t_i) \left[t_i + \frac{F(t_i)}{f(t_i)} + \mu_{2:i} - \mu_{2:i-1}\right], \quad \forall t_i \le t_{i-1}.
\end{aligned} \tag{16}$$

Other than the trivial roots $t_i = 0$ and $t_i = 1$ (which are local maximizers), there is a single root for this derivative that determines $\hat{t}_i$ as follows:

$$\hat{t}_i + \frac{F(\hat{t}_i)}{f(\hat{t}_i)} = \mu_{2:i-1} - \mu_{2:i}. \tag{17}$$

Therefore, we have $\hat{t}_i \ge \hat{t}_j$ for all $i \le j$, since the left-hand side is a monotone increasing function and the right-hand side is a constant, monotone non-increasing in i. Also note that (17) makes $\hat{t}_i$ independent of n (as long as $n \ge i$). This concludes the proof. □

Now, we prove that the canonical thresholds provide the global optimal solution for minimizing the expected cost of the auction. In Lemma 2 we show that when the previous thresholds are set to 1, as we increase threshold t_i from zero to its canonical value $\hat{t}_i$, the expected cost $C_n(t_2, ..., t_n)$ decreases; and as we increase t_i beyond $\hat{t}_i$, the cost increases again.

[4] For consistency of notation, we define $\mu_{2:1} = 1$. This is because when we save i agents in round one, the expected cost of the second round would be $\mu_{2:i}$ for $i \ge 2$, and 1 if $i = 1$.

Lemma 2. *If $t_k = 1$ for all $k \le i-1$, then $\frac{\partial}{\partial t_i} C(t_2, ..., t_n)$ is non-positive for $t_i \in (0, \hat{t}_i)$, zero at $t_i = \hat{t}_i$, and non-negative for $t_i \in (\hat{t}_i, 1)$.*[5]

Proof. Since $t_{i-1} = 1$, we can again use the simplified version of the derivative (16) instead of the general version (7) for all t_i. Since $P_{i,n}$, $F(t_i)$, and $f(t_i)$ are all non-negative, we have to show that $t_i + \frac{F(t_i)}{f(t_i)} + \mu_{2:i} - \mu_{2:i-1}$ is non-positive for $t_i \in (0, \hat{t}_i)$, zero at $t_i = \hat{t}_i$, and non-negative for $t_i \in (\hat{t}_i, 1)$. Assuming that the virtual cost is monotone non-decreasing, it suffices to show that $t_i + \frac{F(t_i)}{f(t_i)} + \mu_{2:i} - \mu_{2:i-1} = 0$ at $t_i = \hat{t}_i$, which is true due to (17). (Note that $t_i + \frac{F(t_i)}{f(t_i)} + \mu_{2:i} - \mu_{2:i-1}$ is strictly negative/positive at 0/1, therefore $\hat{t}_i$ is a fractional point.) □

The previous lemma shows that the canonical threshold $\hat{t}_i$ is the global minimizer of the cost when $t_2 = ... = t_{i-1} = 1$, independent of the values of the remaining thresholds $t_{i+1}, ..., t_n$. However, in the following lemma and its corollary, we show that this holds even if we lower the value of the previous thresholds from 1 to their canonical values.

Lemma 3. *If $t_k = \hat{t}_k$ for all $k \le i-1$, then $\frac{\partial}{\partial t_i} C(t_2, ..., t_n)$ is non-positive for $t_i \in (0, \hat{t}_i)$, zero at $t_i = \hat{t}_i$, and non-negative for $t_i \in (\hat{t}_i, 1)$.*

Proof. Note that compared to the previous lemma, we only lowered the value of $t_2, ..., t_{i-1}$ from 1 to their canonical value $t_k = \hat{t}_k$. One can argue from (7) that this lowering of thresholds does not change the derivative for any $t_i \in [0, \hat{t}_{i-1}]$. This is true because we can use equation (16) in this region, which shows that the derivative is independent of $t_2, ..., t_{i-1}$, whenever $t_i \le t_{i-1}$ (remember that $P_{i,n}$ is independent of $t_2, ..., t_{i-1}$). Figure 2 shows an example of how lowering the thresholds $t_2, ..., t_{i-1}$ affects the derivative with respect to t_i.

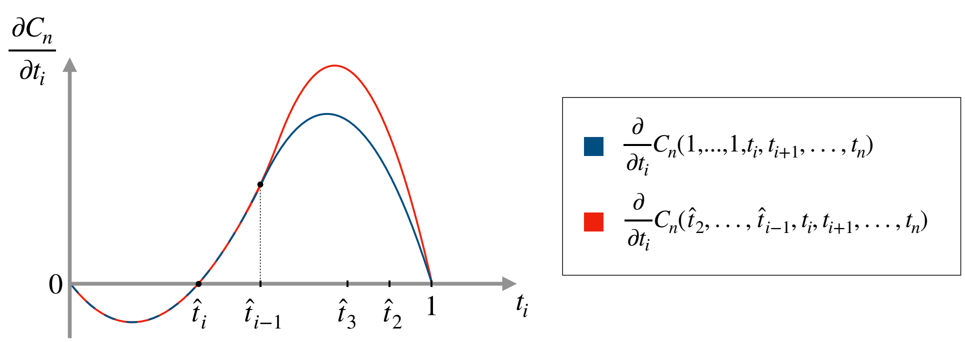


Figure 2: Derivative of the cost with respect to t_i when: (blue) the previous thresholds are set to one, (red) the previous thresholds are lowered to their canonical values.

Note that from Lemma 1 we know that $\hat{t}_i \le \hat{t}_{i-1}$. This immediately implies that $\frac{\partial}{\partial t_i} C(t_2, ..., t_n)$ is non-positive for $t_i \in (0, \hat{t}_i)$, zero at $t_i = \hat{t}_i$, and non-negative for $t_i \in (\hat{t}_i, \hat{t}_{i-1})$. Therefore, we only need to show that the derivative is non-negative for $t_i \ge \hat{t}_{i-1}$. To do this, we show that the lowering of thresholds $t_2, ..., t_{i-1}$ indeed increases the derivative in this region, i.e.,

$$\frac{\partial}{\partial t_i} C(\hat{t}_2, ..., \hat{t}_{i-1}, t_i, t_{i+1}, ..., t_n) \ge \frac{\partial}{\partial t_i} C(1, ..., 1, t_i, t_{i+1}, ..., t_n), \quad \forall t_i \ge \hat{t}_{i-1} \tag{18}$$

[5]Note that whenever $t_{i-1} = 1$, the previous thresholds $t_2, ..., t_{i-2}$ are irrelevant. Therefore, this lemma holds even if we only had $t_{i-1} = 1$. However, we state the lemma as is for the sake of the next lemma.

which implies the non-negativity of the derivative, since the right hand side is non-negative due to Lemma 2. To prove (18), note that from (7), comparing the above two derivatives is equivalent to showing that

$$\mathbb{E}_{\mathbf{M}}\big[g(\mathbf{M}, \hat{t}_2, ..., \hat{t}_{i-1}, t_i, t_{i+1}, ..., t_n) \mid A\big] \leq \mathbb{E}_{\mathbf{M}}\big[g(\mathbf{M}, 1, ..., 1, t_i, t_{i+1}, ..., t_n) \mid A\big],$$

where A is the event that there are exactly $i-1$ bids below t_i and we save at most those $i-1$ agents. Note that this expected cost is exactly equal to the situation if we had only $i-1$ agents in the auction, and we knew that their bids are upper bounded by t_i. In other words, it suffices to show that

$$\tilde{C}_{i-1}(\hat{t}_2, ..., \hat{t}_{i-1}) \leq \tilde{C}_{i-1}(1, ..., 1), \tag{19}$$

where $\tilde{C}_{i-1}$ is the expected cost in a game with $i-1$ agents with distribution $\tilde{F}$ which is obtained from truncating F to have the support $[0, t_i]$ (note that distribution $\tilde{F}$ only applies to the first day, and on day 2 the bids are again drawn according to the original distribution F).

To show (19), we use the following set of inequalities:

$$\begin{aligned}
\tilde{C}_{i-1}(\hat{t}_2, \hat{t}_3, \hat{t}_4, ..., \hat{t}_{i-1}) &\leq \tilde{C}_{i-1}(1, \hat{t}_3, \hat{t}_4, ..., \hat{t}_{i-1}) \\
\tilde{C}_{i-1}(1, \hat{t}_3, \hat{t}_4, ..., \hat{t}_{i-1}) &\leq \tilde{C}_{i-1}(1, 1, \hat{t}_4, ..., \hat{t}_{i-1}) \\
&\vdots \\
\tilde{C}_{i-1}(1, 1, ..., 1, \hat{t}_{i-1}) &\leq \tilde{C}_{i-1}(1, 1, ..., 1, 1)
\end{aligned}$$

Each of the above inequalities is implied by Lemma 2, since this lemma says that the derivative with respect to any t_k is non-negative for $t_k \geq \hat{t}_k$, as long as the previous thresholds are all equal to one. Therefore, increasing any t_k from $\hat{t}_k$ to 1 cannot decrease the cost. The only concern here is that the thresholds $\hat{t}_k$ were calculated for the auction with n agents and distribution F, while we are using the same thresholds here for the auction with $i-1$ agents and truncated distribution $\tilde{F}$. The reason why we are allowed to do this is that neither changing the number of agents nor truncating the distribution can affect the optimality of $\hat{t}_k$. This is because $\hat{t}_k$ is the solution of the following equation:

$$t_k + \frac{F(t_k)}{f(t_k)} + \mu_{2:k} - \mu_{2:k-1} = 0$$

In addition to being independent of n, this equation is invariant to conditioning F from above. In other words, for the truncated distribution $\tilde{F}$ we have:

$$\tilde{F}(t) = \frac{F(t)}{F(t_i)}, \qquad \tilde{f}(t) = \frac{f(t)}{F(t_i)}, \quad \forall t \leq t_i$$

Hence $\frac{\tilde{F}(t)}{\tilde{f}(t)} = \frac{F(t)}{f(t)}$, which implies having $t_k = \hat{t}_k$ (for $k = 2, ..., i-1$) gives a lower cost compared to $t_k = 1$, regardless of having distribution F or $\tilde{F}$.[6] $\square$

Since we showed that the derivative (with respect to t_i) is non-positive up to $\hat{t}_i$ and non-negative afterwards, we arrive at the optimality of $\hat{t}_i$.

[6] This is similar to revenue maximization where if we condition F to be above a certain value v and obtain the conditional distribution $\tilde{F}$, we have that $1 - \tilde{F}(x) = \frac{1-F(x)}{1-F(v)}$ and $\tilde{f}(x) = \frac{f(x)}{1-F(v)}$. This implies that the inverse hazard rate and as a result the virtual value functions of these distributions remain the same.

Corollary 1. *If $t_k = \hat{t}_k$ for all $k \leq i-1$, then $C_n(t_2, ..., t_n)$ is minimized at $t_i = \hat{t}_i$, independent of the values of the remaining thresholds $t_{i+1}, ..., t_n$.*

So far we showed that as long as the previous thresholds are set to their canonical values, $\hat{t}_i$ is the global optimal value for t_i. To achieve the global optimal values for the entire set of thresholds $(t_k, k = 2, ..., n)$ it suffices to use the previous lemma in an inductive manner.

Theorem 3. *If distribution F satisfies the regularity condition (Definition 3), and its second order statistic has the diminishing returns property (Definition 5), then the global optimal thresholds that minimize $C_n(t_2, ..., t_n)$ are*

$$t_k^* = \hat{t}_k, \quad \forall k.$$

Proof. Let us assume that this is not true and there exists another set of thresholds $(t'_2, ..., t'_n)$ with cost smaller than $C_n(\hat{t}_2, ..., \hat{t}_n)$. Looking at t_2, Corollary 1 can be used without any condition on the remaining thresholds, which immediately implies that either $t'_2 = \hat{t}_2$, or we can change it to $\hat{t}_2$ without increasing the cost. Given $t'_2 = \hat{t}_2$, we can now use this argument again for t_3 and conclude that $t'_3 = \hat{t}_3$. Repeating this argument, we arrive at $C_n(t'_2, ..., t'_n) = C_n(\hat{t}_2, ..., \hat{t}_n)$, which contradicts our starting assumption. □

6 Extensions

Our results extend to several generalizations, as long as the agents remain homogeneous. In particular, the following extensions hold individually or in combination with the others:

Symmetric capacities. Throughout this paper we assumed that each agent is able to meet the entire demand, or equivalently there are no capacities on the agents. However, in practice we can have the constraint that the assignments should satisfy $x_i \leq \overline{x}_i$, where $\overline{x}_i$ is the capacity of agent i. We argue that our results hold if the agents have the same capacity, i.e., $\overline{x}_i = \overline{x}$ for all i. To see this, let $m = \lceil 1/\overline{x} \rceil$ be the minimum number of agents that we need to keep in all rounds. To guarantee that we meet the demand on day 2, we have to set $t_2 = ... = t_m = 1$ in day 1. For the remaining thresholds t_k ($k > m$), note that if we save k agents from day 1 to day 2, the expected cost on day 2 would be $m \times \mu_{m+1:k}$. Therefore, similarly to (17), the optimal value of t_k is where the virtual cost matches the savings of having k agents versus having $k-1$ agents, i.e., $t_k + F(t_k)/f(t_k) = m(\mu_{m+1:k-1} - \mu_{m+1:k})$. Note that in this case we need the $(m+1)$-st order statistics to satisfy the diminishing returns property of Definition 5.

Changing distributions. Note that the analysis of our threshold algorithm did not use the fact that the distribution in each round was the same. Let F_1 and F_2 be the distributions for the overhead costs for day 1 and 2 respectively. We define the canonical threshold to satisfy

$$\hat{t}_i + \frac{F_1(\hat{t}_i)}{f_1(\hat{t}_i)} = \mu_{2:i-1} - \mu_{2:i}. \tag{20}$$

where μ is defined according to the order statistics of F_2. As long as the regularity condition holds for F_1, and F_2 has the diminishing returns property of the second order statistic, our theorem still holds since our proof only requires that $\mu_{2:i-1} - \mu_{2:i}$ is a decreasing function of i.

Non-zero service costs. Assume that the agents incur a cost of c^1 and c^2 for providing a unit of demand on day 1 and 2, respectively (but still have the same cost as other agents each round). In this case, the definition of a threshold mechanism changes slightly compared to Definition 1. After finding the highest predicate T_k that is satisfied, the first k agents with lowest bids are allocated an assignment of $x_i = 1/k$, while receiving a payment of $p \cdot x_i$, except that the payment now increases to $p = k \cdot \min\{t_k, M_{k+1}\} + c$ (where c is the service cost of the corresponding day). This is required to ensure that for any agent with strictly positive allocation we have $x_i(p - c_i) \geq M_i$.

Now we argue that the canonical thresholds are still optimal for the case of non-zero costs. This is because any feasible allocation will result in an additional cost of $\sum_i x_i \cdot c = c$ to the mechanism designer. As a result, the optimal threshold mechanism for non-zero service cost ($c > 0$) corresponds to designing the optimal mechanism for $c = 0$, and adding the cost c to the payment rate p so as to ensure that the utilities of the agents are the same. This shows that our assumption of $c_i^j = 0$ (for all i, j) throughout the paper was without loss of generality.

7 Conclusion and Future Work

In this paper we studied a dynamic procurement auction for n symmetric agents. We assumed 3 different properties for the agents that were crucial to achieve the optimality of the canonical thresholds: (i) we assumed a common distribution F for the overhead costs, (ii) we assumed that the per-unit cost of providing the service is the same for all agents, and (iii) we assumed that each agent can provide the entire demand. Relaxing any of these assumptions breaks the symmetry of the agents and opens a new research question for future work.

If we assume a different distribution F_i for each agent's overhead cost, then the savings from allocating the service to k agents and having them participate in the future rounds depend on the identity of those agents. This could potentially lead to having a different threshold for any subset of agents, which would make the problem computationally intractable. On the other hand, if we assume that agents have different per-unit costs, the optimal assignment would not be trivial, even if the set of agents with non-zero assignments are known. In other words, if we want to save a particular set of k agents, the optimal assignment is not necessarily $1/k$, and it depends on the per-unit costs of those particular k agents. The same challenge holds when we consider different capacities for the agents, as equal assignments of $1/k$ may not even be feasible in that setting.

References

[1] Shipra Agrawal, Constantinos Daskalakis, Vahab Mirrokni, and Balasubramanian Sivan. Robust repeated auctions under heterogeneous buyer behavior. In *19th ACM conference on Economics and Computation*, 2018.

[2] James J Anton and Dennis A Yao. Second sourcing and the experience curve: price competition in defense procurement. *The RAND Journal of Economics*, pages 57–76, 1987.

[3] Itai Ashlagi, Constantinos Daskalakis, and Nima Haghpanah. Sequential mechanisms with ex-post participation guarantees. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 213–214, 2016.

[4] Dirk Bergemann and Juuso Välimäki. Dynamic mechanism design: An introduction. *Journal of Economic Literature*, 57(2):235–74, 2019.

[5] Erica Bosio and Simeon Djankov. How large is public procurement?, 2020. `https://blogs.worldbank.org/developmenttalk/how-large-public-procurement`, accessed June 5, 2020.

[6] Shuchi Chawla, David L Malec, and Azarakhsh Malekian. Bayesian mechanism design for budget-constrained agents. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 253–262, 2011.

[7] Fangruo Chen. Auctioning supply contracts. *Management Science*, 53(10):1562–1576, 2007.

[8] Herbert Aron David and Haikady Navada Nagaraja. *Order statistics*. John Wiley Sons, 2004.

[9] Joseph Farrell and Paul Klemperer. Coordination and lock-in: Competition with switching costs and network effects. *Handbook of industrial organization*, 3:1967–2072, 2007.

[10] Jon Feldman, Shanmugavelayutham Muthukrishnan, Evdokia Nikolova, and Martin Pál. A truthful mechanism for offline ad slot scheduling. In *International Symposium on Algorithmic Game Theory*, pages 182–193. Springer, 2008.

[11] Daniel F. Garrett. Intertemporal price discrimination: Dynamic arrivals and changing values. *American Economic Review*, 106(11):3275–99, November 2016.

[12] Daniel F. Garrett. Dynamic mechanism design: Dynamic arrivals and changing values. *Games and Economic Behavior*, 104:595 – 612, 2017.

[13] Jason D Hartline. Mechanism design and approximation. 2011.

[14] Jimmy Horn, Yutong Wu, Ali Khodabakhsh, Evdokia Nikolova, and Emmanouil Pountourakis. The long-term cost of energy generation. In *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*, pages 74–85, 2020.

[15] Vijay Krishna. *Auction theory*. Academic press, 2009.

[16] Tracy R Lewis and Huseyin Yildirim. Managing dynamic competition. *American Economic Review*, 92(4):779–797, 2002.

[17] Tracy R Lewis and Huseyin Yildirim. Managing switching costs in multiperiod procurements with strategic buyers. *International Economic Review*, 46(4):1233–1269, 2005.

[18] Shuren Liu, Changgeng Liu, and Qiying Hu. Optimal procurement strategies by reverse auctions with stochastic demand. *Economic Modelling*, 35:430–435, 2013.

[19] Christos H. Papadimitriou, George Pierrakos, Christos-Alexandros Psomas, and Aviad Rubinstein. On the complexity of dynamic mechanism design. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1458–1475, 2016.

[20] Alessandro Pavan, Ilya Segal, and Juuso Toikka. Dynamic mechanism design: A myersonian approach. *Econometrica*, 82(2):601–653, 2014.

[21] Viplav Saini. Reserve prices in a dynamic auction when bidders are capacity-constrained. *Economics Letters*, 108(3):303–306, 2010.

[22] Viplav Saini. Endogenous asymmetry in a dynamic procurement auction. *The RAND Journal of Economics*, 43(4):726–760, 2012.

[23] Garrett Van Ryzin and Gustavo Vulcano. Optimal auctioning and ordering in an infinite horizon inventory-pricing system. *Operations Research*, 52(3):346–367, 2004.

[24] Huseyin Yildirim. Piecewise procurement of a large-scale project. *International Journal of Industrial Organization*, 22(8-9):1349–1375, 2004.

A Bid-oblivious vs Bid-sensitive Mechanisms

In this section, we show an example of how bid-sensitive mechanisms can achieve a lower cost compared to bid-oblivious mechanisms. Remember that our threshold mechanisms determine the number of agents that serve the demand in a dynamic way, after observing the bids. On the other hand, bid-oblivious mechanisms determine this number a-priori, before the bids are submitted.

Example 3 (Bid-oblivious mechanism)**.** Consider 2 rounds and n agents with overhead costs drawn from the uniform distribution, $M_i \sim U[0,1]$. We only have to decide on the number of agents to be saved from round 1 to round 2. If we save 1 agent, the expected cost would be $\frac{2}{n+1}$ for round 1, which is the expected value of the second lowest bid out of n i.i.d. uniform bids in $[0,1]$. However, in that case, there will be no competition in round 2, and we have to pay the upper-bound of the distribution (which is 1) to the single agent left in round 2. On the other hand, if we saved 2 agents in round 1, the expected cost for that round would be $2 \times \frac{3}{n+1}$, because each of the agents gets the third lowest bid. We also have to pay $2/3$ in expectation in round 2. Comparing the two cases ($\frac{2}{n+1} + 1$ versus $2 \times \frac{3}{n+1} + \frac{2}{3}$), we conclude that it is beneficial to save 2 agents if there are more than $n = 11$ agents initially. Similarly, we can compute the number of agents that justify saving 3,4,... agents.

Example 4 (Bid-oblivious vs bid-sensitive)**.** Consider 2 rounds and 2 agents with uniform $[0,1]$ distribution for overhead costs. According to the previous example, the bid-oblivious mechanism would save only 1 agent in the first round, and therefore pay a total cost of $\frac{2}{3} + 1$ in expectation. However, by setting a threshold of $t_2 = \frac{1}{6}$ (as calculated in Example 1), the expected cost reduces to $\frac{539}{324} \approx 1.663$.

B Dealing with Ties

Allowing for ties in the bids introduces some slight technicalities which we present in the updated definition and proof of truthfulness of our mechanism below.

Definition 7. *A single threshold mechanism using thresholds $t_1, \ldots, t_n \in [0,1]$ is defined as follows: Assume $M_1 \le M_2 \le \cdots \le M_n$ and let us define the predicate $T_k(M_1, \ldots, M_n) = 1$ if and only if $M_k \le t_k$, in other words the k^{th} smallest value is less than the k^{th} threshold. Let k be the highest index such that $T_k = 1$ and let ℓ be the largest index such that $M_k = M_\ell$. Then the mechanism allocation is:*

$$x_i = \begin{cases} 1/\ell & \text{if } i \le \ell, \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

and the payment to agent i is $x_i \cdot p$, where p is the total mechanism payment (also the per unit cost of providing the demand) defined as $p = \ell \cdot \min\{t_k, M_{k+1}\}$.

Proposition 4. *Any threshold mechanism is truthful in the corresponding single-shot game and each agent that has non-zero allocation has non-negative utility.*

Proof. If an agent i is not allocated the service, she receives utility of $-\infty$. Bidding a lower overhead cost may result in her being allocated some part of the demand. There are two scenarios in which this may happen: (1) If there exists some k such that T_k is the highest true predicate both before and after agent i lowered her bid. In this case, it must be that her lower bid is less than or equal to M_k, which is strictly less than M_i since if $M_i = M_k$ then agent i should have been allocated. This results in a payment equal to M_k, which makes her utility $-\infty$ again. (2) If T_k is not the highest true predicate after agent i lowers her bid. Assume that the new highest predicate satisfied is T_w for some $w > k$. Since T_w was not true before, it must be that the threshold t_w is now the critical value, therefore each agent receives payment equal to t_w. But since T_w was false before, we know that $t_w < M_i$, meaning that agent i will receive $-\infty$ utility.

If agent i is allocated the service, notice that her payment is independent of her actual overhead cost. Reporting a lower overhead cost does not change her allocation nor payment. Similarly, if she reports a higher amount, she will receive the same payment, as long as she is still being allocated the service. If her increase makes her not being allocated, then her utility becomes $-\infty$. In neither case is deviating from reporting the true overhead cost profitable. $\square$

C Distributional Assumptions

Throughout the paper, we assumed two important properties for the underlying distribution F: (i) we assumed that the virtual cost is monotone increasing, and (ii) we assumed that the second order statistics have the diminishing returns property, $\mu_{2:n-1} - \mu_{2:n} \geq \mu_{2:n} - \mu_{2:n+1}$. One sufficient condition for (i) is that F is log-concave. For (ii), we can show that: (see for example, [8])

$$\mu_{r:n-1} - \mu_{r:n} = \frac{r}{n}\binom{n}{r}\int_0^1 [F(x)]^r[1-F(x)]^{n-r}dx. \tag{22}$$

Therefore, we have to show that the above expression is monotone decreasing in n, for $r = 2$. Since the above integral is hard to compute for arbitrary distributions F, here we show numerically that many distributions satisfy this property. We also compute the above integral for polynomial distributions and show theoretically that they satisfy both of our assumptions.

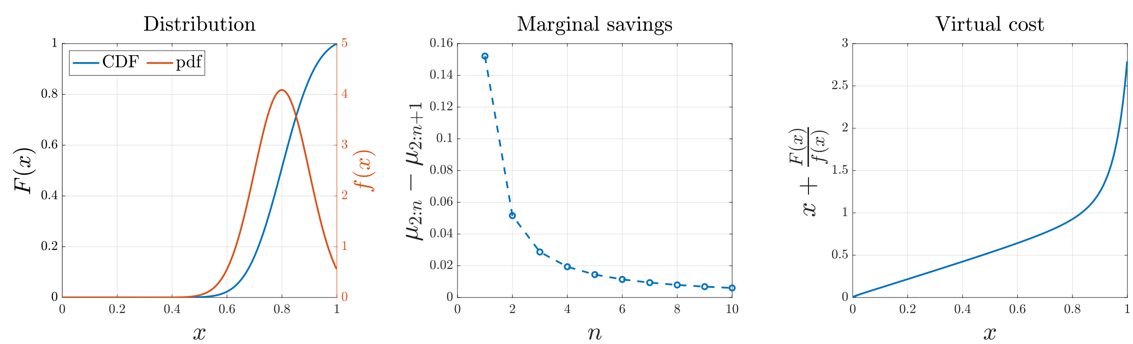


Figure 3: Regularity of a truncated normal distribution with $\mu = 0.8$ and $\sigma = 0.1$: (Left) Distribution F, (Middle) Marginals of second order statistic, (Right) Virtual cost function.

Figure 3-(left) shows a truncated normal distribution with $\mu = 0.8$ and $\sigma = 0.1$ (i.e., this is a probability distribution derived from a normal distribution $\mathcal{N}(\mu, \sigma^2)$ by bounding the random variable between 0 and 1). Figure 3-(middle) shows the marginal decrease in the second order statistic as we increase the number of samples from n to $n+1$, i.e., $\mu_{2:n} - \mu_{n+1}$. As we can see, the

marginal change diminishes as we have more samples. Finally, Figure 3-(right) shows the virtual cost which confirms its monotonicity.

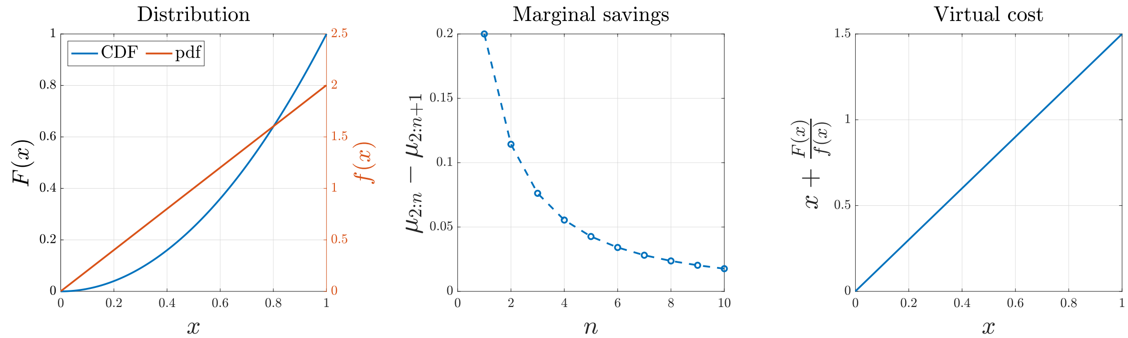


Figure 4: Regularity of the polynomial distribution with exponent $p = 2$: (Left) Distribution $F(x) = x^2$, (Middle) Marginals of second order statistic, (Right) Virtual cost function.

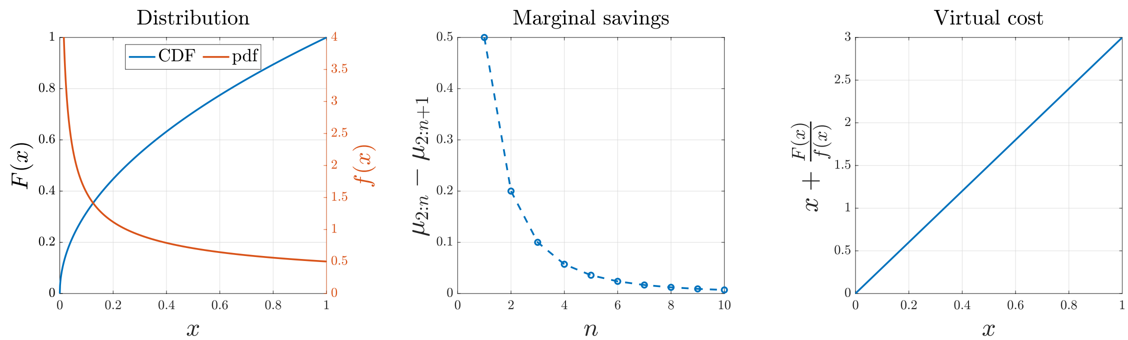


Figure 5: Regularity of the polynomial distribution with exponent $p = 1/2$: (Left) Distribution $F(x) = \sqrt{x}$, (Middle) Marginals of second order statistic, (Right) Virtual cost function.

Figures 4, 5 confirm our distributional assumptions for a polynomial distribution $F(x) = x^p$, where $p > 0$. Figure 4 corresponds to the case of $p = 2$, meaning that the CDF is a quadratic function, $F(x) = x^2$; and Figure 5 corresponds to $p = 1/2$, meaning that $F(x) = \sqrt{x}$. Note that for polynomial distributions, since $f(x) = px^{p-1}$, we have $\frac{F(x)}{f(x)} = x/p$, which is why the virtual costs are linear in both figures.

Finally, Figure 6 shows a beta distribution for which we have $f(x) = x^{\alpha-1}(1-x)^{\beta-1}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$, where Γ is the Gamma function. For this figure we have the shape parameters set to $\alpha = 2$ and $\beta = 5$.

We now prove that polynomial distributions satisfy the distributional assumptions needed for our results.

Theorem 4 (polynomial distributions). *For any distribution $F(x) = x^p, x \in [0,1], p > 0$ we have:*

(a) $x + \frac{F(x)}{f(x)}$ is monotone increasing.

(b) $\mu_{r:n-1} - \mu_{r:n}$ is monotone decreasing in n for any r.

Proof. First, $x + \frac{F(x)}{f(x)} = x + \frac{x^p}{px^{p-1}} = x(1 + \frac{1}{p})$ which is monotone increasing. For the second part,

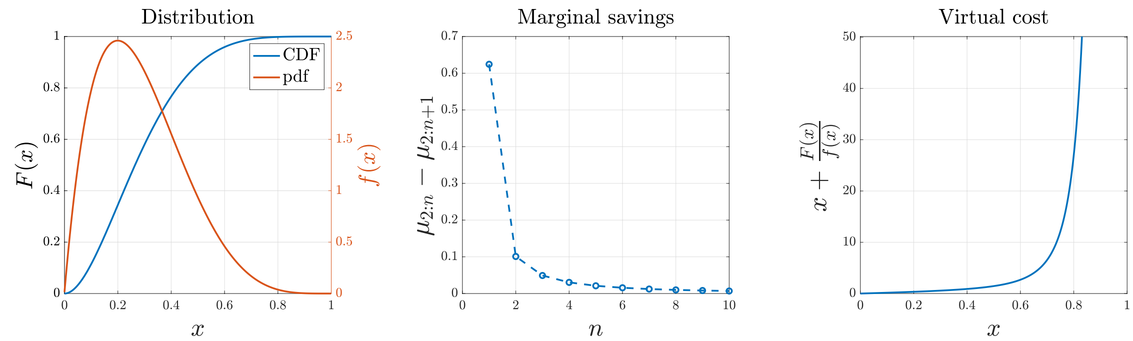


Figure 6: Regularity of beta distribution with shape parameters $\alpha = 2$ and $\beta = 5$: (Left) Distribution $F(x) = I_x(\alpha, \beta)$, where I denotes the regularized incomplete beta function, (Middle) Marginals of second order statistic, (Right) Virtual cost function.

using equation (22), we have:

$$\mu_{r:n-1} - \mu_{r:n} = \frac{r}{n}\binom{n}{r}\int_0^1 (x^p)^r(1-x^p)^{n-r}dx = \frac{r}{n}\binom{n}{r}\frac{\Gamma(n+1-r)\Gamma(r+\frac{1}{p})}{p\cdot\Gamma(n+\frac{1}{p}+1)}, \tag{23}$$

where Γ is the Gamma function, and satisfies $\Gamma(z+1) = z\Gamma(z)$. Comparing this expression with the same expression for $\mu_{r:n} - \mu_{r:n+1}$, we need to show that:

$$\frac{(n-1)!}{(r-1)!(n-r)!} \times \frac{\Gamma(n+1-r)\Gamma(r+\frac{1}{p})}{p\cdot\Gamma(n+\frac{1}{p}+1)} \geq \frac{n!}{(r-1)!(n+1-r)!} \times \frac{\Gamma(n+2-r)\Gamma(r+\frac{1}{p})}{p\cdot\Gamma(n+\frac{1}{p}+2)},$$

or equivalently

$$\frac{\Gamma(n+\frac{1}{p}+2)}{\Gamma(n+\frac{1}{p}+1)} \geq \frac{n}{n+1-r} \times \frac{\Gamma(n+2-r)}{\Gamma(n+1-r)}.$$

Now using the above-mentioned property of the Gamma function $\Gamma(z+1) = z\Gamma(z)$, notice that the left hand side is equal to $n+1+1/p$, while the right hand side is equal to n; therefore, the inequality is true for any $p > 0$. $\square$

D Non-monotone Example

Here we give an example where the global optimal thresholds are non-monotone, in particular for this example $t_2^* < t_3^*$. We proved via Lemma 1 and Theorem 3 that this cannot happen if the underlying distribution F satisfies our two assumptions of regularity (Definition 3) and diminishing returns of order statistics (Definition 5). In fact, the distribution F in the following example fails both of these assumptions. It has a point mass probability which breaks the monotonicity of the virtual cost, and more importantly, its second order statistic does not have the diminishing returns property. More precisely, the savings of having the 3rd agent participating in the second day $(\mu_{2:2} - \mu_{2:3})$ is more than the savings from the 2nd agent $(1 - \mu_{2:2})$. This makes the buyer want to save the 3rd agent even at a higher value than he is willing to pay for the 2nd agent.

Example 5 (Non-monotone thresholds)**.** Consider a distribution F that is $U[0, \frac{1}{2}]$ with probability $1/2$ and equal to 1 with probability $1/2$. In other words,

$$M = \begin{cases} X & \text{w.p. } \frac{1}{2}, \\ 1 & \text{w.p. } \frac{1}{2}, \end{cases}$$

where $X \sim U[0, \frac{1}{2}]$. Now suppose we have three agents with overhead costs drawn independently from F. If we save 1, 2, or 3 agents, the expected cost we incur on day 2 would be 1, $\mu_{2:2} = \frac{5}{6}$, and $\mu_{2:3} = \frac{21}{32}$, respectively. Additionally, the optimal thresholds on day 1 are

$$t_2^* = \frac{1}{2}(1 - \mu_{2:2}) = \frac{1}{12} \approx 0.083,$$

$$t_3^* = \frac{1}{2}(\mu_{2:2} - \mu_{2:3}) = \frac{17}{192} \approx 0.088,$$

which show that the optimal thresholds can be non-monotone ($t_3^* > t_2^*$), if we do not have constraints on the underlying distribution F.