Simple Delegated Choice

Ali Khodabakhsh,[*] Emmanouil Pountourakis,[†] Samuel Taggart[‡]

July 20, 2022

Abstract

This paper studies delegation in a model of discrete choice. In the delegation problem, an uninformed principal must consult an informed agent to make a decision. Both the agent and principal have preferences over the decided-upon action which vary based on the state of the world, and which may not be aligned. The principal may commit to a mechanism, which maps reports of the agent to actions. When this mechanism is deterministic, it can take the form of a menu of actions, from which the agent simply chooses upon observing the state. In this case, the principal is said to have delegated the choice of action to the agent.

We consider a setting where the decision being delegated is a choice of a utility-maximizing action from a set of several options. We assume the shared portion of the agent's and principal's utilities is drawn from a distribution known to the principal, and that utility misalignment takes the form of a known bias for or against each action. We provide tight approximation analyses for simple threshold policies under three increasingly general sets of assumptions. With independently-distributed utilities, we prove a 3-approximation. When the agent has an outside option the principal cannot rule out, the constant-approximation fails, but we prove a $\log \rho / \log \log \rho$-approximation, where ρ is the ratio of the maximum value to the optimal utility. Finally we give a weaker but tight bound that holds for correlated values. One special case of our model is utility-based assortment optimization, for which our results are new.

[*]University of Texas, Austin
[†]Drexel University
[‡]Oberlin College

1 Introduction

This paper considers a ubiquitous scenario in economic decisionmaking. A decisionmaker or *principal* faces a choice in which the appropriate action is dependent on the state of the world. For example, a firm seeks to replace a piece of expensive equipment, or a national health service must choose which treatment to provide to a patient who might display a range of symptoms. For practical reasons, however, the principal is unable to directly observe the state, and must rely on an *agent* to observe the state instead. Equipment operators know their needs better than managers, and the health service relies on doctors to observe patients. In such arrangements, the agent and principal tend not to have preferences which are perfectly aligned: the firm must pay for new equipment (while the operator does not), and doctors tend to overdiagnose certain conditions. How should the uninformed principal interpret information from the agent to manage the misalignment of incentives and choose the appropriate action for the state?

When the principal has commitment power, the selection of a mediating device becomes a problem of mechanism design. The agent observes the state and reports a message to the mechanism, which chooses a possibly randomized action. In this setting, deterministic mechanisms hold special appeal. The *taxation principle* states that every deterministic mechanism is equivalent to a menu: the principal selects the set of allowable actions, and the agent simply chooses their preferred action upon observing the state. Such mechanisms eliminate the need for communication between the agent and principal, and are therefore so common in practice that they are often taken for granted as a managerial tool. The problem of menu design for a better-informed agent is often referred to as *delegation*, coined by Holmstrom (1978).

In the examples above, the alignment of the agent and principals' preferences is well-structured. The principal's main uncertainty in the choice problem is payoff-relevant for both parties: in replacing equipment, both the operator and firm want to purchase the right tool for the job. Meanwhile, misalignment of preferences is predictable – the firm will pay for the new purchase, and prices are likely known in advance. Under these conditions, a particularly salient family of mechanisms is *threshold* mechanisms, which restrict the agent to actions where the misalignment of preferences is not too great. For our firm and operator, this would take the form of a budget.

Our Contributions This work gives a model for delegated choice scenarios like those discussed above. In our model, the agent and principals' preferences for a particular action are captured by two quantities. First, each action i has a shared *value* v_i, which is unknown to the principal (but distributed according to a known prior) but observable to the agent. Second, each action has a commonly-known and fixed *bias* b_i, which captures the amount the agents' utility differs from that of the principal. The agent may also have outside options which the principal cannot prohibit; we extend our model to capture this issue as well.

Our analysis takes a simple-versus-optimal perspective. That is, we completely characterize the performance of threshold mechanisms with respect to the often less intuitive optimal mechanism. We do so under three increasingly general regimes, distinguished by the correlation or independence of the value distributions and the absence or presence of an outside option:

- With independently distributed values and no outside option, we show that threshold mechanisms are a 3-approximation to the optimal mechanism.[1]

- With independently distributed values and an outside option, threshold mechanisms cannot obtain any nontrivial approximation in general. However, we show a parametrized

[1] Our results also hold with an outside option if that action has a fixed value, which we make precise subsequently.

$\log \rho / \log\log \rho$-approximation, where ρ is the ratio of largest possible value to OPT.

- With correlated values, we give a $\log p_{\min}^{-1}$ approximation, where $p_{\min}$ is the probability of the least likely value profile. We also prove a supplementary hardness of approximation result.

We match all three approximation analyses with tight examples. A well-studied special case of our model is utility-based assortment optimization, discussed in Section 3. Our results are new to that literature.

Roadmap In Section 2, we give our formal model. We then survey existing work on delegation in Section 3, and make specific comparisons to existing models for delegated search and assortment optimization. Section 4 contains our constant-approximation under independence and lays the groundwork for our parametrized analysis with an outside option in Section 5. Finally, we analyze delegation with correlated values in Section 6.

2 Model

We now give our model of delegated choice. The principal seeks to choose from a discrete set Ω of n actions. The principal's utility for action i is given by a random *value* $v_i \geq 0$, which the principal is unable to observe. To assist in selecting an action, the principal may consult an agent, who observes all actions' values, and may communicate with the principal after observation. We decompose the agent's utility for action i into its value, shared with the principal, and an unshared *bias* term. That is, the agent's utility is given by $u_i = v_i + b_i$. Throughout the paper, we assume each bias b_i is constant and known to the principal.

We assume the principal has the power to commit ex ante to a mechanism for communicating with the agent and selecting an action, and study deterministic mechanisms. By the taxation principle, it suffices to consider mechanisms described by menus over actions. The agent observes all actions' values and selects their utility-maximizing action from the menu — which may differ from the principal's preferred action. Taking this perspective, we consider the algorithmic problem of selecting a menu A to maximize the principal's expected utility when the agent selects their preferred action according to the observed values. We further assume the existence of an outside option for the agent, denoted action 0, with value v_0 and bias b_0. We assume that regardless of the principal's choice of A, the agent may always select this action.

Formally, when presented with action set $A \subseteq \Omega$ and after observing the vector of values $\mathbf{v}$, denote the agent's preferred choice by $g(A, \mathbf{v})$. That is, $g(A, \mathbf{v}) = \operatorname{argmax}_{a \in A \cup \{0\}} (v_i + b_i)^2$. The principal is faced with a set function optimization problem. We assume the principal has a prior distribution F over the values $\mathbf{v}$, and must select a menu A for the agent which maximizes their own expected utility. That is, the principal solves:

$$\underset{A \subseteq \Omega}{\text{maximize}} \;\; f(A) := \int_{\mathbf{v}} v_{g(A,\mathbf{v})} \, dF(\mathbf{v}).$$

The model above captures applications such those described in Section 1. Note that we allow the agent's utility to be negative, and that the model is invariant to additive shifts in the agent's bias for every action. We will study a particularly simple set of mechanisms, namely *threshold mechanisms*. The threshold mechanism with bias t is given by $A_t = \{i \,|\, b_i \leq t\}$. Note that since the number of threshold policies is at most the number of actions, the principal may compute an optimal threshold efficiently. We analyze the approximation ratio between the best threshold menu and the optimal menu overall.

[2] We assume the agent breaks ties in the principal's favor, then lexicographically.

3 Related Work

Simple Versus Optimal Mechanisms. A primary contribution of computer science to the study of mechanism design is the use of approximation to explain the prevalence of simple mechanisms. For example, Hartline and Roughgarden (2009) prove that the simple auctions often observed in practice can obtain a constant factor of the sometimes-complicated, rarely-used optimal mechanisms. Hartline (2013) surveys similar results for auctions. Recently, Dütting et al. (2019) and Castiglioni et al. (2021) make similar forays into contract theory, characterizing the power of simple linear contracts. Our work initiates the study of delegated choice through a similar lens.

Real-Valued Delegation. Delegated decisionmaking is a canonical problem in microeconomic theory and managerial science. Much of the literature subsequent to Holmstrom (1978) has focused on the special case where the state and action space are continuous and real-valued, and where the preferences of both the agent and principal are single-peaked. Notable examples include Melumad and Shibano (1991), Martimort and Semenov (2006), Alonso and Matouschek (2008), and Amador and Bagwell (2010), who characterize the structure of optimal mechanisms under increasingly general variants of the single-peaked model. The main conclusions from these papers are necessary and sufficient conditions for the optimal delegation set to be an interval on the real line.

Additional work on similar models includes Kovác and Mylovanov (2009), who study the gap in performance between randomized and deterministic mechanisms, and Ambrus and Egorov (2017), who study a principal who can add additional nonmonetary costs to incentivize more preferred decisions. Aghion and Tirole (1997) and Szalay (2005) consider models in which one or more of the principal and the agent may expend effort to observe a signal about the state. Frankel (2014) considers maxmin robust delegation in the face of multiple decisions. With the exception of Armstrong and Vickers (2010) and followup works, though, the economics literature has focused on the real-valued model for decisions. Our work considers the mathematically incomparable but similarly common setting of discrete choice. In the latter setting, the structure of the problem renders exact characterization of optimal mechanisms difficult, and motivates the use of a simple-versus-optimal approach.

Delegated Search. The model of delegated project choice from Armstrong and Vickers (2010) is perhaps closest to ours. The authors consider an agent who chooses between n discrete actions. The principal is able to verify the utilities provided by the selected action, and restrict the agent's choice based on this information. Subsequent followups by Kleinberg and Kleinberg (2018), Bechtel and Dughmi (2021), and Bechtel et al. (2022) note a strong connection between the Armstrong and Vickers (2010) model and well-studied online stochastic optimization problems. They upperbound the *delegation gap*: they show that even when the agent must pay a search cost to discover each action's utility, the principal can obtain utility within a constant factor of the first-best solution, where they solve the search problem themselves. Braun et al. (2022) also give a version where the agent searches online, and make similar comparisons to first-best.

Our model differs from the delegated search literature in two notable ways. First is the absence of search. Our agent can perfectly observe the values of all actions. More significantly, our principal is unable to verify the utilities provided by the agent's selected action; they may only rule actions in or out completely. In our model, the first-best solution is $\mathbb{E}[\max_i v_i]$. The following example shows that no delegation set may approximate the first-best to a factor better than n. This contrasts with the constant-approximation results from the work cited above.

Example 3.1. Consider an instance with n independently-distributed actions. Action i has a value v_i which is $1-\epsilon$ with probability $1/n$ and 0 otherwise. Each action i has bias $b_i = i$. The first-best expected utility is constant, while in any delegation set, the agent will always pick the highest-indexed action, yielding expected utility $(1-\epsilon)/n$.

Assortment Optimization. Our model captures special cases of the well-studied *assortment optimization* problem. In assortment optimization, a seller must decide which among a set of fixed-price items to offer. A variety of models are common for the buyer's purchase choice, including nested logit models (Davis et al., 2014; Li et al., 2015) and Markov chain-based choice (Feldman and Topaloglu, 2017), along with equivalent models based on random buyer utility (Berbeglia, 2016; Aouad et al., 2018), which includes the especially prevalent multinomial logit model as a special case. Our model subsumes assortment optimization with utility-based choice. To see this, consider n items, where the buyer utility w_i for each item i is random, and the revenue r_i for item i is known to the seller. Taking $v_i = r_i + \epsilon w_i$ and $b_i = -(1+\epsilon)r_i$ for sufficiently small $\epsilon > 0$ yields an equivalent delegation problem. Under this transformation, an outside option with $v_0 = 0$ corresponds to the no-buy option, and the option to buy elsewhere with positive utility can be captured with a randomized outside option.

Threshold mechanisms in our model correspond to *revenue-ordered assortments*, a well-studied class of solutions for assortment optimization. Talluri and Van Ryzin (2004) show that revenue-ordered assortments are optimal for several choice models including the well-studied multinomial logit model. Rusmevichientong et al. (2014) study revenue-ordered assortments under the mixed multinomial logit model and give conditions under which they are optimal or approximately optimal, and Berbeglia and Joret (2020) give general, parametrized analyzes. Both our bounds with random outside options, in Sections 5 and Section 6, offer complementary bounds based on different parametrizations. Other work on computational hardness or approximation in assortment optimization includes Désir et al. (2020), who hardness of approximation under a knapsack-constrained version of the problem, and Immorlica et al. (2018), who study a version where the buyer has combinatorial preferences over bundles.

4 Threshold Delegation with Independent Values

We now present our strongest bound, a constant-approximation assuming the principal's prior F over values is a product distribution, and hence, actions' values are independent. We further assume that the outside option's value, v_0, is deterministic, which subsumes the no-outside-option case, as we could have $b_0 = -\infty$.

Theorem 4.1. *Under independent values and a deterministic outside option, there always exists a threshold mechanism with expected utility that is a 3-approximation to the optimal deterministic mechanism.*

Theorem 4.1 holds regardless of choice of the outside option's fixed value and bias. Before giving the details of the proof, we derive two technical results which will facilitate analysis. In Section 4.1 for any delegation set, we give a decomposition of the principal's utility into two quantities, one aligned with the agent's utility and one not. Then, in Section 4.2 we use independence obtain a lower bound on the value from threshold sets which will prove useful for both this and the next section's analyses.

4.1 Utility Decomposition

The principal's task is to balance two sources of utility. On the one hand, when some action has very high value, preferences are aligned: the principal benefits from giving the agent the flexibility to select this action. On the other hand, when actions have smaller values, the principal must control misalignment: they may benefit from restricting the agent away from actions with higher bias, inducing the agent to take actions that provide better value. We now decompose the principal's utility for the optimal delegation set into two quantities, Sur and BDif, which roughly correspond to the value from each of these two cases.

To make the decomposition precise, note that for the optimal delegation set A^*, there are two lower bounds imposed on the agent utility from any selection: first, the chosen action must be preferred to the outside option, action 0, which gives utility at least b_0. Second, the agent's utility is at least the bias of the most-biased action in A^*. Denote the better of these bounds by $\underline{u}$. We can therefore think of the contribution of any action $i \in A^*$ as decomposing into a *bias difference* $\underline{u} - b_i$ and a *surplus* $v_i - (\underline{u} - b_i)$. Intuitively, the surplus captures the principal's utility from giving the agent latitude to pick high-valued actions, and the bias difference captures the misaligned portion of the principal's utility. Formally, the decomposition is the following.

Lemma 4.1. *Let A^* denote the optimal delegation set, and let $b_{max} = \max\{b_i \,|\, i \in A^* \cup \{0\}\}$. Further let $\underline{u} = \max(b_{max}, b_0)$. Define* Sur *and* BDif *as follows:*

$$\begin{aligned} \textsc{Sur} &= \int_{\mathbf{v}} v_{g(A^*,\mathbf{v})} - (\underline{u} - b_{g(A^*,\mathbf{v})})\, dF(\mathbf{v}) \\ \textsc{BDif} &= \int_{\mathbf{v}} \underline{u} - b_{g(A^*,\mathbf{v})}\, dF(\mathbf{v}). \end{aligned}$$

Then we can write $f(A^*) = \textsc{Sur} + \textsc{BDif}$.

To verify the intuition that Sur captures the aligned portion of the principal's utility, note that choosing the smallest threshold set containing all of $A^* \cup \{0\}$ secures Sur for the principal. Formally:

Lemma 4.2. *Let $A_{\underline{u}} = \{i \,|\, b_i \leq \underline{u}\}$. Then $f(A_{\underline{u}}) \geq \textsc{Sur}$.*

Proof. We will argue pointwise for each value profile $\mathbf{v}$. The action chosen by the agent under $A_{\underline{u}}$ is $g(A_{\underline{u}}, \mathbf{v})$, which has $b_{g(A_{\underline{u}},\mathbf{v})} \leq \underline{u}$. Since $g(A_{\underline{u}}, \mathbf{v})$ is the agent's favorite, we have $v_{g(A_{\underline{u}},\mathbf{v})} + b_{g(A_{\underline{u}},\mathbf{v})} \geq v_{g(A^*,\mathbf{v})} + b_{g(A^*,\mathbf{v})}$. Hence,

$$\begin{aligned} v_{g(A_{\underline{u}},\mathbf{v})} &\geq v_{g(A^*,\mathbf{v})} + b_{g(A^*,\mathbf{v})} - b_{g(A_{\underline{u}},\mathbf{v})} \\ &\geq v_{g(A^*,\mathbf{v})} - (\underline{u} - b_{g(A^*,\mathbf{v})}). \end{aligned}$$

Taking expectation over $\mathbf{v}$ yields the lemma. ☐

Lemma 4.2 implies that the main difficulty for obtaining approximately-optimal delegation sets is managing the misaligned portion of the principal's utility. Section 4.3 gives this analysis for the case with v_0 fixed, yielding a 3-approximation. Note Lemmas 4.1 and 4.2 hold even when the outside option's value v_0 is randomized. We will therefore make further use of them in our analysis of that case in Section 5.

4.2 Lower Bounds via Partial Derandomization

To compare the performance of a threshold set A_t to the optimal set A^*, we will show that threshold sets can retain sufficient value from $A_t \cap A^*$ without introducing actions in $A_t \setminus A^*$ which overly distort the agent's choices. Independence allows us to summarize the interference of $A_t \setminus A^*$ with a single deterministic action. This will greatly simplify subsequent analyses. This section focuses on the case of fixed outside options, but we state our lemma for possibly randomized outside options. We will reuse the result in Section 5.

Lemma 4.3. *Assume values are independently distributed. Then for any threshold set A_t, there exists a single action $a(t)$ with bias $b_{a(t)} = t$ and deterministic value $v_{a(t)}$ such that $f(A_t) \geq f(A_t \cap A^* \cup \{a(t)\})$.*

Note that $a(t)$ need not be an action from the original delegation instance. The proof will follow from picking the worst realization of actions in $A_t \setminus A^*$ for the principal. Note further that $a(t)$ may differ for every threshold t: hence our lower bounds correspond not to one derandomized delegation instance, but to one per threshold.

Proof. For brevity, denote $A_t \setminus A^*$ by B_t. Actions in B_t may have randomized values. The principal's expected utility $f(A_t)$ can be computed by first realizing v_i for all $i \in B_t$, then computing the principal's expected utility over the values of actions in $G_t = A_t \cap A^* \cup \{0\}$. Hence, there must exist a joint realization of values $\hat{v}_i$ for each $i \in B_t$ for which this latter expectation is at most $f(S_t)$. Let $\hat{B}_t$ denote a new set of actions consisting of the actions $i \in B_t$ with v_i fixed as $\hat{v}_i$. We have $f(G_t \cup B_t) \geq f(G_t \cup \hat{B}_t)$. Any actions which are not selected in any realization of the values for G_t may be removed from $\hat{B}_t$ without consequence. However, since values are fixed for each $i \in \hat{B}_t$, the agent consistently prefers some particular action $\hat{i} \in \hat{B}_t$ over the others in $\hat{B}_t$. Hence, we may remove all actions but $\hat{i}$ from $\hat{B}_t$ without changing the principal's utility.

We finally use this remaining action $\hat{i}$ to construct $a(t)$. Let $v_{\hat{i}}$ and $b_{\hat{i}}$ denote the value and bias of $\hat{i}$. Define $a(t)$ to have bias t and value $v_{\hat{i}} - (t - b_{\hat{i}})$. Note that $v_{\hat{i}} + d_{\hat{i}} = v_{a(t)} + d_{a(t)}$. Hence, the agent will choose $a(t)$ from $G_t \cup \{a(t)\}$ if and only if they would choose $\hat{i}$ from $G_t \cup \{\hat{i}\}$. Moreover, $v_{a(t)} = v_{\hat{i}} \leq v_{\hat{i}}$. Hence, $f(G_t \cup \{a(t)\}) \leq f(G_t \cup \{\hat{i}\}) \leq f(A_t)$. □

4.3 Proof of Theorem 4.1

We now show how to obtain a 3-approximation to the optimal delegation utility using threshold mechanisms, assuming v_0 is fixed. Lemma 4.1 decomposes the optimal utility into an aligned portion, Sur, and a misaligned portion, BDif. Furthermore, Lemma 4.2 states that Sur can be 1-approximated using a threshold set. Hence, it will suffice to obtain a 2-approximation to BDif using thresholds. To do so, we use the derandomization of Lemma 4.3 to derive an even stronger lower bound which holds when v_0 is fixed. We then select a threshold for which this lower bound is guaranteed to be large.

Lemma 4.4. *For any threshold set A_t:*

$$f(A_t) \geq \min\Big(\underline{u} - t, \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})})\mathbb{I}[g(A^*, \mathbf{v}) \in A_t \cup \{0\}]\, dF(\mathbf{v})\Big).$$

To understand our lower bound, note two pitfalls a threshold set could face. First, a too-expansive threshold could include high-bias actions which attract the agent while providing little value. Second, a too-restrictive threshold could leave the agent with too few options. Lemma 4.4 states that these are the only two problems: if a threshold t is sufficiently low and includes enough of the actions providing BDif for A^*, t will perform well.

Proof of Lemma 4.4. We argue with respect to the derandomized action $a(t)$. For brevity, write $\underline{A}_t = A_t \setminus A^* \cup \{a(t)\}$. Depending on $v_{a(t)} + b_{a(t)}$, we have two cases, each of which produces a lower bound on $f(\underline{A}_t)$,

- *Case 1:* $v_{a(t)} + b_{a(t)} < \underline{u}$. In this case, any time $g(A^*, \mathbf{v}) \in A_t \cup \{0\}$, we have $g(A^*, \mathbf{v}) = g(\underline{A}_t, \mathbf{v})$. Since every choice from A^* gives the agent utility at least $\underline{u}$, we have $v_{g(\underline{A}_t,\mathbf{v})} + b_{g(\underline{A}_t,\mathbf{v})} \geq \underline{u}$, and hence $v_{g(\underline{A}_t,\mathbf{v})} \geq \underline{u} - b_{g(\underline{A}_t,\mathbf{v})}$. Integrating over all $\mathbf{v}$ yields

$$f(A_t) \geq f(\underline{A}_t) \geq \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[g(A^*, \mathbf{v}) \in A_t \cup \{0\}]\, dF(\mathbf{v}).$$

- *Case 2:* $v_{a(t)} + b_{a(t)} \geq \underline{u}$. Then regardless of $\mathbf{v}$, we have $v_{g(\underline{A}_t,\mathbf{v})} + b_{g(\underline{A}_t,\mathbf{v})} \geq \underline{u}$. Since the agent breaks ties in the principal's favor, we also have that $g(\underline{A}_t, \mathbf{v}) \neq 0$, so $b_{g(\underline{A}_t,\mathbf{v})} \leq t$. We may conclude that for all $\mathbf{v}$, $v_{g(A_t,\mathbf{v})} \geq \underline{u} - b_{g(A_t,\mathbf{v})} \geq \underline{u} - t$, and hence

$$f(A_t) \geq f(\underline{A}_t) \geq \underline{u} - t.$$

□

The lower bound in Lemma 4.4 is a minimum of two terms. We will now study the threshold $\hat{t} = \underline{u} - \textsc{BDif}/2$, and observe that both terms in the convex combination are at least $\textsc{BDif}/2$. In particular, we can lower bound the second term as follows:

Lemma 4.5. *Let* $\hat{t} = \underline{u} - \textsc{BDif}/2$. *Then we have:*

$$\int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[g(A^*, \mathbf{v}) \in A_t \cup \{0\}]\, dF(\mathbf{v}) \geq \textsc{BDif}/2.$$

Proof. Let $\mathcal{E}_2$ denote the event that $g(A^*, \mathbf{v}) \in A_t \cup \{0\}$. The following sequence of inequalities, explained below, implies the lemma:

$$\begin{aligned}
\textsc{BDif} &= \int_{\mathbf{v}} \underline{u} - b_{g(A^*,\mathbf{v})}\, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_2]\, dF(\mathbf{v}) + \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\overline{\mathcal{E}_2}]\, dF(\mathbf{v}) \\
&< \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_2]\, dF(\mathbf{v}) + (\underline{u} - \hat{t}) \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_2]\, dF(\mathbf{v}) - \textsc{BDif}/2.
\end{aligned}$$

The first equality is the definition of $\textsc{BDif}$. The third equality follows from the fact that under $\overline{\mathcal{E}_2}$, $b_{g(A^*,\mathbf{v})} > \hat{t}$, and from the fact that this occurs with probability at most 1. The last equality follows from the definition of $\hat{t}$. □

Proof of Theorem 4.1. By combining Lemma 4.4, with the definition of $\hat{t}$ and Lemma 4.5, we have:

$$\begin{aligned}
f(A_{\hat{t}}) &\geq p_{\hat{t}}(\underline{u} - \hat{t}) + (1 - p_{\hat{t}}) \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[g(A^*, \mathbf{v}) \in A_{\hat{t}} \cup \{0\}]\, dF(\mathbf{v}) \\
&= p_{\hat{t}} \tfrac{\textsc{BDif}}{2} + (1 - p_{\hat{t}}) \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[g(A^*, \mathbf{v}) \in A_{\hat{t}} \cup \{0\}]\, dF(\mathbf{v}) \\
&\geq p_{\hat{t}} \tfrac{\textsc{BDif}}{2} + (1 - p_{\hat{t}}) \tfrac{\textsc{BDif}}{2} = \tfrac{\textsc{BDif}}{2}.
\end{aligned}$$

The theorem now follows from noting that $f(A^*) = \textsc{Sur} + \textsc{BDif} \leq f(A_{\underline{u}}) + 2f(A_{\hat{t}})$. □

The proof of Theorem 4.1 used independence once, in the derandomization step of Lemma 4.3. Nevertheless, we show in Section 6 that independence is critical to guaranteeing the performance of threshold mechanisms by giving a super-constant lower bound in its absence. With independence, the following example matches the upper bound exactly:

Example 4.1. Our example will have five actions, with biases and value distributions given below. The outside option will have $b_0 = -\infty$, and therefore can be ignored. Take two small numbers, δ and ϵ, with δ much smaller than ϵ. Actions will be as follows:

- $b_1 = 0$. v_1 is $1 + 2\delta$ with probability ϵ, and 0 otherwise.
- $b_2 = 1 - \epsilon - \delta$. $v_2 = 4\delta + \epsilon$.
- $b_3 = 1 - \epsilon$. $v_3 = \epsilon + \delta$.
- $b_4 = 1 - \delta$. $v_4 = 5\delta$.
- $b_5 = 1$. v_5 is 1 with probability ϵ, and 0 otherwise.

We may analyze the instance neglecting δ terms, which only serve to break ties for the agent. The optimal delegation set is $\{1, 3, 5\}$, with principal utility $(1 - (1 - \epsilon)^2) + \epsilon(1 - \epsilon)^2$, where the first term comes from the event that either actions 1 or 5 realize their high values (in which case they are chosen), and the second term comes from the event that 1 and 5 are low-valued, in which case the agent prefers action 3. As $\epsilon \to 0$, the optimal value goes to 0 as $\approx 3\epsilon$. Meanwhile, no threshold obtains expected value better than ϵ. This yields an approximation ratio of 3 in the limit.

5 Randomized Outside Options

In Section 4, we showed that with a fixed (or non-existent) outside option, a simple delegation set secures a constant fraction of the utility from the optimal delegation set. We now consider the case where the outside option's value is randomized. This may be more realistic in scenarios such as assortment optimization, where the agent's outside option is taking an action (i.e. buying a good) somewhere else. In this regime, we again give tight bounds. In Section 5.1, we show that no nontrivial multiplicative approximation is possible with threshold sets: there are examples where thresholds give no better than an $\Omega(n)$-approximation, which can be matched trivially. However, in Section 5.2 we show that such lower bound examples are necessarily unnatural. In particular, we parametrize our analysis by the ratio $\rho = v_{\max}/\mathrm{OPT}$, where OPT is the optimal principal utility and $v_{\max}$ the highest value in any action's support. We prove that the worst-case approximation is $\Theta(\log \rho / \log \log \rho)$: hence, whenever thresholds perform poorly, it is because the optimal solution relies on exponentially large, exponentially rare values.

5.1 Unparametrized Analysis: Impossibility

This section gives an unparametrized analysis of threshold delegation with a randomized outside option. We show that it is not possible to guarantee a nontrivial approximation factor which holds across all instances.

Our constant-approximation in Section 4 relied on our ability to separate the optimal utility into two parts, BDif and Sur. Approximating the bias difference BDif was the crux of the analysis. The following example shows that with a random outside option value v_0, this analysis — and in particular the approximation of BDif — fails. We will choose our distribution over v_0 to streamline exposition, but the example that follows could be adjusted so that the distribution over v_0 satisfies nearly any desired regularity condition.

Example 5.1. Our example will feature two sets of actions: *good* actions, which are taken by the optimal delegation set, and *bad* actions, which are not. We will index the actions so that the ith good action is $g(i)$, and the bad action between good actions $g(i-1)$ and $g(i)$ is $b(i)$. For $i \in \{1, \ldots, n\}$ $g(i)$ will have:

- bias $n^{n-1} - n^{n-i}$:
- value $n^{n-i} + i\epsilon$ with probability $1/n$, and 0 otherwise.[3]

The bad actions will be indexed by $b(i)$ for $i \in \{2, \ldots, n\}$. Bad action $b(i)$ will have

- bias $n^{n-1} - n^{n-i}$.
- value $n^{n-i} + (i-1)\epsilon + \delta$, for $\delta \ll \epsilon$.

The outside option outside option will have bias n^{n-1} and value v_0 distributed according to a discrete distribution. We will set $\Pr[v_0 = \epsilon/2] = n^{-(n-1)}$. For $i > 1$, we will choose probability mass function $\Pr[v_0 = i\epsilon - \epsilon/2] = n^{-(n-i)} - n^{-(n-i+1)}$. Note that we have picked these probabilities so that $\Pr[v_0 < i\epsilon] = n^{-(n-i)}$. The values of all actions described above are independent.

A solution to the delegation instance we just described is to take only good actions. The probability that at least one good action takes its high value is $1 - (1 - 1/n)^n \geq 1 - 1/e$. Assume this event has occurred, and that the agent's preferred good action is $g(i)$. Then $g(i)$ is preferred to the outside option with probability $n^{-(n-i)}$. Hence, the principal's expected utility from choosing only good actions is at least:

$$f(\{g(1), \ldots, g(n)\}) \geq (1 - 1/e)n^{-(n-i)}(n^{n-i} + \epsilon i) \geq 1 - 1/e.$$

Now consider a threshold set A_t. It is without loss of generality to consider $t = n^{n-1} - n^{n-j}$ for some j, which implies that the highest-bias actions in A_t are $g(j)$ and $b(j)$. For any good action $g(i)$, with $i < j$, the agent's utility for $g(i)$ on a high-valued realization is $n^{n-1} + i\epsilon < n^{n-1} + (j-1)\epsilon + \delta$. Hence, the agent ignores all actions other than $g(j)$, $b(j)$, and the outside option. If $g(j)$ draws its high value, the principal gets utility $n^{n-j} + j\epsilon$ utility if and only if $g(j)$ survives the outside option, which happens with probability $n^{-(n-j)}$. Otherwise, the agent looks to action $a(j)$, and takes it over the outside option with probability $n^{-(n-j+1)}$. Ignoring value from the outside option, which goes to 0 as $\epsilon \to 0$, we can account for the utility from A_t as follows:

$$\begin{aligned} f(A_t) &= \tfrac{1}{n}n^{-(n-j)}(n^{n-j} + j\epsilon) + (1 - \tfrac{1}{n})(n^{n-j} + (j-1)\epsilon + \delta)n^{-(n-j+1)}. \\ &\approx \tfrac{1}{n} + (1 - 1/n)\tfrac{1}{n}, \end{aligned}$$

where the latter approximation holds for ϵ and δ sufficiently small. This implies that every threshold incurs a loss which is $\Omega(n)$.

An upper bound of n for threshold mechanisms is trivial, by the following lemma. Hence, up to a constant, the lower bound in Example 5.1 is tight.

Lemma 5.1. *For any set A and $i \in A \cup \{0\}$, let $A^i = \int_{\mathbf{v}} v_i \, \mathbb{I}[g(A, \mathbf{v}) = i] \, d\mathbf{v}$ denote the contribution to $f(A)$ from action i. Then $f(A_{b_i}) \geq A^i$.*

[3] It is equivalent for this example to use distribution $n^{n-i} + i\epsilon$ with probability $1/n$, and $n^{n-i} - \epsilon$ otherwise. Under this distribution, the example becomes an instance of assortment optimization, as described in Section 3. This makes our parametrized analysis in Section 5.2 tight even for that special case.

Proof. Consider any $\mathbf{v}$ where $g(A, \mathbf{v}) = i$. The action chosen by the agent under A_{b_i} is $g(A_{b_i}, \mathbf{v})$, which has $b_{g(A_{b_i}, \mathbf{v})} \leq b_i$. Since $g(A_{b_i}, \mathbf{v})$ is the agent's favorite, we have $v_{g(A_{b_i}, \mathbf{v})} + b_{g(A_{b_i}, \mathbf{v})} \geq v_{g(A, \mathbf{v})} + b_{g(A, \mathbf{v})} = v_i + b_i$. Hence, $v_{g(A_{b_i}, \mathbf{v})} \geq v_i + b_i - b_{g(A_{b_i}, \mathbf{v})} \geq v_i$.Taking expectation over $\mathbf{v}$, we obtain:

$$f(A_{b_i}) = \int_{\mathbf{v}} v_{g(A_{b_i}, \mathbf{v})} \, dF(\mathbf{v}) \geq \int_{\mathbf{v}} v_{g(A_{b_i}, \mathbf{v})} \, \mathbb{I}[g(A, \mathbf{v}) = i] \, dF(\mathbf{v}) \geq \int_{\mathbf{v}} v_i \, \mathbb{I}[g(A, \mathbf{v}) = i] \, dF(\mathbf{v}),$$

where the first inequality follows from the nonnegativity of v_i. □

An n-approximation then follows from noting that for any set A, $f(A) = \sum_i A^i$.

Corollary 5.1. *With independent values (and possibly randomized outside option), the best threshold is an n-approximation to the optimal delegation set.*

5.2 Parametrized Approximation

In the previous section, we gave an example where no threshold set was better than an $\Omega(n)$-approximation. However, this example was extreme, in the sense that while the optimal solution obtained $O(1)$ utility, some actions had values which were as large as n^{n-1}. We now show that this is no coincidence: any example where threshold mechanisms perform poorly must be unnatural in this way.

Theorem 5.1. *Let $\rho = v_{\max}/\mathrm{OPT}$, where OPT is the optimal principal utility and $v_{\max}$ the highest value in any action's support. Then with independent values (and a possibly randomized outside option), the best threshold is a $O(\log \rho / \log \log \rho)$-approximation to OPT.*

Theorem 5.1 is of particular interest for the application of assortment optimization. For an instance of the latter problem, each item i yields revenue p_i for the seller, and value w_i for the buyer. Framed as a delegation problem, we have $v_i = p_i + \epsilon w_i$, for sufficiently small ϵ. Hence, Theorem 5.1 implies that the prices p_i must be extreme whenever revenue-ordered assortments perform poorly. Another consequence of Theorem 5.1 is a bicriteria approximation when values lie in $[0, 1]$: either some threshold obtains a small multiplicative approximation, or it is trivial to obtain an additive approximation.

To prove Theorem 5.1, we argue the contrapositive. In particular, we analyze a collection of carefully-selected thresholds. If no threshold performs well, we show that some extremely high value must contribute to the optimal utility. In arguing Theorem 5.1, we will make use of our decomposition and derandomization from Lemmas 4.1 and 4.3, respectively.

Proof of Theorem 5.1. We will argue with respect to some integer $\alpha \geq 2$, and assume that no threshold obtains a β-approximation for any $\beta < 8\alpha$. Under this assumption, we show that there must be an action with value at least $(\alpha - 2)^{\alpha - 1}\mathrm{OPT}/8\alpha$ with positive probability. The analysis will proceed in three steps. First, we partition the optimal solution into α sets with roughly equal contribution to OPT. We then consider the thresholds based on each of these sets, and compare their utility to that from the sets themselves. Finally, we combine the resulting inequalities to show that the only way all can hold simultaneously is if the bias of one of these thresholds is extreme. This will imply the existence of a comparably high value.

Before partitioning A^*, we decompose the optimal utility. Assume every action in A^* is selected with positive probability. Following Lemma 4.1, write $\mathrm{OPT} = f(A^*) = \textsc{Sur} + \textsc{BDif}$, and write $b_{\max} = \max\{b_i \,|\, i \in A^* \cup \{0\}\}$ and $\underline{u} = \max(b_{\max}, b_0)$. By Lemma 4.2, it must be that $\textsc{Sur} < \mathrm{OPT}/\alpha$, or else the grand threshold $A_{\underline{u}}$ would be an α-approximation, contradicting the nonexistence of any $\beta < 4\alpha$-approximation. We may therefore focus our analysis on $\textsc{BDif} = \int_{\mathbf{v}} \underline{u} - b_{g(A^*,\mathbf{v})}\, dF(\mathbf{v})$. It must again be that no threshold obtains better than an 4α-approximation to $\textsc{BDif} > (1 - 1/\alpha)\mathrm{OPT} \geq \mathrm{OPT}/2$.

Next, note that no one action can comprise a large fraction of $\textsc{BDif}$. More precisely, let

$$\mathrm{OPT}^i = \int_{\mathbf{v}} v_i \mathbb{I}[g(A^*, \mathbf{v}) = i]\, dF(\mathbf{v})$$
$$\textsc{BDif}^i = \int_{\mathbf{v}} (\underline{u} - b_i) \mathbb{I}[g(A^*, \mathbf{v}) = i]\, dF(\mathbf{v})$$

denote the contribution of action i to OPT and $\textsc{BDif}$, respectively. Since $g(A^*, \mathbf{v}) = i$ only if $v_i \geq (\underline{u} - b_i)$, we must have $\mathrm{OPT}^i \geq \textsc{BDif}^i$. Lemma 5.1 implies that we may obtain OPT^0 from a threshold set. We must therefore have that $\textsc{BDif}^0 < \textsc{BDif}/2\alpha$. This in turn implies that no threshold can obtain better than a 2α-approximation to $\textsc{BDif} - \textsc{BDif}^0 \geq (1 - 1/2\alpha)\textsc{BDif} \geq \textsc{BDif}/2$. Since Lemma 5.1 allows us to obtain $\mathrm{OPT}^i \geq \textsc{BDif}^i$ from a threshold set for any i, we must also have that $\textsc{BDif}^i < (\textsc{BDif} - \textsc{BDif}^0)/2\alpha$ for all $i \in A^*$.

We now obtain a sequence of candidate thresholds by partitioning the actions in A^* based on their contribution to $\textsc{BDif}$. Let $m = |A^*|$, and relabel the actions in A^* as $1, \ldots, m$, with $b_i \leq b_{i+1}$ for all $i \in \{1, \ldots, m-1\}$. Actions not in A^* will be indexed $m+1, \ldots, n$ (with 0 keeping its label). We will partition A^* greedily to produce α subsets of roughly equal contribution to $\textsc{BDif}$. More precisely, define $k_0 = 0$, and for $j \in \{1, \ldots, \alpha\}$, define k_j recursively as the smallest $k \in \{k_{j-1} + 1, \ldots, m\}$ such that $\sum_{i=1}^{k} \textsc{BDif}^i \geq j(\textsc{BDif} - \textsc{BDif}^0)/\alpha$. Define the partition as $A^*(j) = \{k_{j-1} + 1, \ldots, k_j\}$ for all j. Since $\textsc{BDif}^i < (\textsc{BDif} - \textsc{BDif}^0)/2\alpha$ for all i, it must be that $A^*(j)$ is nonempty for $j \in \{1, \ldots, \alpha\}$. Further define $\textsc{BDif}(j) = \sum_{i \in A^*(j)} \textsc{BDif}^i$. We must also have $\textsc{BDif}(j) \geq (\textsc{BDif} - \textsc{BDif}^0)/2\alpha$ for all j.

Define our candidate thresholds $t_j = b_{k_j}$ (with $t_0 = \min_{i \in A^*} b_i$), and define the threshold sets $A_j = \{i \,|\, b_i \leq b_{k_j}\}$ for all j. To lower bound the principal utility from A_j, we apply Lemma 4.3 and consider $f(A_j \cap A^* \cup \{a(t_j)\})$ for some suitably constructed action $a(t_j)$ with bias t_j. We write $\underline{A}_j = A_j \cap A^* \cup \{a(t_j)\}$. Our analysis will lower bound the utility from $\underline{A}_j$ and upper bound the utility from $A^*(j)$. To this end, we will define several events:

- Let $\mathcal{E}_j$ be the event that the agent's preferred action from A^* is in $A^*(j)$. Note that the agent may still ultimately choose action 0 under either $\mathcal{E}_j$ or $\overline{\mathcal{E}}_j$, and that $\sum_j \Pr[\mathcal{E}_j] = 1$.
- Let $\mathcal{E}_j^{=}$ denote the event that the agent's favorite action in $A^*(j)$ is also their favorite in $\underline{A}_j$. Again, the agent may still ultimately choose action 0.
- Let $\mathcal{E}_j^{+}$ be the event that $g(\underline{A}_j, \mathbf{v}) \neq 0$.
- Let $\mathcal{E}_j^{\leq}$ be the event that $v_0 + b_0 \leq v_{a(t_j)} + b_{a(t_j)}$.

For all j, we must have $v_{a(t_j)} + b_{a(t_j)} \geq \underline{u}$. If this inequality was violated for some j, then

$\mathcal{E}_j \subseteq \mathcal{E}_j^{=}$. This would imply

$$\begin{aligned}
f(\underline{A}_j) &= \int_{\mathbf{v}} v_{g(\underline{A}_j,\mathbf{v})}\, dF(\mathbf{v}) \\
&\geq \int_{\mathbf{v}} v_{g(\underline{A}_j,\mathbf{v})}\mathbb{I}[\mathcal{E}_j]\, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} v_{g(A^*,\mathbf{v})}\mathbb{I}[\mathcal{E}_j]\, dF(\mathbf{v}) \\
&\geq \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})})\mathbb{I}[\mathcal{E}_j]\, dF(\mathbf{v}) \\
&\geq \textsc{BDif}(j).
\end{aligned}$$

Since $\textsc{BDif}(j) \geq (\textsc{BDif} - \textsc{BDif}^0)/2\alpha$, the threshold t_j would be better than a 2α-approximation to $\textsc{BDif} - \textsc{BDif}^0$, contradicting our initial assumption. This in turn implies that in $\mathcal{E}_j^+$, since $v_{a(t_j)} + b_{a(t_j)} \geq \underline{u}$, we must have $v_{g(\underline{A}_j,\mathbf{v})} + b_{g(\underline{A}_j,\mathbf{v})} \geq v_{a(t_j)} + b_{a(t_j)} \geq \underline{u}$. It follows that $v_{g(\underline{A}_j,\mathbf{v})} \geq \underline{u} - b_{g(\underline{A}_j,\mathbf{v})} \geq \underline{u} - t_j$ in $\mathcal{E}_j^+$.

We lowerbound $f(\underline{A}_j)$ via the following inequalities, explained after their statement:

$$\begin{aligned}
f(\underline{A}_j) &= \int_{\mathbf{v}} v_{g(\underline{A}_j,\mathbf{v})}\, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} v_{g(\underline{A}_j,\mathbf{v})}\mathbb{I}[\mathcal{E}_j]\, dF(\mathbf{v}) + \int_{\mathbf{v}} v_{g(\underline{A}_j,\mathbf{v})}\mathbb{I}[\overline{\mathcal{E}}_j]\, dF(\mathbf{v}) \\
&\geq \int_{\mathbf{v}} v_{g(\underline{A}_j,\mathbf{v})}\mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^+ \cap \mathcal{E}_j^{=}]\, dF(\mathbf{v}) + \int_{\mathbf{v}} v_{g(\underline{A}_j,\mathbf{v})}\mathbb{I}[\overline{\mathcal{E}}_j \cap \mathcal{E}_j^+]\, dF(\mathbf{v}) \\
&\geq \int_{\mathbf{v}} (\underline{u} - b_{g(\underline{A}_j,\mathbf{v})})\mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^+ \cap \mathcal{E}_j^{=}]\, dF(\mathbf{v}) + \int_{\mathbf{v}} (\underline{u} - t_j)\mathbb{I}[\overline{\mathcal{E}}_j \cap \mathcal{E}_j^+]\, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})})\mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^+ \cap \mathcal{E}_j^{=}]\, dF(\mathbf{v}) + \int_{\mathbf{v}} (\underline{u} - t_j)\mathbb{I}[\overline{\mathcal{E}}_j \cap \mathcal{E}_j^+]\, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})})\mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^+ \cap \mathcal{E}_j^{=}]\, dF(\mathbf{v}) + (\underline{u} - t_j)\Pr[\overline{\mathcal{E}}_j \cap \mathcal{E}_j^+] \\
&\geq \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})})\mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^+ \cap \mathcal{E}_j^{=}]\, dF(\mathbf{v}) + (\underline{u} - t_j)\Pr[\overline{\mathcal{E}}_j \cap \mathcal{E}_j^{\leq}] \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})})\mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^+ \cap \mathcal{E}_j^{=}]\, dF(\mathbf{v}) + (\underline{u} - t_j)\Pr[\overline{\mathcal{E}}_j]\Pr[\mathcal{E}_j^{\leq}].
\end{aligned}$$

The second inequality follows from the fact that under $\mathcal{E}_j^+$, $v_{g(\underline{A}_j,\mathbf{v})} \geq \underline{u} - b_{g(\underline{A}_j,\mathbf{v})} \geq \underline{u} - t_j$. The third inequality follows from fact that $\mathcal{E}_j^{\leq} \subseteq \mathcal{E}_j^+$, and the final equality from the independence of $\overline{\mathcal{E}}_j$ and $\mathcal{E}_j^{\leq}$. All other lines follow immediately from definitions or basic probability.

Next, we upperbound the contribution of $A^*(j)$ to $\textsc{BDif}$. Further define two more events. Let $\mathcal{E}_j^*$ denote the event that the agent prefers $a(t_j)$ over all actions in $A^*(j)$, and let $\mathcal{E}^*$ denote the event that $g(A^*, \mathbf{v}) \neq 0$. Then we may write:

$$\begin{aligned}
\textsc{BDif}(j) &= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}^*] \, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}^* \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}^* \cap \overline{\mathcal{E}_j^{=}}] \, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}^* \cap \overline{\mathcal{E}_j^{=}}] \, dF(\mathbf{v}) \\
&\leq \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + (\underline{u} - t_{j-1}) \int_{\mathbf{v}} \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}^* \cap \overline{\mathcal{E}_j^{=}}] \, dF(\mathbf{v}) \\
&\leq \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + (\underline{u} - t_{j-1}) \int_{\mathbf{v}} \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{\leq} \cap \overline{\mathcal{E}_j^{=}}] \, dF(\mathbf{v}) \\
&\leq \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + (\underline{u} - t_{j-1}) \int_{\mathbf{v}} \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{\leq}] \, dF(\mathbf{v}) \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + (\underline{u} - t_{j-1}) \Pr[\mathcal{E}_j \cap \mathcal{E}_j^{\leq}] \\
&= \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + (\underline{u} - t_{j-1}) \Pr[\mathcal{E}_j] \Pr[\mathcal{E}_j^{\leq}].
\end{aligned}$$

The first equality follows from the fact that conditioned on $\mathcal{E}_j \cap \mathcal{E}_j^{=}$, it holds that $\mathcal{E}^*$ and $\mathcal{E}_j^{\leq}$ are the same event. The first inequality follows from the fact that all actions in $A^*(j)$ have bias at least t_{j-1}. The second inequality follows from the fact that conditioned on $\mathcal{E}_j \cap \overline{\mathcal{E}_j^{=}}$, it must be that the agent prefers $a(t_j)$ to all actions in $A^*(j)$. Hence, conditioned on $\mathcal{E}_j \cap \overline{\mathcal{E}_j^{=}}$, it must be that any time the agent prefers an action in $A^*(j)$ to the outside option, it must be that they also prefer $a(t_j)$. The last equality follows from the independence of $\mathcal{E}_j$ and $\mathcal{E}_j^{\leq}$. The remaining lines follow from definitions and basic probability.

Since no threshold is better than a 2α-approximation to $\textsc{BDif} - \textsc{BDif}^0$, and since $\textsc{BDif}(j) \geq (\textsc{BDif} - \textsc{BDif}^0)/2\alpha$, it must hold that our upper bound on $\textsc{BDif}(j)$ exceeds our lower bound on $f(\underline{A}_j)$. That is:

$$\begin{aligned}
&\int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + (\underline{u} - t_j) \Pr[\overline{\mathcal{E}}_j] \Pr[\mathcal{E}_j^{\leq}] \\
&\qquad \leq \int_{\mathbf{v}} (\underline{u} - b_{g(A^*,\mathbf{v})}) \mathbb{I}[\mathcal{E}_j \cap \mathcal{E}_j^{+} \cap \mathcal{E}_j^{=}] \, dF(\mathbf{v}) + (\underline{u} - t_{j-1}) \Pr[\mathcal{E}_j] \Pr[\mathcal{E}_j^{\leq}].
\end{aligned}$$

We may rearrange this as

$$\frac{\underline{u} - t_{j-1}}{\underline{u} - t_j} \geq \frac{1 - \Pr[\mathcal{E}_j]}{\Pr[\mathcal{E}_j]}.$$

Taking the product over all $j \leq \alpha - 1$ and canceling yields:

$$\frac{\underline{u} - t_0}{\underline{u} - t_{\alpha-1}} \geq \prod_{j=1}^{\alpha-1} \frac{1 - \Pr[\mathcal{E}_j]}{\Pr[\mathcal{E}_j]}.$$

Note that the righthand side is a convex, symmetric function of the $\Pr[\mathcal{E}_j]$s. Moreover, we have $\sum_{j=1}^{\alpha-1} \Pr[\mathcal{E}_j] \leq \sum_{j=1}^{\alpha} \Pr[\mathcal{E}_j] = 1$. Hence, minimizing the righthand side as a function of the $\Pr[\mathcal{E}_j]$s yields a minimum at $\Pr[\mathcal{E}_j] = 1/(\alpha-1)$ for all j, and hence $(\underline{u} - t_0)/(\underline{u} - t_{\alpha-1}) \geq (\alpha-2)^{\alpha-1}$. Note also that $\textsc{BDif}(\alpha) \geq (\textsc{BDif} - \textsc{BDif}^0)/2\alpha$. Since $\textsc{BDif}(\alpha) \leq (\underline{u} - t_{\alpha-1}) \Pr[\mathcal{E}_\alpha] \leq (\underline{u} - t_{\alpha-1})$, we obtain:

$$\underline{u} - t_0 \geq \frac{(\alpha-2)^{\alpha-1}}{2\alpha} (\textsc{BDif} - \textsc{BDif}^0).$$

Since every action in A^* is selected with positive probability, and since $t_0 = \min_{i \in A^*} b_i$, it must be that some action in A^* has value at least $\underline{u} - t_0$ with positive probability. Since $\text{BDIF} - \text{BDIF}^0 \geq \text{BDIF}/2 \geq \text{OPT}/4$, we obtain the desired lower bound on $v_{\max}/\text{OPT}$.

□

6 Threshold Delegation with Correlated Values

In the previous sections, we showed that under independently-distributed values, simple threshold rules obtain a close approximation the optimal principal utility. We now allow arbitrarily correlated values and show that the situation worsens considerably. Assuming the value distribution is discrete, prove tight a approximation guarantee for the principal's best threshold policy, showing that it is a $\Theta(\log p_{\min}^{-1})$-approximation, where $p_{\min}$ denotes the lowest probability mass of any value profile realization. Hence, absent independence, threshold policies still perform well under low levels of uncertainty, but their performance gradually degrades as the uncertainty grows more extreme. We state our results formally below, starting with our upper bound.

Theorem 6.1. *There always exists a threshold policy which is a $4\log(p_{min}^{-1})$-approximation where p_{min} is the mass of the least likely value profile.*

Proof. Let OPT be the optimal delegation set, and let t_0 be the maximum bias across actions in OPT. Let B be the random variable that corresponds to the bias of the action chosen in OPT. Let t_1 be the bias threshold such that $\Pr[B \in [t_1, t_0]] \leq \frac{1}{2}$ and $\Pr[B \in [0, t_1]] \geq \frac{1}{2}$. Let $OPT([t_1, t_0])$ be the principal utility generated conditioned on $B \in [t_1, t_0]$. We claim that the principal utility generated by using a best threshold out A_{t_1} or A_{t_0} achieves principal utility at least $\frac{1}{4}OPT([t_1, t_0])$.

First consider the threshold A_{t_0}. Consider any realization where OPT picks an action with bias at least t_0. Let v, b be the value and bias of the action chosen by OPT and let v', b' be the value and bias of the action chosen by A_{t_0} respectively. Since the action chosen by OPT is available in A_{t_0} it must be that

$$v' + b' \geq v + b \Leftrightarrow v - v' \leq b' - b.$$

Note that from our assumptions $b' \leq t_0$ and $b \geq t_1$, therefore the pointwise loss of A_{t_0} compared to OPT is at most $t_0 - t_1$ in this event. As a result we can lower bound the principal utility of A_{t_0} as follows:

$$f(A_{t_0}) \geq \Pr[B \in [t_1, t_0]](OPT([t_1, t_0]) - t_0 + t_1) \geq \frac{1}{2}(OPT([t_1, t_0]) - t_0 + t_1).$$

Second, consider the threshold A_{t_1}. Every time OPT chooses an action with bias less than or equal to t_1 this action is also available to A_{t_1}. Note that if such action is chosen its agent utility must be at least t_0 otherwise the action with the maximum bias would have been chosen instead. The action chosen in A_{t_1} therefore must have at least agent utility t_0. Since the bias is at most t_1 this means that the principal utility from that action is at least $t_1 - t_0$. We conclude that

$$f(A_{t_1}) \geq (t_0 - t_1) \Pr[B \in [0, t_1] \geq \frac{1}{2}(t_0 - t_1).$$

As a result,

$$f(A_{t_0}) + f(A_{t_1}) \geq \frac{1}{2}[OPT([t_1, t_0]) - t_0 + t_1 + t_0 - t_1] = \frac{1}{2}OPT([t_1, t_0])].$$

Hence, the best out of both of these sets provides at least $\frac{1}{4}OPT([t_1, t_0])$ principal utility. Note that OPT gets at most $\Pr[B \in [t_1, t_0]]OPT([t_1, t_0])$ utility from this event therefore the best of these threshold provides a 4-approximation to the events' contribution to OPT's principal utility.

We have shown that there exists a threshold that approximates the utility of OPT conditioned that $B \in [t_1, t_0]$ that is

$$\Pr[B \in [t_1, t_0]]OPT([t_1, t_0]).$$

Let us focus on the remaining principal utility that is obtained by OPT in the event that $B \in [0, t_2]$ where t_2 bias of the most-biased action smaller than t_1. One key observation is that this event happens with probability at most 1/2 by our choice of t_1 since $\Pr[B \in [t_1, t_0]] \geq \frac{1}{2}$.

If we consider the conditional distributions on this event we can repeat the same analysis to prove that there exists bias threshold $t_3 < t_2$ such that $Prob[B \in [t_3, t_2) \mid B \in [0, t_2]] \geq 1/2$ and also

$$\max\{f(A_{t_2}), f(A_{t_3})\} \geq \frac{1}{4}Prob[B \in [0, t_2]]OPT([t_3, t_2]) \geq \frac{1}{4}\Pr[B \in [t_3, t_2]]OPT([t_3, t_2]),$$

which corresponds to 4-approximation to the contribution to OPT solution's principal utility in this interval. Note that

$$\Pr[B \in [0, t_3) \mid B \in [0, t_2]] \leq 1/2 \Rightarrow \Pr[B \in [0, t_3]] \leq 1/4.$$

Repeating this process shrinks the probability of the remaining probability space by half. Let m be the maximum number of times we can repeat this process. There are two ways this process can stop. Either we are left with a single action or the probability that B (OPT bias) is strictly below the last used threshold is 0. Since the minimum probability of any realization is p_{min} and each time we repeat this process the probability is shrunk by half m cannot be larger than $\log p_{min}^{-1}$.

This process generates m disjoint events $B \in [t_{2i+1}, t_{2i}]$ for $i = 0, \ldots, m-1$ such that

$$OPT = \sum_{i=0}^{m-1} OPT([t_{2i+1}, t_{2i}])Prob[B \in [t_{2i+1}, t_{2i}]]$$

and in addition

$$\max\{f(A_{t_{2i+1}}, A_{t_{2i}}\} \geq \frac{1}{4}OPT([t_{2i+1}, t_{2i}])\Pr[B \in [t_{2i+1}, t_{2i}]]$$

If we combined these two equations together we get that

$$\max_{i \in \{0, \ldots, 2m-1\}} f(A_{t_i}) \geq \frac{m}{4}OPT$$

Since $m \leq \log p_{min}^{-1}$ we get that the best possible threshold is at least a $4 \log p_{min}^{-1}$ approximation.

□

Matching lower bound. The above analysis is tight. We show in Appendix A.1 that our analysis in Theorem 6.1 is tight up to a constant factor. We do so by providing instances where no threshold policy can outperform the logarithmic approximation ratio.

Theorem 6.2. *There exists a family of instances where no threshold policy is better than a* $\Omega(\log p_{min}^{-1})$*-approximation.*

In Appendix A.2, we also prove the following supplementary hardness result for the case with discretely-distributed, correlated values:

Theorem 6.3. *With correlated, discrete values, there exists a constant c such that it is NP-hard to compute a mechanism with approximation factor better than c.*

To prove this result, we reduce from bounded-degree vertex cover, which is similarly hard to approximate.

Acknowledgments

Acknowledgments removed for the sake of blind review.

A Supplementary materials

A.1 Proof of Theorem 6.2

In this appendix, we show that the logarithmic approximation upper bond of Theorem 6.1 is tight, up to a constant factor. That is, no threshold algorithm can perform better than $\log p_{min}^{-1}$. To prove the tightness of our analysis in Section 6, we construct an infinite family of instances.

For any $k \geq 2$, consider an instance with $n = 2k - 1$ actions. The correlated distribution has $m = 2^k - 1$ value profile realizations. We construct a value matrix where each row correspond to an action and each column corresponds to a realization. Therefore, the value at cell $V_{i,j}$ gives the value of action i at realization value profile j. The distribution over value profiles simply selects and value realization uniformly at random.

$$V = \left[\begin{array}{c|c|c|c|c|c|c|c|c|c|c}
2^k & & & & & & & & & & \\ \hline
2^{k-1}+2\epsilon & & & & & & & & & & \\ \hline
 & 2^{k-1} & 2^{k-1} & & & & & & & & \\ \hline
2^{k-2}+3\epsilon & 2^{k-2}+3\epsilon & 2^{k-2}+3\epsilon & & & & & & & & \\ \hline
 & & & 2^{k-2} & 2^{k-2} & 2^{k-2} & 2^{k-2} & & & & \\ \hline
2^{k-3}+4\epsilon & 2^{k-3}+4\epsilon & 2^{k-3}+4\epsilon & 2^{k-3}+4\epsilon & 2^{k-3}+4\epsilon & 2^{k-3}+4\epsilon & 2^{k-3}+4\epsilon & & & & \\ \hline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & & \\ \hline
2+O(k\epsilon) & 2+O(k\epsilon) & 2+O(k\epsilon) & 2+O(k\epsilon) & 2+O(k\epsilon) & 2+O(k\epsilon) & 2+O(k\epsilon) & \dots & & & \\ \hline
 & & & & & & & \dots & \multicolumn{3}{c}{\underbrace{2 \quad \dots \quad 2}_{2^{k-1}}}
\end{array}\right] \tag{1}$$

Note that all the empty entries in the above matrix are zero, and are removed to make the structure of the matrix more apparent. Also, every solution corresponds to eliminating a a set of rows. For each realization (column) the row with maximum agent utility is selected. The optimal solution is to select set of odd actions (with size k). The colored entries indicate (realization,action) pairs that contribute to the optimal principal's utility (OPT). The optimal utility is equally divided between the odd rows, leaving 2^k for each one: the first row has 2^k in the first column, the third row has 2^{k-1} in columns 2 and 3, the fifth row has 2^{k-2} over the next 4 columns and so on. In the example, the even actions are constructed to lower the principal's utility whenever they are included in a threshold solution. In every state (column), we divide the colored utility by 2 to find the utility of the next row, and keep dividing by 2 to complete the subsequent even rows. The ϵ terms are added to break the ties and are of little importance.

Next, we define the bias: we set $b_1 = 0$, and the rest of actions have the following bias:

$$b_{2i+1} = \sum_{j=1}^{i} 2^{k-j}, \quad b_{2i} = b_{2i+1} - \epsilon, \qquad i \in \{1, ..., k-1\}. \tag{2}$$

Now that all the parameters are set, it is easy to verify that given the set of odd actions (rows), the agent will indeed pick the colored entries. This generates the optimal utility, since it is optimal in every single realization. Since each value profiled is realized with probability $\frac{1}{m}$ the optimal expected utility is equal to:

$$OPT = \frac{k \times 2^k}{m}.$$

However, the best threshold solution in the constructed instance is to allow the entire set of actions (Ω). To see this, assume that the principal allows actions with bias less than or equal to $b_{2\ell-1}$ for some $\ell \leq k$. (Thresholds set at even-indexed actions can be easily shown to be suboptimal.) Note that every even action is preferred by the agent to any other action with less bias. Therefore, the only actions chosen by the agent are $2\ell - 1$ or $2\ell - 2$ In this case, the principal will get utility of $2^{k-\ell+1} + O(\ell\epsilon)$ from the first $2^\ell - 1$ states, and zero from the remaining states.

Observe that the overall utility $(2^\ell - 1) \times 2^{k-\ell+1}$ is an increasing function in ℓ, meaning that the best strategy for the principal is to not limit the agent. In this case, the agent will pick the penultimate action in the first half of columns, and the last action for the second half, generating utility of (almost) 2 for principal in every state. More precisely, we have:

$$APX = 2 + O(k\epsilon).$$

We get the desired lower bound by dividing the above objectives:

$$\frac{OPT}{APX} = \frac{k \times 2^k}{2n + O(nk\epsilon)} \cong \frac{k}{2} \cong \frac{\log n}{2}.$$

Example A.1. In order to make sure that the above construction is clear, here we present the full matrices for the case of $k = 3$, which translates into $n = 5$ actions and $m = 7$ realizations. The value matrix in this case is

$$V = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4+2\epsilon & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 4 & 0 & 0 & 0 & 0 \\ 2+3\epsilon & 2+3\epsilon & 2+3\epsilon & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 & 2 & 2 \end{bmatrix}$$

Calculating the bias in (2) results in:

$$\mathbf{b} = (0, 4-\epsilon, 4, 6-\epsilon, 6)$$

It is clear that the value matrix V is non-negative, and the agent's utility $V + B$ will be:

$$V + B = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 8+\epsilon & 4-\epsilon & 4-\epsilon & 4-\epsilon & 4-\epsilon & 4-\epsilon & 4-\epsilon \\ 4 & 8 & 8 & 4 & 4 & 4 & 4 \\ 8+2\epsilon & 8+2\epsilon & 8+2\epsilon & 6-\epsilon & 6-\epsilon & 6-\epsilon & 6-\epsilon \\ 6 & 6 & 6 & 8 & 8 & 8 & 8 \end{bmatrix}$$

Observe that $OPT = 24/7$ by the set of odd actions $\{1, 3, 5\}$, while $APX = (14 + 9\epsilon)/7$ from the entire set of actions $\Omega = \{1, 2, 3, 4, 5\}$.

A.2 Proof of Theorem 6.3

Proof. We give a reduction from the bounded degree vertex cover problem, i.e., the vertex cover problem on graphs with degree at most B (constant). This problem is known to be APX-hard Clementi and Trevisan (1999). Consider an instance of the bounded degree vertex cover problem $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\tilde{n}$ nodes and $\tilde{m}$ edges (where $\tilde{m} \leq B \cdot \tilde{n}/2 = O(\tilde{n})$).[4]

We construct an instance of the delegation problem with $\tilde{n}+1$ actions with action a_i corresponding to node i and an additional "default" action a_0. All actions have 0 bias apart from a_0 which has bias -1. The correlated distribution of the actions values is defined as follows: we pick an edge $e = \{i, j\}$ or some node i uniformly at random, i.e., each element with probability $(\tilde{m} + \tilde{n})^{-1}$

If we picked some edge $e = \{i, j\}$ then we assign value 5 to actions a_i and a_j, 2 to the default action a_0, and 0 for all other actions. If we picked a node i we assign value 2 to a_i and a_0 (default action) and 0 for all other actions.

We claim that the optimal solution of the delegation problem produces a utility of $(5\tilde{m} + 3\tilde{n} - \tilde{k})/(\tilde{m} + \tilde{n})$ for the principal, where $\tilde{k}$ is the size of the smallest vertex cover of $\mathcal{G}$. To see this, first note that any solution $\mathcal{S} \subseteq \mathcal{V}$ can be improved by including a_0, since a_0 has a negative bias. Any time the agent would choose a_0, it is the optimal choice for the principal as well. We therefore only consider solutions containing a_0.

Now if $\mathcal{S}$ is a vertex cover of $\mathcal{G}$ with $|\mathcal{S}| = \tilde{k}$, consider the corresponding delegation set where the principal allows actions $\{a_i : i \in \mathcal{S}\} \cup \{a_0\}$. If we generate the values by picking an edge, the agent will pick the action corresponding to one end of that edge (one is guaranteed to be in the cover $\mathcal{S}$) to get a utility of 5 compared to $2 - 1$ achievable from the default action. This choice will also generate utility of 5 for the principal, which makes $5\tilde{m}$ in total. If the utility is generated by picking node i the agent will pick action a_i which generates the utility of 2 for both principal and agent. This will make $2\tilde{k}$ in total. Finally, if the utilities are generated using some node $i \in \mathcal{V} \backslash \mathcal{S}$ the agent picks the default action which generates a utility of $2 - 1$ for the agent but $2 + 1$ for the principal. This will give $3(\tilde{n} - \tilde{k})$ in total. As a result the principal utility in expectation is $(5\tilde{m} + 3\tilde{n} - \tilde{k})/(\tilde{m} + \tilde{n})$.

For the converse, consider an optimal solution A to the delegation problem. We show that the nodes corresponding to the actions in A (excluding the default action) induce a vertex cover; otherwise the solution can be improved. Assume that there exists an edge $e = \{i, j\}$ where neither a_i nor a_j is allowed in A. If we add action a_i to A, the principal gets a utility of 5 if the utilities are generated from pick edge e, compared to current utility of 3 from the default action. On the other hand, the utility of the principal decreases from 3 to 2 if the values are generated by action i. So the total utility of $A \cup \{a_i\}$ is more than A which contradicts the optimality of A. Therefore A should be a vertex cover (plus default action). This in turn implies that the utility is at most $(5\tilde{m} + 3\tilde{n} - \tilde{k})/(\tilde{m} + \tilde{n})$ where $\tilde{k}$ is the size of the minimum vertex cover.

Since $\tilde{m} = \Theta(\tilde{n})$ and the minimum vertex cover has size at least $\tilde{m}/B = \Omega(\tilde{n})$, a constant factor gap in the bounded degree vertex cover problem translates into a constant factor gap in the optimal solution of the delegation problem, which yields the desired hardness result. □

References

Philippe Aghion and Jean Tirole. 1997. Formal and real authority in organizations. *Journal of political economy* 105, 1 (1997), 1–29.

[4] To distinguish between the parameters of the vertex cover instance and the delegation instance, we use tilde ($\sim$) for the graph instance.

Ricardo Alonso and Niko Matouschek. 2008. Optimal delegation. *The Review of Economic Studies* 75, 1 (2008), 259–293.

Manuel Amador and Kyle Bagwell. 2010. On the optimality of tariff caps. *Unpublished working paper* (2010).

Attila Ambrus and Georgy Egorov. 2017. Delegation and nonmonetary incentives. *Journal of Economic Theory* 171 (2017), 101–135.

Ali Aouad, Vivek Farias, Retsef Levi, and Danny Segev. 2018. The approximability of assortment optimization under ranking preferences. *Operations Research* 66, 6 (2018), 1661–1669.

Mark Armstrong and John Vickers. 2010. A model of delegated project choice. *Econometrica* 78, 1 (2010), 213–244.

Curtis Bechtel and Shaddin Dughmi. 2021. Delegated Stochastic Probing. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Curtis Bechtel, Shaddin Dughmi, and Neel Patel. 2022. Delegated Pandora's box. *arXiv preprint arXiv:2202.10382* (2022).

Gerardo Berbeglia. 2016. Discrete choice models based on random walks. *Operations Research Letters* 44, 2 (2016), 234–237.

Gerardo Berbeglia and Gwenaël Joret. 2020. Assortment optimisation under a general discrete choice model: A tight analysis of revenue-ordered assortments. *Algorithmica* 82, 4 (2020), 681–720.

Pirmin Braun, Niklas Hahn, Martin Hoefer, and Conrad Schecker. 2022. Delegated Online Search. *arXiv preprint arXiv:2203.01084* (2022).

Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. 2021. Bayesian Agency: Linear versus Tractable Contracts. *arXiv preprint arXiv:2106.00319* (2021).

Andrea EF Clementi and Luca Trevisan. 1999. Improved non-approximability results for minimum vertex cover with density constraints. *Theoretical Computer Science* 225, 1-2 (1999), 113–128.

James M Davis, Guillermo Gallego, and Huseyin Topaloglu. 2014. Assortment optimization under variants of the nested logit model. *Operations Research* 62, 2 (2014), 250–273.

Antoine Désir, Vineet Goyal, Danny Segev, and Chun Ye. 2020. Constrained assortment optimization under the Markov chain–based choice model. *Management Science* 66, 2 (2020), 698–721.

Paul Dütting, Tim Roughgarden, and Inbal Talgam-Cohen. 2019. Simple versus optimal contracts. In *Proceedings of the 2019 ACM Conference on Economics and Computation*. 369–387.

Jacob B Feldman and Huseyin Topaloglu. 2017. Revenue management under the Markov chain choice model. *Operations Research* 65, 5 (2017), 1322–1342.

Alexander Frankel. 2014. Aligned delegation. *American Economic Review* 104, 1 (2014), 66–83.

Jason D Hartline. 2013. Mechanism design and approximation. *Book draft. October* 122 (2013), 1.

Jason D Hartline and Tim Roughgarden. 2009. Simple versus optimal mechanisms. In *Proceedings of the 10th ACM conference on Electronic commerce*. 225–234.

Bengt Robert Holmstrom. 1978. *On Incentives and Control in Organizations*. Ph.D. Dissertation.

Nicole Immorlica, Brendan Lucier, Jieming Mao, Vasilis Syrgkanis, and Christos Tzamos. 2018. Combinatorial Assortment Optimization. In *International Conference on Web and Internet Economics*. Springer, 218–231.

Jon Kleinberg and Robert Kleinberg. 2018. Delegated Search Approximates Efficient Search. In *Proceedings of the 2018 ACM Conference on Economics and Computation*. ACM, 287–302.

Eugen Kovác and Tymofiy Mylovanov. 2009. Stochastic mechanisms in settings without monetary transfers: The regular case. *Journal of Economic Theory* 144, 4 (2009), 1373–1395.

Guang Li, Paat Rusmevichientong, and Huseyin Topaloglu. 2015. The d-level nested logit model: Assortment and price optimization problems. *Operations Research* 63, 2 (2015), 325–342.

David Martimort and Aggey Semenov. 2006. Continuity in mechanism design without transfers. *Economics Letters* 93, 2 (2006), 182–189.

Nahum D Melumad and Toshiyuki Shibano. 1991. Communication in settings with no transfers. *The RAND Journal of Economics* (1991), 173–198.

Paat Rusmevichientong, David Shmoys, Chaoxu Tong, and Huseyin Topaloglu. 2014. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management* 23, 11 (2014), 2023–2039.

Dezsö Szalay. 2005. The economics of clear advice and extreme options. *The Review of Economic Studies* 72, 4 (2005), 1173–1198.

Kalyan Talluri and Garrett Van Ryzin. 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science* 50, 1 (2004), 15–33.