Auction-Based Queue Disciplines

Thomas Kittsteiner and Benny Moldovanu*

March 16, 2003

Abstract

Each one of several impatient agents has a job that needs to be processed by a server. The server can process the jobs sequentially, one at a time. Agents are privately informed about the realization of a random variable representing processing time. If the cost of delay is represented by a concave function of waiting time till job completion, the efficient *shortest processing time first* schedule arises in the equilibrium of a simple auction where agents bids for slots in the queue. If the cost function is convex, the equilibrium yields the anti-efficient *longest processing time first* schedule. In this case, the performance of the auction (both efficiency and revenue) can be improved by capping bids from above. Finally, we show that the ex-post incentive compatible mechanisms that minimize the expected total waiting cost cannot depend on the private information available to the agents.

1 Introduction

Imagine that you arrive with a bulky package at the department's copier machine together with one of your esteemed colleagues. He/She says to you: "I have only a few pages to copy. May I copy first ?" Most of us usually agree to this request, and indeed this courtesy is well-founded in economic theory: total waiting cost is minimized if shorter jobs are processed before longer ones. Moreover, it seems intuitive that your colleague should be willing to pay more than you for the right to be first since he/she can thus avoid a longer delay than you. But sometimes the colleague's job turns out to be longer than the announced few pages ("I just noticed that I also need this chapter"), and in other situations everyone claims to have the shorter job... Are these announcements sincere? Should we base the queue discipline upon them? This paper is about the design of simple pricing mechanisms (e.g., auctions or lotteries) for allocating slots in a queue among impatient, privately informed agents.

Queuing theory[1] builds models in order to predict the behavior of systems providing service for randomly fluctuating customer demand. Besides a very large and valuable theoretical literature, there

*Kittsteiner: Nuffield College, Oxford University; Moldovanu: Department of Economics, University of Bonn; mold@wiwi.uni-bonn.de

[1] see for example the recent textbook by Gross and Harris (1998).

are many practical applications to traffic flow (e.g., vehicles at toll booths, aircraft at airport landing or take-off gates, network communication), scheduling (e.g., patients in hospitals, jobs on machines, programs on computers) and facility design (e.g., banks, post offices, amusement parks and fast-food restaurants). A standard task for the queuing analyst is to determine an appropriate measure for system performance (which depends of course on the customers' and system's characteristics) and to design an "optimal system" according to such a measure.

A queuing system is generally described by several basic characteristics pertaining to the stochastic arrival pattern of possibly impatient customers (e.g., inter-arrival times, the possibility to balk before or renege after entering the queue, the possibility to jockey for position, etc..) and to the pattern of service (e.g., system capacity, number of service channels, queue discipline, etc...).

Almost the entire queuing literature views customers as non-strategic agents endowed with some (possibly) random characteristics. Several papers consider the agents' decision on whether to queue or not, given a fixed queue discipline such as first-come first-served. This literature originated with Naor (1969). His main result is that a first-come first-served discipline leads to inefficient entry decisions, due to the externalities joining agents exert on their successors. If agents can submit a bid or bribe, and priority is given to agents with higher bids, Hassin (1995) showed that agents' joining decisions are socially optimal (see also the book of Hassin and Haviv, 2002 for a survey of these topics).

In the queueing literature the process generating the agents' characteristics is assumed to be common knowledge. In particular, the theory does not consider the interplay between system design (e.g., queue discipline) and the strategic incentives to manipulate accessible information arising from the desire to improve one's position in the queue and thus to increase one's utility.

In this paper we consider the most basic scheduling problem with one server and with impatient customers all arriving at the same time. Each customer has a job that needs to be processed, and the server can sequentially process one job at a time[2]. The design problem reduces to determining the allocation of slots in the waiting line. Each customer is privately informed about her needed processing time, and agents incur costs of delay. From the point of view of other agents, the processing time of a specific agent is a random variable governed by a common-knowledge distribution. The total waiting cost born by an agent i that is allocated the $j'th$ slot depends both on $i's$ privately known processing time and on the processing times of the agents scheduled to slots $1, 2, ..j-1$. Thus agent $i's$ utility depends both on the allocation of slots to other agents and on information that is ex-ante available to other agents. In the language of mechanism design, we obtain a setting with both allocative and informational externalities. An important assumption that is implicit in our analysis is that processing times are too costly to monitor ex-post or, equivalently, that no additional fees can be imposed after processing times have been realized and service has been completed.

In the above framework we first consider a natural auction procedure in which agents bid for slots: the highest bidder gets the first slot, the second-highest bidder gets the second slot, and so

[2] We assume here that all agents can be served, but our results do not qualitatively change if the number of slots is less than the number of jobs to be processed.

on till all slots are allocated. All agents pay their own bid. It turns out that the performance of this auction crucially depends on the form of the function describing the costs of delay. If this function is concave, agents with a shorter processing time bid in a Bayes-Nash equilibrium more than agents with a longer processing time and, for any realization of the stochastic processing times, the auction implements the efficient *shortest processing time first* (SEPT) schedule. In contrast, if the cost function is convex, agents with longer processing times bid in equilibrium more than agents with shorter times, and the auction implements the "anti-efficient" *longest processing time first* (LEPT) schedule.

Since the case of convex cost functions is the more pertinent one for most applications (e.g., consider the ubiquitous exponential cost functions or the presence of deadlines) it makes sense to inquire whether there are mechanisms that perform better than the auction in this case. We first analyze auctions with bid caps, i.e., auctions where agents are constrained to make bids that are lower than a pre-determined maximum. Since with convex cost functions the high bids come from agents with long processing times, constraining such bidders to a maximum bid implies that the allocation of slots among these bidders will contain a random element. Such a lottery necessarily improves upon the welfare attained by the LEPT schedule implemented by the unconstrained auction.

In many situations of interest the server is owned by an agent that needs to raise revenue in order to maintain operation and/or to make a profit (think about a data processing firm or an airport operator). At first sight it seems that, whenever revenue raising is important, constraining the high bidders via a bid cap will lower expected revenue. But this intuition is misleading since in the constrained auction there are in fact agents with low processing times that bid higher than their respective bids in the unconstrained auction. While bid caps are shown never to be revenue enhancing in an auction with concave cost functions, we find that bid caps may raise revenue if the cost function is convex (in some cases it is even revenue maximizing to impose an extremely low cap such that the auction degenerates to a pure lottery where agents pay a fixed fee). Thus, with convex cost functions, schedule auctions constrained by bid caps may perform better than unconstrained auctions on both efficiency and revenue measures.

Since we found that we can enhance performance by using procedures that do not, or only partially condition on privately available information (lotteries and capped auctions, respectively) we are next interested to characterize mechanisms that achieve the highest possible welfare for the customers (i.e., minimize total expect waiting costs) subject to the incentive compatibility constraint (i.e., subject to the requirement that their outcome arises in an equilibrium of a game played by strategic, privately informed agents). It turns out that, in an important class of mechanisms, lotteries are indeed optimal if the agents have the same expected processing time. If more information about the respective distribution of processing time is available (e.g, if it is known that one agent's distribution stochastically dominates another), this information can be used to determine optimal schedules. Surprisingly, even if it is the case that the individual distributions of processing times are stochastically ordered in the usual sense, it is not necessarily the case that the random variable governing the total costs of delay associated with the SEPT schedule (based on the expectations of processing times) is stochastically dominated by all other schedules. But such a result holds if the individual processing times are ordered in the likelihood ratio sense.

The basic scheduling problem with interdependent costs has been introduced by Hain and Mitra (2002)[3]. These authors show that concavity of the cost function is a necessary condition for the implementability of the efficient schedule in ex-post equilibria (note that any ex-post equilibrium is Bayes-Nash). Their main result is that, for cost functions that are concave polynomials of degree less than or equal to $n-2$ (where n is the number of agents and slots), a generalized Clarke-Groves-Vickrey mechanism can be constructed[4] that is efficient and ex-post budget balanced. Note that in our framework the auction's designer (who has no private information) is a residual claimant and budget balancedness is satisfied per definition.

Wellmann et. al. (2001) study private values scheduling problems without waiting costs: agents derive the same utility if their jobs are completed early or late as long as this is done before a deadline. These authors apply insights gained from the theory of matching markets. Holt and Sherman (1982) model a waiting-line as an auction. In their model buyers that differ in their opportunity costs for waiting queue to purchase one of several goods which are sold at fixed and known prices and points in time. The buyers' arrival time at the queue, which determines their probability of receiving one of the goods, can be interpreted as a bid in an all-pay auction[5].

Gavious, Moldovanu and Sela (2002) analyze a private value all-pay auction for a single object (there are no externalities of any kind) where the seller can impose bid caps. While a bid-cap is disadvantageous if the function describing the bid cost is linear or concave, it is shown that a bid cap may increase revenue if this function is convex. Note that in their framework the concavity/convexity issue pertains to the cost of bids, while in the present paper it pertains to the determination of values themselves. This crucial difference yields some contrasting insights.

This paper is organized as follows: In Section 2 we describe the design problem arising from a scheduling problem with waiting costs. In Section 3 we derive Bayes-Nash equilibria of a multi-object auction that allocates slots in the queue based on the respective bids. In Section 4 we study auctions with bid caps and analyze the effects of these on auction efficiency and revenue. In Section 5 we focus on the case of convex cost functions and show, for the case of two agents, that the welfare maximizing ex-post incentive compatible mechanism (which minimizes expected processing time) does not condition on the agents' private information. Finally, we connect this result to well-known insights about the application of stochastic orders to queuing problems.

[3] Mitra (2001) focuses on efficiency and budget-balancedness for simpler scheduling problems with private values.

[4] The construction of the CGV mechanism is based on general insights about efficient implementation for multi-object auctions with interdependent valuations due to Dasgupta and Maskin (2000) and Jehiel and Moldovanu (2001).

[5] Or in a discriminatory price auction if buyers can observe whether it is already too late to queue.

2 The Model

A processing unit (e.g., a computer, a landing/take-off gate, a repair tool) is available to n risk-neutral agents. Each agent needs to perform a task (or job), and the agents' jobs can be processed sequentially, one job at a time. Each agent i has information about her own needed processing time t_i, which is the realization of a random variable with support $[\underline{t}, \overline{t}]$. The distribution of t_i is given by a strictly increasing and continuous distribution function F with density f. F is common knowledge and the individual processing times are independently distributed[6].

A *schedule* is a permutation $\sigma \in \Sigma_n$, $\sigma = (\sigma(1), \ldots, \sigma(n))$ where Σ_n is the set of all one to one mappings $\sigma : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$. In particular, $\sigma(1)$ is the index of the agent who is served first, $\sigma(2)$ the index of the agent served second, and so forth. All agents bear a cost of waiting given by a strictly increasing function $C : \mathbb{R}^+ \mapsto \mathbb{R}^+$, i.e. an agent waiting for t time units for her job to be finished bears a cost $C(t)$. Given a schedule $\sigma \in \Sigma_n$, the cost borne by an agent who is served j'th is given by $C\left(\sum_{k=1}^{j} t_{\sigma(k)}\right)$,

All agents derive an utility V from the completion of their job (net of time costs). This utility is assumed to be common knowledge. We assume that $V \geq E_{(t_2,\ldots,t_n)} C\left(\overline{t} + \sum_{k=2}^{n} t_k\right)$ to ensure that expected waiting costs never exceed V.

Given a realization of types $(t_1, \ldots, t_n)$ and a schedule σ, the utility of an agent i with $\sigma(i) = j$ is given by $V - C\left(\sum_{k=1}^{j} t_{\sigma(k)}\right)$ since her waiting time is given the sum of processing times for all predecessors and her own processing time. Hence we have a model with *interdependent valuations*: an agent's utility is directly influenced by private information available only to other agents (the processing time of her predecessors in the queue).

Denote by $\Pi(\Sigma_n)$ the set of all probability distributions on Σ_n. In a *direct revelation mechanism*[7] (ϕ, p) each agent i reports a processing time $\widehat{t}_i \in [\underline{t}, \overline{t}]$. Given reports $\widehat{t} = (\widehat{t}_1, \ldots, \widehat{t}_n)$, the designer implements a schedule $\sigma \in \Sigma_n$ according to the scheduling rule $\phi : [\underline{t}, \overline{t}]^n \mapsto \Pi(\Sigma_n)$, and receives payments $p = (p_1, \ldots, p_n)$,where $p_i(\widehat{t}_1, \ldots, \widehat{t}_n)$ denotes the payment of agent i .

A mechanism (ϕ, p) *truthfully implements* the rule ϕ if truth-telling is a Bayes-Nash equilibrium in the game induced by (ϕ, p) and the agents' utility functions. Such a mechanism is called *incentive compatible* (IC).

The interim utility of agent i given her type t_i, her announcement $\widehat{t}_i$ and given truthtelling of the other agents is given by:

$$\begin{aligned} U_i\left(t_i, \widehat{t}_i\right) &:= E_{t_{-i}}\left[u_i\left(t_i, \widehat{t}_i, t_{-i}\right)\right] \\ &:= V - E_{t_{-i}}\left[\sum_{\sigma\in\Sigma_n} \phi\left(\widehat{t}_i, t_{-i}\right)(\sigma)\left[C\left(\sum_{k=1}^{\sigma(i)} t_{\sigma(k)}\right)\right] + p_i\left(\widehat{t}_i, t_{-i}\right)\right]. \end{aligned}$$

We use the following notation:

$$U_i(t_i) := U_i(t_i, t_i).$$

[6] We will relax this assumption in Section 5.

[7] As usual, a revelation principle applies here.

The overall performance of a mechanism (ϕ, p) is measured by the total expected cost due to waiting, i.e. by

$$P(\phi) = E_t \left[\sum_{i=1}^{n} \sum_{\sigma \in \Sigma_n} \phi(t_i, t_{-i})(\sigma) \left[C \left(\sum_{k=1}^{\sigma(i)} t_{\sigma(k)} \right) \right] \right].$$

We call an incentive compatible mechanism *optimal* if it minimizes $P(\phi)$ in the class of all incentive compatible mechanisms. Note that, for a given and fixed vector of processing times, total waiting cost is minimized by the well known *shortest processing time first (SEPT)* schedule where agents with a shorter processing time are served before those with longer processing time. A scheduling rule ϕ^* is called *ex-post efficient* if it yields the SEPT schedule for any realization of the random vector of processing times. Obviously, an incentive compatible mechanism that implements an ex-post efficient allocation rule is optimal.

3 Auction-based queue disciplines

In this section we analyze the equilibrium schedules arising in the following auction: agents simultaneously submit sealed bids; the highest bidder is served first, the second highest bidder is served second, and so forth; finally, each bidder has to pay his bid. This multi-object auction for n slots constitutes a simple and natural way of implementing a schedule based on the agents' information about processing times. Obviously, bidders do not make losses by participating in this auction since they can assure themselves a positive payoff by bidding zero and being queued last. The next result shows that equilibrium behavior crucially depends on the structure of the cost function C. We define:

$$\begin{aligned}
\overline{C}_1(x) &:= \underline{C}_1(x) = C(2x) - C(x), \\
\overline{C}_l(x) &:= E_{(t_2,\ldots,t_l)} \left[C\left(2x + \sum_{j=2}^{l} t_j\right) - C\left(x + \sum_{j=2}^{l} t_j\right) \mid t_j \le x,\ j = 2,\ldots,l \right],\ l \ge 2, \\
\underline{C}_l(x) &:= E_{(t_2,\ldots,t_l)} \left[C\left(2x + \sum_{j=2}^{l} t_j\right) - C\left(x + \sum_{j=2}^{l} t_j\right) \mid t_j \ge x,\ j = 2,\ldots,l \right],\ l \ge 2.
\end{aligned}$$

Theorem 1

1. *If C is strictly concave, the unique symmetric equilibrium of the slot auction is given by the strictly decreasing function*

$$b_{cave}(t) = (n-1) \sum_{l=1}^{n-1} \binom{n-2}{l-1} \int_{t}^{\bar{t}} (1 - F(x))^{n-l-1} F^{l-1}(x) \overline{C}_l(x) f(x)\, dx \tag{1}$$

   Hence, the auction implements the efficient SEPT discipline.

2. *If C is strictly convex, the unique symmetric equilibrium of the slot auction is given by the strictly increasing function*

$$b_{vex}(t) = (n-1) \sum_{l=1}^{n-1} \binom{n-2}{l-1} \int_{\underline{t}}^{t} F(x)^{n-l-1} (1 - F(x))^{l-1} \underline{C}_l(x) f(x)\, dx \tag{2}$$

Hence, the auction implements the anti-efficient LEPT discipline.

Proof. The proof is given for strictly concave C, the case of strictly convex C is similar and therefore omitted.

Assume that all agents other than agent 1 bid according to a strictly decreasing function b. The expected utility of agent 1 with processing time t_1 who bids as if she were of type $\widehat{t}_1$ is given by

$$\begin{aligned} U\left(t_1,\widehat{t}_1\right) &= V-\sum_{l=1}^{n}\binom{n-1}{l-1}\left(1-F\left(\widehat{t}_1\right)\right)^{n-l}F\left(\widehat{t}_1\right)^{l-1} \\ &\quad E_{(t_2,\ldots,t_l)}\left[C\left(\sum_{j=1}^{l}t_j\right)\mid t_j\leq\widehat{t}_1,\ j=2,\ldots,l\right]-b\left(\widehat{t}_1\right). \end{aligned} \tag{3}$$

Differentiating with respect to $\widehat{t}_1$ yields

$$\begin{aligned} &\frac{\partial U\left(t_1,\widehat{t}_1\right)}{\partial\widehat{t}_1} = -\left(n-1\right)\left(1-F\left(\widehat{t}_1\right)\right)^{n-2}f\left(\widehat{t}_1\right)C\left(t_1\right)- \\ &\sum_{l=2}^{n-1}\binom{n-1}{l-1}\Bigg[\left(l-1\right)\left(1-F\left(\widehat{t}_1\right)\right)^{n-l}F\left(\widehat{t}_1\right)^{l-2}f\left(\widehat{t}_1\right) \\ &E_{(t_2,\ldots,t_l)}\left[C\left(\widehat{t}_1+\sum_{j=1}^{l-1}t_j\right)\mid t_j\leq\widehat{t}_1,\ j=2,\ldots,l-1\right] \\ &-\left(n-l\right)\left(1-F\left(\widehat{t}_1\right)\right)^{n-l-1}F\left(\widehat{t}_1\right)^{l-1}f\left(\widehat{t}_1\right)E_{(t_2,\ldots,t_l)}\left[C\left(\widehat{t}_1+\sum_{j=1}^{l}t_j\right)\mid t_j\leq\widehat{t}_1,\ j=2,\ldots,l\right]\Bigg] \\ &-\left(n-1\right)F\left(\widehat{t}_1\right)^{n-2}f\left(\widehat{t}_1\right)E_{(t_2,\ldots,t_{n-1})}\left[C\left(\widehat{t}_1+\sum_{j=1}^{n-1}t_j\right)\mid t_j\leq\widehat{t}_1,\ j=2,\ldots,n-1\right]-\frac{db\left(\widehat{t}_1\right)}{d\widehat{t}_1} \\ = &-\left(n-1\right)\sum_{l=1}^{n-1}\binom{n-2}{l-1}\left(1-F\left(\widehat{t}_1\right)\right)^{n-l-1}F\left(\widehat{t}_1\right)^{l-1}f\left(\widehat{t}_1\right) \\ &E_{(t_2,\ldots,t_l)}\left[C\left(\widehat{t}_1+\sum_{j=1}^{l}t_j\right)-C\left(\sum_{j=1}^{l}t_j\right)\mid t_j\leq\widehat{t}_1,\ j=2,\ldots,l\right]-\frac{db\left(\widehat{t}_1\right)}{d\widehat{t}_1}. \end{aligned}$$

The first order condition $\left.\frac{\partial U\left(t_1,\widehat{t}_1\right)}{\partial\widehat{t}_1}\right|_{\widehat{t}_1=t_1}=0$ is obviously fulfilled for b_{cave} as given by (1). The sufficient condition $\frac{\partial^2U\left(t_1,\widehat{t}_1\right)}{\partial\widehat{t}_1\partial t_1}>0$ for all $t_1,\widehat{t}_1$ is satisfied, since C is strictly concave (which implies that $C'=\frac{d}{dt}C\left(t\right)$ is strictly decreasing) and since

$$\begin{aligned} \frac{\partial^2U\left(t_1,\widehat{t}_1\right)}{\partial\widehat{t}_1\partial t_1} &= -\left(n-1\right)\sum_{l=1}^{n-1}\binom{n-2}{l-1}\left(1-F\left(\widehat{t}_1\right)\right)^{n-l-1}F\left(\widehat{t}_1\right)^{l-1}f\left(\widehat{t}_1\right) \\ &\quad E_{(t_2,\ldots,t_l)}\left[C'\left(\widehat{t}_1+\sum_{j=1}^{l}t_j\right)-C'\left(\sum_{j=1}^{l}t_j\right)\mid t_j\leq\widehat{t}_1,\ j=2,\ldots,l\right]. \end{aligned}$$

Q.E.D. ■

The above result shows that the efficient SEPT schedule is implemented by the auction if the cost function C is concave. In the case of a convex cost function the equilibrium bidding function is strictly increasing in processing time. Hence, agents are queued in the reverse order meaning that those with longer processing time are served first (this is the well known LEPT policy). In this case the auction yields the worst possible schedule since it maximizes $P(\phi)$.

The reason for these contrasting results is as follows: Agent's i cost from being delayed for a period of time t, $C(t+t_i) - C(t_i)$, is increasing in t_i if C is convex and decreasing if C is concave. Hence, if C is convex, it is more costly for an agent with a longer processing time to queue for some time t (before her own job is processed) than for an agent with a shorter processing time. The need to avoid the higher cost is expressed by a higher bid in the auction, yielding the increasing bid function for the case of convex cost functions. The opposite occurs for a concave cost function. Note that a linear cost function (as used in much of the queuing literature) yields a "degenerate" model where the difference $C(t+t_i) - C(t_i) = t$ does not depend at all on t_i. In this case the auction admits both an increasing and a decreasing symmetric equilibrium.

The "anti-efficient" auction outcome for convex cost functions suggests to look for mechanisms that perform better in this case. We study such mechanisms in the next section.

4 Slot auctions with bid caps

Assume that bidders are not allowed to submit bids that are above a predetermined common-knowledge maximum bid $\widetilde{b}$. If $m > 1$ agents submit bids equal to $\widetilde{b}$, then each of these bidders is given any of the first m slots in the queue with probability $\frac{1}{m}$.

Theorem 2 *1) Let C be concave and assume that the bid cap $\widetilde{b}$ satisfies*

$$E_{(t_2,\dots,t_n)}\left[C\left(\bar{t}+\sum_{j=2}^{n} t_j\right) - \frac{1}{n}\sum_{k=0}^{n-1} C\left(\bar{t}+\sum_{j=2}^{n-k} t_j\right)\right] < \widetilde{b} < b\left(\underline{t}\right). \quad (4)$$

The schedule auction with bid cap $\widetilde{b}$ has a unique symmetric equilibrium characterized by a type $\widetilde{t}$ $= \widetilde{t}(\widetilde{b})$, strictly decreasing in $\widetilde{b}$, such that

$$\widetilde{b}_{cave}\left(\widetilde{b},\ t\right) = \begin{cases} b_{cave}(t) & \text{if} \quad t > \widetilde{t} \\ \widetilde{b} & \text{if} \quad t \leq \widetilde{t}., \end{cases} \quad (5)$$

2) Let C be convex and assume that the bid cap $\widetilde{b}$ satisfies

$$E_{(t_2,\dots,t_n)}\left[C\left(\underline{t}+\sum_{j=2}^{n} t_j\right) - \frac{1}{n}\sum_{k=0}^{n-1} C\left(\underline{t}+\sum_{j=2}^{n-k} t_j\right)\right] < \widetilde{b} < b\left(\bar{t}\right). \quad (6)$$

The schedule auction with bid cap $\widetilde{b}$ has a unique symmetric equilibrium characterized by a type $\widetilde{t}$ $= \widetilde{t}(\widetilde{b})$, strictly increasing in $\widetilde{b}$, such that

$$\widetilde{b}_{vex}\left(\widetilde{b},\ t\right) = \begin{cases} b_{vex}(t) & \text{if} \quad t \leq \widetilde{t} \\ \widetilde{b} & \text{if} \quad t > \widetilde{t}, \end{cases} \quad (7)$$

Proof. See Appendix. ■

By choosing an appropriate bid cap the auctioneer is able to determine the types of agents that pool in equilibrium. Whereas the auctioneer can only decrease efficiency by introducing a bid cap $\widetilde{b} < b(\underline{t})$ in the case of a concave cost function (note that performance is monotonically decreasing in $\widetilde{b}$), an effective bid cap $\widetilde{b} < b(\overline{t})$ necessarily increases performance in the case of a convex cost function since, in that case, the auction performs worst among all possible scheduling mechanisms. Moreover, overall performance is increasing in $\widetilde{b}$ if C is convex. In particular, by setting a bid cap equal to the lower bound in (6) all agents bid $\widetilde{b}$ and hence are given a certain position in the queue with probability $\frac{1}{n}$. In this case all agents pay a fixed fee, and the schedule is determined by an equal chance lottery among all agents.

4.1 Revenue considerations

We now turn to an analysis of bid caps on the designer's revenue. We first show that, if the cost function is concave, the designer cannot profit by setting an effective bid cap. Thus, the ensuing efficiency loss in this case has necessarily a negative effect on her revenue. The result is not trivial since some bidders bid in the constrained auction more than in the unconstrained one. Hence, an auctioneer interested in revenue needs to balance this positive effect of bid caps with the negative effect caused by the fact that some other bidders cannot bid more than $\widetilde{b}$.

Theorem 3 *Assume that C is concave. Then the designer's revenue in an auction with an effective bid cap is lower than her revenue in the auction without bid cap.*

Proof. see Appendix ■

In the case of a convex cost function it is interesting to note that the introduction of a bid cap causes low-type bidders to bid $\widetilde{b}$ even if they submitted lower bids in the unrestricted auction (because this significantly increases their probability of getting a better slot). The next result shows that if the cost function is not "too convex", then it is optimal for a revenue-maximizing designer to set the lowest relevant effective a bid cap. Thus, increasing the efficiency of the auction is also beneficial in terms of revenue, and the optimum is achieved with a lottery that does not condition on the agents' private information

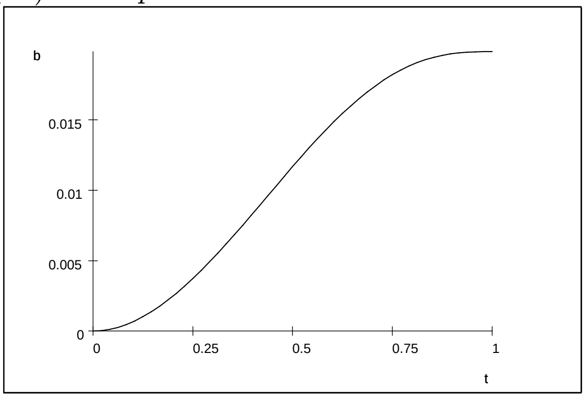
Theorem 4 *Assume that C is convex. For any distribution function F and for any number of jobs n, there exists a constant $K_n > 0$ such that if $C'' < K_n$ the revenue maximizing bid cap is given by*

$$\widetilde{b}_M = E_{(t_2,\ldots,t_n)}\left[C\left(\underline{t} + \sum_{j=2}^{n} t_j\right) - \frac{1}{n}\sum_{k=0}^{n-1} C\left(\underline{t} + \sum_{j=2}^{n-k} t_j\right)\right].$$

Proof. see Appendix ■

In contrast to the finding above, there exist environments with strongly convex cost functions where the unrestricted auction yields more revenue than any auction with an effective bid cap. The intuition is that more convexity leads to more aggressive bidding by agents with high processing costs. The revenue loss resulting from limiting these bidders cannot be anymore compensated by the low-type bidders' willingness to increase their bids up to the bid cap.

Example 5 *Assume $n = 2$, $C(t) = e^t$ and $t_i \sim U[0,1]$. The following graph shows the dependence of a bidder's expected bid on the marginal type $\widetilde{t}$. Since the marginal bid increases in $\widetilde{t}$ (which, in turn, increases in the bid cap $\widetilde{b}$) it is optimal not to restrict the auction with an effective bid cap.*



Finally, note that in many standard auctions the auctioneer can increase her revenue by rationing the auctioned good through the imposition of a reserve price. Apart from the fact that this has a negative effect on overall efficiency and hence is not desirable if efficiency is the auctioneer's concern, it might be difficult to commit to an exclusion of bidders ex-post. In contrast, we analyzed here the effect of an instrument (bid caps) that ensures that all agents are served.

5 Optimal Mechanisms

In the previous section we showed that, for the case of convex cost functions, it may be advantageous (both for revenue and for efficiency purposes) not to use any private information about processing times, i.e., both the revenue maximizing and the performance maximizing auction may sometime coincide with a lottery where slots in the queue are allocated at random. But, it is conceivable that scheduling mechanisms other than the above described auctions do perform better. In this section we show that mechanism that do not depend on the available private information do minimize expected waiting costs in an important class of mechanisms.

Definition 6 *A direct revelation mechanism (ϕ, p) is called ex-post incentive compatible if truthful announcement of processing times constitutes an ex-post equilibrium, i.e. for any realization of types, it is optimal for any agent i to announce his true type given that all other agents also announce their processing times truthfully.*

Ex-post equilibria do not depend on the underlying distribution functions (in particular, they do not depend on the assumption of independency of agents' processing times).

A mechanism (ϕ, p) that minimizes $P(\phi)$ in the class of all ex-post incentive compatible mechanisms is called an *optimal ex-post incentive compatible mechanism.*

The following result shows that, for the case of two agents, the best allocation that can be implemented in ex-post equilibrium is given by a lottery giving the first position in the queue to both agents with probability $\frac{1}{2}$. In fact, any mechanism that does not depend on reports is such an optimal mechanism. More generally, if the types t_1, t_2 are distributed on $[\underline{t}_1, \overline{t}_1]$ and $[\underline{t}_2, \overline{t}_2]$ according to different distribution functions F_1 and F_2 (with densities f_1 and f_2) the best allocation in ex-post equilibrium is given by the ex-ante efficient allocation, i.e., by scheduling first the agent with the lowest expected processing cost. In particular, an agent with a stochastically smaller distribution of processing time should be served first. In other words, it is optimal to choose the SEPT schedule based on the ex-ante expected processing time.

Theorem 7 *Let $n = 2$ and assume that C is convex. The optimal ex-post incentive compatible mechanism is to always schedule first the agent with the smaller ex-ante expected processing costs $\int_{\underline{t}_i}^{\overline{t}_i} C(t_i) f_i(t_i)\, dt_i$. If the ex-ante expected processing costs are the same for both agents, then any random schedule is optimal.*

Proof. Appendix ■

While we believe that an analogous result holds any number of agents, a formal proof seems difficult because general ex-post incentive compatible mechanisms may have the complex feature that information available to one agent influences also the service order of other agents, i.e a change in $i's$ report leads to a different permutation on agents other than i. This problem does not appear if there are only two agents.

5.1 Delay Costs and Stochastic Dominance

In many situations the designer is not only interested in minimizing expected total waiting costs for the agents, but additionally prefers that this cost stays below a given threshold as often as possible. Such a constraint arises for example if there is an overall budget constraint with costly overdraws.

The following stylized example (which can be approximated in our model) shows that these two goals might not be simultaneously attainable.

Example 8 *Assume that there are two agents and that the (convex) cost function is given by $C(t) = t^2$. Consider the following distribution of processing times:*

$$\mathbf{t}_1 = \left\{ \begin{array}{ll} 1 & \textit{with probability } \frac{1}{5} \\ 3 & \textit{with probability } \frac{4}{5} \end{array} \right| , \quad \mathbf{t}_2 = \left\{ \begin{array}{ll} 2 & \textit{with probability } \frac{1}{5} \\ 5 & \textit{with probability } \frac{4}{5} \end{array} \right| .$$

Note that $\mathbf{t}_2$ stochastically dominates $\mathbf{t}_1$ hence, according to Theorem 7, it is optimal to serve agent 1 first. On the other hand, the probability that overall processing cost is below 30 is higher *if agent 2 is queued first, i.e.*

$$\Pr\{C(t_1) + C(t_1 + t_2) \leq 30\} = \frac{4}{25} < \frac{1}{5} = \Pr\{C(t_2) + C(t_1 + t_2) \leq 30\}.$$

We now proceed to offer a sufficient condition guaranteeing that the distribution of overall processing cost achieved by the SEPT discipline where agents with stochastically lower processing times are served first is stochastically dominated by the distribution of processing time in any other discipline. We use the following definition:

Definition 9 *The random variable* $\mathbf{t}_2$ *is larger then* $\mathbf{t}_1$ *in the sense of likelihood ratio,* $\mathbf{t}_2 \geq_{LR} \mathbf{t}_1$*, if* $\frac{f_2(t)}{f_1(t)}$ *is (weakly) increasing in* t*.*

Note that the likelihood ratio order implies standard stochastic dominance. In particular, $\mathbf{t}_2 \geq_{LR} \mathbf{t}_1$ implies that $E(C(\mathbf{t}_2)) \geq E(C(\mathbf{t}_1))$.

The following is a well-known property of the likelihood ratio ordering:

Proposition 10 *Assume that* $\mathbf{t}_2 \geq_{LR} \mathbf{t}_1$*, and let* $h(x, y)$ *be a real-valued function satisfying* $h(t_2, t_1) \geq h(t_1, t_2)$ *for all* $t_2 \geq t_1$*. Then the random variable* $\mathbf{h}(t_1, t_2)$ *is stochastically dominated by* $\mathbf{h}(t_2, t_1)$[8].

Proof. See Ross (1983), Proposition 8.4.2 on page 268.[9] ■

It is immediate that the above condition is satisfied for the function h that measures total waiting costs, i.e., $h(t_1, t_2) = C(t_1) + C(t_1 + t_2)$. Hence, if $\mathbf{t}_2 \geq_{LR} \mathbf{t}_1$, it follows that the probability of overall costs being below an arbitrary threshold is higher if agent 1 is served before agent 2 than the other way round. In other words, if in a deterministic scheduling problem with known processing times it is optimal to interchange two jobs, then it is also stochastically optimal to interchange two jobs with random processing times that are likelihood ratio ordered. In conjunction with Theorem 7, these observations yield the following result:

Theorem 11 *Assume that* $\mathbf{t}_2 \geq_{LR} \mathbf{t}_1$ *and that* C *is convex. The random total cost in the optimal ex-post incentive compatible mechanism is then stochastically dominated by the random cost in any mechanism that does not depend on announcements.*

6 Conclusion

We have combined a simple queuing problem with an incentive problem arising if impatient agents are privately informed about the processing time needed to complete their respective jobs. We have shown that auction-based queue disciplines are efficient if the delay cost function is concave, and anti-efficient if it is convex. For the later case, the auction's performance can be enhanced by imposing bid caps. Finally, for the case of convex cost functions, we have shown that the best performance is attained by mechanisms that do not attempt to condition on the private information. These mechanisms take into account only ex-ante available information.

[8] This result easily generalizes to more than two players.

[9] For an extensive analysis of the uses of stochastic orders in queueing see Chang and Yao (1993) and Shanthikumar and Yao (1991).

While the model analyzed here is very simple, the vast queuing literature has considered much more complicated models with random arrivals, multiple servers, preemptive service, multi-stage service, etc...In principle, the performance of various pricing mechanisms similar to those studied here can be studied in such frameworks. We think that the combination of queuing and incentive models constitutes a fruitful avenue and that such studies will have many significant real-life applications.

7 Appendix

We use the following abbreviation: $dF(t_{l+1},\ldots,t_n) := f(t_{l+1})\ldots f(t_n)\,dt_{l+1},\ldots dt_n$.

Proof of Theorem 2: The proof is performed for the case of a convex cost function. The case of a concave cost functions follows along the same line.

Given that other bidders bid according to a strictly increasing bidding function b if their type is smaller than $\widetilde{t}$ and bid $\widetilde{b}$ otherwise, the interim expected utility of a bidder with type t_1 who bids according to b as if he were of type $\widehat{t}_1 < \widetilde{t}$ is identical to the utility in the case without bid caps, i.e. it is given by

$$U\left(t_1, b\left(\widehat{t}\right)\right) = V - \sum_{l=1}^{n} \binom{n-1}{l-1} F^{l-1}\left(\widehat{t}\right) \underbrace{\int_{\widehat{t}}^{\bar{t}} \cdots \int_{\widehat{t}}^{\bar{t}}}_{n-l} C\left(t_1 + \sum_{j=l+1}^{n} t_j\right) dF(t_{l+1},\ldots,t_n) - b\left(\widehat{t}\right).$$

As in the proof of Theorem 1 , it can be shown that the necessary and sufficient optimality conditions (for $\widehat{t}_1 < \widetilde{t}$) are fulfilled by $b(t_1) = b_{vex}(t_1)$.

If a bidder with type t_1 bids $\widetilde{b}$, his interim expected utility is given by:

$$U\left(t_1, \widetilde{b}\right) = V - \sum_{l=1}^{n} \binom{n-1}{l-1} F^{l-1}\left(\widetilde{t}\right) \underbrace{\int_{\widetilde{t}}^{\bar{t}} \cdots \int_{\widetilde{t}}^{\bar{t}}}_{n-l} \frac{1}{n-l+1} \sum_{k=0}^{n-l} C\left(t_1 + \sum_{j=l+1}^{n-k} t_j\right) dF(t_{l+1},\ldots,t_n) - \widetilde{b}.$$

Given $\widetilde{b}$, the marginal type $\widetilde{t}$ is determined by the condition $U\left(\widetilde{t}, b\left(\widetilde{t}\right)\right) = U\left(\widetilde{t}, \widetilde{b}\right)$, i.e. by

$$\begin{aligned} \widetilde{b} &= b\left(\widetilde{t}\right) + \sum_{l=1}^{n-1} \binom{n-1}{l-1} F^{l-1}\left(\widetilde{t}\right) \\ & \underbrace{\int_{\widetilde{t}}^{\bar{t}} \cdots \int_{\widetilde{t}}^{\bar{t}}}_{n-l} \left[C\left(\widetilde{t} + \sum_{j=l+1}^{n} t_j\right) - \frac{1}{n-l+1} \sum_{k=0}^{n-l} C\left(\widetilde{t} + \sum_{j=l+1}^{n-k} t_j\right) \right] dF(t_{l+1},\ldots,t_n). \end{aligned} \tag{8}$$

For $\widetilde{b} = b\left(\bar{t}\right)$ we get $\widetilde{t} = \bar{t}$ whereas for

$$\widetilde{b}_M = \underbrace{\int_{\underline{t}}^{\bar{t}} \cdots \int_{\underline{t}}^{\bar{t}}}_{n-1} \left[C\left(\underline{t} + \sum_{j=2}^{n} t_j\right) - \frac{1}{n} \sum_{k=0}^{n-1} C\left(\underline{t} + \sum_{j=2}^{n-k} t_j\right) \right] dF(t_2,\ldots,t_n)$$

we obtain $\widetilde{t} = \underline{t}$.

We show next that the derivative with respect to $\widetilde{t}$ of the r.h.s. of (8) is strictly positive. This implies that $\widetilde{t}$ is uniquely defined ,and that a higher bid cap $\widetilde{b}$ yields a higher marginal type $\widetilde{t}$. Using

$$\begin{aligned}\frac{d}{dt}b\left(t\right) &= (n-1)\sum_{l=1}^{n-1}\binom{n-2}{l-1}F^{l-1}\left(t\right)f\left(t\right) \\ &\underbrace{\int_{t}^{\bar{t}}\dots\int_{t}^{\bar{t}}}_{n-l-1}\left[C\left(2t+\sum_{j=l+2}^{n}t_j\right)-C\left(t+\sum_{j=l+2}^{n}t_j\right)\right]dF(t_{l+2},\dots,t_n).\end{aligned}$$

we obtain

$$\begin{aligned}\frac{d}{d\widetilde{t}}\widetilde{b} &= \frac{d}{d\widetilde{t}}b\left(\widetilde{t}\right)+\sum_{l=2}^{n-1}\binom{n-1}{l-1}\left(l-1\right)F^{l-2}\left(\widetilde{t}\right)f\left(\widetilde{t}\right) \\ &\underbrace{\int_{\widetilde{t}}^{\bar{t}}\dots\int_{\widetilde{t}}^{\bar{t}}}_{n-l}\left[C\left(\widetilde{t}+\sum_{j=l+1}^{n}t_j\right)-\frac{1}{n-l+1}\sum_{k=0}^{n-l}C\left(\widetilde{t}+\sum_{j=l+1}^{n-k}t_j\right)\right]dF(t_{l+1},\dots,t_n) \\ &-\sum_{l=1}^{n-1}\binom{n-1}{l-1}F^{l-1}\left(\widetilde{t}\right)f\left(\widetilde{t}\right)\underbrace{\int_{\widetilde{t}}^{\bar{t}}\dots\int_{\widetilde{t}}^{\bar{t}}}_{n-l-1}\left[\left(n-l\right)C\left(2\widetilde{t}+\sum_{j=l+2}^{n}t_j\right)\right. \\ &\left.-\sum_{k=0}^{n-l-1}\left(\left(\frac{n-l-k}{n-l+1}C\left(2\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)+\frac{k+1}{n-l+1}C\left(\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)\right)\right)\right]dF(t_{l+2},\dots,t_n) \\ &+\sum_{l=1}^{n-1}\binom{n-1}{l-1}F^{l-1}\left(\widetilde{t}\right) \\ &\underbrace{\int_{\widetilde{t}}^{\bar{t}}\dots\int_{\widetilde{t}}^{\bar{t}}}_{n-l}\left[\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n}t_j\right)-\frac{1}{n-l+1}\sum_{k=0}^{n-l}\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n-k}t_j\right)\right]dF(t_{l+1},\dots,t_n) \\ &= \frac{d}{d\widetilde{t}}b\left(\widetilde{t}\right)+(n-1)\sum_{l=1}^{n-2}\binom{n-2}{l-1}F^{l-1}\left(\widetilde{t}\right)f\left(\widetilde{t}\right) \\ &\underbrace{\int_{\widetilde{t}}^{\bar{t}}\dots\int_{\widetilde{t}}^{\bar{t}}}_{n-l-1}\left[C\left(\widetilde{t}+\sum_{j=l+2}^{n}t_j\right)-\frac{1}{n-l}\sum_{k=0}^{n-l-1}C\left(\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)-C\left(2\widetilde{t}+\sum_{j=l+2}^{n}t_j\right)\right. \\ &\left.+\frac{1}{n-l}\sum_{k=0}^{n-l-1}\left(\left(\frac{n-l-k}{n-l+1}C\left(2\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)+\frac{k+1}{n-l+1}C\left(\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)\right)\right)\right]dF(t_{l+2},\dots,t_n) \\ &-\left(n-1\right)F^{n-2}\left(\widetilde{t}\right)f\left(\widetilde{t}\right)\frac{1}{2}\left[C\left(2\widetilde{t}\right)-C\left(\widetilde{t}\right)\right]+\sum_{l=1}^{n-1}\binom{n-1}{l-1}F^{l-1}\left(\widetilde{t}\right) \\ &\underbrace{\int_{\widetilde{t}}^{\bar{t}}\dots\int_{\widetilde{t}}^{\bar{t}}}_{n-l}\left[\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n}t_j\right)-\frac{1}{n-l+1}\sum_{k=0}^{n-l}\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n-k}t_j\right)\right]dF(t_{l+1},\dots,t_n)\end{aligned}$$

$$
\begin{aligned}
= \quad & (n-1)\sum_{l=1}^{n-2}\binom{n-2}{l-1}F^{l-1}\left(\widetilde{t}\right)f\left(\widetilde{t}\right) \\
& \underbrace{\int_{\widetilde{t}}^{\bar{t}}\cdots\int_{\widetilde{t}}^{\bar{t}}}_{n-l-1}\frac{1}{n-l}\left[\sum_{k=0}^{n-l-1}\left(\left(\frac{n-l-k}{n-l+1}C\left(2\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)+\frac{k+1}{n-l+1}C\left(\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)\right)\right)\right. \\
& \left.-\sum_{k=0}^{n-l-1}C\left(\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)\right]dF(t_{l+2},\ldots,t_n) \\
& +(n-1)F^{n-2}\left(\widetilde{t}\right)f\left(\widetilde{t}\right)\frac{1}{2}\left(C\left(2\widetilde{t}\right)-C\left(\widetilde{t}\right)\right)+\sum_{l=1}^{n-1}\binom{n-1}{l-1}F^{l-1}\left(\widetilde{t}\right) \\
& \underbrace{\int_{\widetilde{t}}^{\bar{t}}\cdots\int_{\widetilde{t}}^{\bar{t}}}_{n-l}\left[\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n}t_j\right)-\frac{1}{n-l+1}\sum_{k=0}^{n-l}\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n-k}t_j\right)\right]dF(t_{l+1},\ldots,t_n).
\end{aligned}
$$

The above last expression is (strictly) positive since C is strictly increasing and convex. For all $t_2,\ldots,t_n$ and $l=1,\ldots,n-2$ we thus have

$$
\begin{aligned}
\sum_{k=0}^{n-l-1}\left(\frac{n-l-k}{n-l+1}C\left(2\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)+\frac{k+1}{n-l+1}C\left(\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)-C\left(\widetilde{t}+\sum_{j=l+2}^{n-k}t_j\right)\right) &> 0, \\
\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n}t_j\right)-\frac{1}{n-l+1}\sum_{k=0}^{n-l}\frac{\partial}{\partial\widetilde{t}}C\left(\widetilde{t}+\sum_{j=l+1}^{n-k}t_j\right) &> 0.
\end{aligned}
$$

Obviously, it is not optimal to deviate to a bid $b<\widetilde{b}$ if $t_1>\widetilde{t}$ or to a bid $b>b\left(\widetilde{t}\right)$ if $t_1<\widetilde{t}$. Q.E.D.

Proof of Theorem 3: In order to have a transparent proof, we restrict attention here to the case $n=2$. Given a bid cap $\widetilde{b}$, and given the uniquely defined corresponding marginal type $\widetilde{t}$, the auctioneer's expected revenue per agent is:

$$
R(\widetilde{b}):=\widetilde{b}F\left(\widetilde{t}\right)+\int_{\widetilde{t}}^{\bar{t}}b_{cave}\left(x\right)f\left(x\right)dx.
$$

It suffices to show that $\frac{d}{d\widetilde{t}}R(\widetilde{b})=\frac{d}{d\widetilde{t}}\widetilde{b}F\left(\widetilde{t}\right)+\left(\widetilde{b}-b_{cave}\left(\widetilde{t}\right)\right)f\left(\widetilde{t}\right)<0$ for $\widetilde{t}<\bar{t}$. We have that :

$$
\begin{aligned}
U\left(\widetilde{t},b_{cave}\left(\widetilde{t}\right)\right) &= V-\int_{\underline{t}}^{\widetilde{t}}C\left(\widetilde{t}+x\right)f\left(x\right)dx-\left(1-F\left(\widetilde{t}\right)\right)C\left(\widetilde{t}\right)-b_{cave}\left(\widetilde{t}\right), \\
U\left(\widetilde{t},\widetilde{b}\right) &= V-\frac{1}{2}\int_{\underline{t}}^{\widetilde{t}}\left[C\left(\widetilde{t}+x\right)-C\left(\widetilde{t}\right)\right]f\left(x\right)dx-\left(1-F\left(\widetilde{t}\right)\right)C\left(\widetilde{t}\right)-\widetilde{b}
\end{aligned}
$$

and hence we obtain

$$
\widetilde{b}-b_{cave}\left(\widetilde{t}\right)=\frac{1}{2}\int_{\underline{t}}^{\widetilde{t}}\left[C\left(\widetilde{t}+x\right)-C\left(\widetilde{t}\right)\right]f\left(x\right)dx.
$$

Using

$$
\frac{d}{d\widetilde{t}}\widetilde{b}=f\left(\widetilde{t}\right)\left[C\left(\widetilde{t}\right)-\frac{1}{2}\left(C\left(\widetilde{t}\right)+C\left(2\widetilde{t}\right)\right)\right]+\frac{1}{2}\int_{\underline{t}}^{\widetilde{t}}\left[\frac{d}{d\widetilde{t}}C\left(\widetilde{t}+x\right)-\frac{d}{d\widetilde{t}}C\left(\widetilde{t}\right)\right]f\left(x\right)dx
$$

we obtain

$$\frac{d}{d\widetilde{t}}R(\widetilde{b}) = f\left(\widetilde{t}\right)\frac{1}{2}\int_{\underline{t}}^{\widetilde{t}}\left[C\left(\widetilde{t}+x\right) - C\left(2\widetilde{t}\right)\right]f\left(x\right)dx + \frac{1}{2}F\left(\widetilde{t}\right)\int_{\underline{t}}^{\widetilde{t}}\left[\frac{d}{d\widetilde{t}}C\left(\widetilde{t}+x\right) - \frac{d}{d\widetilde{t}}C\left(\widetilde{t}\right)\right]f\left(x\right)dx.$$

The statement follows since C is strictly increasing and concave. Q.E.D.

Proof of Theorem 4: It suffices to show that a bid cap $\widetilde{b}_M$ leads to a strictly higher revenue than any other bid cap for the linear cost function $C\left(x\right) = x$. The result follows then by a continuity argument.

Given a bid cap $\widetilde{b}$, and given the uniquely defined corresponding marginal type $\widetilde{t}$, the auctioneer's expected revenue per agent is:

$$R(\widetilde{b}) := \widetilde{b}\left(1 - F\left(\widetilde{t}\right)\right) + \int_{\underline{t}}^{\widetilde{t}} b_{vex}\left(x\right)f\left(x\right)dx.$$

We show below that $\frac{d}{d\widetilde{t}}R(\widetilde{b}) = \frac{d}{d\widetilde{t}}\widetilde{b}\left(1 - F\left(\widetilde{t}\right)\right) + \left(b_{vex}\left(\widetilde{t}\right) - \widetilde{b}\right)f\left(\widetilde{t}\right) < 0$ for $\widetilde{t} < \overline{t}$. Thus, the optimal marginal type is $\underline{t}$, and, accordingly, the optimal bid cap is $\widetilde{b}_M$.

For $C\left(x\right) = x$ we have:

$$\begin{aligned}
\frac{d}{d\widetilde{t}}R(\widetilde{b}) &= (n-1)\sum_{l=1}^{n-2}\binom{n-2}{l-1}F^{l-1}\left(\widetilde{t}\right)f\left(\widetilde{t}\right)\left(1-F\left(\widetilde{t}\right)\right) \\
&\quad \underbrace{\underbrace{\int_{\widetilde{t}}^{\overline{t}}\cdots\int_{\widetilde{t}}^{\overline{t}}}_{n-l-1}\frac{1}{n-l}\left[\sum_{k=0}^{n-l-1}\frac{n-l-k}{n-l+1}\widetilde{t}\right]dF(t_{l+2},\ldots,t_n)}_{\left(1-F\left(\widetilde{t}\right)\right)^{n-l-1}\frac{1}{2}\widetilde{t}} \\
&\quad + (n-1)F^{n-2}\left(\widetilde{t}\right)f\left(\widetilde{t}\right)\frac{1}{2}\widetilde{t}\left(1-F\left(\widetilde{t}\right)\right) \\
&\quad - f\left(\widetilde{t}\right)\sum_{l=1}^{n-1}\binom{n-1}{l-1}F^{l-1}\left(\widetilde{t}\right)\underbrace{\int_{\widetilde{t}}^{\overline{t}}\cdots\int_{\widetilde{t}}^{\overline{t}}}_{n-l}\left[\sum_{j=l+1}^{n}t_j - \frac{1}{n-l+1}\sum_{k=0}^{n-l}\sum_{j=l+1}^{n-k}t_j\right]dF(t_{l+1},\ldots,t_n)
\end{aligned}$$

$$
\begin{aligned}
&= (n-1)\sum_{l=1}^{n-1}\binom{n-2}{l-1}F^{l-1}\left(\widetilde{t}\right)f\left(\widetilde{t}\right)\left(1-F\left(\widetilde{t}\right)\right)^{n-l}\frac{1}{2}\widetilde{t} \\
&\quad -f\left(\widetilde{t}\right)\sum_{l=1}^{n-1}\binom{n-1}{l-1}F^{l-1}\left(\widetilde{t}\right)\underbrace{\int_{\widetilde{t}}^{\bar{t}}\cdots\int_{\widetilde{t}}^{\bar{t}}}_{n-l}\left[\sum_{j=l+1}^{n}\frac{j-l}{n-l+1}t_j\right]dF(t_{l+1},\ldots,t_n) \\
&= \sum_{l=1}^{n-1}\frac{(n-1)!}{(n-l)!\,(n-l-1)!}F^{l-1}\left(\widetilde{t}\right)f\left(\widetilde{t}\right) \\
&\quad \left(\left(1-F\left(\widetilde{t}\right)\right)^{n-l}\frac{1}{2}\widetilde{t}-\frac{1}{n-l}\underbrace{\int_{\widetilde{t}}^{\bar{t}}\cdots\int_{\widetilde{t}}^{\bar{t}}}_{n-l}\left[\sum_{j=l+1}^{n}\frac{j-l}{n-l+1}t_j\right]dF(t_{l+1},\ldots,t_n)\right) \\
&< \sum_{l=1}^{n-1}\frac{(n-1)!}{(n-l)!\,(n-l-1)!}F^{l-1}\left(\widetilde{t}\right)f\left(\widetilde{t}\right)\left(\left(1-F\left(\widetilde{t}\right)\right)^{n-l}\frac{1}{2}\widetilde{t}-\underbrace{\int_{\widetilde{t}}^{\bar{t}}\cdots\int_{\widetilde{t}}^{\bar{t}}}_{n-l}\frac{1}{2}\,\widetilde{t}\,dF(t_{l+1},\ldots,t_n)\right) \\
&= 0.
\end{aligned}
$$

Q.E.D.

Proof of Theorem 7: Since the suggested mechanisms do not depend on announcements, truthtelling is an ex-post equilibrium. To simplify notation, let $k_i\left(\widehat{t}_1,\widehat{t}_2\right)$ denote the probability that agent i is served first, given announcements $\widehat{t}_1,\widehat{t}_2$. Obviously we must have $k_2 = 1-k_1$. For an allocation (k_1,k_2) to be implementable in ex-post equilibrium we must have for all $t_i \geq \widehat{t}_i$:

$$
u_i\left(t_i,\widehat{t}_i,t_{-i}\right)-u_i\left(\widehat{t}_i,\widehat{t}_i,t_{-i}\right)\leq u_i\left(t_i,t_i,t_{-i}\right)-u_i\left(\widehat{t}_i,t_i,t_{-i}\right) \qquad \text{for all } t_{-i}.
$$

For $i=1$ this yields for all t_2:

$$
\begin{aligned}
& \left[C\left(t_1\right)-C\left(\widehat{t}_1\right)\right]k_1\left(\widehat{t}_1,t_2\right)+\left[C\left(t_2+t_1\right)-C\left(t_2+\widehat{t}_1\right)\right]\left(1-k_1\left(\widehat{t}_1,t_2\right)\right) \\
\geq\; & \left[C\left(t_1\right)-C\left(\widehat{t}_1\right)\right]k_1\left(t_1,t_2\right)+\left[C\left(t_2+t_1\right)-C\left(t_2+\widehat{t}_1\right)\right]\left(1-k_1\left(t_1,t_2\right)\right)
\end{aligned}
$$

This is equivalent to:

$$
\begin{aligned}
& \left[C\left(t_2+t_1\right)-C\left(t_1\right)-C\left(t_2+\widehat{t}_1\right)+C\left(\widehat{t}_1\right)\right]k_1\left(\widehat{t}_1,t_2\right) \\
\leq\; & \left[C\left(t_2+t_1\right)-C\left(t_1\right)-C\left(t_2+\widehat{t}_1\right)+C\left(\widehat{t}_1\right)\right]k_1\left(t_1,t_2\right).
\end{aligned}
$$

Since C is convex, the above condition is equivalent to $k_1\left(\widehat{t}_1,t_2\right)\leq k_1\left(t_1,t_2\right)$ for all t_2. This last condition implies in turn that ex-post implementability requires that

$$
\int_{\underline{t}_2}^{\bar{t}_2} k_1\left(t_1,t_2\right)f_2\left(t_2\right)dt_2 \;\text{ is increasing in } t_1. \tag{9}
$$

Equivalently, by looking at $i=2$ we get the requirement that

$$
\int_{\underline{t}_1}^{\bar{t}_1} k_1\left(t_1,t_2\right)f_1\left(t_1\right)dt_1 \text{ is decreasing in } t_2. \tag{10}
$$

We are now looking for the solution of the following problem:

$$\min_{(k_1,k_2)} \int_{\underline{t}_1}^{\bar{t}_1} \int_{\underline{t}_2}^{\bar{t}_2} [C(t_1) k_1(t_1,t_2) + C(t_2) k_2(t_1,t_2)) + C(t_1+t_2)] f_1(t_1) f_2(t_2)\, dt_1 dt_2$$

subject to incentive compatibility constraints.

Since $k_2 = 1 - k_1$, and since the cost $C(t_1+t_2)$ is incurred for sure in any allocation, the above problem becomes:

$$\min_{k_1} \int_{\underline{t}_1}^{\bar{t}_1} \int_{\underline{t}_2}^{\bar{t}_2} (C(t_1) - C(t_2)) k_1(t_1,t_2) f_1(t_1) f_2(t_2)\, dt_1 dt_2 \quad (11)$$
$$s.t. \ (9), (10).$$

We have that

$$\begin{aligned} & \int_{\underline{t}_1}^{\bar{t}_1} \int_{\underline{t}_2}^{\bar{t}_2} (C(t_1) - C(t_2)) k_1(t_1,t_2) f_1(t_1) f_2(t_2)\, dt_1 dt_2 \\ = & \int_{\underline{t}_1}^{\bar{t}_1} C(t_1) \int_{\underline{t}_2}^{\bar{t}_2} k_1(t_1,t_2) f_2(t_2)\, dt_2 f_1(t_1)\, dt_1 \\ & - \int_{\underline{t}_2}^{\bar{t}_2} C(t_2) \int_{\underline{t}_1}^{\bar{t}_1} k_1(t_1,t_2) f_1(t_1)\, dt_1 f_2(t_2)\, dt_2. \end{aligned}$$

For the solution k_1 of (11) define:

$$\int_{\underline{t}_1}^{\bar{t}_1} \int_{\underline{t}_2}^{\bar{t}_2} k_1(t_1,t_2) f_1(t_1) f_2(t_2)\, dt_1 dt_2 := K \in [0,1].$$

Since C is increasing and because of (9) and (10) we obtain that

$$\int_{\underline{t}_1}^{\bar{t}_1} C(t_1) \int_{\underline{t}_2}^{\bar{t}_2} k_1(t_1,t_2) f_2(t_2)\, dt_2 f_1(t_1)\, dt_1$$

is minimized if $\int_{\underline{t}_2}^{\bar{t}_2} k_1(t_1,t_2) f_2(t_2)\, dt_2 = K$ and that

$$\int_{\underline{t}_2}^{\bar{t}_2} C(t_2) \int_{\underline{t}_1}^{\bar{t}_1} k_1(t_1,t_2) f_1(t_1)\, dt_1 f_2(t_2)\, dt_2$$

is maximized if $\int_{\underline{t}_1}^{\bar{t}_1} k_1(t_1,t_2) f_1(t_1)\, dt_1 = K$. Hence, for the solution of (11), we must have:

$$\begin{aligned} & \int_{\underline{t}_1}^{\bar{t}_1} \int_{\underline{t}_2}^{\bar{t}_2} (C(t_1) - C(t_2)) k_1(t_1,t_2) f_1(t_1) f_2(t_2)\, dt_1 dt_2 \\ = & K \left(\int_{\underline{t}_1}^{\bar{t}_1} C(t_1) f_1(t_1)\, dt_1 - \int_{\underline{t}_2}^{\bar{t}_2} C(t_2) f_2(t_2)\, dt_1 dt_2 \right). \end{aligned}$$

This last expression is minimized at:

$$\begin{array}{lll} K = 0 & \text{if} & \int_{\underline{t}_1}^{\bar{t}_1} C(t_1) f_1(t_1)\, dt_1 \geq \int_{\underline{t}_2}^{\bar{t}_2} C(t_2) f_2(t_2)\, dt_1 dt_2 \\ K = 1 & \text{if} & \int_{\underline{t}_1}^{\bar{t}_1} C(t_1) f_1(t_1)\, dt_1 \leq \int_{\underline{t}_2}^{\bar{t}_2} C(t_2) f_2(t_2)\, dt_1 dt_2 \\ \text{any } K \in [0,1] & \text{if} & \int_{\underline{t}_1}^{\bar{t}_1} C(t_1) f_1(t_1)\, dt_1 = \int_{\underline{t}_2}^{\bar{t}_2} C(t_2) f_2(t_2)\, dt_1 dt_2. \end{array}$$

The mechanism suggested in the statement of the theorem obviously fulfills these requirements. Q.E.D.

8 References

Chang, C. and Yao, D. (1993): "Rearrangement, majorization and stochastic scheduling", *Mathematics of Operations Research **18**, 658-684*

Dasgupta, P., and Maskin, E. (2000): "Efficient Auctions", *Quarterly Journal of Economics* **115**, 341-388.

Gavious, A., Moldovanu, B., and Sela, A. (2002) "Bid Costs and Endogenous Bid Caps," *Rand Journal of Economics **33**, 709-722.*

Gross, D. and Harris, C. (1998): *Fundamentals of Queueing Theory,* New York: John Wiley&Sons.

Hain, R. and Mitra, M. (2002): "Simple Sequencing Problems with Interdependent Costs" mimeo, University of Bonn.

Hassin, R. (1995): "Decentralized Regulation of a Queue", *Management Science* **41,** 163-173.

Hassin, R. and Haviv, M. (2002): *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems,* Boston: Kluwer Academic Publishers.

Holt, C. A. and Sherman, R. (1982): "Waiting-Line Auctions", *Journal of Political Economy* **90**, 280-294.

Jehiel, P., and Moldovanu, B. (2001) "Efficient Design with Interdependent Valuations", *Econometrica* **69**, 1237-1259.

Mitra, M. (2001): "Mechanism Design in Queueing Problems", *Economic Theory* **17**, 277-305.

Naor, P. (1969): "The Regulation of Queue Size by Levying Tolls", *Econometrica* **37**, 15-24.

Ross, S. (1983): *Stochastic Processes,* New York: John Wiley&Sons

Shanthikumar, J. and Yao, D. (1991): "Bivariate Characterizations of Some Stochastic Order Relations", *Advances in Applied Probability* **23**, 642-659.

Wellman, P., Walsh, E., Wurman, P. and MacKie-Mason, J. (2001): "Auction Protocols for Decentralized Scheduling", *Games and Economic Behavior* **35**, 271-303.