Optimal Mechanisms with Costly Information Acquisition∗

Jingfeng Lu†  Lixin Ye‡

August 2018

Abstract

We study optimal mechanisms in an auction environment where bidders are endowed with original estimates (“types”) about their private values and can further learn their true values of the object for sale by incurring an information acquisition cost. We demonstrate that the optimality of the generalized Myerson allocation rule is robust to our setting. When shortlisting is restricted to one single round, we show that optimal entry is to admit the set of most efficient bidders that maximizes expected highest virtual surplus adjusted by both the second-stage signal and entry cost. When shortlisting can be done sequentially, optimal entry requires that at each round, at most the bidder with the highest type be shortlisted provided that her expected contribution to the virtual surplus is positive conditional on all the information revealed up to this round. We construct specific payment rules such that our optimal shortlisting and allocation rules are both IR and IC implementable. Our analytical framework is also general enough to encompass many existing models in the literature on auctions with costly entry.

Keywords: Two-stage auctions, information acquisition, entry, sequential screening, optimal shortlisting, optimal mechanisms.
JEL Classification: D44, D80, D82.

1 Introduction

In high-valued asset sales, buyers often need to go through a due diligence process before developing final bids. Due diligence is usually a process to update or acquire information about the value of the asset for sale or to prepare for the bidding process (e.g., to establish qualifications to bid). This process is costly

∗We thank seminar participants at University of Michigan, University of British Columbia, MilgromFest at Stanford University, Hong Kong University of Science and Technology, City University of Hong Kong, University of New South Wales, University of Technology Sydney, the Midwest Economic Theory Conference, ICETA (Chengdu), NSF/CEME Decentralization Conference, and in particular, Dirk Bergemann, Tilman Börgers, Yeon-Koo Che, Jeff Ely, Peter Esö, Daniel Li, Hao Li, Jiangtao Li, Yunan Li, Bin Liu, Wei He, Audrey Hu, Preston McAfee, David Miller, Alessandro Pavan, Ilya Segal, Xianwen Shi, Vasiliki Skreta, Nora Szech, Juuso Toikka, Balázs Szentes, Jun Zhang, Mengxi Zhang, Charles Zheng for very helpful comments and suggestions. All remaining errors are our own.
†Department of Economics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260. Tel: (65)65166026, Email: ecsljf@nus.edu.sg.
‡Department of Economics, The Ohio State University, 449A Arps Hall, 1945 North High Street, Columbus, OH 43210. Tel.: (614)292-6883. Email: ye.45@osu.edu.
and is usually modeled as entry as it is closely monitored by the auctioneer. For a sale of an asset worth billions of dollars, the entry cost can run from tens of thousands to millions of dollars.\textsuperscript{1}

Given the substantial entry cost, how to coordinate agents’ costly information acquisition becomes one central issue in optimal mechanism design. The importance of coordinating bidders’ entry for the purpose of enhancing seller revenue as well as total surplus has been revealed as early as in Levin and Smith (1994), who find that seller revenue and total surplus can decrease with the number of potential bidders when bidder entry is symmetric. Clearly, the success of a sale very much depends on whether the adopted mechanism attracts the most qualified bidders conducting the due diligence process and participating in the final sale. Mainly motivated by the need for entry screening and coordination, variants of two-stage selling mechanisms have emerged in the real world. A leading example of the two-stage auction procedure is known as indicative bidding, which is commonly used in sales of complicated business assets with very high values. It works as follows: the auctioneer actively markets the assets to a large group of potentially interested buyers. The potential buyers are then asked to submit non-binding bids, based on which a final set of bidders is shortlisted to advance to the second stage. The auctioneer then communicates only with these final bidders, providing them with extensive access to information about the assets,\textsuperscript{2} and finally runs the auction (typically using binding sealed bids). The use of this two-stage auction procedure is quite widespread. For example, in response to the restructuring of the electric power industry in the U.S. – which was designed to separate power generation from transmission and distribution – billions of dollars of electrical generating assets were divested through this two-stage auction procedure over the last two decades.\textsuperscript{3} This two-stage auction procedure is also commonly used in privatization, takeover, and merger and acquisition contests.\textsuperscript{4} Finally, it is commonly used in the institutional real estate market, which has an annual sales volume in the order of $60 to $100 billion.\textsuperscript{5}

Ye (2007) studies this two-stage auctions based on the assumption of costly information acquisition.\textsuperscript{6} Ye’s analysis suggests that the current design of indicative bidding cannot reliably select the most qualified bidders for the final sale, as there does not exist a symmetric, strictly increasing equilibrium bid function in the indicative bidding stage. In a more recent paper, by restricting indicative bids to a finite discrete domain, Quint and Hendricks (2018) show that a symmetric equilibrium exists in weakly-monotone strategies. But again, the highest-value bidders are not always selected, as bidder types “pool” over a finite number of bids. Without safely selecting the most qualified bidders for the final sale, the mechanism is unlikely optimal in maximizing expected revenue. What the optimal mechanism is in

\textsuperscript{1}A more detailed description of a typical due diligence process is provided in Section 5. According to Kleiberg, Waggoner and Weyl (2018), due diligence costs for the acquisition by a large technology company of a start-up are typically 20-40\% of the size of a deal.

\textsuperscript{2}Data rooms, which are described in Section 5, are typically set up to facilitate bidders’ due diligence process.

\textsuperscript{3}A list of industry examples using this two-stage auction design can be found in Ye (2007).

\textsuperscript{4}Leading examples include the privatization of the Italian Oil and Energy Corporation (ENI), the acquisition of Ireland’s largest cable television provider Cablelink Limited, and the takeover contest for South Korea’s second largest conglomerate Daewoo Motors.

\textsuperscript{5}See Foley (2003) for a detailed account.

\textsuperscript{6}Boone and Goeree (2009) provide an analysis of pre-qualifying auctions, which are similar to indicative bidding.
this two-stage auction environment remains an open question in the literature, and this paper seeks to provide an answer.

We model the situation as follows. Before entry, each potential bidder is endowed with a private signal, $\alpha_i$, which can be regarded as her pre-entry “type.” After entry (by incurring a common entry cost, $c$), each bidder $i$ fully observes her (private) value $v_i$, which is positively correlated with her pre-entry type. Given costly entry, it is not optimal for all potential bidders to be included in the final sale. As such, a general mechanism must consist of an entry-right allocation stage to shortlist bidders into the sale and the final stage to allocate the asset. The entry-right allocation stage may potentially consist of multiple rounds, depending on whether the shortlisting is conducted in one single round or multiple rounds. The shortlisting rule in a subsequent stage would depend on all information revealed in previous stages. Despite the potential complication due to both sequential screening and endogenous information acquisition, we are able to completely characterize the optimal revenue-maximizing selling mechanism with sequential information acquisition. Our analysis benefits greatly from recent developments in the literature of sequential screening (e.g., Courty and Li, 2000; Esö and Szentes, 2007; Pavan, Segal, and Toikka, 2014; and Bergemann and Wambach, 2015). In particular, our model resembles that of Esö and Szentes and we follow their main approaches including the orthogonalization technique in characterizing optimal dynamic mechanisms. Our paper differs from theirs in that buyer information acquisition is costly and endogenous in our model, and our analysis focuses on identifying the shortlisting mechanism that optimally orchestrates information acquisition of buyers.

Given the widespread use of two-stage auctions, we start our analysis with the case where shortlisting is completed simultaneously in one single round. In effect, we restrict our search of optimal mechanisms to the class of two-stage mechanisms, with the first stage allocating entry rights and the second stage allocating the asset. We first derive an integral form of the envelope formula as a necessary condition for incentive compatibility for our two-stage mechanisms, which extends the validity of the envelope theorem to dynamic auctions with costly and endogenous information acquisition. Based on the derived envelope formula, we are able to show that the optimal allocation rule of the asset requires that the asset be allocated to the bidder with the highest virtual value adjusted by the second-stage signal, same finding as identified by Esö and Szentes. Our analysis thus suggests that the optimality of the generalized Myerson optimal allocation rule (adjusted by second-round signals) is robust to the dynamic auction setting with costly and endogenous entry. The first-stage entry right allocation mechanism is new to the original Esö-Szentes framework, and we show that the optimal entry rule is to admit the set of bidders that gives rise to the maximum expected virtual surplus (adjusted by both the second-stage signal and entry cost). Alternatively, given the regularity assumption and that buyers are ex ante symmetric in our model, the optimal entry rule is to admit the bidders in descending order of their pre-entry “types”, the highest type first, the second highest type second, etc., provided that their marginal contribution to

---

7Early work on dynamic contracting with a single agent are due to Baron and Besanko (1984) and Riordan and Sappington (1987).
the expected virtual surplus is positive. Therefore, the optimal number of shortlisted bidders typically
depends on the reported type profile from the potential bidders, which is endogenously determined. We
then show that specific payment rules can be constructed in each stage to implement both optimal entry
and allocation rules truthfully.

In Section 4, we relax the restriction of single-round shortlisting and consider the case with poten-
tially multiple-round sequential shortlisting. The seller may now select a single bidder or any subset of
bidders at each round to go through due diligence and submit final bids, and if the seller is not satisfied
with any offer, he can go back to the unselected bidders and invite another bidder or another subset of
bidders to go through due diligence and submit final bids. This process can then repeat itself, until the
seller finds a satisfactory offer. Such mechanisms can be much more complicated. First of all, the seller
will need to determine the order of bidders to invite for conducting due diligence (i.e., who should be in-
vited first and who second, etc.). Given that bidders are heterogeneous before entry, it is desirable to make
the optimal “ordering” or “sequencing” of entry contingent on their pre-entry types. Our main results
with such a general analysis are as follows. First, the optimal final good allocation rule is the same as
characterized previously, that is, the object is allocated to the shortlisted bidder with the highest virtual
value \( w(\alpha_i, s_i) \), provided that it is positive. The optimal shortlisting rule should be modified, however,
in a way that at each round, at most one bidder (the one with the highest pre-entry type among all the
bidders outside the auction) is shortlisted, and a new bidder is shortlisted at a given round if and only if
conditional on all the information revealed up to this round, her expected contribution to the virtual sur-
plus is positive. The bidders are approached sequentially in the order of their first stage types, starting
from the highest type.\(^8\)

Other than the connection with sequential screening and dynamic auctions mentioned above, our
paper is related to the literature on auctions and mechanism design with information acquisition (see,
for example, Persico, 2000; Compte and Jehiel, 2001; Shi (2012); Rezende, 2013; Li (2018); Gershkov,
Moldovanu and Strack (2018); and Zhang (2018)). These papers either study bidders’ incentives to ac-
quire information in different specific auction formats or consider single stage optimal mechanism design.
Our paper differs from theirs in that we follow the normative approach to identify optimal dynamic mech-
anism with information acquisition. Our paper is also closely related to Krähmer and Strausz (2011),
who study procurement contracts with pre-project information acquisition, and Halac, Kartik, and Liu
(2016), who consider optimal dynamic contracts with experimentation. Unlike in our model, information
acquisition in their models is unobservable and thus not contractible, so they have to deal with both ad-
verse selection and moral hazard in their analysis. In our model, since information acquisition is modeled
as entry, moral hazard is absent from our analysis. Our paper differs from theirs also in that we work
with multiple agents/bidders, while there is only one agent in their model.

\(^8\)In the environment where bidders are not endowed with pre-entry private information, Crémer, Spiegel, and Zheng (2009)
find that an ex post efficient auction with sequential entry is both ex ante efficient and revenue maximizing. In our setting,
bidders are endowed with both pre-entry and post-entry private information, which dramatically complicates the analysis due
to additional incentive compatibility conditions.
To the extent that information acquisition is modeled as entry, our paper is closely related to the growing literature on auctions with costly entry. This literature can be summarized into three branches. In the first branch, bidders are assumed to possess no private information before entry and they learn their private values or signals only after entry (see, for example, McAfee and McMillan, 1987; Engelbrecht-Wiggans, 1993; Tan, 1992; Levin and Smith, 1994; Ye, 2004; and Jehiel and Lamy (2015)). In the second branch, it is assumed that bidders are endowed with private information about their values but have to incur entry costs to participate in an auction (see, for example, Samuelson, 1985; Stegeman, 1996; Campbell, 1998; Menezes and Monteiro, 2000; Tan and Yilankaya, 2006; Cao and Tian, 2009; and Lu, 2009). Finally, in the third branch, bidders are endowed with some private information before entry, and are able to acquire additional private information after entry (Ye, 2007; Quint and Hendricks, 2018). The framework in this current paper nests all the models mentioned above as special cases. Our paper thus characterizes optimal mechanisms for a very general framework in the literature on auctions with costly entry.

Our paper is also related to a literature on search, which is originally inspired by Weitzman (1979) who studies the so-called Pandora’s problem of infinite sequential search with recall and establishes the well-known Pandora rule. Crémer, Spiegel, and Zheng (2009) extends this model to an auction context. Szech (2011) and Lee and Li (2018) study the cases with a search deadline. Olszewski and Weber (2015) study a generalized version of the Pandora’s problem by allowing more general form of utility functions. Doval (2018) further allows the possibility of uninspected box. Our analysis differs from these studies by allowing bidders to be endowed with pre-entry private information about their private values. In this sense, our setting resembles that of Kleiberg, Waggoner, and Weyl (2018) who find that when bidders are endowed with pre-entry private information, descending price auction (among standard auctions) optimally coordinates buyer search. Unlike their work, we adopt a full-fledged dynamic mechanism approach to identify the optimal shortlisting (entry coordinating) rule and asset allocation rule jointly.

Our research is also related to a small literature on auctions of entry rights. Fullerton and McAfee (1999) introduce auctions for entry rights to shortlist contestants for a final tournament. Ye (2007) extends their approach to the setting of two-stage auctions described above. Our current approach differs from theirs in the way the set of finalists is determined: while in their approach the number of finalists to be selected is fixed and pre-announced, in our entry right allocation mechanism the selection of shortlisted bidders is contingent on the reported bid profile, making the number of finalists endogenously determined. For this reason the entry right allocation mechanism examined in this research is more general.

In another relevant paper, Lu and Ye (2013) explore optimal two-stage mechanisms in an environment where bidders are characterized by heterogeneous and private information acquisition costs before entry. In that setting the pre-entry “type” is the entry cost, which is neither correlated to nor part of the

---

9See Bergemann and Välimäki (2006) for a thoughtful survey of this literature.
10In fact, it resembles multi-unit auctions with endogenously determined supply (see, e.g., McAdams, 2007).
value of the asset for sale. As such, there is no benefit to make the second-stage mechanism contingent on the reports of the pre-entry types, resulting in a much simpler characterization of optimal mechanisms. The setting in this current paper is different, as the pre-entry “type” is correlated to the value of the asset, hence there are potential gains to make the second-stage mechanism contingent on first-stage reports. Indeed, in our current setting, the optimal allocation and payment rules in a subsequent stage does depend on report(s) from the previous stage(s). Therefore the characterization of optimal mechanisms is more demanding in this work, and the implementation of the optimal mechanism is also more sophisticated.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 characterizes the optimal two-stage mechanism when shortlisting is restricted to one single round. Section 4 characterizes the optimal multi-stage mechanisms when multi-stage sequential shortlisting is allowed. Section 5 discusses main assumptions/restrictions in our analysis and the robustness of our results. Section 6 concludes.

2 The Model

The information structure in our model is closest to that of Esö and Szentes (2007). The main differences are that in Esö and Szentes, the additional information is controlled by the seller, and they focus on the seller’s incentive to disclose (without observing) additional signals to the buyers. In our setting, however, it is costly for the bidders to acquire additional information, and we focus on the bidders’ incentive for information acquisition (entry). In addition, all buyers are included in the final sale in Esö and Szentes, but due to costly entry in our setting, generally not all buyers should be included in the final auction. As such, we will additionally consider entry or shortlisting mechanisms – which is the major difference from the analysis in Esö and Szentes.

Formally, a single indivisible asset is offered for sale to \( N \) potentially interested buyers. The seller and bidders are assumed to be risk neutral. The seller’s own valuation for the asset is normalized to 0. Buyer \( i \)'s true valuation for the asset is \( v_i \). However, initially she only observes a noisy signal of it, \( \alpha_i \), which is her private information and can be interpreted as her original “type”. After incurring a common information acquisition cost (or entry cost) of \( c(>0) \), bidder \( i \) fully observes her ex post value, \( v_i \). The pairs \((\alpha_i, v_i)\) are assumed to be independent across \( i \).\(^{11}\)

Ex ante, \( \alpha_i \) follows distribution \( F(\cdot) \) with its associated density \( f(\cdot) \) on support \([\underline{\alpha}, \overline{\alpha}]\). We assume that \( f \) is positive on the interval \([\underline{\alpha}, \overline{\alpha}]\) and satisfies the monotone hazard rate condition; that is, \( f/(1-F) \) is weakly increasing. Given \( \alpha_i \), the ex post value \( v_i \) follows distribution \( H_{\alpha_i} \equiv H(\cdot|\alpha_i) \) with its density \( h_{\alpha_i} \equiv h(\cdot|\alpha_i) \) over support \([\underline{v}, \overline{v}] \subset \mathbb{R} \).\(^{12}\) The values \( N \) and \( c \) and distributions \( F \) and \( H_{\alpha_i} \) are all common

---

\(^{11}\)As in Esö and Szentes (2007) and Pavan, Segal, and Toikka (2014), this assumption rules out the possibility of full rent extraction (Crémer and McLean, 1988).

\(^{12}\)Following the dynamic mechanism design literature, we assume that the support of \( v_i \) is independent of the first-stage signal \( \alpha_i \).
knowledge.

Following the signal orthogonalization technique introduced by Esö and Szentes (2007), there exist functions $u$ and $s_i$, such that $u(\alpha_i, s_i) \equiv v_i$, where $u$ is strictly increasing in both arguments, and $s_i$ is independent of $\alpha_i$. In particular, $s_i$ can be constructed as follows:

$$s_i = H(v_i|\alpha_i),$$

which is the percentile of the value realization to bidder $i$. Thus given type $\alpha_i$ and signal $s_i$, the valuation can be computed as

$$v_i = H^{-1}_a(s_i) \equiv u(\alpha_i, s_i).$$

We will denote the c.d.f. of $s_i$ by $G_i$.

We maintain the following assumptions that are adopted in Esö and Szentes (2007):

**Assumption 1.** $(\partial H_a(v)/\partial \alpha)/h_a(v)$ is increasing in $v$.

**Assumption 2.** $(\partial H_a(v)/\partial \alpha)/h_a(v)$ is increasing in $\alpha$.

Esö and Szentes show that Assumption 1 is equivalent to $u_{12} \leq 0$ and Assumption 2 is equivalent to $u_{11}/u_1 \leq u_{12}/u_2$. Assumption 1 thus states that the marginal impact of the new information on buyer $i$’s value is decreasing in her type $\alpha_i$. Assumption 2 implies that an increase in $\alpha_i$, holding $u(\alpha_i, s_i)$ constant, weakly decreases the marginal value of $\alpha_i$. Assumptions 1 and 2 can thus be interpreted as a kind of substitutability in buyer $i$’s posterior valuation between $\alpha_i$ and $s_i$.

**Example 1.** (Ye, 2007): Each potential bidder is endowed with a private value component $\alpha_i$ before entry; after entry, each buyer learns another private value component $s_i$, where $s_i$ is independent of $\alpha_i$. The ex post value $u(\alpha_i, s_i) = \alpha_i + s_i$. By the linearity of $u(\alpha_i, s_i)$, Assumptions 1 and 2 hold.

**Example 2.** (Adapted from Esö and Szentes, 2007): $v_i$ is drawn from a normal distribution with mean $\mu$ and precision (inverse variance) $\tau_0$. The pre-entry type, $\alpha_i$, is normally distributed with mean $\nu_i$ and precision $\tau_\nu$. After entry, the buyer can observe her true value, $v_i$. It is clear that $v_i$ and $\alpha_i$ are strictly affiliated. The distribution of $\alpha_i$, which is normal, satisfies the hazard rate condition. The c.d.f. of $v_i$ conditional on $\alpha_i$, $H_{\alpha_i}$, is normal with mean $(\tau_0 \mu + \tau_\nu \alpha_i)/(\tau_0 + \tau_\nu)$ and precision $\tau_0 + \tau_\nu$. Define $s_i = H_{\alpha_i}(v_i)$ and let $u(\alpha_i, s_i) = H^{-1}_{\alpha_i}(s_i) \equiv v_i$. Obviously $u$ is strictly increasing in $s_i$. It can be verified that $u_1(\alpha_i, s_i) = \tau_\nu/(\tau_0 + \tau_\nu)$, which is a constant. Therefore, $u$ is linear and strictly increasing in $\alpha_i$. Hence Assumptions 1 and 2 hold.

---

13 The use of this technique has become standard in the literature (see, e.g., Pavan, Segal, and Toikka, 2014, and Bergemann and Wambach, 2015).

14 It is easily seen that $s_i$ is uniformly distributed over $[0,1]$, and is hence statistically independent of the initial information $\alpha_i$.

15 $G_i$ could be assumed to be uniform on $[0,1]$. More generally, all $s_i$’s satisfying $u(\alpha_i, s_i) \equiv v_i$ are positive monotonic transformation of each other (Lemma 1 in Esö and Szentes).
Since information acquisition is modeled as entry in our setting, we consider a mechanism design framework in which the seller exercises entry control. In Section 3, we restrict our analysis to two-stage mechanisms: the first stage is the entry right allocation mechanism, and the second stage is the private good provision mechanism. In Section 4, we will extend our analysis to multi-stage mechanisms allowing for sequential shortlisting.

We restrict our analysis to direct mechanisms where agents report their types truthfully at each stage on the equilibrium path. We assume that all shortlisted bidders are disclosed and the first-stage reported profile \( \alpha \) is revealed to all admitted bidders so that the first-stage entry allocation and payments are immediately verifiable.\(^{16}\) This revelation policy turns out to be “optimal”, in the sense that no other revelation policy (e.g., not revealing or partially revealing \( \alpha \)) can generate a higher expected revenue to the seller. For this reason, our restriction to fully revealing \( \alpha \) is without loss of generality in our search for optimal mechanisms. A detailed discussion is relegated to Section 5. In our paper, the principal has no control over the ways in which new information in a subsequent stage is revealed to bidders. A shortlisted bidder will be fully informed about her true value \( v_i \) after incurring the entry cost. As such, we are not concerned about the discriminatory information disclosure issue studied in Li and Shi (2017).

As in Esö and Szentes, we can focus on equivalent direct mechanisms that require bidders to report \( s_i \)'s, rather than \( v_i \)'s. Note that reporting \((\alpha'_i, v'_i)\) is equivalent to reporting \((\alpha'_i, s'_i = H_{\alpha'_i}(v'_i))\).

Let \( N = \{1, 2, ..., N\} \) denote the set of all the potential buyers and \( 2^N \) denote the collection of all the subsets (subgroups) of \( N \), including the empty set, \( \phi \).

3 ANALYSIS WITH SINGLE-ROUND SHORTLISTING

The first-stage mechanism is characterized by the shortlisting rule \( A^g(\alpha) \) and payment rule \( x_i(\alpha), i = 1, 2, ..., N \). Given the reported profile \( \alpha \), the shortlisting rule, \( A^g : [\alpha, \overline{\alpha}]^N \to [0, 1] \), assigns a probability to each subgroup \( g \in 2^N \), where \( \sum_{g \in 2^N} A^g(\alpha) = 1 \). The payment rule \( x_i : [\alpha, \overline{\alpha}]^N \to \mathbb{R} \), specifies bidder \( i \)'s first-stage payment given the reported profile \( \alpha \).

Given the first-stage reported profile \( \alpha \), and that group \( g \) is shortlisted, the second-stage mechanism is characterized by \( p_i^g(\alpha, s^g) \), the probability that the asset is allocated to buyer \( i \in g \), and \( t_i^g(\alpha, s^g) \), the payment to the seller made by buyer \( i \in g, \forall g \in 2^N \).

We start with the second stage. Suppose group \( g \) is shortlisted, and the profile \( \tilde{\alpha} \) reported in the first stage is revealed as public information to the shortlisted bidders.

First, suppose \( \alpha \) is truthfully reported at the first stage and group \( g \) is shortlisted. Assume that they follow the recommendation and incur the information acquisition cost \( c \) to discover \( s^g \).\(^{17}\)

Given the announced \( \alpha \) and \( s_i \), define the interim winning probability and expected payment rule as

---

\(^{16}\)In Esö and Szentes, there is no such need for interim verification, as their allocation and payment rules are executed at the end of the mechanism.

\(^{17}\)As will be shown, the equilibrium expected profit from going forward is positive for a buyer upon entry, so in equilibrium, a bidder does have an incentive to follow the recommendation to acquire (costly) information and participate in the final auction once admitted (as dropping out only results in zero profit).
\[ P_i^g(\alpha, s_i) = E_{s^g_i} P_i^g(\alpha, s^g_i) \] and \[ T_i^g(\alpha, s_i) = E_{s^g_i} T_i^g(\alpha, s^g_i) \], where \[ s^g_i = s^g \setminus \{s_i\} \], \( i \in g \) and \( g \in 2^\mathbb{N} \). Then bidder \( i \)'s second-stage interim expected payoff when she observes \( s_i \) but reports \( \hat{s}_i \) is as follows:

\[ \tilde{\pi}_i^g(\alpha; s_i, \hat{s}_i) = E_{s^g_i} [u(\alpha, s_i)p_i^g(\alpha, s^g_i, \hat{s}_i^g) - t_i^g(\alpha, \hat{s}_i, s^g_i)] = u(\alpha, s_i)p_i^g(\alpha, \hat{s}_i) - T_i^g(\alpha, \hat{s}_i). \]

The second-stage incentive compatibility (IC) conditions require that

\[ \tilde{\pi}_i^g(\alpha; s_i, \hat{s}_i) \leq \tilde{\pi}_i^g(\alpha; s_i, s_i), \forall g, \alpha, s_i, \hat{s}_i. \] (1)

First, the following lemma is standard in the traditional screening literature:

**Lemma 1.** Suppose \( \alpha \) is truthfully revealed from the first stage and \( P_i^g(\alpha, s_i), \forall i \in g \), is continuous and weakly increasing in \( s_i \) where \( g \) denotes the group being shortlisted, then the second-stage incentive compatibility condition (1) holds if and only if

\[ \tilde{\pi}_i^g(\alpha; s_i, \hat{s}_i) = \tilde{\pi}_i^g(\alpha; \hat{s}_i, \hat{s}_i) + \int_{\hat{s}_i}^{s_i} u_2(\alpha, \tau) P_i^g(\alpha, \tau) d\tau, \forall s_i > \hat{s}_i, \forall i \in g. \] (2)

(2) is an integral form of the envelope formula. Next, we consider the case when \( \hat{\alpha}_i \) instead of \( \alpha_i \) is reported by bidder \( i \) while others report their types truthfully. As demonstrated in Esö and Szentes (2007), whenever a bidder had misreported her type in the first stage, she would “correct” her lie in the second stage. Formally in our setting, suppose \( \alpha_{-i} \) is truthfully revealed from the first stage and the second-stage mechanism is incentive-compatible given a truthfully revealed \( \alpha \). Then buyer \( i \) of type \( \alpha_i \) who reported \( \hat{\alpha}_i \) in the first round will report \( \hat{s}_i = \sigma_i(\alpha_i, \hat{\alpha}_i, s_i) \) if she observes \( s_i \) in the second stage such that

\[ u(\alpha_i, s_i) = u(\hat{\alpha}_i, \sigma_i(\alpha_i, \hat{\alpha}_i, s_i)). \] (3)

Reporting \( \hat{s}_i \) after a lie \( \hat{\alpha}_i \) is equivalent to revealing \( v_i \) truthfully regardless of the first-stage report. The optimality of this strategy has been established in general for Markov environments by Pavan, Segal, and Toikka (2014). Our two-stage setting resembles the Markov environment defined in Pavan, Segal, and Toikka since the agents’ payoffs depend only on their second-stage true types (\( v_i \)’s) and the allocation outcome, but not on their first-stage true types. For this reason, an agent’s reporting incentive in the second stage depends only on her current type and her first-stage report, but not on her first-stage true type.

Note that \( \hat{s}_i \) does not depend on \( \alpha_{-i}, g, \) or \( s^g_i \). Define

\[
\tilde{\pi}_i^g(\alpha, \hat{\alpha}_i; s_i, \hat{s}_i) = E_{s^g_i} [u(\alpha, s_i)p_i^g(\alpha_{-i}, \hat{\alpha}_i, \hat{s}_i, s^g_i) - t_i^g(\alpha_{-i}, \hat{\alpha}_i, \hat{s}_i, s^g_i)] \\
= u(\alpha, s_i)p_i^g(\alpha_{-i}, \hat{\alpha}_i, \hat{s}_i) - T_i^g(\alpha_{-i}, \hat{\alpha}_i, \hat{s}_i); 
\]

18The existence of \( \sigma_i(\cdot, \cdot, \cdot) \) relies on the assumption that the support of \( v_i \) does not depend on the first-stage signal \( a_i \).
Lemma 3. \[ \tilde{\pi}^B_i(\alpha_i, \hat{\alpha}_i; \alpha_{-i}) = E_{s_i} \tilde{\pi}^B_i (\alpha, \hat{\alpha}_i; s_i, \hat{s}_i = \sigma_i(\alpha_i, \hat{\alpha}_i, s_i)). \]

\( \tilde{\pi}^B_i(\alpha_i, \hat{\alpha}_i; \alpha_{-i}) \) is the expected second-stage payoff for the type-\( \alpha_i \) bidder if she reported \( \hat{\alpha}_i \) in the first stage (and everyone else reported truthfully) given her opponents' types being \( \alpha_{-i} \). Parallel to Lemma 5 in Esö and Szentes, we can show the following lemma:

**Lemma 2.** Suppose \( \alpha_{-i} \) is truthfully revealed from the first stage and the second-stage mechanism is incentive-compatible given a truthfully revealed \( \alpha \). If buyer \( i \) of type \( \alpha_i \) who reported \( \hat{\alpha}_i \) in the first stage is shortlisted in group \( g_i \), her expected payoff from the second stage is given by

\[ \tilde{\pi}^{g_i}_i(\alpha_i, \hat{\alpha}_i; \alpha_{-i}) = \tilde{\pi}^{g_i}_i(\hat{\alpha}_i, \hat{\alpha}_i; \alpha_{-i}) + \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P^{g_i}_i(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) dy dG_i(s_i). \]  
(4)

Throughout, \( g_i \) will be used to denote the group including bidder \( i \). (4) should again be regarded as an integral form of the envelope formula: the winning probability (\( P^{g_i}_i \)) is now obtained when evaluating at \( \hat{s}_i = \sigma_i(y, \hat{\alpha}_i, s_i) \) (which is optimal given the first-round "lie" \( \hat{\alpha}_i \)). We are now ready to consider the first-stage IC mechanism.

Let \( \pi_i(\alpha_i, \hat{\alpha}_i) \) be the expected payoff (net of the entry cost) for a type-\( \alpha_i \) bidder who reports \( \hat{\alpha}_i \) in the first stage. By (3), we have

\[ \pi_i(\alpha_i, \hat{\alpha}_i) = E_{\alpha_{-i}} \left\{ \sum_{g_i} A^{g_i}(\hat{\alpha}_i, \alpha_{-i}) \left[ \tilde{\pi}^{g_i}_i(\alpha_i, \hat{\alpha}_i; \alpha_{-i}) - c \right] - x_i(\hat{\alpha}_i, \alpha_{-i}) \right\} \]  
\[ = E_{\alpha_{-i}} \left\{ \sum_{g_i} A^{g_i}(\hat{\alpha}_i, \alpha_{-i}) \left[ E_{s_i} \left\{ u(\alpha_i, s_i) P^{g_i}_i(\alpha_{-i}, \hat{\alpha}_i, \hat{s}_i) - T^{g_i}_i(\alpha_{-i}, \hat{\alpha}_i, \hat{s}_i) \right\} - c \right] - x_i(\hat{\alpha}_i) \right\}, \]  
(5)

where \( \hat{s}_i = \sigma_i(\alpha_i, \hat{\alpha}_i, s_i) \) and \( x_i(\hat{\alpha}_i) = E_{\alpha_{-i}} x_i(\hat{\alpha}_i, \alpha_{-i}) \).

The following lemma characterizes the bidder's expected payoff in an IC two-stage mechanism with costly entry.

**Lemma 3.** If the two-stage mechanism is incentive compatible and \( E_{\alpha_{-i}} A^{g_i}(\alpha_i, \alpha_{-i}) P^{g_i}_i(\alpha_i, \alpha_{-i}, s_i) \) is continuous in \( \alpha_i \) then buyer \( i \)'s expected payoff (as a function of her pre-entry type) can be expressed as

\[ \pi_i(\alpha_i, \alpha_i) = \pi_i(\alpha, \alpha) + \int_{\alpha}^{\alpha_i} \int u_1(y, s_i) \sum_{g_i} \left[ E_{\alpha_{-i}} A^{g_i}(y, \alpha_{-i}) P^{g_i}_i(y, \alpha_{-i}, s_i) \right] dG_i(s_i) dy. \]  
(6)

**Proof.** See Appendix. \( \square \)

Note that \( \sum_{g_i} \left[ E_{\alpha_{-i}} A^{g_i}(y, \alpha_{-i}) P^{g_i}_i(y, \alpha_{-i}, s_i) \right] \) is buyer \( i \)'s equilibrium probability of eventually winning the asset with signals \( (y, s_i) \) in our setting. Thus (6) is also an integral form of the envelope formula.
Under a set of regularity conditions, which basically require that each agent’s expected utility be a sufficiently well behaved function of her private information, Pavan, Segal, and Toikka (2014) show that the envelope formula continues to hold in the dynamic mechanism design setting. Lemma 3 can be regarded as an extension of their result to a dynamic mechanism design setting with costly information acquisition.

3.1 The Optimal Two-stage Mechanisms

We are now ready to derive the seller’s expected payoff from an IC two-stage mechanism. By Lemma 3, we have

$$E \pi_i(a_i, \alpha_i) = \pi_i(\alpha_i) + \int_a^{\alpha_i} \int_a^{\alpha_i} u_1(y, s_i) \cdot \sum_{g_i} \left[ E_{a_{-i}} A^g_i(y, \alpha_{-i}) P^g_i(y, \alpha_{-i}, s_i) \right] dG_i(s_i) dy F(a_i)$$

$$= \pi_i(\alpha_i) + \int_a^{\alpha_i} \left[ 1 - F(a_i) \right] \int u_1(a_i, s_i) \cdot \sum_{g_i} \left[ E_{a_{-i}} A^g_i(a_i, \alpha_{-i}) P^g_i(a_i, \alpha_{-i}, s_i) \right] dG_i(s_i) dF(a_i)$$

$$= \pi_i(\alpha_i) + E_a \left\{ \sum_{g_i} A^g_i(a_i, \alpha_{-i}) \left[ \int \left[ 1 - F(a_i) \right] u_1(a_i, s_i) P^g_i(a_i, \alpha_{-i}, s_i) dG_i(s_i) \right] \right\}.$$ 

The second equality above is due to Fubini’s Theorem. Thus

$$\sum_{i=1}^{N} E \pi_i(a_i, \alpha_i) = \sum_{i=1}^{N} \pi_i(\alpha_i) + E_a \left\{ \sum_{g_i} A^g(a) E_s \left[ \sum_{i \in g} p^g_i(\alpha_i, s^g_i) \left[ \frac{1 - F(a_i)}{f(a_i)} u_1(a_i, s_i) \right] \right] \right\}.$$ 

The total expected surplus from the two-stage mechanism is

$$TS = E_a \left\{ A^g(a) E_s \left[ \sum_{i \in g} p^g_i(\alpha_i, s^g_i) u(a_i, s_i) - |g|^c \right] \right\}.$$ 

The seller’s expected revenue is thus given by

$$ER = TS - \sum_{i=1}^{N} E \pi_i(a_i, \alpha_i)$$

$$= E_a \left\{ A^g(a) E_s \left[ \sum_{i \in g} p^g_i(\alpha_i, s^g_i) \left( u(a_i, s_i) - \frac{1 - F(a_i)}{f(a_i)} u_1(a_i, s_i) \right) - |g|^c \right] \right\} - \sum_{i=1}^{N} \pi_i(\alpha_i), \quad (7)$$

where \( A^g(a) \) is the shortlisting rule and \( p^g_i(\alpha_i, s^g_i) \) is the second-stage allocation rule. To maximize \( ER \) subject to IC and IR (individual rationality), the seller sets \( \pi_i(\alpha_i) = 0 \) for all \( i = 1, 2, ..., N \); i.e., no rent should be given to the buyer with the lowest possible (pre-entry) type.
Define the virtual value adjusted by the second-stage signal as follows:

\[
w(a_i, s_i) = u(a_i, s_i) - \frac{1 - F(a_i)}{f(a_i)} w(a_i, s_i).
\] (8)

From the expression of the expected revenue, we can derive the optimal allocation rules in both stages as follows, provided that some suitable monotonicity conditions hold. At the second stage, given the revealed \( \alpha \) and the shortlisted group \( g \), \( \forall s^g \).

\[
p^*_i(\alpha, s^g) = \begin{cases} 
1 & \text{if } i = \arg \max_{j \in g} [w(a_j, s_j)] \text{ and } w(a_i, s_i) \geq 0 \\
0 & \text{otherwise} 
\end{cases} \quad \forall g, \forall i \in g. \tag{9}
\]

So consistent with the results identified by Esö and Szentes, the asset should be awarded to the bidder with the highest non-negative virtual value adjusted by the second-stage signal, which is a generalization of the optimal allocation rule in Myerson (1981). Our analysis thus shows that the generalized Myerson allocation rule is robust to settings with costly entry, which affects the final allocation only through its effect on the entry right allocation rule.

Define the expected virtual surplus (the virtual value less the entry cost) as follows:

\[
w^*g(a) = E_g \left[ \sum_{i \in g} p^*_i(\alpha, s^g) w(a_i, s_i) - |g|c \right].
\]

Then at the first stage, contingent on the revealed \( \alpha \), the optimal shortlisting rule is as follows: \( \forall g \).

\[
A^*g(a) = \begin{cases} 
1 & \text{if } g = \arg \max_{\hat{g}} [w^*g(a)] \text{ and } w^*g(a) \geq 0 \\
0 & \text{otherwise} 
\end{cases} \quad \forall g. \tag{10}
\]

The optimal shortlisting rule admits the set of bidders that gives rise to the maximal expected virtual surplus. Alternatively, the optimal shortlisting rule admits the bidders in descending order of their marginal contribution to the expected virtual surplus – the bidder with the highest contribution first, the bidder with the second-highest contribution second, etc. – provided that their marginal contribution is positive. Let \( g^*(a) \) denote the set of bidders admitted under the optimal shortlisting rule.

Similarly to Esö and Szentes, following Assumptions 1 and 2, we can establish the following properties of the optimal second-stage allocation rule: \( \forall g \).

\textbf{Corollary 1.} (i) \( p^*_i(a, s^g_i) \) increases in both \( \alpha_i \) and \( s_i \), \( \forall i \in g_i, \forall g_i, \alpha_{-i}, \text{ and } s^g_{-i} \), which implies that \( p^*_i(a, \alpha_{-i}, s_i) \) increases in both \( \alpha_i \) and \( s_i \), \( \forall g_i, \alpha_{-i} \); (ii) If \( \alpha_i > \hat{\alpha}_i, s_i < \hat{s}_i \) and \( u(a_i, s_i) = u(\hat{\alpha}_i, \hat{s}_i) \), then \( p^*_i(\alpha_i, \alpha_{-i}, s_i, s^g_{-i}) \geq p^*_i(\hat{\alpha}_i, \alpha_{-i}, \hat{s}_i, s^g_{-i}) \), which implies \( P^*_i(a_i, \alpha_{-i}, s_i) \geq P^*_i(\hat{\alpha}_i, \alpha_{-i}, \hat{s}_i), \forall g_i, \alpha_{-i} \).

---

\footnote{19}Ties occur with probability zero and are hence ignored.
\footnote{20}Again ties occur with probability zero and are hence ignored.
\footnote{21}Assumption 2 is used to show property (ii).
Property (ii) above suggests that whenever \( \alpha_i > \hat{\alpha}_i, s_i < \hat{s}_i \) and \( u(\alpha_i, s_i) = u(\hat{\alpha}_i, \hat{s}_i) \), the optimal allocation rule favors the “truth-telling” pair \((\alpha_i, s_i)\).

Given \( \alpha_i \), let \( s(\alpha_i) \) be defined such that \( w(\alpha_i, s(\alpha_i)) = 0 \). To identify properties of the optimal shortlisting rule, we define a truncated random variable as follows:

\[
w^+_{i}(a_i, s_i) = \begin{cases} 
w(a_i, s_i) & \text{if } w(a_i, s_i) \geq 0 \text{ or equivalently } s_i \geq s(\alpha_i) \forall i. \\
0 & \text{otherwise} \end{cases}
\]

Note that conditional on \( \alpha \), \( w^+_i \)'s are independent across \( i \in g \).

Let \( \Delta \hat{S}^g(\alpha_i; \alpha_{-i}) \) denote buyer \( i \)'s marginal contribution to the expected virtual surplus, \( i \in g \), then

\[
\Delta \hat{S}^g(\alpha_i; \alpha_{-i}) = \hat{S}(a^g) - \hat{S}(a^{-g}_{\alpha_i}), i \in g, \forall a^g,
\]

where \( a^{-g}_{\alpha_i} = a^g \setminus \{\alpha_i\} \) and

\[
\hat{S}(a^g) = \mathbb{E}_{a^g} \max_{i \in g} \{w^+_i(a_i, s_i)\}, \forall g, \forall a^g.
\]

The following two properties are obvious:

(1) \( \Delta \hat{S}^g(\alpha_i; \alpha_{-i}) \) increases with \( \alpha_i \), and decreases with \( \alpha_j, \forall j \neq i, \forall i \in g, \forall g \).

(2) \( \Delta \hat{S}^g(\alpha_i; \alpha_{-i}) \geq \Delta \hat{S}^g(\alpha_i; \alpha_{-i}'), \forall \alpha_{-i}, \forall i \in g, \forall g \subseteq g' \).

The revenue-optimal shortlisting rule can be alternatively described as follows. For given \( \alpha \), we can rank all \( \alpha_i \)'s from the highest to the lowest. The seller admits bidders one by one in descending order of \( \alpha_i \)'s as long as the bidder’s marginal contribution to the expected virtual surplus is nonnegative, i.e.

\[
\Delta \hat{S}^g(\alpha_i; \alpha_{-i}) - c = \hat{S}(a^g) - \hat{S}(a^{-g}_{\alpha_i}) - c \geq 0,
\]

where \( g \) denotes the group of bidders with the highest \(|g|\) types before entry.

Two properties follow immediately from the optimal shortlisting rule \( A^{*g} \):

**Corollary 2.** (i) Given \( \alpha_{-i} \), if bidder \( i \) with \( \alpha_i \) is shortlisted, then she would also be shortlisted with a higher type \( \hat{\alpha}_i(> \alpha_i) \); (ii) Given \( \alpha_{-i} \), bidder \( i \) will be shortlisted as long as \( \alpha_i \) is higher than a threshold \( \hat{\alpha}_i(\alpha_{-i}) \). As \( \alpha_i \) increases, the shortlisted group weakly shrinks. As \( \alpha_i \) increases from \( \hat{\alpha}_i(\alpha_{-i}) \), the bidders in \( g^*(a) \setminus \{i\} \) would be excluded one by one (with the lowest type originally shortlisted being excluded first).

We are now ready to show that the optimal final good allocation and entry right allocation rules (9) and (10) are truthfully implementable by some well constructed payment rules in both stages.

**Theorem 1.** Under Assumptions 1 and 2, the optimal final good allocation and entry right allocation rules (9) and (10) are IR and IC implementable.
Proof. \( u(a_i, s_i) \) increases with \( s_i \) and by Assumption 1, \( u_1(a_i, s_i) \) (weakly) decreases with \( s_i \). This implies that \( w(a_i, s_i) \) increases with \( s_i \). By the final good allocation rule (9), the winning probability \( P^{x^*g}_i(a, s_i) \) is weakly increasing in \( s_i \). By Lemma 1, the second-stage mechanism is incentive compatible (given \( a \) and \( g \)). Thus, given the truthfully revealed \( a \) and shortlisted group \( g \), a second-stage payment rule, say, \( t^{x^*g}_i(\alpha, s_i), \forall i \in g, \forall g \), can be constructed to truthfully implement the second-stage allocation rule \( P^{x^*g}_i(a, s_i), \forall i \in g, \forall g \) while maintaining the second-stage IR constraints (to participate in the second-stage mechanism), i.e. \( \bar{\pi}^g_i(\alpha, a_i; s_i, s_i) \geq 0 \) on equilibrium path. This resembles the Myerson (1981) setting with asymmetric bidders.

We use \( \bar{\pi}^{g_i}_i(\alpha_i, \hat{\alpha}_i; a_i) \) to denote the second-stage expected payoff to buyer \( i \) of type \( a_i \) if she announces \( \hat{\alpha}_i \) and is shortlisted in group \( g_i \), given that everyone else announces \( a_{-i} \) truthfully at the first stage. \( \bar{\pi}^{g_i}_i(\alpha_i, \hat{\alpha}_i; a_i) \) is well defined given Lemma 2. Therefore, when buyer \( i \) of type \( a_i \) announces \( \hat{\alpha}_i \) while others reveal \( a_{-i} \) truthfully, her first-stage expected payoff can be written as follows:

\[
\pi^*_i(\alpha_i, \hat{\alpha}_i) = E_{a_{-i}} \left\{ \sum_{g_i} A^{g_i}(\hat{\alpha}_i, a_{-i}) [\bar{\pi}^{g_i}_i(\alpha_i, \hat{\alpha}_i; a_{-i}) - c] - x^*_i(\hat{\alpha}_i, a_{-i}) \right\},
\]

where \( x^*_i \) is the first-stage payment rule.

Next, we will show that the optimal shortlisting rule (10) is truthfully implementable by a properly chosen first-stage payment rule \( x^*_i \), together with the second-stage payment rules \( t^{x^*g}_i \) chosen above.

Note that by (5), we have

\[
\pi^*_i(\alpha_i, a_i) = E_{a_{-i}} \left\{ \sum_{g_i} A^{g_i}(\alpha_i, a_{-i}) [\bar{\pi}^{g_i}_i(\alpha_i, a_i; a_{-i}) - c] - x^*_i(\alpha_i, a_{-i}) \right\}. \tag{11}
\]

Construct the first-stage payment rule as follows:

\[
x^*_i(\alpha) = \sum_{g_i} A^{g_i}(\alpha_i, a_{-i}) [\bar{\pi}^{g_i}_i(\alpha_i, a_i; a_{-i}) - c] - \int_{\alpha} \int u(y, s_i) \cdot \sum_{g_i} [E_{a_{-i}} A^{g_i}(y, a_{-i}) P^{g_i}_i(y, a_{-i}, s_i)] dG_i(s_i) dy \tag{12}
\]

Substituting (12) into (11), we can verify that

\[
\pi^*_i(\alpha_i, a_i) = \int_{\alpha} \int u(y, s_i) \cdot \sum_{g_i} [E_{a_{-i}} A^{g_i}(y, a_{-i}) P^{g_i}_i(y, a_{-i}, s_i)] dG_i(s_i) dy,
\]

which is precisely equation (6) with \( \pi^*_i(\alpha_i, \alpha) = 0 \) (the optimality requirement). Note that \( \pi^*_i(\alpha_i, a_i) \geq 0 \), so IR is satisfied in the first stage.

Suppose that all buyers except \( i \) report their types \( a_{-i} \) truthfully. Consider buyer \( i \) with \( a_i \) contemplating to misreport \( \hat{\alpha}_i < a_i \). The deviation payoff is

\[
\Delta = \pi^*_i(\alpha_i, \hat{\alpha}_i) - \pi^*_i(\alpha_i, a_i) = [\pi^*_i(\alpha_i, \hat{\alpha}_i) - \pi^*_i(\hat{\alpha}_i, \hat{\alpha}_i)] + [\pi^*_i(\hat{\alpha}_i, \hat{\alpha}_i) - \pi^*_i(a_i, a_i)].
\]
Since (6) is satisfied by the construction of \( x^*_i(\alpha) \), we have

\[
\pi_i^*(\hat{\alpha}_i, \hat{\alpha}_i) - \pi_i^*(\alpha_i, \alpha_i) = -\int_{\hat{\alpha}_i}^{\alpha_i} \int u_1(y, s_i) \cdot \sum_{g_i} \left[ E_{\alpha_{-i}} A^{g_i}(y, \alpha_{-i}) P_i^{g_i}(y, \alpha_{-i}, s_i) \right] dG_i(s_i) dy.
\]

Recall the definition of \( \pi_i^*(\alpha_i, \hat{\alpha}_i) \) above, we have from Lemma 2 that

\[
\pi_i^*(\alpha_i, \hat{\alpha}_i) - \pi_i^*(\hat{\alpha}_i, \hat{\alpha}_i) = \int_{\hat{\alpha}_i}^{\alpha_i} \int u_1(y, s_i) \cdot \sum_{g_i} \left[ E_{\alpha_{-i}} A^{g_i}(\hat{\alpha}_i, \alpha_{-i}) P_i^{g_i}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) \right] dG_i(s_i) dy.
\]

Therefore, we have

\[
\Delta = \int_{\hat{\alpha}_i}^{\alpha_i} E_{\alpha_{-i}} \sum_{g_i} A^{g_i}(y, \alpha_{-i}) \int u_1(y, s_i) \left[ P_i^{g_i}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) - P_i^{g_i}(y, \alpha_{-i}, s_i) \right] dG_i(s_i) dy
\]

\[
+ \int_{\hat{\alpha}_i}^{\alpha_i} E_{\alpha_{-i}} \sum_{g_i} \left[ A^{g_i}(\hat{\alpha}_i, \alpha_{-i}) - A^{g_i}(y, \alpha_{-i}) \right] \int u_1(y, s_i) P_i^{g_i}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) dG_i(s_i) dy.
\]

From Corollary 1 (ii), we have \( P_i^{g_i}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) - P_i^{g_i}(y, \alpha_{-i}, s_i) \leq 0 \), which implies that the first term in \( \Delta \) is negative.

We now consider the second term in \( \Delta \) when \( y > \hat{\alpha}_i \). By Corollary 2, the optimal shortlisting rule implies that given \( \alpha_{-i} \), when buyer \( i \) is admitted with a higher \( \alpha_i \), she must be admitted to a group with a weakly smaller size. If \( y \) and \( \hat{\alpha}_i \) are admitted in the same group, then \( A^{g_i}(\hat{\alpha}_i, \alpha_{-i}) = A^{g_i}(y, \alpha_{-i}) \) and this term in \( \Delta \) is zero.

We now turn to the case where \( g^*(\hat{\alpha}_i, \alpha_{-i}) \supset g^*(y, \alpha_{-i}) \supset \{ i \} \). Note that \( A^{g_i}(\cdot, \alpha_{-i}) \) is 1 for the shortlisted group, and 0 for all other groups. Therefore,

\[
\sum_{g_i} \left[ A^{g_i}(\hat{\alpha}_i, \alpha_{-i}) - A^{g_i}(y, \alpha_{-i}) \right] u_1(y, s_i) P_i^{g_i}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i))
\]

\[
= u_1(y, s_i) \left[ P_i^{g^*(\hat{\alpha}_i, \alpha_{-i})}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) - P_i^{g^*(y, \alpha_{-i})}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) \right]
\]

\[
\leq 0,
\]

which implies that the second term in \( \Delta \) is negative. Since \( g^*(\hat{\alpha}_i, \alpha_{-i}) \supset g^*(y, \alpha_{-i}) \supset \{ i \} \), we must have \( P_i^{g^*(\hat{\alpha}_i, \alpha_{-i})}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) \leq P_i^{g^*(y, \alpha_{-i})}(\hat{\alpha}_i, \alpha_{-i}, \sigma_i(y, \hat{\alpha}_i, s_i)) \), i.e. entrant \( i \) wins with a smaller probability if a strictly bigger group is shortlisted.

A similar argument can be used to rule out deviating to \( \hat{\alpha}_i > \alpha_i \). \( \square \)

It is worth noting that Assumptions 1 and 2 are sufficient but not necessary for the optimal entry rule to be truthfully implementable: the necessary and sufficient condition is that \( \Delta \) defined in (13) is non-positive, which is also the integral monotonicity condition characterized by Pavan, Segal, and Toikka (2014).
3.2 Applications

Our optimal mechanism analysis is general enough to encompass many existing models in the literature on auctions with costly entry. Below we demonstrate how we can apply our general optimal mechanism to special models previously studied.

1. Bidders do not have pre-entry types and only learn about their values after entry (e.g., McAfee and McMillan, 1987; Tan, 1992; and Levin and Smith, 1994). In this case, \( u(\alpha_i, s_i) = s_i \). Hence \( w(\alpha_i, s_i) = s_i \), which implies that the optimal auction is ex post efficient, and the optimal entry is to select a set of bidders that results in the maximal expected social surplus. Since bidders are identical before entry, optimal entry is entirely characterized by \( n^* \), the optimal number of bidders to be selected. The implementation is somewhat simple: the second round is a standard auction (first-price, second-price, or English auction). The first round (entry stage) is to select exactly \( n^* \) bidders, and whomever selected is required to pay an upfront entry fee \( e^* \), which is set so that no rent is left for the entrants ex ante.

2. Bidders know their values before entry, and entry is merely a bid preparation process (without value updating) (e.g. Samuelson, 1985; Stegeman, 1996; Campbell, 1998; Menezes and Monteiro, 2000; Tan and Yilankaya, 2006; Cao and Tian, 2009; and Lu, 2009). In this setting, \( u(\alpha_i, s_i) = \alpha_i \) and \( s_i \equiv 0 \), and hence \( w(\alpha_i, s_i) = \alpha_i - (1 - F(\alpha_i))/f(\alpha_i) \). It is easily verified that according to Theorem 1, the optimal allocation rules can be described as follows: the bidder with the highest “type” \( \alpha_i \) is admitted as the sole entrant to win the item, as long as her contribution to the virtual surplus \( w(\alpha_i, s_i) - c \) is positive.

3. Each bidder is endowed with pre-entry type \( \alpha_i \), and learns an additional private value component \( s_i \) (e.g., Ye, 2007; Quint and Hendricks, 2018). The total value is given by \( u(\alpha_i, s_i) = \alpha_i + s_i \). We assume that \( \alpha_i \) is distributed uniformly over \([0, 1]\) and \( s_i \) follows a standard Normal distribution, and let \( c \in (0, 1) \). Hence \( w(\alpha_i, s_i) = \alpha_i + s_i - (1 - F(\alpha_i))/f(\alpha_i) = 2\alpha_i + s_i - 1 \). The optimal second-stage allocation rule thus requires that the asset be allocated to the entrant bidder with the highest virtual value \( w(\alpha_i, s_i) \) provided that it is nonnegative. The optimal entry rule requires that bidders be admitted in descending order of their pre-entry types, as long as their contribution to the expected virtual surplus is nonnegative. If only one buyer (the one with the highest type \( \alpha(1) \)) is admitted, the expected virtual surplus is given by \( w_1 = E_{s_{i_1}}[(2\alpha_{(1)} + s_{i_1} - 1) \vee 0] - c \). So the optimal number of entrants \( n^* \geq 1 \) if \( w_1 \geq 0 \). If two top buyers are admitted, the expected virtual surplus is given by

\[
 w_2 = E_{(s_{i_1}, s_{i_2})}[\max\{2\alpha_{(1)} + s_{i_1} - 1, 2\alpha_{(2)} + s_{i_2} - 1, 0\}] - 2c.
\]

\[22\] Therefore, there is no issue of reporting \( s_i \) in the second stage.

\[23\] Let \( i_k \) be the index of bidder who possesses \( a_k \), the \( k \)th highest value of \( a \).
So the optimal number of entrants $n^* \geq 2$ if the incremental expected virtual surplus

$$w_2 - w_1 = E(s_{i_1}, s_{i_2}) \left[ \max \left\{ 2\alpha(1) + s_{i_1} - 1, 2\alpha(2) + s_{i_2} - 1, 0 \right\} \right] - E(s_{i_1}) \left[ (2\alpha(1) + s_{i_1} - 1) \lor 0 \right] \geq c.$$ 

Continuing this procedure of calculation, it can be verified that $n^* \geq n$ if $w_n - w_{n-1} \geq 0$, or

$$E(s_{i_1}, s_{i_2}, ..., s_{i_n}) \left[ \max_{k=1, ..., n} \left\{ 2\alpha(k) + s_{i_k} - 1 \right\} \lor 0 \right] - E(s_{i_1}, s_{i_2}, ..., s_{i_{n-1}}) \left[ \max_{k=1, ..., n-1} \left\{ 2\alpha(k) + s_{i_k} - 1 \right\} \lor 0 \right] \geq c. \ (14)$$

### 4 Analysis with Sequential Shortlisting

Now we move to the setting where the seller may conduct sequential shortlisting, which consists of $M \geq 2$ shortlisting rounds/stages and one final allocation stage $M + 1$. Specifically, the procedure is described as follows.\footnote{We again use the null message $\phi$ to denote a buyer’s report if she is not required to report her $s_i$.}

At stage 1, all bidders report their initial types of $a_i$. Denote the reports by $m_1 = (m_{1,i})$ where $m_{1,i} \in [g, \bar{\alpha}]$, $\forall i \in N$ is $i$’s report. Let $g_0 = \emptyset$. $\forall g_1 \subset 2^N$, the probability that $g_1$ is shortlisted is $A^{g_1}(m_1|g_0)$. The payment of $i \in N$ is $t_{1,i}(m_1)$. All bidders in shortlisted group $g_1$ incur their information acquisition cost $c$ to discover their ex post values, $v_i, i \in g_1$.

At stage 2, if group $g_1$ is shortlisted and discover their values $v_i$’s at stage 1, they are asked to report their $s_i$’s. The bidders’ second-stage reports about $s_i$’s are denoted by $m_2 = (m_{2,i})$ where $m_{2,i} \in [0, 1]$ if $i \in g_1$ and $m_{2,i} = \phi$ if $i \notin g_1$. $\forall g_2 \subset 2^{N \setminus g_1}$, the probability for $g_2$ to be shortlisted is $A^{g_2}(m_1, m_2|g_0, g_1)$ and $i$’s payment is $t_{2,i}(m_1, m_2)$ for any $i \in N$. All bidders in shortlisted group $g_2$ incur their information acquisition cost $c$ to discover their ex post values, $v_i, i \in g_2$.

At stage 3, if group $g_2 \in 2^{\Omega \setminus g_1}$ is shortlisted and discover their values $v_i$’s at stage 2, they are asked to report their $s_i$’s. The bidders’ stage-3 reports about $s_i$’s are denoted by $m_3 = (m_{3,i})$ where $m_{3,i} \in [0, 1]$ if $i \in g_2$ and $m_{3,i} = \phi$ if $i \notin g_2$. $\forall g_3 \subset 2^{N \setminus (g_1 \cup g_2)}$, the probability for $g_3$ to be shortlisted is $A^{g_3}(m_1, m_2, m_3|g_0, g_1, g_2)$, and $i$’s payment is $t_{3,i}(m_1, m_2, m_3)$ for any $i \in N$. All bidders in shortlisted group $g_3$ incur their information acquisition cost $c$ to discover their ex post values, $v_i, i \in g_3$.

The procedure proceeds analogously up to stage $M$. At stage $M$, if group $g_{M-1}$ is shortlisted and the bidders in $g_{M-1}$ discover their values $v_i$’s at stage $M - 1$, they are asked to report their $s_i$’s. The bidders’ stage-$M$ reports about $s_i$’s are denoted by $m_M = (m_{M,i})$ where $m_{M,i} \in [0, 1]$ if $i \in g_{M-1}$ and $m_{M,i} = \phi$ if $i \notin g_{M-1}$. $\forall g_M \subset 2^{N \setminus (g_1 \cup \ldots \cup g_{M-1})}$, the probability for $g_M$ to be shortlisted is $A^{g_M}(m_1, m_2, ..., m_M|g_0, g_1, g_2, ..., g_{M-1})$, and $i$’s payment is $t_{M,i}(m_1, m_2, ..., m_M)$ for any $i \in N$. All bidders in shortlisted group $g_M$ incur their information acquisition cost $c$ to discover their ex post values, $v_i, i \in g_M$.

At the final stage, i.e. stage $M + 1$, if $g_M$ is shortlisted and discover their values $v_i$’s at the end of stage $M$, their reports are denoted by $m_{M+1} = (m_{M+1,i})$ where $m_{M+1,i} \in [0, 1]$ if $i \in g_M$, $m_{M+1,i} = \phi$ if $i \notin g_M$. \footnote{We again use the null message $\phi$ to denote a buyer’s report if she is not required to report her $s_i$.}
Denote the sequence of shortlisting outcome by vector \( g = (g_1, g_2, \ldots, g_M) \) with \(|g| = M\). Let \( G_g \) denote the set of all agents shortlisted in sequence \( g \). Given the final shortlisted group \( G_g \), let \( p_i^{G_g}(m_1, m_2, \ldots, m_{M+1}) \) be agent \( i \)'s winning probability, for agent \( i \in G_g \), and agent \( i \)'s payment is \( t_{M+1,i}(m_1, m_2, \ldots, m_M, m_{M+1}) \) for any \( i \in N \).

We use \( (A, p, t, M) \) to denote the procedure specified above. Without loss of generality, we can focus on the cases where \( M \geq N \). The mechanisms with \( M < N \) can be trivially duplicated by a mechanism with \( M = N \), in which case the shortlisting stops at stage \( M \).

Our analysis proceeds as follows. We first consider a relaxed environment where the agents are only endowed with private information \( \alpha \), where \( s_i \)'s become public once they are discovered. The optimal solution for this relaxed environment provides an upper bound for the seller's expected revenue in the original environment where the discovered \( s_i \)'s are private information for the shortlisted. We will establish that this upper bound is actually achievable in the original environment.

### 4.1 The Relaxed Environment

For a given mechanism \( (A, p, t, M) \), and message sequence \( (m_k, k = 1, 2, \ldots, M) \), the probability of a shortlisting outcome \( g = (g_1, g_2, \ldots, g_M) \) is given by

\[
\Pr(g|\{m_k\}_{k=1}^M) = \Pi_{k=1}^M A^S_i(m_1, m_2, \ldots, m_k|g_0, g_1, g_2, \ldots, g_{k-1}).
\]

As \( s_i \) becomes public once discovered in the relaxed environment, we have for \( k \geq 2 \), \( m_{k,i} = s_i, i \in g_{k-1} \), and \( m_{k,i} = \phi, i \notin g_{k-1} \). We use \( m^*_k, k \geq 2 \) to denote these true types from stages 2 to \( M + 1 \).

Agent \( i \)'s expected payoff when \( i \) is endowed with \( \alpha_i \) but announces \( \hat{a}_i \) is given by:

\[
\pi_i(a_i, \hat{a}_i) = E_{a_{-i}}E_s\left\{ -\sum_{g} \Pr(g|\{\hat{a}_i, a_{-i}\}, m^*_2, \ldots, m^*_M) \sum_{k=1}^{M} t_{k+1,i}((\hat{a}_i, a_{-i}), m^*_2, \ldots, m^*_{k+1}) \right\} + \sum_{g:s.t. i \in G_g} \Pr(g|\{\hat{a}_i, a_{-i}\}, m^*_2, \ldots, m^*_M) u(a_i, s_i) p_i^{G_g}(\{\hat{a}_i, a_{-i}\}, m^*_2, \ldots, m^*_M) - c \}
\]

Incentive compatibility together with the envelop theorem gives:

\[
\frac{d\pi_i(a_i, a_i)}{d\alpha_i} = E_{a_{-i}}E_s\left\{ \sum_{g:s.t. i \in G_g} \frac{\partial u(a_i, s_i)}{\partial \alpha_i} \Pr(g|\{a_i, a_{-i}\}, m^*_2, \ldots, m^*_M) p_i^{G_g}(\{a_i, a_{-i}\}, m^*_2, \ldots, m^*_M) \right\}.
\]

\(^{25}\)We will show that in equilibrium, the agents who are shortlisted have incentives to incur their information acquisition costs to discover their values.
Thus, we have

\[ \pi_i(\alpha_i, \alpha_i) = \pi_i(\alpha, \alpha) + E_{a_i} \int_a^{a_i} E_s \left\{ \sum_{g \in N \text{s.t. } i \in G_g} u_1(y, s_i) [Pr(g|\alpha_i, m_{2}^{s}, ..., m_{M+1}^{s}) \cdot p_i^{G_g}((y, \alpha_{-i}), m_{2}^{s}, ..., m_{M+1}^{s})] \right\} dy. \]

The expected social surplus given \( \alpha \) is as follows:

\[ TS(\alpha) = E_s \left\{ \sum_{g \in N} \left[ \sum_{i \in G_g} \left[ p_i^{G_g}(\alpha, m_{2}^{s}, ..., m_{M+1}^{s}) \cdot u_1(\alpha_i, s_i) \right] \right] - c \right\}. \]

The seller seeks to maximize the expected revenue:

\[ ER = E_a \left[ TS(\alpha) - \sum_{i \in N} \pi_i(\alpha_i, \alpha_i) \right]. \]

Define \( Pr(g|\alpha, s) = Pr(g|\alpha, m_{2}^{s}, ..., m_{M+1}^{s}) \), and for any \( G \in 2^N \), define

\[ Pr(G|\alpha, s) = \sum_{g \in N \text{s.t. } G_g = G} Pr(g|\alpha, s), \]

where, as before, \( G_g \) denotes the set of all agents shortlisted in sequence \( g \). Note we have

\[ \sum_{g \in N \text{s.t. } i \in G_g} Pr(g|\alpha, s) = \sum_{G \in 2^N \text{s.t. } i \in G} Pr(G|\alpha, s). \]

By the standard procedure, we can rewrite the seller's objective as follows.

**Lemma 4.** The seller's objective is to maximize:

\[ ER = E_a E_s \sum_{g \in N} \left[ Pr(g|\alpha, m_{2}^{s}, ..., m_{M+1}^{s}) \cdot \sum_{i \in G_g} \left[ p_i^{G_g}(\alpha, m_{2}^{s}, ..., m_{M+1}^{s}) w(\alpha_i, s_i) - c \right] - \sum_{i} \pi_i(\alpha, \alpha) \right]. \]

where \( w(\alpha_i, s_i) = u(\alpha_i, s_i) - u_1(\alpha_i, s_i) \cdot \frac{1 - F(\alpha_i)}{f(\alpha_i)} \).

**Proof.** See Appendix.

From Lemma 4, it is clear that we have the following results at the optimum.

**Lemma 5.** For any \( \{Pr(G), \forall G \in 2^\Omega\} \) derived from any shortlisting rule, to maximize the expected revenue \( ER \), the seller sets \( \pi_i(\alpha, \alpha) = 0 \) and allocates the object to the shortlisted bidder whose virtual value is the
We now turn to the optimal sequential shortlisting rule. Before searching for the optimal shortlisting rule, we first establish the following lemma.

**Lemma 6.** Shortlisting one agent at a stage (the shortlisted agent can be a different agent with different probability) until the last is shortlisted yields weakly higher seller revenue than other rules.

**Proof.** See Appendix.

According to Lemma 6, there is no loss of generality in focusing on shortlisting rules under which the seller shortlists one agent at each stage before stopping the shortlisting. It further implies that to search for the optimal shortlisting rule, without loss of generality, we can focus on the rules where at each stage the seller either shortlists an agent with probability 1 or stops shortlisting. The reason is as follows.

At stage 1, the principal has \(N + 1\) choices: shortlisting no one or shortlisting any \(i\). By Lemma 6, shortlisting no one implies that the shortlisting process stops at stage 1. According to Lemma 5, we have \(ER = 0\). If the seller opts to shortlist agent \(i\), he would continue to adopt the best shortlisting decisions in all subsequent stages. Suppose this leads to an optimal expected revenue of \(ER_i, i = 1, 2, \ldots, N\). Clearly, there is no loss of generality for the seller to adopt an option which delivers the maximum revenue with probability 1. If this option is \(\emptyset\), then the shortlisting process stops. If it is to shortlist agent \(i\), then we move to the next stage with the new information revealed by \(i\). At stage 2, similarly, the seller has \(N\) options: shortlisting no one or shortlisting any \(j \in N \setminus \{i\}\). Similarly, the seller should go for an option that delivers the maximum revenue with probability 1. The process continues until the seller runs out of agents or he decides to stop shortlisting. We further have the following result.

**Lemma 7.** At any stage, the seller should shortlist the agent who has the highest \(\alpha_i\) among the remaining bidders or stop shortlisting.

Based on Lemma 7, we have the following optimal shortlisting rule.

**Proposition 1.** Without loss of generality, we assume \(\alpha_i\) decreases with \(i\). For \(k = 1, 2, \ldots, N\), bidder \(k\) is shortlisted if and only if \(ER_k((w_i^*(\alpha_k, s_k) - \max\{w_i^+(\alpha_l, s_l)\}_{1 \leq l < k}) \vee 0) \geq c\). If bidder \(k\) is not shortlisted in stage \(k\), then no buyer is shortlisted in subsequent stages.

By Lemma 7, in stage 1, the seller considers whether he should shortlist agent 1. Recall the expression for expected revenue in Lemma 4. If he does not shortlist agent 1, then the shortlisting process ends and he gets 0 revenue. If he shortlists agent 1 and simply ends the process, then his expected revenue is
revenue is $E_s[w_1^+(a_1,s_1)] - c$. If he shortlists agent 1 and follows the optimal shortlisting strategy after that, his expected revenue can be higher than $E_s[w_1^+(a_1,s_1)] - c$. This means that the seller should definitely shortlist agent 1 in stage 1 if $E_s[w_1^+(a_1,s_1)] \geq 0$. If $E_s[w_1^+(a_1,s_1)] < 0$, then we must have $E_s[(w_2^+(a_k,s_k) - \max(w_1^+(a_1,s_1))]_{1 \leq k \leq N} < c, \forall k = 2, ..., N$. This means that for $k = N, N-1, ..., 2$, even if the first $k-1$ bidders are shortlisted, the next bidder should not be shortlisted. This further implies that agent 1 should not be shortlisted in stage 1.

In stage 2, suppose agent 1 is shortlisted in stage 1, and $s_1$ is revealed. If the seller does not shortlist agent 2, he gets a revenue of $w_1^+(a_1,s_1) - c$. If he shortlists agent 2 and simply ends the process, then his expected revenue is $E_s[w_2^+(a_1,s_1)] - c$. If he shortlists agent 2 and follows the optimal shortlisting strategy after that, his expected revenue can be higher than $E_s[w_2^+(a_1,s_1)] - \sum_{i \in [1,2]} c \geq w_1^+(a_1,s_1) - c$ if and only if $E_s[(w_2^+(a_2,s_2) - w_1^+(a_1,s_1))]_{1 \leq k \leq N} \geq c$. If he shortlists agent 2 and follows the optimal shortlisting strategy after that, his expected revenue can be higher than $E_s[w_2^+(a_1,s_1)] - \sum_{i \in [1,2]} c$. This means that the seller should definitely shortlist agent 2 in stage 2 if $E_s[(w_2^+(a_2,s_2) - w_1^+(a_1,s_1))]_{1 \leq k \leq N} \geq c$. If $E_s[(w_2^+(a_2,s_2) - w_1^+(a_1,s_1))]_{1 \leq k \leq N} < c$, then we must have $E_s[(w_2^+(a_k,s_k) - \max[w_1^+(a_1,s_1))]_{1 \leq k \leq N} < c, \forall k = 3, ..., N$. This means that for $k = N, N-1, ..., 3$, even if the first $k-1$ bidders are shortlisted, the next bidder should not be shortlisted. This further implies that agent 2 should not be shortlisted in stage 2. The reasoning for other stages is similar.

4.2 Incentive Compatibility in the Original Setting

We use $(\hat{a}, m_2, ..., m_{M+1})$ to denote the announcements of agents at different stages. We denote the shortlisting rule of Proposition 1 by $A^* = (A^*g_1(\hat{a}, m_2, ..., m_{k-1}, g_1, g_2, ..., g_{k-1}), k = 1, 2, ..., M, \forall g = (g_1, g_2, ..., g_M))$, and denote the allocation rule of (5) by $p^* = (p^*_i g_1(\hat{a}, m_2, ..., m_{M+1}), i \in N, \forall g = (g_1, g_2, ..., g_M))$. In addition,

$$Pr^*(g|\{m_i\}_{i=1}^M) = \prod_{k=1}^M A^*g_1(\hat{a}, m_2, ..., m_k|g_0, g_1, g_2, ..., g_{k-1}),$$

which is the probability that sequence $g$ is shortlisted given messages reported $(m_i|_{i=1}^M$.

Then $(A^*, p^*)$ imposes an upper bound for seller expected revenue in the relaxed environment where $s_i$’s are public information. While we utilize the incentive compatibility condition in stage 1 when deriving these rules, the stage-1 incentive compatibility has not yet been established even in the relaxed environment. In this section, we will establish that these rules are indeed incentive compatible even in the original environment where both $(a_i, s_i)$ are the private information of agent $i$. For this purpose, we will construct payment rules $\tau^*$ that together with $(A^*, p^*)$ induce truthful information revelation on the equilibrium path. In addition, we will have $\pi_i(g, \hat{a}) = 0, \forall i$.

Given stage-1 report $\hat{a}$, we use $i_k$ to denote the agent whose report is ranked the $k$th highest. We first look at the reporting incentive at stages $k \geq 2$. Let $g_{k, h} = (i_1, g_2 = (i_2), ..., g_{h-1} = (i_{h-1}), ..., g_h = (i_h), g_{h+1} = \emptyset, ..., g_M = \emptyset$, $h \geq k \geq 2$ be a sequence of shortlisted where $h$ agents (including $i_k$) are shortlisted in total. Let $\hat{m}_h$ denote the stage-$h$ announcement (so its $i_{h-1}$th element is $\hat{s}_{i_{h-1}}$ and all the other elements are $\emptyset$). Given the announcement in the history $(\hat{a}, m_2, ..., m_{k-1})$, assuming agents $i_l, l \geq k$
truthfully reveal their information at stage \( l + 1 \), agent \( i_{k-1} \)'s conditional expected payoff is

\[
\pi_{i_{k-1}}(s_{i_{k-1}}, \hat{s}_{i_{k-1}} | \hat{a}, m_2, \ldots, m_{k-1}) = \mathbb{E}_{(s_{i_k}, \ldots, s_{i_M})} \left[ - \sum_{h=k}^{N} \Pr^* (g_{k-1,h} | \hat{a}, m_2, \ldots, m_{k-1}, \hat{m}_h, m_{k+1}^*, \ldots, m_{M+1}^*) \cdot \right.
\]

\[
\sum_{l=k+1}^{M} t_{i,l-1} (\hat{a}, m_2, \ldots, m_{k-1}, \hat{m}_l, m_{k+1}^*, \ldots, m_{M+1}^*)]
\]

\[
+ u(a_{i_{k-1}}, s_{i_{k-1}}) \sum_{h=k-1}^{N} \Pr^*(g_{k-1,h} | \hat{a}, m_2, \ldots, m_{k-1}, \hat{m}_h, m_{k+1}^*, \ldots, m_{M+1}^*)
\]

\[
p_{i_{k-1}}^* (\hat{a}, m_2, \ldots, m_{k-1}, \hat{m}_h, m_{k+1}^*, \ldots, m_{M+1}^*)
\]

\[
- t_{k,i_{k-1}} (\hat{a}, m_2, \ldots, m_{k-1}, \hat{m}_h).
\]

We next identify the \( t_{k,i} (\hat{a}, m_2, \ldots, m_{k-1}, \hat{m}_h), i \in \mathbb{N} \) that induces truthful revelation at stage \( k \) whenever \( \hat{a}_{i_l} = a_{i_l}, l \geq k - 1 \). In particular, \( t_{k,i_{k-1}} \) can be constructed following the standard Myersonian procedure with \( \pi_{i_{k-1}}(0,0 | \hat{a}, m_2, \ldots, m_{k-1}) = 0 \), and \( \forall k \geq 3, t_{k-1,i_{k-1}} \) can be set at \( -c \) to induce information discovery of the shortlisted agent \( i_{k-1} \) at stage \( k - 1 \). All other \( t_{k,i} \) are set to zero for \( i \neq i_{k-1}, i_k, \forall k \geq 2 \). This means at each stage \( k \geq 2 \), transfers are nonzero only for the agents shortlisted in stages \( k - 1 \) and \( k \).

We first consider stage \( N + 1 \). Suppose \( i_N \) is shortlisted in stage \( N \). At stage \( N \), we have

\[
\pi_{i_N}(s_{i_N}, \hat{s}_{i_N} | \hat{a}, m_2, \ldots, m_N) = u(a_{i_N}, s_{i_N}) p_{i_N}^* (\hat{a}, m_2, m_{N+1}, m_{N+2}^*, \ldots, m_{M+1}^*) - t_{N+1,i_N} (\hat{a}, m_2, \ldots, m_N, \hat{m}_{N+1}).
\]

By the envelop theorem, optimality of truthful revelation requires

\[
\frac{d \pi_{i_N}(s_{i_N}, s_{i_N} | \hat{a}, m_2, \ldots, m_N)}{ds_{i_N}} = u(a_{i_N}, s_{i_N}) p_{i_N}^* (\hat{a}, m_2, m_{N+1}, m_{N+2}^*, \ldots, m_{M+1}^*).
\]

Recall that we set \( \pi_{i_N}(0,0 | \hat{a}, m_2, \ldots, m_N) = 0 \). We thus have

\[
\pi_{i_N}(s_{i_N}, s_{i_N} | \hat{a}, m_2, \ldots, m_N) = \int_{0}^{s_{i_N}} u(a_{i_N}, y) p_{i_N}^* (\hat{a}, m_2, m_{N+1}(y), m_{N+2}^*, \ldots, m_{M+1}^*) dy,
\]

where in \( m_{N+1}^*(y), s_{i_N} \) is replaced by \( y \).

Given (16) and (17), we define

\[
t_{N+1,i_N}^* (\hat{a}, m_2, \ldots, m_N, m_{N+1}^*) = u(a_{i_N}, s_{i_N}) p_{i_N}^* (\hat{a}, m_2, \ldots, m_N, m_{N+1}^*, m_{N+2}^*, \ldots, m_{M+1}^*)
\]

\[
- \int_{0}^{s_{i_N}} u(a_{i_N}, y) p_{i_N}^* (\hat{a}, m_2, \ldots, m_N, m_{N+1}^*(y), m_{N+2}^*, \ldots, m_{M+1}^*) dy.
\]
Therefore,

\[ t_{N+1,iN}^*(\hat{a}, m_2, \ldots, m_N, m_{N+1}^*) = 0 \] if agent \( i_N \) loses; and

\[ t_{N+1,iN}^*(\hat{a}, m_2, \ldots, m_N, m_{N+1}^*) = u(a_{iN}, \hat{s}_{iN}) \] if she wins,

where \( \hat{s}_{iN} > 0 \) is agent \( i_N \)'s minimum winning type in stage \( N \) based on winning rule \( p^* \). Note that under Assumptions 1 and 2, we have that \( u(a_{iN}, s_{ik}) \) increases with both \( a_{ik} \) and \( s_{ik} \), \( \forall k \). Thus \( \hat{s}_{iN} \) is well defined, which is determined by

\[
u(a_{iN}, \hat{s}_{iN}) - u_1(a_{iN}, \hat{s}_{iN}) \frac{1 - F(a_{iN})}{f(a_{iN})} = \max_{i,k < N} \{u(\hat{a}_{ik}, s_{ik}) - u_1(\hat{a}_{ik}, s_{ik}) \frac{1 - F(\hat{a}_{ik})}{f(\hat{a}_{ik})} \}..\]

It is clear that (1) \( \pi_{iN}(s_{iN}, \hat{s}_{iN}) | \hat{a}, m_2, \ldots, m_N \) satisfies the strict and smooth single crossing differences property in \( (s_{iN}, \hat{s}_{iN}) \); (2) \( p_{iN}^*(\hat{a}, m_2, \ldots, m_N, \hat{m}_{N+1}, m_{N+2}^*, \ldots, m_{M+1}^*) \) increases in \( \hat{s}_{iN} \); (3) \( t_{N+1,iN}^* \) defined above verifies the envelope theorem. By the constraint simplification theorem,\(^{26}\) truthful revelation at stage \( N + 1 \) is incentive compatible for \((p^*, t_{N+1,iN}^*)\).

We now turn to stage \( N \). First note that, given \( t_{N+1,iN}^* \) constructed above, \( i_N \) would reveal truthfully at stage \( N + 1 \) if shortlisted, regardless of \( i_{N-1} \)'s announcement at stage \( N \). Let \( \hat{m}_N \) denote the stage-\( N \) announcement with the \( i_{N-1} \)-th element being \( \hat{s}_{iN-1} \) and all other elements being \( \emptyset \). Given announcement in the history \((\hat{a}, m_2, \ldots, m_{N-1})\), agent \( i_{N-1} \)'s conditional expected payoff is

\[
\pi_{i_{N-1}}(s_{i_{N-1}}, \hat{s}_{i_{N-1}} | \hat{a}, m_2, \ldots, m_{N-1}) = E_{s_{iN}} [u(a_{i_{N-1}, s_{i_{N-1}}}) \sum_{h=N-1}^{N} Pr^*(g_{N-1,h} | \hat{a}, m_2, \ldots, m_{N-1}, \hat{m}_N, m_{N+1}^*, \ldots, m_{M+1}^*)]
- t_{N,i_{N-1}}(\hat{a}, m_2, \ldots, m_{N-1}, \hat{m}_N).
\]

By the envelop theorem, optimality of truthful revelation requires

\[
\frac{d \pi_{i_{N-1}}(s_{i_{N-1}}, \hat{s}_{i_{N-1}} | \hat{a}, m_2, \ldots, m_{N-1})}{d \hat{s}_{i_{N-1}}}
= u_2(a_{i_{N-1}, s_{i_{N-1}}})E_{s_{iN}} \{ \sum_{h=N-1}^{N} Pr^*(g_{N-1,h} | \hat{a}, m_2, \ldots, m_{N-1}, m_{N+1}^*, \ldots, m_{M+1}^*)]
- t_{N,i_{N-1}}(\hat{a}, m_2, \ldots, m_{N-1}, \hat{m}_N, m_{N+1}^*, \ldots, m_{M+1}^*)\}.
\]

Recall that we set \( \pi_{i_{N-1}}(0,0 | \hat{a}, m_2, \ldots, m_{N-1}) = 0 \). We thus have

\[
\pi_{i_{N-1}}(s_{i_{N-1}}, \hat{s}_{i_{N-1}} | \hat{a}, m_2, \ldots, m_{N-1})
\]

\(^{26}\)A version of the constraint simplification theorem can be seen from Theorem 4.3 in Milgrom (2004, page 13).
where in \( m_N^s(y) \), \( s_{iN-1} \) is replaced by \( y \).

Based on (19) and (20), we define

\[
\begin{align*}
t_{N,iN-1}^* (\hat{a}, m_2, ..., m_{N-1}, m_N^s) \\
= E_{s_{iN}} (u(a_{iN-1}, s_{iN-1}) \sum_{\forall h=1}^{N} \Pr^* (g_{N-1,h} | \hat{a}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, m_{M+1}^s) p_{iN-1}^{sG_{N-1,k}^N}(\hat{a}, m_2, ..., m_{N-1}, m_N^s(\hat{a}, m_2, ..., m_{N-1}, m_N^s(y), m_{N+1}^s, m_{M+1}^s)) dy, \tag{21}
\end{align*}
\]

Let

\[
\Phi(s_{iN-1}) \equiv \sum_{\forall h=1}^{N} \Pr^* (g_{N-1,h} | \hat{a}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, m_{M+1}^s) p_{iN-1}^{sG_{N-1,k}^N}(\hat{a}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, m_{M+1}^s)
\]

be the probability that \( i_{N-1} \) is shortlisted and wins the asset at the end. We first show that this probability is increasing in \( s_{iN-1} \).

**Lemma 8.** \( \Phi(s_{iN-1}) \) increases in \( s_{iN-1} \).

**Proof.** See Appendix.

Thus we have (1) \( \pi_{iN-1}(s_{iN-1}, \hat{s}_{iN-1}, s_{iN-1}) | \hat{a}, m_2, ..., m_N \) satisfies the strict and smooth single crossing differences property in \( (s_{iN-1}, \hat{s}_{iN-1}) \); (2) \( \Phi(s_{iN-1}) \) increases in \( \hat{s}_{iN-1} \); (3) \( t_{N,iN-1}^* \) defined above verifies the envelope theorem. By the constraint simplification theorem again, it is incentive compatible for agent \( i_{N-1} \) to reveal \( s_{iN-1} \) truthfully at stage \( N \).

Following the same procedure and going backward stage by stage, for \( k = N - 2, N - 3, ..., 3, 2 \), we can construct payment rules for the corresponding stages to establish that agent \( i_{k-1} \) would reveal her type \( s_{i_{k-1}} \) truthfully given allocation rules \( (A^*, p^*) \) and the constructed payments rule \( t^* \) as long as \( \hat{a}_{i_l} = a_{i_l} \), for \( l \geq k + 1 \). In particular, the constructed transfers are given by

\[
\begin{align*}
t_{k,k-1}^* (\hat{a}, m_2, ..., m_{k-1}, m_k^s) \\
= E_{(s_{ik+1}, ..., s_{iN})} (u(a_{ik+1}, s_{ik+1}) \sum_{\forall h=1}^{N} \Pr^* (g_{N-1,h} | \hat{a}, m_2, ..., m_{k-1}, m_k^s, m_{k+1}^s, m_{M+1}^s) p_{ik-1}^{sG_{k-1,k}^N}(\hat{a}, m_2, ..., m_{k-1}, m_k^s, m_{k+1}^s, m_{M+1}^s))
\end{align*}
\]
where in \( m_k^s(y) \), \( s_{i,N-1} \) is replaced by \( y \).

These results are formally stated in the following proposition.

**Proposition 2.** Suppose \( \hat{a} = a \). Under shortlisting and allocation rules \((A^*, p^*)\) and payments rule \( t^* \) constructed above, in stage \( k \in \{2, \ldots, N+1\} \), regardless of others’ reports \((m_2, \ldots, m_{k-1})\) in the history, the most recently shortlisted agent \( i_{k-1} \) would reveal her \( s_{i_{k-1}} \) truthfully.

In fact, in the preceding arguments we show that so long as \( \hat{a}_{i_l} = a_{i_l}, l \geq k + 1 \), truthful revelation for stages \( l \geq k + 1 \) also follows. We now turn to stage 1. We will show that given \( \hat{a}_{i_l} = a_{i_l} \), under shortlisting and allocation rule \((A^*, p^*)\), we can construct stage-1 payments such that agent \( i \) reveals \( a_i \) truthfully and \( \pi_i(a, a) = 0, \forall i \). In addition, the shortlisted agent has the incentive to incur cost \( c \).

By the same logic as in Lemma 4 of Esö and Szentes (2007), when agent \( i \) reports \( \hat{a}_i \) at stage 1, she will report \( \sigma_i(a_i, \hat{a}_i, s_i) \) when she is shortlisted later and asked to report her second type, where \( u(a_i, s_i) = u(\hat{a}_i, \sigma_i(a_i, \hat{a}_i, s_i)) \). In the proof of Corollary 1 in Esö and Szentes (2007), they show that \( w_i(a_i, s_i) \leq w_i(\hat{a}_i, \sigma_i(a_i, \hat{a}_i, s_i)) \) if and only if \( a_i \leq \hat{a}_i \).

Let \( r(\hat{a}_i, a_{i-1}) \) denote the rank of \( \hat{a}_i \) in \((\hat{a}_i, a_{i-1})\), and \( m_{r(\hat{a}_i, a_{i-1})+1} \) denote the stage \( r(\hat{a}_i, a_{i-1}) \) reports in which agent \( i’s \) report is \( \sigma_i(a_i, \hat{a}_i, s_i) \). Further assume that all shortlisted agents get a subsidy of \( c \) from the seller besides the stage-1 transfer \( t_{1,i}(\cdot) \) to make sure that they have the incentive to conduct the due diligence. At stage 1, agent \( i’s \) expected payoff when \( i \) is of type \( a_i \) but announces \( \hat{a}_i \) is:

\[
\pi_i(a_i, \hat{a}_i) = E_{a_{i-1}} E_s \left\{ -\sum_{\forall h=r(\hat{a}_i, a_{i-1})}^N \Pr^* (g_{1, h} | (\hat{a}_i, a_{i-1}), m_{r(\hat{a}_i, a_{i-1})+1}^s, \ldots, m_{M+1}^s) \\
\quad \times \left[ t_{r(\hat{a}_i, a_{i-1})+1}^*(\hat{a}_i, a_{i-1}, m_{r(\hat{a}_i, a_{i-1})+1}^s, \ldots, m_{M+1}^s) - c \right] \\
+ \sum_{\forall h=r(\hat{a}_i, a_{i-1})}^N \Pr^* (g_{1, h} | (\hat{a}_i, a_{i-1}), m_{r(\hat{a}_i, a_{i-1})+1}^s, \ldots, m_{M+1}^s) \\
\quad \times \left[ u(a_i, s_i) \Pr_{l_{1,i}}^* (g_{1, h} | (\hat{a}_i, a_{i-1}), m_{r(\hat{a}_i, a_{i-1})+1}^s, \ldots, m_{M+1}^s) - c \right] \\
- E_{a_{i-1}} E_{t_{1,i}} \left[ -t_{1,i}^*(\hat{a}_i, a_{i-1}) \right] \right\}
\]  

We are now ready to pin down the transfer \( t_{1,i}^*(\cdot) \) (net of the entry subsidy \( c \)) that induces truthful revelation in stage 1.

By the envelop theorem, optimality of truthful revelation requires

\[
\frac{d\pi_i(a_i, a_i)}{da_i} = 0
\]
Recall that we set $\pi_i(\alpha_i, \alpha_i) = 0$. We thus have

$$\pi_i(\alpha_i, \alpha_i) = \int_a^\alpha E_s u_i(y, s_i) \left[ \sum_{\forall h = r(\alpha_i, \alpha_i)} \Pr^* (g_{1, h} | (y, s_i), m^s_2, \ldots, m^s_{r(\alpha_i, \alpha_i) + 1}, \ldots, m^s_{M+1}) \right] dy. \quad (24)$$

By (23) and (24), we define

$$t^*_{1,i}(\alpha_i) = E_s \left[ - \sum_{\forall h = r(\alpha_i, \alpha_i)} \Pr^* (g_{1, h} | (\hat{\alpha}_i, \hat{\alpha}_i - 1), m^s_2, \ldots, m^s_{r(\hat{\alpha}_i, \hat{\alpha}_i - 1) + 1}, \ldots, m^s_{M+1}) \right]$$

$$+ \sum_{\forall h = r(\hat{\alpha}_i, \hat{\alpha}_i - 1)} \Pr^* (g_{1, h} | (\hat{\alpha}_i, \hat{\alpha}_i - 1), m^s_2, \ldots, m^s_{r(\hat{\alpha}_i, \hat{\alpha}_i - 1) + 1}, \ldots, m^s_{M+1})$$

$$\left[ u(\hat{\alpha}_i, s_i) p^*_i (g_{1, h} | (\hat{\alpha}_i, \hat{\alpha}_i - 1), m^s_2, \ldots, m^s_{r(\hat{\alpha}_i, \hat{\alpha}_i - 1) + 1}, \ldots, m^s_{M+1}) \right]$$

$$- \int_a^\alpha E_s u_1(y, s_i) \left[ \sum_{\forall h = r(\alpha_i, \alpha_i)} \Pr^* (g_{1, h} | (y, \alpha_i), m^s_2, \ldots, m^s_{r(\alpha_i, \alpha_i) + 1}, \ldots, m^s_{M+1}) \right] dy. \quad (25)$$

To show that it is incentive compatible for agent $i$ to reveal truthfully at stage 1 under $(A^*, p^*, t^*)$, we need to show

$$\pi_i(\alpha_i, \alpha_i) \geq \pi_i(\alpha_i, \hat{\alpha}_i), \forall \alpha_i, \hat{\alpha}_i. \quad (26)$$

It is not readily clear that $\pi_i(\alpha_i, \hat{\alpha}_i)$ satisfies the single crossing property. As such, to establish IC we will turn to an alternative argument other than the constraint simplification theorem.

**Proposition 3.** Under $(A^*, p^*, t^*)$, we have

$$\pi_i(\alpha_i, \alpha_i) \geq \pi_i(\alpha_i, \hat{\alpha}_i), \forall \alpha_i, \hat{\alpha}_i. \quad (27)$$

**Proof.** See Appendix. □

The proof follows arguments paralleling to those used to establish Lemma 5 in Esö and Szentes,
which relies on both Assumptions 1 and 2. Given Propositions 2 and 3, the (sequential) shortlisting and allocation rule \((A^*, p^*)\) is IC implementable. Note that by reporting a higher first-stage signal, a bidder would be shortlisted with a higher probability, and would be shortlisted in a smaller group with a higher probability, given \((a_{-i}, s)\). Given any shortlisted group, reporting a higher first-stage signal and correcting the lie later would raise the bidder's virtual value, and thus the winning probability. We thus conclude that reporting a higher first-stage type leads to a higher winning probability at the end. This suggests that the monotonicity of the winning probability, which is typically required for incentive compatibility, is also satisfied.

In the following proposition, we establish that the proposed mechanism \((A^*, p^*, t^*)\) achieves the revenue bound identified in the relaxed environment in Section 4.1.

**Proposition 4.** Mechanism \((A^*, p^*, t^*)\) generates the same seller expected revenue as in the relaxed environment of Section 4.1, in which the bidders' second-stage additional information is public.

This result is clear since at the optima of both settings, the allocation rule (including the shortlisting rule and the object allocation rule) is exactly the same, and the expected payoff for bidders with the lowest type is set to be zero. This implies that the total expected surplus in both environments is the same. In addition, given (15) and (24), a bidder with the same type enjoys the same expected payoff across the two scenarios. Since seller expected revenue is the difference between the total expected surplus and expected bidder payoff, the seller revenue must be the same in the two environments.

By Proposition 4, we conclude that mechanism \((A^*, p^*, t^*)\) must be optimal when there is \(M = N\) stages of shortlisting. Any \(M(\geq N)\) would induce the same expected revenue at the optimum, and any \(M(< N)\) is dominated by \(M = N\). We can thus set the optimal number of shortlisting stages \(M^* = N\).

5 Discussion

5.1 Revelation Policy

In our preceding analysis, we have focused on the revelation policy so that the first-stage reports are fully revealed to the shortlisted bidders. Due to this particular revelation policy, one concern is that there might be some loss of generality in identifying optimal mechanisms. To address this concern, we focus on the case when shortlisting is restricted to a single round (the analysis and conclusion can be carried over to the case with sequential shortlisting). Our strategy is to identify an upper bound for the expected revenue that can be achieved by examining a relaxed setting by dropping the IC and IR constraints for the shortlisted bidders in the second stage so that all shortlisted bidders must incur entry costs to learn their second-stage signals as in our original setup, and regardless of their second-stage signals, they must participate in the second-stage selling mechanism and report truthfully their second stage signals. As a result, regardless of the disclosure policy of the first-stage reports, the highest possible expected

\(^{27}\)In Section 4, we have established the irrelevance of ex post information in the sequential shortlisting setting.
revenue achievable in this relaxed setting should impose an upper bound for the expected revenue that can be obtained in our original setup, where the bidders’ second-stage IC and IR constraints must both be satisfied. A useful observation is that in the relaxed setting, bidders can only misreport their first-stage signals, and the shortlisted bidders’ beliefs on bidders’ first-stage type profiles have no impact on their second-stage decisions (as shortlisted bidders must enter and truthfully report their second-stage signals). This observation implies that the revelation policy of the first-stage reports is not relevant to the mechanism design in the relaxed setting. Consequently, the highest expected revenue attainable in this relaxed setting does not depend on the prevailing disclosure policy of the first-stage signals. We next proceed to identify this bound.

In the relaxed setting, the mechanisms are specified exactly the same as in Section 2. All potential bidders report their types $\alpha_i$, giving rise to a reported type profile $\alpha$. The mechanism specifies the first-stage shortlisting rule $A^\theta(\alpha)$ and payment rule $x_i(\alpha_i, \alpha_{-i})$. Every shortlisted bidder $j$ incurs cost $c$ to discover her second-stage signal $s_j$. The second-stage selling mechanism specifies the winning probability $p^\theta_i(\alpha, s^g)$ and payment rule $t^\theta_i(\alpha, s^g)$, $\forall i \in g$, $\forall g \in 2^N$.

Recall that $P^g_i(\alpha, s_i) = E_{s^g_j} p^\theta_i(\alpha, s^g)$ and $T^g_i(\alpha, s_i) = E_{s^g_j} t^\theta_i(\alpha, s^g)$. For shortlisted bidder $i \in g_i$ with type $\alpha_i$, her interim expected payoff when she reports $\hat{\alpha}_i$ and others report truthfully is given by

$$\pi_i(\alpha_i, \hat{\alpha}_i) = E_{\alpha_{-i}} \left[ \sum_{g^i} A^{g_i}(\hat{\alpha}_i, \alpha_{-i}) \left( E_{s_i} \left[ u(\alpha_i, s_i) P^g_i(\hat{\alpha}_i, \alpha_{-i}, s_i) - T^g_i(\hat{\alpha}_i, \alpha_{-i}, s_i) \right] - c \right) - x_i(\hat{\alpha}_i, \alpha_{-i}) \right]. \quad (26)$$

Applying the envelope theorem, the IC condition $\pi_i(\alpha_i, \alpha_i) \geq \pi_i(\alpha_i, \hat{\alpha}_i)$ leads to the following necessary condition:

$$\frac{d\pi_i(\alpha_i, \alpha_i)}{d \alpha_i} = \frac{\partial \pi_i(\alpha_i, \hat{\alpha}_i)}{\partial \alpha_i} |_{\hat{\alpha}_i = \alpha_i} = E_{\alpha_{-i}} \left[ \sum_{g^i} A^{g_i}(\alpha_i, \alpha_{-i}) E_{s_i} [u_1(\alpha_i, s_i) P^g_i(\alpha_i, \alpha_{-i}, s_i)] \right].$$

Therefore, we have

$$\pi_i(\alpha_i, \alpha_i) = \pi_i(\alpha, \alpha) + E_{\alpha_{-i}} \int_\alpha^{\alpha_i} \sum_{g^i} A^{g_i}(y, \alpha_{-i}) E_{s_i} \left[ u_1(y, s_i) P^g_i(y, \alpha_{-i}, s_i) \right] dy = \pi_i(\alpha, \alpha) + E_{\alpha_{-i}} \int_\alpha^{\alpha_i} \int u_1(y, s_i) \sum_{g^i} A^{g_i}(y, \alpha_{-i}) P^g_i(y, \alpha_{-i}, s_i) dG_i(s_i) dy.$$

Note that the above expression is exactly the same as (6), which implies that the seller’s expected revenue must be the same as in (7); in other words, the upper bound of the expected revenue in the relaxed setting is achieved in our original setting. In this sense, there is no loss of generality to derive optimal mechanisms by only considering mechanisms that fully reveal the buyers’ first-stage reports to all admitted bidders.
Another important aspect in our analysis is that we model information acquisition as entry. An implication is that information acquisition is mandatory, in the sense that a bidder is not allowed to bid without going through the “due diligence” process. This assumption is due to the specific institutional setup we are trying to model. For example, “data rooms” are usually provided by the selling party to disclose a large amount of confidential data to bidders during the due diligence process. A typical data room is a continually monitored space that the bidders and their advisers will visit in order to inspect and report on the various documents and data made available. Often only one bidder at a time will be allowed to enter a data room. Teams involved in large due diligence processes will typically remain available throughout the process. Such teams often consist of a number of experts in different fields, hence the overall cost of keeping such groups on call near to the data room is often extremely high. In a typical electrical generating asset sale as studied by Ye (2007), before submitting a final bid, each bidder (more precisely, bidding team) usually needs to go through the due diligence process to meet with senior management and personnel, study equipment conditions and operating history, evaluate supply contracts and employment agreements, etc. This process is strictly controlled and closely monitored by the auctioneer (typically an investment banker serving as the financial advisor for the selling party). Given the complexity and high-stakes nature of the sale, it is very unlikely that a seller would be comfortable accepting a bid from someone who did not go through such an important information acquisition process. As such, we believe that it is appropriate to model information acquisition as entry for such an environment. From both theoretical and practical perspectives, it would be interesting to identify optimal mechanisms in environments where bidders are allowed to bid without having to go through information acquisition (and information acquisition may not be observable or contractible). Such an analysis would be more involved, however, as we will need to worry that the informed and uninformed buyers may mimic each other, a moral hazard issue also studied in Krähmer and Strausz (2011).

6 Concluding Remarks

Our paper contributes to the literature on two fronts. First, it characterizes optimal mechanisms, either when shortlisting is restricted or unrestricted to a single round, for an environment of multi-stage auctions, which are commonly employed in sales of complicated and high-valued business assets, procurements, privatization, takeover, and merger and acquisition contests. Our analysis is general enough to nest many existing studies in the literature of auctions with costly entry. Second, our paper contributes to the literature on sequential screening by introducing costly and endogenous information acquisition into a dynamic auction framework. Information acquisition makes the optimal mechanism design more challenging, as now it must balance bidders’ information acquisition incentives and information elicitation in

28See Vallen and Bullinger (1999) for a detailed description of the due diligence process in a typical electric power plant sale in the US.
the final good allocation stage, which are interdependent.

Since single-round shortlisting can be trivially replicated by sequential shortlisting, the optimal two-stage mechanism characterized in Section 3 must be revenue-dominated by the optimal mechanism allowing for sequential shortlisting characterized in Section 4. This is true when there is no time discounting. When time discounting is taken into account, however, an obvious drawback of running a multi-stage mechanism is the potential of delay, which would be too costly and therefore favors a more time-efficient two-stage mechanism. We believe that this consideration, along with the practical difficulty in administering multiple rounds of the due diligence process, leads to the “norm” of the two-stage auction format widely used in the real world.

Implementation of the optimal mechanism characterized in this paper may face some practical obstacles. First, the industry may not be comfortable with the idea of paying entry fees before knowing the auction outcome, and this is the major reason, we believe, that contributes to the common use of nonbinding indicative bidding. Second, the optimal mechanism is so complicated that the industry bidders might face great difficulties in developing bidding strategies for different rounds (although such a concern is alleviated to some extent if professional or sophisticated experts are hired to help). For these reasons the nature of our analysis is primarily normative, offering a “market design” approach to guide a potential refinement of an extremely important transaction procedure widely used in the industry. Despite this limitation, our analysis does conform to the “norm” of business in at least two aspects. First, a defining feature of our optimal mechanism is the shortlisting rule, which is also central in the two-stage auction practices. Second, we demonstrate that the optimal number shortlisted is endogenously determined, which is also consistent with the fact that in real sales, the number of finalists is often not pre-determined.

Our analysis offers a theoretical benchmark for evaluating various two-stage or multi-stage auctions currently used in the real world. The information structure modeled in this research has recently received attention not only from theorists but also from econometricians and empiricists. For example, Marmer, Shneyerov, and Xu (2013) and Gentry and Li (2014) have successfully proposed nonparametric specification tests on a so-called affiliated-signal (AS) model with entry, and Roberts and Sweeting (2013) estimate a parametric variant of the AS model using data on California timber auctions. The affiliated-signal models can be regarded as a special case in the framework studied in our paper, and the optimal mechanism characterized in this paper may potentially serve as a calibration benchmark for counter-factual simulations for related empirical work to come.

---

29Just imagine, for example, the hassle of arranging multiple meetings with senior management.

30For example, in the sale of PGW (Philadelphia Gas Works), a recent application of two-stage auctions, a “smaller number” of firms were invited to submit final bids after the first round – although this number was neither pre-announced nor disclosed (CBS Phily, November 19, 2013, “Sell-off of Philadelphia’s Natural Gas Utility Goes To Binding Bidding,” by Mike Dunn).
Proof of Lemma 3: Let $g_i$ denote any subset that includes $i$. By (5) and Lemma 2, we have

$$\pi_t(\alpha_i, \hat{\alpha}_i)$$

$$= \pi_t(\hat{\alpha}_i, \hat{\alpha}_i) + E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\hat{\alpha}_i, \alpha_{-i}) \left[ \bar{\pi}_{g_i}^{R}(\hat{\alpha}_i, \hat{\alpha}_i; \alpha_{-i}) - \bar{\pi}_{g_i}^{R}(\hat{\alpha}_i, \hat{\alpha}_i; \alpha_{-i}) \right] \right\}$$

$$= \pi_t(\hat{\alpha}_i, \hat{\alpha}_i) + E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\hat{\alpha}_i, \alpha_{-i}) \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P_{g_i}^{R}(\hat{\alpha}_i, \alpha_{-i}, \sigma(y, \hat{\alpha}_i, s_i)) dy dG_i(s_i) \right\}.$$

Thus for $\hat{\alpha}_i < \alpha_i$, $\pi_t(\alpha_i, \hat{\alpha}_i) \leq \pi_t(\alpha_i, \alpha_i)$ implies that

$$\pi_t(\alpha_i, \alpha_i) - \pi_t(\hat{\alpha}_i, \hat{\alpha}_i) \geq E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\alpha_i, \alpha_{-i}) \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P_{g_i}^{R}(\alpha_i, \alpha_{-i}, \sigma(y, \alpha_i, s_i)) dy dG_i(s_i) \right\}.$$

Similarly,

$$\pi_t(\hat{\alpha}_i, \alpha_i)$$

$$= \pi_t(\alpha_i, \alpha_i) + E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\alpha_i, \alpha_{-i}) \left[ \bar{\pi}_{g_i}^{R}(\hat{\alpha}_i, \alpha_i; \alpha_{-i}) - \bar{\pi}_{g_i}^{R}(\alpha_i, \alpha_i; \alpha_{-i}) \right] \right\}$$

$$= \pi_t(\alpha_i, \alpha_i) - E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\alpha_i, \alpha_{-i}) \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P_{g_i}^{R}(\alpha_i, \alpha_{-i}, \sigma(y, \alpha_i, s_i)) dy dG_i(s_i) \right\}.$$

Thus for $\hat{\alpha}_i < \alpha_i$, $\pi_t(\hat{\alpha}_i, \alpha_i) \leq \pi_t(\alpha_i, \alpha_i)$ implies that

$$\pi_t(\alpha_i, \alpha_i) - \pi_t(\hat{\alpha}_i, \hat{\alpha}_i) \leq E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\alpha_i, \alpha_{-i}) \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P_{g_i}^{R}(\alpha_i, \alpha_{-i}, \sigma(y, \alpha_i, s_i)) dy dG_i(s_i) \right\}.$$

So

$$E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\alpha_i, \alpha_{-i}) \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P_{g_i}^{R}(\alpha_i, \alpha_{-i}, \sigma(y, \hat{\alpha}_i, s_i)) dy \right\}$$

$$\leq \frac{\pi_t(\alpha_i, \alpha_i) - \pi_t(\hat{\alpha}_i, \hat{\alpha}_i)}{\alpha_i - \hat{\alpha}_i} \leq E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\alpha_i, \alpha_{-i}) \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P_{g_i}^{R}(\alpha_i, \alpha_{-i}, \sigma(y, \alpha_i, s_i)) dy \right\}.$$

By Fubini's Theorem, we have

$$E_{a_{-i}} \left\{ \sum_{g_i} A_{g_i}(\hat{\alpha}_i, \alpha_{-i}) \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) P_{g_i}^{R}(\hat{\alpha}_i, \alpha_{-i}, \sigma(y, \hat{\alpha}_i, s_i)) dy \right\}$$

$$= \sum_{g_i} \int_{\hat{\alpha}_i}^{\alpha_i} u_1(y, s_i) E_{a_{-i}} \left\{ A_{g_i}(\hat{\alpha}_i, \alpha_{-i}) P_{g_i}^{R}(\hat{\alpha}_i, \alpha_{-i}, \sigma(y, \hat{\alpha}_i, s_i)) \right\} dy dG_i(s_i).$$
Since $A^{g_i}, P_i^{g_i} \leq 1$, and $u$ is concave in $\alpha_i$, we have
\[
\frac{\int_{\hat{a}_i} u(y, s_i)E_{\alpha_i} \left[ A^{g_i}(\hat{a}_i, \alpha_{-i})P_i^{g_i}(\hat{a}_i, \alpha_{-i}, \sigma_i(y, \hat{a}_i, s_i)) \right] dy}{\alpha_i - \hat{a}_i} \leq \frac{\int_{\hat{a}_i} u(y, s_i)dy}{\alpha_i - \hat{a}_i} \leq u_1(\hat{a}_i, s_i).
\]

By assumption $u_1(\hat{a}_i, s_i)$ has a finite expectation with respect to $s_i$. Hence, by the Lebesgue convergence theorem,
\[
\lim_{\hat{a}_i \to a_i} E_{\alpha_i} \left\{ \sum_{g_i} A^{g_i}(\hat{a}_i, \alpha_{-i}) \int \frac{\int_{\hat{a}_i} u(y, s_i)P_i^{g_i}(\hat{a}_i, \alpha_{-i}, \sigma_i(y, \hat{a}_i, s_i)) dy}{\alpha_i - \hat{a}_i} dG_i(s_i) \right\}
= \sum_{g_i} \int \lim_{\hat{a}_i \to a_i} \left\{ u_1(\hat{a}_i, s_i)E_{\alpha_i} \left[ A^{g_i}(\hat{a}_i, \alpha_{-i})P_i^{g_i}(\hat{a}_i, \alpha_{-i}, s_i) \right] \right\} dG_i(s_i)
= \sum_{g_i} \{ u_1(\alpha_i, s_i)E_{\alpha_i} \left[ A^{g_i}(\alpha_i, \alpha_{-i})P_i^{g_i}(\alpha_i, \alpha_{-i}, s_i) \right] \} dG_i(s_i)
= E_{\alpha_i} \left\{ \sum_{g_i} A^{g_i}(\alpha_i, \alpha_{-i}) \int u_1(\alpha_i, s_i)P_i^{g_i}(\alpha_i, \alpha_{-i}, s_i) dG_i(s_i) \right\}.
\]

The third equality above is due to the assumption that $E_{\alpha_{-i}} \left[ A^{g_i}(\hat{a}_i, \alpha_{-i})P_i^{g_i}(\hat{a}_i, \alpha_{-i}, s_i) \right]$ is continuous in $\hat{a}_i$ (which is guaranteed as long as both $A^{g_i}$ and $P_i^{g_i}$ are continuous a.e. in $[\alpha, \alpha']$).

Analogously, we can show that
\[
\lim_{\hat{a}_i \to a_i} E_{\alpha_i} \left\{ \sum_{g_i} A^{g_i}(\alpha_i, \alpha_{-i}) \int \frac{\int_{\hat{a}_i} u(y, s_i)P_i^{g_i}(\alpha_i, \alpha_{-i}, \sigma_i(y, \alpha_i, s_i)) dy}{\alpha_i - \hat{a}_i} dG_i(s_i) \right\}
= E_{\alpha_i} \left\{ \sum_{g_i} A^{g_i}(\alpha_i, \alpha_{-i}) \int u_1(\alpha_i, s_i)P_i^{g_i}(\alpha_i, \alpha_{-i}, s_i) dG_i(s_i) \right\}.
\]

Thus the left derivative of $\pi_i(\alpha_i, a_i)$ is given by
\[
\frac{d\pi_i^-}{d\alpha_i}(\alpha_i, a_i) = E_{\alpha_i} \left\{ \sum_{g_i} A^{g_i}(\alpha_i, \alpha_{-i}) \int u_1(\alpha_i, s_i)P_i^{g_i}(\alpha_i, \alpha_{-i}, s_i) dG_i(s_i) \right\}.
\]

Working with the case $\hat{a}_i > a_i$, we can obtain the right derivative of $\pi_i(\alpha_i, a_i)$, which is given by
\[
\frac{d\pi_i^+}{d\alpha_i}(\alpha_i, a_i) = E_{\alpha_i} \left\{ \sum_{g_i} A^{g_i}(\alpha_i, \alpha_{-i}) \int u_1(\alpha_i, s_i)P_i^{g_i}(\alpha_i, \alpha_{-i}, s_i) dG_i(s_i) \right\}.
\]

Therefore, we conclude that $\pi_i(\alpha_i) = \pi_i(\alpha_i, a_i)$ is differentiable everywhere, and
\[
\pi'_i(\alpha_i) = E_{\alpha_i} \left\{ \sum_{g_i} A^{g_i}(\alpha_i, \alpha_{-i}) \int u_1(\alpha_i, s_i)P_i^{g_i}(\alpha_i, \alpha_{-i}, s_i) dG_i(s_i) \right\}.
\]
\[= \int u_1(a_i, s_i) \cdot \sum_{g_i} \left[ E_{a_i} A^g_i(a_i, a_{-i}) P^g_i(a_i, a_{-i}, s_i) \right] dG_i(s_i) \]

Since \( \pi'_i(a_i) \) is bounded over \([a, \bar{a}]\), \( \pi_i \) satisfies a Lipschitz condition and hence it can be recovered from its derivative, which gives rise to (6).

**Proof of Lemma 4:** Note that

\[ E_a[\pi_i(a_i, a_i)] \]

\[= \pi_i(a_i, a_i) + \]

\[ E_{a \rightarrow a_i} \int_{\alpha}^{a_i} \sum_{\forall g \text{ s.t. } i \in G_g} u_1(\alpha, s_i) \left[ \Pr(g | (\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \cdot p^g_i((\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \right] d\alpha \]

\[= \pi_i(a_i, a_i) + \]

\[ E_{a \rightarrow a_i} \int_{\alpha}^{a_i} \sum_{\forall g \text{ s.t. } i \in G_g} u_1(\alpha, s_i) \left[ \Pr(g | (\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \cdot p^g_i((\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \right] f(\alpha_i) d\alpha_i \]

\[= \pi_i(a_i, a_i) + \]

\[ E_{a \rightarrow a_i} \int_{g}^{a_i} \sum_{\forall g \text{ s.t. } i \in G_g} u_1(\alpha, s_i) \left[ \Pr(g | (\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \cdot p^g_i((\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \right] \int_{y}^{\pi} f(\alpha_i) d\alpha_i \]

\[= \pi_i(a_i, a_i) + \]

\[ E_{a \rightarrow a_i} \int_{\alpha}^{a_i} \sum_{\forall g \text{ s.t. } i \in G_g} u_1(\alpha, s_i) \left[ \Pr(g | (\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \cdot p^g_i((\alpha, a_{-i}), m^s_2, ..., m^s_{M+1}) \right] \frac{1 - F(\alpha_i)}{f(\alpha_i)} d\alpha_i \]

\[= \pi_i(a_i, a_i) - E_{a \rightarrow a_i} \int_{\alpha}^{a_i} \sum_{\forall g \text{ s.t. } i \in G_g} u_1(\alpha, s_i) \left[ \Pr(g | (\alpha, a_{-i}, m^s_2, ..., m^s_{M+1}) \cdot p^g_i((\alpha, a_{-i}, m^s_2, ..., m^s_{M+1}) \right] \frac{1 - F(\alpha_i)}{f(\alpha_i)} d\alpha_i \]

Therefore, the seller's expected revenue is given by

\[ ER = E_a \left[ TS(a) - \sum_{i \in \mathbb{N}} \pi_i(a_i, a_i) \right] \]

\[= E_a \sum_{g} \left[ \Pr(g | (a, m^s_2, ..., m^s_{M+1}) \cdot \sum_{i \in G_g} p^g_i(a, m^s_2, ..., m^s_{M+1}) u(a_i, s_i) - c) \right] \]

\[- \sum_{i \in \mathbb{N}} \pi_i(a_i, a_i) \]

\[E_a \sum_{g} \Pr(g | (a, m^s_2, ..., m^s_{M+1}) \cdot \sum_{i \in G_g} p^g_i(a, m^s_2, ..., m^s_{M+1}) \frac{1 - F(\alpha_i)}{f(\alpha_i)} \]

\[- \sum_{i \in \mathbb{N}} \pi_i(a_i, a_i) - E_a \sum_{g} \Pr(g | (a, m^s_2, ..., m^s_{M+1}) \cdot \sum_{i \in G_g} p^g_i(a, m^s_2, ..., m^s_{M+1}) u(a_i, s_i) \frac{1 - F(\alpha_i)}{f(\alpha_i)} \]

33
We can rank these agents in

\[ G = \{ G \} \]  

can be generated by the one-agent-per-stage-until-the-last shortlisting rule described by Lemma 6.

\[ A \]  

shortlisted in subsequent stages.

\[ \]  

the single agent in \( G \) for \( G \) is the sum of the probabilities of all the final shortlisted groups that contain \( A \) as the smallest indexed within the group. If agent \( A \) is shortlisted is simply the ratio between the sum of probabilities of \( j \) and focus on all \( G^+ \)’s that contain the single agent in \( g_1 \) as the smallest indexed. Let \( G^+(g_1) \) denote this set of \( G^+ \)’s. We are now ready to define \( A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) \), \( \forall g_2 \in 2^{N \backslash g_1} \). For \( g_2 = \emptyset \), we define \( A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) = 0 \) if \( g_1 \notin G^+(g_1) \), and \( A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) = \frac{Pr(G^+=g_2)\alpha_{\alpha}}{A^+(a|g_0)} \) if \( g_1 \in G^+(g_1) \). We let \( A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) = 0 \) for any group \( g_2 \in 2^{N \backslash g_1} \) that contains more than two agents. We use \( N(g_1) \) to denote the pool of the second smallest indexed agents in the \( G^+ \)’s in \( G^+(g_1) \). If \( N(g_1) \) is empty, then we let \( A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) = 0 \), \( \forall j \notin N(g_1) \). If \( N(g_1) \) is not empty, we let

\[ A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) = \begin{cases} \sum_{g_1=\alpha} A^+(a|g_0) & \text{if } j \in N(g_1), \\
0 & \text{if } j \notin N(g_1). \end{cases} \]

In words, for any agent \( j \in N(g_1) \), the probability that \( j \) is shortlisted conditional on that \( g_1 \) has been shortlisted is simply the ratio between the sum of probabilities of \( G^+ \)’s that contain the single agent in \( g_1 \) as the smallest indexed and \( j \) as the second smallest indexed and the sum of probabilities of \( G^+ \)’s that contain the single agent in \( g_1 \). By construction, we have \( A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) + \sum_{j \in N(g_1)} A_{g_2=(a,m_1^s,\ldots,m_1^{s}\alpha)}|g_0,g_1) = 1 \).
Similarly, for stage 3, we only need to consider these shortlisting histories \((g_1, g_2)\) where both \(g_1\) and \(g_2\) are nonempty single-element groups. We define \(\mathcal{G}^+ (g_1, g_2)\) as the set of all \(G^+\)'s that contain the single element in \(g_1\) as the smallest indexed, and the single element in \(g_2 \in \mathcal{G}^+ (g_1)\) as the second smallest indexed. We are now ready to define \(A^{g_3}(a, m_1^g, m_2^g | g_0, g_1, g_2)\), \(\forall g_3 \in 2^{N \setminus \{g_1 \cup g_2\}}\). For \(g_3 = \emptyset\), we define \(A^{g_3=\emptyset}(a, m_1^g, m_2^g | g_0, g_1, g_2) = 0\) if \(g_1 \cup g_2 \notin \mathcal{G}^+ (g_1, g_2)\) and \(A^{g_3=\emptyset}(a, m_1^g, m_2^g | g_0, g_1, g_2) = \frac{\Pr(G^+=g_1\cup g_2 | a, s)}{A^{g_1}(a, m_1^g|g_0, g_1) A^{g_2}(a, m_2^g|g_0, g_1)}\) if \(g_1 \cup g_2 \in \mathcal{G}^+ (g_1, g_2)\). We let \(A^{g_3}(a, m_1^g, m_2^g | g_0, g_1, g_2) = 0\) for any group \(g_3 \in 2^{N \setminus \{g_1 \cup g_2\}}\) that contains more than two agents. We use \(N(g_1, g_2)\) to denote the pool of the third smallest indexed agents in the \(G^+\)'s in \(\mathcal{G}^+ (g_1, g_2)\). If \(N(g_1, g_2)\) is empty, then we let \(A^{g_3=\emptyset}(a, m_1^g, m_2^g | g_0, g_1, g_2) = 0\), \(\forall j \in N\setminus \{g_1 \cup g_2\}\). If \(N(g_1, g_2)\) is not empty, then \(\forall j \in N(g_1, g_2)\), we let \(A^{g_3=\emptyset}(a, m_1^g, m_2^g | g_0, g_1, g_2) = \frac{\sum_{\alpha \in \mathcal{G}^+ (g_1, g_2)} \alpha_{g_1} a^* \frac{\Pr(G^+ = g_1 \cup g_2 \setminus g_j | a, s)}{A^{g_1}(a, m_1^g | g_0, g_1) A^{g_2}(a, m_2^g | g_0, g_1)} }{\sum_{\alpha \in \mathcal{G}^+ (g_1, g_2)} \alpha_{g_1} a^* \frac{\Pr(G^+ = g_1 \cup g_2 | a, s)}{A^{g_1}(a, m_1^g | g_0, g_1) A^{g_2}(a, m_2^g | g_0, g_1)}}\); and \(\forall j \notin N(g_1, g_2)\), we let \(A^{g_3=\emptyset}(a, m_1^g, m_2^g | g_0, g_1, g_2) = 0\). By construction, we have \(A^{g_3=\emptyset}(a, m_1^g, m_2^g | g_0, g_1, g_2) + \sum_{j \in N(g_1, g_2)} A^{g_3=\emptyset}(a, m_1^g, m_2^g | g_0, g_1, g_2) = 1\).

This process continues analogously until we exhaust all agents in every group \(G^+ \in \mathcal{G}^+(a, s)\). By construction, it is clear that this process of shortlisting one agent at each stage generates the same probabilities of \((\Pr(G^+ | a, s), \forall G \in 2^N, a, s)\). As a result, the above constructed shortlisting rule would generate the same expected revenue for the seller.

**Proof of Lemma 7:** Consider any given \(a\) and any Lemma 6 shortlisting rule \(A^{g_3}(a, m_1^g, m_2^g, ..., m_{k-1}^g | g_0, g_1, g_2, ..., g_{k-1})\).

\(\Pr(G^+ | a, s)\) denotes the conditional shortlisting probability of group \(G\). Without loss of generality, we assume \(a_i\) decreases with \(i\). In stage \(k \geq 1\), which agent \(i_k\) is shortlisted with probability 1 depends on \((a, m_1^g, m_2^g, ..., m_{k-1}^g | g_0, g_1, g_2, ..., g_{k-1})\). Note signals \(s_i, i \in N\) are i.i.d. and they are independent of \(a_i, i \in N\). We construct a new shortlisting rule under which the seller shortlists one agent at each stage before stopping the shortlisting, as follows. In stage 1, when \(i_1\) is shortlisted, we replace her by agent 1. In stage 2, we use the same shortlisting rule \(A^{g_3}(a, m_1^g | g_0, g_1)\) treating as if \(g_1 = \{i_1\}\), however, \(m_1^g\) now stands for agent 1's type instead. Because \(s_{i_1}\) and \(s_1\) follow the same distribution, the rule \(A^{g_3}(a, m_1^g | g_0, g_1)\) would generate the same agent \(i_2 \in N \setminus \{i_1\}\) to be shortlisted in stage 2 while which \(i_2\) is shortlisted might depend on the realization of \(m_1^g\). Whenever an \(i_2\) is shortlisted in stage 2, we replace her by agent 2. In stage \(k \geq 3\), we use the same shortlisting rule \(A^{g_3}(a, m_1^g, m_2^g, ..., m_{k-1}^g | g_0, g_1, ..., g_{k-1})\) treating as if \(g_k = \{i_k\}, 1 \leq h < k\), however, \(m_i^g\) now stands for agent \(i\)'s type instead. Because \(s_{i_k}\) and \(s_k\) follow the same distribution, the rule \(A^{g_3}(a, m_1^g, m_2^g, ..., m_{k-1}^g | g_0, g_1, ..., g_{k-1})\) would generate the same agent \(i_k \in N \setminus \cup_{h=1}^{k-1} \{i_h\}\). Whenever an \(i_k\) is shortlisted in stage \(k\), we replace her by agent \(k\). We use \(\hat{Pr}(G | a, s)\) to denote the conditional shortlisting probability of group \(G\) under the new rule. Let \(G_k = \{1, 2, ..., k\}, k = 1, 2, ..., N\). Note that \(\hat{Pr}(G | a, s) = 0\) for any \(G \in 2^N \setminus \{\emptyset, G_k, k = 1, 2, ..., N\}\).

Based on the above construction, \(\forall a, \text{for any } k = 1, 2, ..., N\) such that \(E_s \sum_{G \text{ s.t. } |G| = k} \hat{Pr}(G | a, s) > 0\), we
must have \( E_s \widehat{\Pr}(G_k | \alpha, s) = E_s \sum_{G \subseteq G, |G| = k} \Pr(G | \alpha, s) \). Therefore,

\[
E_s \left\{ \sum_{G \in 2^N} \Pr(G | \alpha, s) \max \{w_i^+(\alpha_i, s_i) \}_{i \in G} - \sum_{i \in G} c \right\} 
\]

\[= E_s \left\{ \sum_{k=1}^{N} \sum_{G \subseteq G, |G| = k} \Pr(G | \alpha, s) \max \{w_i^+(\alpha_i, s_i) \}_{i \in G} - kc \right\} \]

\[\geq E_s \left\{ \sum_{k=1}^{N} \sum_{G_k} \Pr(G_k | \alpha, s) \max \{w_i^+(\alpha_i, s_i) \}_{i \in G_k} - kc \right\} \]

\[= E_s \left\{ \sum_{G \in 2^N} \Pr(G | \alpha, s) \max \{w_i^+(\alpha_i, s_i) \}_{i \in G} - \sum_{i \in G} c \right\} \].

Therefore, by Lemma 5, the constructed shortlisting rule generates higher expected revenue.

**Proof of Lemma 8:** To simplify notation, given \((\hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s)\), let

\[
a_1(s_{i_{N-1}}) = \Pr^*(g_{N-1,N-1} | \hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s),
\]

\[
a_2(s_{i_{N-1}}) = \Pr^*(g_{N-1,N-1} | \hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s),
\]

\[
P_1(s_{i_{N-1}}) = p_{i_{N-1}}^{G_{N-1,N-1}} (\hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s),
\]

\[
P_2(s_{i_{N-1}}) = p_{i_{N-1}}^{G_{N-1,N}} (\hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s).
\]

In addition, let \(a_0(s_{i_{N-1}}) \) be the probability that \(i_{N-1} \) is not shortlisted given \((\hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s)\). So \(a_0(s_{i_{N-1}}) = 1 - a_1(s_{i_{N-1}}) - a_2(s_{i_{N-1}})\). We have

\[
\Phi(s_{i_{N-1}})
\]

\[= \sum_{h=N-1}^{N} \Pr^*(g_{N-1,N} | \hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s) p_{i_{N-1}}^{G_{N-1,N}} (\hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s) \]

\[= \Pr^*(g_{N-1,N} | \hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s) p_{i_{N-1}}^{G_{N-1,N}} (\hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s) \]

\[+ \Pr^*(g_{N-1,N} | \hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s) p_{i_{N-1}}^{G_{N-1,N}} (\hat{\alpha}, m_2, ..., m_{N-1}, m_N^s, m_{N+1}^s, ..., m_M^s) \]

\[= a_1(s_{i_{N-1}})P_1(s_{i_{N-1}}) + a_2(s_{i_{N-1}})P_2(s_{i_{N-1}}).
\]

When \(s_{i_{N-1}} \) increases, \(a_1(s_{i_{N-1}}) \) increases, \(a_2(s_{i_{N-1}}) \) decreases, and \(a_0(s_{i_{N-1}}) \) also decreases. That is, the probability that \(i_{N-1} \) will be shortlisted as the last one shortlisted increases, while the probability that \(i_{N-1} \) will be both shortlisted decreases and the probability that \(i_{N-1} \) will not be shortlisted also decreases. It is also clear that \(P_1(s_{i_{N-1}}) \) and \(P_2(s_{i_{N-1}}) \) are both increasing in \(s_{i_{N-1}} \). Besides, we have \(P_1(s_{i_{N-1}}) \geq P_2(s_{i_{N-1}}) \) \(i_{N-1} \) has a better chance of winning with a smaller set of entrants. Thus, given \(s'_{i_{N-1}}, s''_{i_{N-1}} \), \( s''_{i_{N-1}} > s'_{i_{N-1}} \), we have

\[36\]
Proof of Proposition 3: We can show that

\[
\pi_i(\alpha, \hat{s}_1) - \pi_i(\hat{s}_1, \hat{\alpha}) = \int_{\hat{\alpha}_i}^{a_i} E_{\alpha_1} E_{u_1(y, s_1)} \left[ \sum_{\forall h = (\hat{s}_1, \alpha_i)}^{N} \Pr^*(g_{1, h}|(\hat{s}_1, \alpha_i), m_2^s, \ldots, m_M^s) \right] \, dy.
\]

Consider

\[
\Delta = \pi_i(\alpha, \hat{s}_1) - \pi_i(\hat{s}_1, \hat{\alpha}) = [\pi_i(\alpha, \alpha_i) - \pi_i(\hat{s}_1, \hat{\alpha})] + [\pi_i(\hat{s}_1, \hat{\alpha}) - \pi_i(\hat{s}_1, \hat{\alpha})]
\]

\[
= \int_{\hat{\alpha}_i}^{a_i} E_{\alpha_1} E_{u_1(y, s_1)} \left[ \sum_{\forall h = (\hat{s}_1, \alpha_i)}^{N} \Pr^*(g_{1, h}|(\hat{s}_1, \alpha_i), m_2^s, \ldots, m_M^s) \right] \, dy
\]

\[
- \int_{\hat{\alpha}_i}^{a_i} E_{\alpha_1} E_{u_1(y, s_1)} \left[ \sum_{\forall h = (\hat{s}_1, \alpha_i)}^{N} \Pr^*(g_{1, h}|(\hat{s}_1, \alpha_i), m_2^s, \ldots, m_M^s) \right] \, dy
\]

\[
= \int_{\hat{\alpha}_i}^{a_i} E_{\alpha_1} E_{u_1(y, s_1)} \left[ \sum_{\forall h = (\hat{s}_1, \alpha_i)}^{N} \Pr^*(g_{1, h}|(\hat{s}_1, \alpha_i), m_2^s, \ldots, m_M^s) \right] \, dy
\]

\[
- \left[ \sum_{\forall h = (\hat{s}_1, \alpha_i)}^{N} \Pr^*(g_{1, h}|(\hat{s}_1, \alpha_i), m_2^s, \ldots, m_M^s) \right] \, dy.
\]

Note that

\[
\sum_{\forall h = (\hat{s}_1, \alpha_i)}^{N} \Pr^*(g_{1, h}|(\hat{s}_1, \alpha_i), m_2^s, \ldots, m_M^s) \, dy = 1.
\]
is agent $i$'s winning probability given her type $y$ and that she reports truthfully; and

$$
\sum_{\forall h = r(\hat{a}_i, \alpha_{-i})}^N \Pr^* (g_{1,h}|(\hat{a}_i, \alpha_{-i}), m_2^s, ..., m_{r(\hat{a}_i, \alpha_{-i})+1}^s, ..., m_{M+1}^s) p_{1}^{*G_{1,h}} ((\hat{a}_i, \alpha_{-i}), m_2^s, ..., m_{r(\hat{a}_i, \alpha_{-i})+1}^s, ..., m_{M+1}^s)$$

is agent $i$'s winning probability given her type $y$ and that she reports $\hat{a}_i$ in stage 1 and corrects her lie when shortlisted. Recall that, by the proof of Corollary 1 in Esö and Szentes (2007), $w_i(y, s_i) \leq w_i(\hat{a}_i, \sigma_i(y, \hat{a}_i, s_i))$ if and only if $y \leq \hat{a}_i$. \(\forall \alpha_{-i}, s\), by definition of $\mathbf{(A^*, p^*)}$, \(\forall \alpha_{-i}, s\), we conclude that the latter is smaller if and only if $y > \hat{a}_i$ because in the latter case a (weakly) bigger group is shortlisted, and $i$'s winning probability is smaller even when the group remains the same. Thus we must have $\Delta \geq 0$. 


REFERENCES


39


[26] Li, Yunan (2018), Efficient Mechanisms with Information Acquisition, Working Paper, City University of Hong Kong.


