

# Optimal Search, Learning, and Implementation

Alex Gershkov and Benny Moldovanu \*

May 15, 2009

## Abstract

We derive conditions on the learning environment - which encompasses both Bayesian and non-Bayesian processes - ensuring that an efficient allocation of resources is achievable in a dynamic allocation environment where impatient, privately informed agents arrive over time, and where the designer gradually learns about the distribution of agents' values. There are two main kind of conditions: 1) Higher observations should lead to more optimistic beliefs about the distribution of future values; 2) The allowed optimism associated with higher observations needs to be carefully bounded. Our analysis reveals and exploits close, formal relations between the problem of ensuring monotone - and hence implementable - allocation rules in our dynamic allocation problems with incomplete information and learning, and between the classical problem of finding optimal stopping policies for search that are characterized by a *reservation price property*

## 1 Introduction

In this paper we derive conditions on the learning process ensuring that an efficient allocation of resources is implementable in a dynamic allocation environment, where impatient, privately informed agents arrive over time, and where the designer gradually learns about the distribution of agents' values.

---

\*We wish to thank Sergiu Hart and Philippe Jehiel for several helpful remarks. We thank participants in seminars at the Paris School of Economics, Hebrew University, and Tel-Aviv University for their comments. We are grateful to the German Science Foundation for financial support. Gershkov, Moldovanu: Department of Economics, University of Bonn, mold@uni-bonn.de; alex.gershkov@uni-bonn.de

We also show that these conditions resemble insights about the *reservation price property* in search models obtained by the literature that followed Rothschild's [16] classical paper.

Although rather rare in the mechanism design literature, the assumption of gradual learning about the environment (which replaces here the standard assumption whereby the agents' values are not known but their distribution is) seems to us descriptive of most real-life dynamic allocation problems. This feature is inconsequential in static models where an efficient allocation is achieved by the dominant-strategy Vickrey-Clarke-Groves construction, but leads to new and interesting phenomena in dynamic settings.

The allocation (or assignment) model studied here is based on a classical model due to Derman, Lieberman and Ross [7] (DLR hereafter). In the DLR model, a finite set of possibly heterogenous, commonly ranked objects needs to be assigned to a set of agents who arrive one at a time. After each arrival, the designer decides which object (if any) to assign to the present agent (in a framework with several homogenous objects the decision is simply whether to assign an object or not).

Both the attribute of the present agent (that determines his value for the various available objects) and the future distribution of attributes are known to the designer in the DLR analysis. Learning in the complete-information DLR model has been first analyzed by Albright [1]. In Gershkov and Moldovanu [8] (GM) we added incomplete information to the learning model of Albright (resulting in a model of the private values type), and we showed that the efficient policy need not be implementable if the designer insists on the simultaneity of physical allocations and monetary payments (such schemes are called "online mechanisms" in the literature). This contrasts available results about efficient dynamic implementation for the standard case where the designer knows the distribution of values (see for example Parkes and Singh, [14], and Bergemann and Valimäki [3]). If all payments can be delayed until a time in the future when no new arrival occurs, the efficient allocation can always be implemented since the payments can be then conditioned on the actual allocation in each instance (see also Athey and Segal [2] who use such schemes in a dynamic mechanism design framework without learning). But such uncoupling of the physical and monetary parts is not realistic in many applications, and we will abstract from it here as well.

When learning about the environment takes place, the information revealed by a strategic agent affects both the current and the option values attached by the designer to various allocations. Since option values for the future serve as proxies for the values of allocating resources to other (future) agents, the private values model with learning indirectly generates informa-

tional externalities. Besides illustrating how these externalities can lead to the impossibility of efficient implementation, GM [8] derived implicit structural condition on the allocation policy ensuring that efficient implementation is possible.

A necessary condition for extracting truthful information about values is the monotonicity of the allocation rule, i.e., agents with higher values should not be worse-off than contemporaneous agents with lower values. This monotonicity will hold when the impact of currently revealed information on today's values is higher than the impact on option values (see Proposition 5 below which is a generalization of our previous result ). This insight translates to the private values dynamic framework with learning well-known results obtained for the static case with interdependent values.<sup>1</sup>

A next natural question is to characterize the learning environments where the monotonicity property holds. Intuitively, monotonicity will be satisfied if the increased optimism about the future distribution of values associated with higher current observation is not too drastic. A drastic optimism may be detrimental for an agent whose information induces it- leading to a failure of truthful revelation- if the designer decides in response to deny present resources in order to keep them for the "sunnier" future.

In the present paper we derive direct conditions on the learning environment - composed of the initial beliefs, and the belief updating process - that bound the afore mentioned optimism, thus allowing efficient dynamic implementation. Our analysis reveals and exploits close, formal relations between the problem of ensuring monotone - and hence implementable - allocation rules in our dynamic allocation problems with incomplete information and learning, and between the older, classical problem of obtaining optimal stopping policies for search that are characterized by a *reservation price property*. It is important to note that, in the relevant search literature, incomplete information and strategic interaction did not play any role.

In a famous paper, Rothschild [16] studied the problem of a consumer who obtains a sequence of price quotations from various sellers, and who must decide when to stop the (costly) search for a lower price. In Rothschild's model, the buyer has only partial information about the price distribution, and she updates (in a Bayesian way) her beliefs after each observation. Under full information about the environment, the optimal stopping rule is characterized by a reservation price  $R$  such that that the searcher accepts (or stops search) at any price less than equal to  $R$ , and rejects (or continues to search)

---

<sup>1</sup>Dasgupta and Maskin [6] and Jehiel and Moldovanu [9] have analyzed efficient implementation in static models with direct informational externalities. Kittsteiner and Moldovanu [10] used these insights in a dynamic model with direct externalities and without learning.

any price higher than  $R$ . One of the appealing features of this policy (see Rothschild's paper for the others) is that, if all customers follow it, a firm in the market will face a well-behaved demand function: expected sales are a non-increasing function of the price it charges. Such regularity conditions are extensively used in theoretical and empirical studies, and thus it is of major interest to find out when they are validated by theory. In the classical search model, price quotations are non-strategic, and the monotonicity requirement behind the reservation price property is a only a convenient, intuitive feature, facilitating the application of structural empirical methods in applied studies. In contrast, implementability is, of course, a "non-plus-ultra" requirement in our strategic, incomplete information model.

In the case studied by Rothschild, stopping prices necessarily change as information changes, and hence the optimal policy cannot be characterized by a single reservation price. But, in order to have expected sales decreasing in price, it is enough to assume that, for each information state, a searcher follows a *reservation price policy*, i.e., for each information state  $s$  there exists a price  $R(s)$  such that prices above are rejected and prices below are accepted. While the optimal Bayesian search rule need not generally have this property, Rothschild gave an example showing that the property holds for a searcher who obtains price quotations from a multinomial distribution which depends on parameters about which prior is Dirichlet<sup>2</sup>. Albright [1] computed several cases of Bayesian learning with conjugate priors where a generalized reservation price property holds in his model with several objects. This requires then that sets of types to whom particular objects are allocated are convex and ordered, with better objects being allocated to higher types. An obvious open problem was to establish some more or less general, sufficient conditions under which optimal search policies have the reservation price property. Various answers to this problem were offered by Rosenfield and Shapiro [15], Morgan [13], Seierstad [18] and Bickchandani and Sharma [4].

The conditions derived in our paper are more stringent than those obtained in the search literature (at least for Bayesian environments), mainly because of the presence here of multiple objects: these induce a more complex structure of the optimal search policies, and more stringent conditions are needed in order to control it<sup>3</sup>. In particular, and letting aside for a while the mechanism design/dynamic efficiency interpretation, our results can also be seen as offering conditions ensuring that the optimal search policy without recall for highest prices for several (possibly heterogenous) objects exhibits

---

<sup>2</sup>The Dirichlet is the conjugate prior of the multinomial distribution, so the posterior is also Dirichlet in this case.

<sup>3</sup>These more stringent conditions are needed even if all objects are homogenous.

the relevant reservation price property. Moreover, by a simple inversion of the interpretation of the optimal policy - better objects are then associated with lower types - our results also hold for the analogous version where a buyer sequentially searches for several lowest prices.

The paper is organized as follows: In Section 2 we present the sequential allocation and learning model. In Section 3 we first recall the result, due to Albright [1], that characterizes the efficient dynamic allocation policy under complete information about the arriving agents' values. Then, we derive several useful structural properties of that policy that are used in the sequel. In Section 4 we add incomplete information about the agents' value (while keeping the assumption that the designer gradually learn about the distribution of values). First, we generalize a result in GM [8] by deriving a simple sufficient condition ensuring that the first-best (i.e., complete information) efficient policy can be implemented also under incomplete information. This condition requires the optimal cutoffs defining the efficient allocation at each stage to satisfy a certain Lipschitz condition when regarded as functions of the current observation (see Proposition 5). The Lipschitz condition ensure that the impact of new information on option values is lower than the impact on current values. We next turn to our main results: these describe structural properties of general learning processes that induce the Lipschitz property. We offer two separate sets of sufficient conditions on general learning process, each set being composed of two requirements. The common requirement is a stochastic dominance condition: higher current observations should lead to more optimistic beliefs about the distribution of future values. The other requirement, respectively, puts a precise bound on the allowed optimism associated to higher observations in each period of search. The two obtained bounds differ in their response to an increase in the number of objects (or search periods): in the first result, Theorem 6, the bound becomes tighter in early search stages, while in the second the bound becomes tighter in later periods. We also offers illustrations for both results within the framework of standard Bayesian learning. In Subsection 4.1 we highlight the similarities and the differences between our results and several earlier results about the reservation price property obtained in the search literature. In Section 5 we turn our attention to two special, non-Bayesian learning models where updating is based on the empirical distribution and on a maximum entropy principle, respectively. Theorems 11 and 13 show that, given these learning models, the Lipschitz condition is always satisfied, and hence the corresponding efficient allocation policy is always implementable. Section 6 concludes.

## 2 The Model

There are  $m$  items and  $n$  agents. Each item  $i$  is characterized by a "quality"  $q_i$ , and each agent  $j$  is characterized by a "type"  $x_j$ . If an item with quality  $q_i \geq 0$  is assigned to an agent with type  $x_j$  and this agent is asked to pay  $p$ , then this agent enjoys a utility given by  $q_i x_j - p$ . Getting no item generates utility of zero. The goal is to find an assignment that maximizes total welfare. In a static problem, total welfare is maximized by assigning the item with the highest quality to the agent with the highest type, the item with the second highest quality to the agent with the second highest type, and so on... This assignment rule is called "*assortative matching*".

Here we assume that agents arrive sequentially, one agent per period of time, that each agent can only be served upon arrival (there is no recall), and that assigned items cannot be reallocated in the future.

Let period  $n$  denote the first period, period  $n-1$  denote the second period, ..., period 1 denote the last period. If  $m > n$  we can obviously discard the  $m - n$  worst items without welfare loss. If  $m < n$  we can add "dummy" objects with  $q_i = 0$ . Thus, we can assume without loss of generality that  $m = n$ .

While the items' properties  $0 \leq q_1 \leq q_2 \dots \leq q_m$  are assumed to be known, the agents' types are assumed to be independent and identically distributed random variables  $X_i$  on  $[0, +\infty)$  with common cumulative distribution function  $F$ .

The function  $F$  is not known to the designer nor to the agents. At the beginning of the allocation process the designer has a prior  $\Phi_n$  over possible distribution functions, and he updates his beliefs after each additional observation. Denote by  $\Phi_k(x_n, \dots, x_{k+1})$  the designer's beliefs about the distribution function  $F$  after observing types  $x_n, \dots, x_{k+1}$ . Given such beliefs, let  $\tilde{F}_k(x|x_n, \dots, x_{k+1})$  denote the distribution of the next type  $x_k$ , conditional on observing  $x_n, \dots, x_{k+1}$ . We assume that the distribution  $\tilde{F}_k(x|x_n, \dots, x_{k+1})$  is symmetric with respect to observed signals. The symmetry of  $\tilde{F}_k(x_k|\chi_{k+1}, x_{k+1})$  may depend on the used updating process. It is satisfied in Bayesian learning models and in non-Bayesian models used below.

Finally, we assume that each agent, upon arrival observes the whole history of the previous play.

### 3 The Efficient Allocation under Complete Information

We start by characterizing the dynamically efficient allocation while disregarding the agents' incentives to truthfully report their types. Alternatively, we assume first that there is complete information, i.e., the agent's type is revealed to the designer upon the agent's arrival (thus there is still uncertainty about the types of future agents). The efficient allocation maximizes at each decision period the sum of the expected utilities of all agents, given all the information available at that period.

Let the history at period  $k$ ,  $H_k$ , be the ordered set of all signals reported by the agents that arrived at periods  $n, \dots, k + 1$ , and of allocations to those agents. Let  $\mathcal{H}_k$  be the set of all histories at period  $k$ . Denote by  $\chi_k$  the ordered set of signals reported by the agents that arrived at periods  $n, \dots, k + 1$ . Finally, denote by  $\Pi_k$  the set of available objects at  $k$  (which has cardinality  $k$  by our convention that equates the number of objects with the number of periods). Note that an initial inventory  $\Pi_n$  and a history  $H_k$  completely determine the set  $\Pi_k$ .

The result below characterizes, at each period, the dynamically efficient policy in terms of cutoffs which are determined by the history of observed signals. This policy can be seen as the dynamic version of the assortative matching policy that is optimal in the static case where all agents arrive simultaneously.

**Theorem 1** (*Albright, 1977*)

1. Assume that types  $x_n, \dots, x_{k+1}$  have been observed, and consider the arrival of an agent with type  $x_k$  in period  $k \geq 1$ . There exist functions  $0 = a_{0,k}(\chi_k, x_k) \leq a_{1,k}(\chi_k, x_k) \leq a_{2,k}(\chi_k, x_k) \dots \leq a_{k,k}(\chi_k, x_k) = \infty$  such that the efficient dynamic policy - which maximizes the expected value of the total reward - assigns the item with the  $i$ -th smallest type if  $x_k \in (a_{i-1,k}(\chi_k, x_k), a_{i,k}(\chi_k, x_k)]$ . The functions  $a_{i,k}(\chi_k, x_k)$  do not depend on the  $q$ 's.
2. Each  $a_{i,k+1}(\chi_{k+1}, x_{k+1})$  equals the expected value of the agent's type to which the item with  $i$ -th smallest type is assigned in a problem with  $k$  periods before the period  $k$  signal is observed. These constants are

related to each other by the following recursive formulae:

$$\begin{aligned}
a_{i,k+1}(\chi_{k+1}, x_{k+1}) &= \int_{A_{i,k}} x_k d\tilde{F}_k(x_k | \chi_{k+1}, x_{k+1}) \\
&+ \int_{\underline{A}_{i,k}} a_{i-1,k}(\chi_k, x_k) d\tilde{F}_k(x_k | \chi_{k+1}, x_{k+1}) \\
&+ \int_{\bar{A}_{i,k}} a_{i,k}(\chi_k, x_k) d\tilde{F}_k(x_k | \chi_{k+1}, x_{k+1})
\end{aligned} \tag{1}$$

where<sup>4</sup>

$$\begin{aligned}
\underline{A}_{i,k} &= \{x_k : x_k \leq a_{i-1,k}(\chi_k, x_k)\} \\
A_{i,k} &= \{x_k : a_{i-1,k}(\chi_k, x_k) < x_k \leq a_{i,k}(\chi_k, x_k)\} \\
\bar{A}_{i,k} &= \{x_k : x_k > a_{i,k}(\chi_k, x_k)\}
\end{aligned}$$

Note that, by the above Theorem, we can write

$$a_{i,k+1}(\chi_{k+1}, x_{k+1}) = E_{x_k | x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \tag{2}$$

where the function  $G_{i,k}(x_k, x_{k+1}, \chi_{k+1})$  is given by:

$$\left\{ \begin{array}{ll} a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k) & \text{if } x_k \leq a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k) \\ x_k & \text{if } a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k) < x_k \leq a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) \\ a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) & \text{if } x_k > a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) \end{array} \right. . \tag{3}$$

In other words  $G_{i,k}(x_k, x_{k+1}, \chi_{k+1})$  is the second-highest order statistic out of the set  $\{a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k), x_k, a_{i,k}(\chi_{k+1}, x_{k+1}, x_k)\}$ . Note also that if  $\tilde{F}_k(x_k | \chi_{k+1}, x_{k+1})$  is symmetric with respect to the observed signals, then  $a_{i,k+1}(\chi_{k+1}, x_{k+1})$  is symmetric as well.

We next prove two structural results that will be used in the proofs of our main Theorems. First, we show that the average of all but the extreme cutoffs equals the expectation about the next type.

**Lemma 2** *For any  $k \leq n$ , it holds that*

$$\sum_{i=1}^{k-1} a_{i,k}(\chi_k, x_k) = (k-1) E_{x_k | \chi_k}(x_k).$$

---

<sup>4</sup>We set  $+\infty \cdot 0 = -\infty \cdot 0 = 0$ .

**Proof.** We prove the claim by induction. For  $k = 2$ ,  $a_{1,2}(\chi_2, x_2) = \int_0^\infty x_1 d\tilde{F}_1(x_1|\chi_1, x_2) = E_{x_1|\chi_2, x_2} x_1$ . Theorem 1 implies that, for any fixed  $x_k$ ,

$$\sum_{i=1}^k \left[ a_{i-1,k}(\chi_k, x_k) \mathbf{1}_{\underline{A}_{i,k}} + a_{i,k}(\chi_k, x_k) \mathbf{1}_{\overline{A}_{i,k}} \right] = \sum_{i=1}^{k-1} a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) \quad (4)$$

where  $\mathbf{1}_s$  is an index function. Using (1) and the previous expression we obtain for period  $k + 1$  that:

$$\begin{aligned} \sum_{i=1}^k a_{i,k+1}(\chi_{k+1}, x_{k+1}) &= \int_0^\infty x d\tilde{F}_k(x|\chi_{k+1}, x_{k+1}) \\ + \sum_{i=1}^{k-1} E_{x_k|\chi_{k+1}, x_{k+1}} a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) &= k E_{x_k|\chi_{k+1}, x_{k+1}} x_k \end{aligned}$$

where the first equality follows from (4), and where the last equality follows from the induction argument. ■

Next, we derive a monotonicity properties of the cutoffs that holds whenever higher observations induce more optimistic beliefs about the distribution of values:

**Lemma 3** *Assume that for any  $k$ , and for any pair of ordered lists of reports  $\chi_k \geq \chi'_k$  that differ only in one coordinate  $\tilde{F}_k(x|\chi_k) \succeq_{FOSD} \tilde{F}_k(x|\chi'_k)$ . Then the cutoff  $a_{i,k}(\chi_k, x_k)$  is non-decreasing in  $x_k$ .*

**Proof.** The proof is by induction on the number of remaining periods. For  $k = 2$  we have

$$\begin{aligned} a_{2,2}(\chi_2, x_2) &= \infty \\ a_{1,2}(\chi_2, x_2) &= \int_0^\infty x_1 d\tilde{F}_1(x_1|\chi_2, x_2) \\ a_{0,2}(\chi_2, x_2) &= 0 \end{aligned}$$

Stochastic dominance immediately implies that the cutoffs are non-decreasing in  $x_2$ . We now apply the induction argument, and assume that, for any  $\chi_k$  and for any  $i$ ,  $a_{i,k}(\chi_k, x_k)$  is non-decreasing in  $x_k$ . This implies that the function  $G_{i,k}(x_k, x_{k+1}, \chi_{k+1})$  is non-decreasing in  $x_k$  and that for any  $i$ ,

$$\begin{aligned} a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) &= a_{i,k}(\chi_{k+1}, x_k, x_{k+1}) \geq \\ a_{i,k}(\chi_{k+1}, x_k, x'_{k+1}) &= a_{i,k}(\chi_{k+1}, x'_{k+1}, x_k) \end{aligned}$$

where both equalities follow from the assumption of symmetry whereby switching the order of the observations does not affect the final beliefs. Therefore we obtain  $G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \geq G_{i,k}(x_k, x'_{k+1}, \chi_{k+1})$  for any  $x_k$ . Moreover we have that

$$\begin{aligned} a_{i,k+1}(\chi_{k+1}, x_{k+1}) &= E_{x_k|x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \\ &\geq E_{x_k|x_{k+1}} G_{i,k}(x_k, x'_{k+1}, \chi_{k+1}) \\ &\geq E_{x_k|x'_{k+1}} G_{i,k}(x_k, x'_{k+1}, \chi_{k+1}) = a_{i,k+1}(\chi_{k+1}, x'_{k+1}) \end{aligned}$$

where the second inequality follows from the assumed stochastic dominance, and from the fact that, by the induction argument,  $G_k(x_k, x'_{k+1}, \chi_{k+1})$  is non-decreasing in  $x_k$ . ■

The stochastic dominance condition employed in order to obtain the monotonicity of the cutoffs defining the optimal policy is, for example, a simple consequence of a standard setting found in the literature: Assume that values  $x$  are drawn according to a density  $f(x|\theta)$  where  $\theta \in \mathbb{R}$ . Denote by  $h(\theta)$  the density of  $\theta$ , and by  $H(\theta)$  the corresponding probability distribution - this is here the prior belief which gets then updated after each observation. Then the following holds:

**Lemma 4** *Assume that  $f(x|\theta)$  has the Monotone Likelihood Ratio (MLR) property. That is, for any  $x > x'$ ,*

$$\frac{\partial}{\partial \theta} \left[ \frac{f(x|\theta)}{f(x'|\theta)} \right] > 0.$$

*Then, for any  $k$ , and for any ordered lists of observations  $\chi_k \geq \chi'_k$  that differ only in one coordinate it holds that*

$$\tilde{F}_k(x|\chi_k) \leq \tilde{F}_k(x|\chi'_k) \text{ for all } x.$$

*where  $\tilde{F}$ , the conditional distribution of the next value, is obtained by Bayesian updating.*

**Proof.** First, recall that the MLR property implies that  $F(x|\theta)$  first order stochastically dominates  $F(x|\theta')$  if  $\theta > \theta'$ . Therefore, for any  $x > 0$ , we obtain that  $F(x|\theta)$  decreases with  $\theta$ . Bayes' rule implies then that

$$\tilde{F}_k(x|\chi_k) = \int_{-\infty}^{\infty} F(x|\theta) h(\theta|\chi_k) d\theta.$$

Consider now two sequences  $\chi_k$  and  $\chi'_k$  that differ only in one observation (say the observation of period  $i$ ) and such that  $\chi_k \geq \chi'_k$ . Since  $F(x|\theta)$

decreases with  $\theta$  for any  $x > 0$ , it is sufficient to show that  $H(\theta|\chi_k)$  first-order stochastically dominates  $H(\theta|\chi'_k)$ . By a result of Milgrom [12], The *MLR* property implies that  $H(\theta|x)$  first-order stochastically dominates  $H(\theta|x')$  if  $x > x'$ . Since Milgrom's result holds for any prior  $h(\theta)$ , we can also apply it to  $h(\theta|\chi_k \setminus x_i)$  which represents the posterior belief about the parameter  $\theta$  after having observed the sequence  $\chi_k$  without signal  $x_i$ . This completes the proof. ■

## 4 Dynamic Efficient Implementation

We now assume that there is incomplete information about values. Without loss of generality, we restrict attention to direct mechanisms where every agent, upon arrival, reports his characteristic  $x_i$  and where the mechanism specifies an allocation (which item, if any, the agent gets) and a payment. The next result, related to a result in GM [8], displays a sufficient condition on the cutoffs of the efficient, complete information allocation, characterized in Theorem 1 above, ensuring that this allocation is implementable also under incomplete information.

**Proposition 5** *Assume that for any  $k, \chi_k, i \in \{0, \dots, k\}$ , the cutoff  $a_{i,k}(\chi_k, x_k)$  is a Lipschitz function of  $x_k$  with constant 1. Then, the efficient dynamic policy is implementable under incomplete information.*

**Proof.** GM [8] showed that the efficient allocation is implementable if and only if for any  $k, i \leq k$  and  $\chi_k$  the set  $\{x : a_{i,k}(\chi_k, x) > x \geq a_{i-1,k}(\chi_k, x)\}$  is convex. The characterization of the complete information efficient allocation provided by Albricht states that for any  $k, i \leq k, x$  and  $\chi_k$  we have  $a_{i,k}(\chi_k, x) \geq a_{i-1,k}(\chi_k, x)$ . Therefore, it is sufficient to show that if there exist  $k, \chi_k$  and  $i \in \{0, \dots, k\}$ , and a signal  $x_k$  with  $a_{i,k}(\chi_k, x_k) < x_k$ , then there is no  $x'_k > x_k$  such that  $a_{i,k}(\chi_k, x'_k) > x'_k$ . Assume that such  $x'_k$  exists. Since  $a_{i,k}$  is Lipschitz with constant 1,  $a_{i,k}(\chi_k, x'_k) \leq x'_k - x_k + a_{i,k}(\chi_k, x_k)$ . Since  $a_{i,k}(\chi_k, x_k) < x_k$ , we obtain  $a_{i,k}(\chi_k, x'_k) < x'_k$ , which yields a contradiction. ■

Due to the learning process, the current information affects both the current value of allocating some object to the arriving agent and the option value of keeping that object and allocating it in the future. The previous result requires the effect of the current information on the current value to be stronger than the effect on the option value. But what conditions on the models' primitives induce the Lipschitz property, and hence the possibility of

implementing the first-best dynamically efficient allocation also under conditions of incomplete information? The next Theorems provide several distinct answers to this question.

**Theorem 6** *Assume that for any  $k$ , and for any pair of ordered lists of reports  $\chi_k \geq \chi'_k$  that differ only in one coordinate, the following conditions hold:*

$$\text{suff1 } \tilde{F}_k(x|\chi_k) \succeq_{FOSD} \tilde{F}_k(x|\chi'_k)$$

$$\text{suff2 } E(x|\chi_k) - E(x|\chi'_k) \leq \frac{\Delta}{k-1} \text{ where } \Delta \text{ is size of the difference between } \chi_k \text{ and } \chi'_k$$

*Then, the efficient dynamic policy can be implemented also under incomplete information.*

**Proof.** Lemma 2 and the second condition in the Theorem's statement imply that

$$\begin{aligned} \sum_{i=1}^{k-1} (a_{i,k}(\chi_k, x_k) - a_{i,k}(\chi_k, x'_k)) &= (k-1) (E_{x_{k-1}|\chi_k, x_k} x_{k-1} - E_{x_{k-1}|\chi_k, x'_k} x_{k-1}) \\ &\leq \frac{k-1}{k-1} (x_k - x'_k). \end{aligned} \quad (5)$$

In other words, the sum of cutoffs  $\sum_{i=1}^{k-1} a_{i,k}(\chi_k, x_k)$  is a Lipschitz function with constant 1 of  $x_k$ . By Lemma 3, and the stochastic dominance condition, we know that the cutoff  $a_{i,k}(\chi_k, x_k)$  is a non-decreasing function of  $x_k$ . Therefore, inequality 5 implies that, for any  $i$ , the function  $a_{i,k}(\chi_k, x_k)$  must also be a Lipschitz function with constant 1 of  $x_k$ . By Proposition 5, the efficient dynamic policy is then implementable. ■

The first condition (stochastic dominance) in the above Theorem says that higher observations should lead to optimism about future observations, while the second condition puts a bound on this optimism. The result is simple, but its disadvantage is that, as the number of objects (or search periods) grows, the second condition gets tougher (i.e., the bound on the optimism associated to higher observation gradually decreases) in the early search periods. Here is an illustration of this phenomenon:

**Example 7** *Assume that with probability  $p$  the arriving agent's type  $x$  is distributed on the interval  $[0, 1]$  with density  $f_1(x) = 1 - \frac{b_1}{2} + b_1 x$ , and with*

probability  $1 - p$  it is distributed on  $[0, 1]$  with density  $f_2(x) = 1 - \frac{b_2}{2} + b_2x$ , where  $b_1, b_2 \in [-2, 2)$ . Note that

$$\begin{aligned} E[F_i] &= \frac{1}{2} + \frac{b_i}{12} \text{ and} \\ E(x|\chi_k) &= \Pr(b_i = b_1|x_n, \dots, x_{k+1}) E[F_1] + \Pr(b_i = b_2|x_n, \dots, x_{k+1}) E[F_2] \end{aligned}$$

Using Bayesian updating we get that

$$\Pr(b_i = b_1|x_n, \dots, x_{k+1}) = \left( 1 + \frac{1-p}{p} \prod_{j=k+1}^n \frac{1 - \frac{b_2}{2} + b_2x_j}{1 - \frac{b_1}{2} + b_1x_j} \right)^{-1}$$

Therefore,

$$E(x|\chi_k) - E(x|\chi'_k) = \frac{b_1 - b_2}{12} [\Pr(b_i = b_1|\chi_k) - \Pr(b_i = b_1|\chi'_k)].$$

Let  $\chi_k$  and  $\chi'_k$  be two sequences of observed signals that differ only in one coordinate, with  $\chi_k \geq \chi'_k$ . Then

$$\begin{aligned} &\Pr(b_i = b_1|\chi_k) - \Pr(b_i = b_1|\chi'_k) \\ &= \frac{\frac{1-p}{p} \prod_{j=k+1}^n \frac{1 - \frac{b_2}{2} + b_2x_j}{1 - \frac{b_1}{2} + b_1x_j} \frac{(b_1 - b_2)(x_i - x'_i)}{(1 - \frac{b_2}{2} + b_2x_i)(1 - \frac{b_1}{2} + b_1x'_i)}}{\left( 1 + \frac{1-p}{p} \prod_{j=k+1}^n \frac{1 - \frac{b_2}{2} + b_2x_j}{1 - \frac{b_1}{2} + b_1x_j} \right) \left( 1 + \frac{1-p}{p} \prod_{j=k+1}^n \frac{1 - \frac{b_2}{2} + b_2x'_j}{1 - \frac{b_1}{2} + b_1x'_j} \right)}. \end{aligned}$$

Since

$$\frac{\frac{1-p}{p} \prod_{j=k+1}^n \frac{1 - \frac{b_2}{2} + b_2x_j}{1 - \frac{b_1}{2} + b_1x_j}}{\left( 1 + \frac{1-p}{p} \prod_{j=k+1}^n \frac{1 - \frac{b_2}{2} + b_2x_j}{1 - \frac{b_1}{2} + b_1x_j} \right) \left( 1 + \frac{1-p}{p} \prod_{j=k+1}^n \frac{1 - \frac{b_2}{2} + b_2x'_j}{1 - \frac{b_1}{2} + b_1x'_j} \right)} < 1$$

we obtain

$$\begin{aligned} E(x|\chi_k) - E(x|\chi'_k) &< \frac{(b_1 - b_2)^2}{12} \frac{(x_i - x'_i)}{\left( 1 - \frac{b_2}{2} + b_2x_i \right) \left( 1 - \frac{b_1}{2} + b_1x'_i \right)} \\ &\leq \frac{(b_1 - b_2)^2}{3(2 - b_2)(2 - b_1)} (x_i - x'_i). \end{aligned}$$

Finally, if

$$\frac{(b_1 - b_2)^2}{3(2 - b_2)(2 - b_1)} \leq \frac{1}{n - 1}, \quad (6)$$

we obtain that

$$E(x|\chi_k) - E(x|\chi'_k) \leq \frac{(x_i - x'_i)}{n-1} \leq \frac{(x_i - x'_i)}{k-1}.$$

as desired. It is obvious that the set of parameters  $\{b_1, b_2\}$  where the condition is satisfied shrinks as  $n$  goes to infinity

In order to obtain conditions on the learning process that hold independently of the number of objects/ periods, we focus now on bounds that, as the number of objects grows, get tighter in late, rather than in early periods. Such conditions are, in principle, easier to satisfy generally since in many learning models (in particular in those learning models where beliefs converge, say, to the true distribution) the impact of later observations on beliefs is significantly lower than that of early observations. Thus, a tighter bound on the allowed optimism associated with higher observations is less likely to be binding in late periods. The proof of the next result is somewhat more involved. For mathematical convenience, we make a mild differentiability assumption that allows us to work with bounds on derivatives rather than with the Lipschitz condition of Proposition 5.

**Theorem 8** *Assume that, for all  $k$ , all  $x$ , and all  $n - k \geq i \geq 1$ , the conditional distribution function  $\tilde{F}_k(x|x_n, \dots, x_{k+1})$  and the density  $\tilde{f}_k(x|x_n, \dots, x_{k+1})$  are continuously differentiable with respect to  $x_{k+i}$ . If for all  $x$ ,  $\chi_k$ , and all  $n - k \geq i \geq 1$ , it holds that*

$$0 \geq \frac{\partial \tilde{F}_k(x|\chi_k)}{\partial x_{k+i}} \geq -\frac{1}{n-k} \tilde{f}_k(x|\chi_k) \quad (7)$$

then the efficient dynamic policy can be implemented also under incomplete information.

**Proof.** Note first that

$$\begin{aligned} \frac{\partial E(x_k|\chi_k)}{\partial x_{k+i}} &= \frac{\partial}{\partial x_{k+i}} \int_0^\infty (1 - \tilde{F}_k(x_k|\chi_k)) dx_k \\ &\leq \frac{1}{n-k} \int_0^\infty \tilde{f}_k(x_k|\chi_k) dx_k = \frac{1}{n-k} \end{aligned} \quad (8)$$

where the inequality follows from the condition of the theorem. By Proposition 5, it is sufficient to show that for any  $k$ , any history of reports  $\chi_k$ ,

and any  $n - k \geq i \geq 1$ , the cutoff  $a_{i,k}(\chi_k, x_k)$  is differentiable and satisfies  $\frac{\partial}{\partial x_k} a_{i,k}(\chi_k, x_k) \leq 1$ . Since  $a_{i,k}(\chi_k, x_k) = E_{x_{k-1}|\chi_k, x_k} G_{i,k-1}(x_{k-1}, x_k, \chi_k)$ , we need to show that  $\frac{\partial}{\partial x_k} E_{x_{k-1}|\chi_k, x_k} G_{i,k-1}(x_{k-1}, x_k, \chi_k)$  exists and that

$$\frac{\partial}{\partial x_k} E_{x_{k-1}|\chi_k, x_k} G_{i,k-1}(x_{k-1}, x_k, \chi_k) \leq 1.$$

We claim now that  $E_{x_k|\chi_{k+1}, x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1})$  is differentiable and that

$$\frac{\partial E_{x_k|\chi_{k+1}, x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \leq \frac{1}{n-k}$$

This yields  $\frac{\partial}{\partial x_{k+1}} a_{i,k+1}(\chi_{k+1}, x_{k+1}) \leq \frac{1}{n-k}$  for any history of signals  $\chi_{k+1}$ , any pair of signals  $x_k, x_{k+1}$ , any period  $k+1 > 1$ , and any item  $i$ .

We prove the claim by induction on the number of the remaining periods  $k$ . For  $k=1$ , note that  $a_{0,1}(\chi_2, x_2, x_1) = 0$  and  $a_{1,1}(\chi_2, x_2, x_1) = \infty$ . Hence, we have  $G_{1,1}(x_1, x_2, \chi_2) = x_1$ . Therefore, inequality (8) implies

$$\begin{aligned} \frac{\partial}{\partial x_2} E_{x_1|\chi_2, x_2} G_{1,1}(x_1, x_2, \chi_2) &\leq \frac{1}{n-1} \text{ and} \\ \frac{\partial}{\partial x_2} a_{1,2}(\chi_2, x_2) &\leq \frac{1}{n-1} \end{aligned}$$

Note also that continuous differentiability of  $\tilde{f}_1(x|x_n, \dots, x_2)$  implies continuous differentiability of  $a_{1,2}(\chi_2, x_2)$ . Assume now that  $a_{i,k}(\chi_k, x_k)$  is continuously differentiable and that

$$\begin{aligned} \frac{\partial E_{x_{k-1}|\chi_k, x_k} G_{i,k-1}(x_{k-1}, x_k, \chi_k)}{\partial x_k} &\leq \frac{1}{n-k+1}, \\ \frac{\partial a_{i,k}(\chi_k, x_k)}{\partial x_k} &\leq \frac{1}{n-k+1} \end{aligned}$$

Since  $a_{i,k}(\chi_k, x_k)$  is continuous, the induction hypothesis implies that for any  $i \in \{1, \dots, k-1\}$  there exists at most one solution to the equation  $a_{i,k}(\chi_k, x) = x$ . Denote this solution by  $a_{i,k}^*(\chi_k)$ . If  $a_{i,k}(\chi_k, x) > x$  for any  $x$ , define  $a_{i,k}^*(\chi_k) = \infty$ , and if  $a_{i,k}(\chi_k, x) < x$  for any  $x$  define  $a_{i,k}^*(\chi_k) = 0$ . Recall

that, by induction, we can rewrite

$$\begin{aligned}
& E_{x_k|\chi_{k+1}, x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \\
= & \int_0^{a_{i-1,k}^*(\chi_k)} a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k) f(x_k|\chi_{k+1}, x_{k+1}) dx_k \\
& + \int_0^{a_{i,k}^*(\chi_k)} x_k f(x_k|\chi_{k+1}, x_{k+1}) dx_k \\
& + \int_{a_{i-1,k}^*(\chi_k)}^\infty a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) f(x_k|\chi_{k+1}, x_{k+1}) dx_k.
\end{aligned}$$

Since  $a_{i,k}(\chi_{k+1}, x_{k+1}, x_k)$  is continuously differentiable in  $x_{k+1}$  for any  $i \in \{1, \dots, k-1\}$  by the induction argument, and since  $\tilde{f}_k(x_k|\chi_{k+1}, x_{k+1})$  is continuously differentiable by assumption, we can invoke the Implicit Function Theorem to deduce that the fixed point  $a_{i,k}^*(\chi_k)$  is continuously differentiable in  $x_{k+1}$ . Thus, we obtain that  $E_{x_k|\chi_{k+1}, x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1})$  is continuously differentiable in  $x_{k+1}$ .

We now show that  $\frac{\partial}{\partial x_{k+1}} E_{x_k|\chi_{k+1}, x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \leq \frac{1}{n-k}$ . We have

$$\begin{aligned}
& \frac{\partial}{\partial x_{k+1}} \int_0^\infty G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \tilde{f}_k(x_k|x_{k+1}, \chi_{k+1}) dx_k \\
= & \int_0^\infty \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \tilde{f}_k(x_k|x_{k+1}, \chi_{k+1}) dx_k \tag{9}
\end{aligned}$$

$$+ \int_0^\infty \frac{\partial \tilde{f}_k(x_k|x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) dx_k. \tag{10}$$

Consider first the term in the sum above (9):

$$\begin{aligned}
& \int_0^\infty \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \tilde{f}_k(x_k | x_{k+1}, \chi_{k+1}) dx_k \\
= & \int_0^{a_{i-1,k}^*(x_{k+1}, \chi_{k+1})} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \tilde{f}_k(x_k | x_{k+1}, \chi_{k+1}) dx_k \\
& + \int_{a_{i,k}^*(x_{k+1}, \chi_{k+1})}^{a_{i-1,k}^*(x_{k+1}, \chi_{k+1})} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \tilde{f}_k(x_k | x_{k+1}, \chi_{k+1}) dx_k \\
& + \int_{a_{i,k}^*(x_{k+1}, \chi_{k+1})}^\infty \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \tilde{f}_k(x_k | x_{k+1}, \chi_{k+1}) dx_k \\
\leq & \frac{1}{n-k+1} - \\
& \frac{1}{n-k+1} \left[ \tilde{F}_k(a_{i,k}^*(x_{k+1}, \chi_{k+1}) | x_{k+1}, \chi_{k+1}) - \tilde{F}_k(a_{i-1,k}^*(x_{k+1}, \chi_{k+1}) | x_{k+1}, \chi_{k+1}) \right]
\end{aligned}$$

where the existence of the fixed points  $a_{i,k}^*(x_{k+1}, \chi_{k+1})$  and  $a_{i-1,k}^*(x_{k+1}, \chi_{k+1})$  follows from the induction argument, while the inequality follows from the induction argument and from the fact that  $\frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} = 0$  if  $x_k \in [a_{i-1,k}^*(x_{k+1}, \chi_{k+1}), a_{i,k}^*(x_{k+1}, \chi_{k+1})]$ .

Consider now the second term in the sum (10):

$$\begin{aligned}
& \int_0^\infty \frac{\partial \tilde{f}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) dx_k \\
&= \left. \frac{\partial \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \right|_{x_k=0}^\infty \\
&\quad - \int_0^\infty \frac{\partial \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_k} dx_k \\
&= - \int_0^\infty \frac{\partial \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_k} dx_k \\
&\leq \frac{1}{n-k+1} \int_0^\infty \frac{\partial [1 - \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})]}{\partial x_{k+1}} dx_k \\
&\quad - \frac{n-k}{n-k+1} \int_{a_{i-1,k}^*(x_{k+1}, \chi_{k+1})}^{a_{i,k}^*(x_{k+1}, \chi_{k+1})} \frac{\partial \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} dx_k \\
&\leq \frac{1}{n-k+1} \frac{1}{n-k} - \frac{n-k}{n-k+1} \int_{a_{i-1,k}^*(x_{k+1}, \chi_{k+1})}^{a_{i,k}^*(x_{k+1}, \chi_{k+1})} \frac{\partial \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} dx_k
\end{aligned}$$

where the first equality follows by integration by parts, and where the second equality follows because  $\lim_{x \rightarrow \infty} \tilde{F}_k(x | x_{k+1}, \chi_{k+1}) = 1$  and  $\tilde{F}_k(0 | x_{k+1}, \chi_{k+1}) = 0$ . The first inequality follows by the induction argument (which implies the existence of the fixed points  $a_{i,k}^*(x_{k+1}, \chi_{k+1})$ ,  $a_{i-1,k}^*(x_{k+1}, \chi_{k+1})$ ) and because

$$\frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_k} \begin{cases} = 1 & \text{if } x_k \in [a_{i-1,k}^*(x_{k+1}, \chi_{k+1}), a_{i,k}^*(x_{k+1}, \chi_{k+1})] \\ \leq \frac{1}{n-k+1} & \text{if } x_k \notin [a_{i-1,k}^*(x_{k+1}, \chi_{k+1}), a_{i,k}^*(x_{k+1}, \chi_{k+1})] \end{cases} .$$

Combining now the two terms 9 and 10 we obtain

$$\begin{aligned}
& \frac{\partial}{\partial x_{k+1}} \int_0^\infty G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \tilde{f}_k(x_k | x_{k+1}, \chi_{k+1}) dx_k \\
\leq & \frac{1}{n-k+1} - \\
& \frac{1}{n-k+1} \left[ \tilde{F}_k(a_{i,k}^*(x_{k+1}, \chi_{k+1}) | x_{k+1}, \chi_{k+1}) - \tilde{F}_k(a_{i-1,k}^*(x_{k+1}, \chi_{k+1}) | x_{k+1}, \chi_{k+1}) \right] \\
& + \frac{1}{n-k+1} \frac{1}{n-k} - \frac{n-k}{n-k+1} \int_{a_{i-1,k}^*(x_{k+1}, \chi_{k+1})}^{a_{i,k}^*(x_{k+1}, \chi_{k+1})} \frac{\partial \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} dx_k \quad (11)
\end{aligned}$$

Recalling the miraculous relation

$$\frac{1}{n-k+1} \frac{1}{n-k} + \frac{1}{n-k+1} = \frac{1}{n-k}$$

it is therefore sufficient to prove that

$$\begin{aligned}
& \frac{1}{n-k} \left[ \tilde{F}_k(a_{i,k}^*(x_{k+1}, \chi_{k+1}) | x_{k+1}, \chi_{k+1}) - \tilde{F}_k(a_{i-1,k}^*(x_{k+1}, \chi_{k+1}) | x_{k+1}, \chi_{k+1}) \right] \\
\geq & - \int_{a_{i-1,k}^*(x_{k+1}, \chi_{k+1})}^{a_{i,k}^*(x_{k+1}, \chi_{k+1})} \frac{\partial \tilde{F}_k(x_k | x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} dx_k.
\end{aligned}$$

Integrating with respect to  $x$  both sides of the assumed inequality

$$-\frac{\partial \tilde{F}_k(x | \chi_k)}{\partial x_{k+i}} \leq \frac{1}{n-k} \tilde{f}_k(x | \chi_k)$$

between the fixed points  $a_{i-1,k}^*(x_{k+1}, \chi_{k+1})$  and  $a_{i,k}^*(x_{k+1}, \chi_{k+1})$  yields the desired result. ■

While the left hand inequality in condition 7 is just another way to express the stochastic dominance condition also employed in Theorem 6, it is worth to explore deeper the right hand side. Putting aside differentiability for a moment, this condition is equivalent to requiring that the function  $\tilde{F}_k(x + \frac{z}{n-k} | x_{k+1}, \cdot, x_{k+i} + z, x_{k+i+1}, \cdot, x_n)$  is non-decreasing in  $z$ . In other words, after having already obtained  $n-k$  observations, a shift to the right - which moves the value of the distribution upwards - is enough to compensate the downward shift in the value of the distribution caused by an  $(n-k)$  times larger upward shift in one of the past observations (recall that, by stochastic dominance, higher observations move the entire distribution downwards).

**Example 9** A simple illustration where the conditions in the above Theorem are satisfied is obtained by considering a normal distribution of values  $\tilde{x} \sim N(\mu, 1)$  with unknown mean  $\mu$ , and prior beliefs about  $\mu$  of the form  $\tilde{\mu} \sim N(\mu_0, 1/\tau)$  where  $\tau > 0$ . After observing  $x_n, \dots, x_{k+1}$  the posterior on  $\tilde{\mu}$  is given by  $N(\bar{\mu}, 1/(\tau + n - k))$  where

$$\bar{\mu} = \frac{\tau\mu_0 + \sum x_i}{\tau + (n - k)}$$

This yields

$$\tilde{F}_k(x|x_n, \dots, x_{k+1}) = N(\bar{\mu}, 1 + 1/(\tau + n - k))$$

Note that

$$\tilde{F}_k\left(x + \frac{z}{\tau + (n - k)} \middle| x_n, \dots, x_i + z, \dots, x_{k+1}\right) = \tilde{F}_k(x|x_n, \dots, x_i, \dots, x_{k+1}) \quad (12)$$

so that the stochastic dominance condition necessarily holds. By differentiating with respect to  $z$  both sides of the identity 12, and by letting  $z$  go to zero, we obtain that

$$\begin{aligned} \frac{\partial \tilde{F}_k(x|x_n, \dots, x_{k+1})}{\partial x_{k+i}} &= -\frac{1}{\tau + n - k} \tilde{f}_k(x|x_n, \dots, x_{k+1}) \Rightarrow \\ \frac{\partial \tilde{F}_k(x|x_n, \dots, x_{k+1})}{\partial x_{k+i}} &\geq -\frac{1}{n - k} \tilde{f}_k(x|x_n, \dots, x_{k+1}) \end{aligned}$$

as desired.

## 4.1 A Connection to Search for the Lowest Price

As mentioned in the Introduction, the first general conditions ensuring that the optimal search policy in Rothschild's search model is characterized by a sequence of reservation prices appear in a subtle paper by Rosenfield and Shapiro [15]. In order to understand the relation between our results and theirs, recall first our condition from Theorem 8: For all  $x$ ,  $\chi_k$ , and all  $n - k \geq i \geq 1$

$$0 \geq \frac{\partial \tilde{F}_k(x|\chi_k)}{\partial x_{k+i}} \geq -\frac{1}{n - k} \tilde{f}_k(x|\chi_k) \quad (13)$$

The first requirement in the paper by Rosenfield and Shapiro is identical to our stochastic dominance condition (the left hand side of condition 13), while their second condition - translated to the differentiable case and to the

case of a searching seller instead of a searching buyer in order to facilitate comparison- reads: For all  $x, k, \chi_k$  and all  $n - k \geq i \geq 1$

$$\int_x^\infty \frac{\partial \tilde{F}_k(y|\chi_k) dy}{\partial x_{k+i}} \geq -\frac{1}{n-k}(1 - \tilde{F}_k(x|\chi_k)) \quad (14)$$

In other words, theirs is simply the "average" version of the right hand side side of our condition 13, and hence it is obviously implied by it.

Seierstad [18] offers another variant. Besides stochastic dominance, his condition reads (again in the differentiable case): For all  $x, k$  and  $\chi_k$

$$\sum_{i=1}^{n-k} \frac{\partial \tilde{F}_k(x|\chi_k)}{\partial x_{k+i}} \geq -\tilde{f}_k(x|\chi_k) \quad (15)$$

which is also clearly implied by our condition 13. The reason why we need stronger conditions than both Rosenfield and Shapiro's and Seierstad's is intimately related to the fact that we do analyze a model with several objects: at each point in time we have several critical cutoffs to control, instead of only one. In particular, the reservation price property is connected in our model to the existence of several fixed points at each period, and we need to control the conditional distribution of future values between any two such fixed points (without a-priori knowing where they will be). In contrast, in the one-object search problem there are only two fixed points to consider at each period, and one of them is trivially equal to either "minus infinity" (for a searching buyer) or "plus infinity" (for a searching seller). This fact allows Rosenfield and Shapiro to use an average bound, and Seierstad to use a bound that aggregates the effect of all past observations.

## 5 Non-Bayesian Learning

In this Section we study two adaptive, non-Bayesian learning process that have been analyzed in the classical one-object search framework by Bickchandani and Sharma [4], and by Chou and Talmain [5], respectively. Both processes are consistent in the sense that they uniformly converge to the true distribution as the number of observations goes to infinity<sup>5</sup>.

In both cases, we prove that the efficient allocation is always implementable<sup>6</sup>. A word of caution is needed here: Our results **do not** imply

<sup>5</sup>In both cases, this is a consequence of the well known Glivenko-Cantelli Theorem

<sup>6</sup>As in the case of Bayesian learning, the efficient allocation maximizes at each decision period the sum of the expected utilities of all agents, given all the available information. The only difference to the Bayesian approach is in the inference made from new information.

that the considered non-Bayesian procedures are "better" than Bayesian updating for the purposes of efficient implementation! They just say that the complete information efficient allocation - whose calculation proceeds **given** an assumed learning procedure - can always be implemented for the particular adaptive processes studied here. Of course, if full rationality is taken to include Bayesian learning, the considered procedures implement the "wrong" allocation to start with.

## 5.1 Learning Based on the Empirical Distribution

Assume that before stage  $n$  (the first stage), the designer's prior belief about the distribution of the first type  $x_n$  is given by a given distribution  $H$ . Then, conditional on sequentially observing  $x_n, x_{n-1}, \dots, x_{k+1}$  at stages  $n, n-1, \dots, k+1$ , the designer's belief about the distribution of the next type  $x = x_k$  is given by:

$$\tilde{F}_k(x|x_n, \dots, x_{k+1}) = (1 - \beta_k^n)H(x) + \beta_k^n \frac{1}{n - k} \sum_{i=k+1}^n \mathbf{1}_{[x_i, \infty)}(x), \quad k = 1, 2, \dots, n-1$$

where  $0 < \beta_k^n < 1$ , and where  $\mathbf{1}_{[z, \infty)}(x)$  denotes the indicator function of the set  $[z, \infty)$ . Thus at each stage, the posterior distribution is given by a convex combination of the prior distribution and of the empirical distribution. Since, by the Glivenko-Cantelli theorem, the empirical distribution uniformly converges to the true underlying distribution, the posterior distribution also converges to the true distribution if the weight on the empirical distribution satisfies:  $\forall k, \lim_{n \rightarrow \infty} \beta_k^n = 1$ .

For the proof of our main result in this Subsection we need the following well-known Lemma:

**Lemma 10** *Let  $u(x)$  be a function on the interval  $[a, b]$  such that there exist a division of the interval  $a = z_0 < z_1 < \dots < z_n = b$  and values  $c_1, \dots, c_n$  with  $u(x) = c_i$  for  $z_i < x < z_{i+1}$ ,  $i = 0, 1, \dots, n-1$ . Then, for any continuous function  $v(x)$  on  $[a, b]$ , it holds that*

$$\int_a^b v(x) du(x) = \sum_{i=0}^{n-1} v(z_i)(c_{i+1} - c_i)$$

where  $\int$  denotes here the Stieltjes integral.

**Theorem 11** *Assume that the designer learns based on the empirical distribution. Then, the efficient dynamic policy can always be implemented under incomplete information.*

**Proof.** Analogously to our proof of Theorem 8 we will show that for any  $i$ ,  $k$ ,  $\chi_k$  and  $x_k$ , the cutoff  $a_{i,k}(\chi_k, x_k)$  is continuously differentiable in  $x_k$ , and that  $\frac{\partial}{\partial x_k} a_{i,k}(\chi_k, x_k) \leq 1$ . Let  $\mathbf{m}x = (x, x, \dots, x)$  denote an  $m$ -vector of  $x$ .

We first show by induction that  $\forall m$ ,  $m \leq n - k + 1$ , the function  $a_{i,k}(x_n, \dots, x_{k+m}, \mathbf{m}x)$  is continuously differentiable in the observed signals and that

$$\forall i, k, \quad \frac{\partial a_{i,k}(x_n, \dots, x_{k+m}, \mathbf{m}x)}{\partial x} < 1.$$

Since the conditional distribution  $\tilde{F}_k(x|x_n, \dots, x_{k+1})$  does not have a well-defined density, we use below the notion of *Stieltjes integral*. In the last but one period  $k = 2$ , the only relevant, non-trivial cutoff is:

$$\begin{aligned} a_{1,2}(x_n, \dots, x_2) &= \int_0^\infty x_1 d\tilde{F}_1(x_1|x_n, \dots, x_2) \\ &, = (1 - \beta_2^n) \int_0^\infty x_1 dH(x_1) + \beta_2^n \int_0^\infty x_1 d\left(\sum_{i=2}^n 1_{[x_i, \infty)}(x_1)\right) \\ &= (1 - \beta_2^n)E(H) + \beta_2^n \frac{1}{n-1} \sum_{i=2}^n x_i \end{aligned}$$

The second equality follows by the additivity property of the Stieltjes integral. The third equality follows by Lemma 10 since  $\sum_{i=2}^n 1_{[x_i, \infty)}(x)$  is a step function. Thus, as required, we obtain that  $a_{1,2}(x_n, \dots, x_2)$  is continuously differentiable and that

$$\frac{\partial a_{1,2}(x_n, \dots, x_{2+m}, \mathbf{m}x)}{\partial x} \leq \frac{m\beta_2^n}{n-1} < 1, \quad m = 1, 2, \dots, n-1$$

Assume now that the statement holds for all periods up to  $k$  (recall that period 1 is the last period, and so on...) and let us look at period  $k+1$ , and at  $m \leq n - k$ . Recalling the definition of the function  $G_{i,k}(\chi_{k+1}, x_{k+1}, x_k)$  in

equation ??, we obtain that

$$\begin{aligned}
& a_{i,k+1}(x_n, \dots, x_{k+m+1}, \mathbf{m}\mathbf{x}) \\
&= \int_0^\infty G_{i,k}(x_n, \dots, x_{k+1+m}, \mathbf{m}\mathbf{x}, x_k) d\tilde{F}_k(x_k | (x_n, \dots, x_{k+1+m}, \mathbf{m}\mathbf{x})) \\
&= (1 - \beta_{k+1}^n) \int_0^\infty G_{i,k}(x_n, \dots, x_{k+1+m}, \mathbf{m}\mathbf{x}, x_k) dH(x_k) \\
&\quad + \frac{m\beta_{k+1}^n}{n-k} [G_{i,k}(x_n, \dots, x_{k+m}, (\mathbf{m} + \mathbf{1})\mathbf{x})] \\
&\quad + \frac{\beta_{k+1}^n}{n-k} \sum_{j=k+m}^n G_{i,k}(x_n, \dots, x_{k+m}, \mathbf{m}\mathbf{x}, x_j)
\end{aligned}$$

where the second equality follows from Lemma 10. Continuous differentiability of  $a_{i,k+1}(x_n, \dots, x_{k+m+1}, \mathbf{m}\mathbf{x})$  follows here from the same argument as in Theorem 8. Hence, for any  $m \leq n - k$ , we obtain that:

$$\begin{aligned}
& \frac{\partial a_{i,k+1}((x_n, \dots, x_{k+m+1}, \mathbf{m}\mathbf{x}))}{\partial x} \\
&= (1 - \beta_{k+1}^n) \int_0^\infty \frac{\partial G_{i,k}(x_n, \dots, x_{n-k-m}, \mathbf{m}\mathbf{x}, x_k)}{\partial x} dH(x_k) \\
&\quad + \frac{m\beta_{k+1}^n}{n-k} \frac{\partial G_{i,k}(x_n, \dots, x_{k+1+m}, (\mathbf{m} + \mathbf{1})\mathbf{x})}{\partial x} \\
&\quad + \frac{\beta_{k+1}^n}{n-k} \sum_{j=k+m}^n \frac{\partial G_{i,k}(x_n, \dots, x_{k+m}, \mathbf{m}\mathbf{x}, x_j)}{\partial x} \\
&< (1 - \beta_{k+1}^n) + \beta_{k+1}^n \left( \frac{m}{n-k} + \frac{n-k-m}{n-k} \right) = 1
\end{aligned}$$

where the inequality follows by the induction hypothesis. By setting  $m = 1$ , we obtain from the above that:

$$\forall i, k \quad \frac{\partial a_{i,k}(x_n, \dots, x_{k+1})}{\partial x_{k+1}} < 1$$

which, together with Proposition 5, proves the result. ■

Recalling the word of caution at the beginning of the Section, it is illustrative to compare Bayesian and non-Bayesian learning in a simple example where the dynamically efficient allocation is **not** implementable under Bayesian learning:

**Example 12** *There are two periods and one indivisible object. Before starting the allocation process, the designer believes that the distribution of values*

is uniform on the interval  $[0, 1]$  with probability 0.5, while with probability 0.5 he believes that it is uniform on  $[1, 2]$ . Under Bayesian learning, the posterior after observing  $x_2 < (>)1$ , is that  $x_1$  is uniformly distributed on  $[0, 1]$  ( $[1, 2]$ ). This yields

$$a_{12}^B(x_2) = \begin{cases} 0.5 & \text{if } x_2 < 1 \\ 1 & \text{if } x_2 = 1 \\ 1.5 & \text{if } x_2 > 1 \end{cases}$$

Thus, the first arriving agent should efficiently get the object if  $x_2 \in [0.5, 1] \cup [1.5, 2]$ . This non-convex allocation policy cannot be implemented (see GM [8])

Consider now the adaptive learning process with weight  $0 < \beta < 1$  on the empirical distribution. Then, after having observed  $x_2$ , the beliefs of the designer are given by  $F(x_1|x_2) = (1 - \beta)U([0, 2]) + \beta\mathbf{1}_{[x_2, 2]}$ , which yields

$$a_{12}^A(x_2) = (1 - \beta) + \beta x_2.$$

Thus, the first arriving agent should get the object if and only if  $x_2 \geq a_{12}^A(x_2) = (1 - \beta) + \beta x_2 \Leftrightarrow x_2 \geq 1$ , which can be implemented by a take-it-or-leave-it offer at a price of 1. Note how the implemented allocation differs here from the one that needs to be implemented under Bayesian learning.

For special prior distributions, the process studied above does in fact coincide with the standard Bayesian learning. This is the case, for example, for a multinomial Dirichlet prior or for a Dirichlet process prior. Thus, for such priors, Theorem 11 asserts the implementability of the efficient dynamic allocation under Bayesian learning.

Bickchandani and Sharma [4] showed that the above learning model induces optimal search with the reservation price property in Rothschild's model. As shown above, this insight continues to hold unchanged for the case with several objects.

## 5.2 Maximum Entropy/Quantile Preserving Learning

For the current purpose we only assume that designer believes that types distribute continuously on a finite interval, which we normalize here to be the interval  $[0, 1]$ . Recall first that the *maximum entropy* distribution among all continuous distributions with support on an interval  $[a, b]$  is the uniform distribution on this interval. More generally, consider a sub-division  $a = a_0 < a_1 < \dots < a_m = b$  and probabilities  $p_1, \dots, p_m$  which add up to one, and consider the class of all continuous distributions supported on  $[a, b]$  such that

$$\Pr\{a_{i-1} \leq X \leq a_i\} = p_i, \quad i = 1, \dots, m$$

Then, the density of the maximum entropy distribution for this class is constant on each of the intervals  $[a_{j-1}, a_j)$ . Guided by this principle, Chou and Talmain [5] looked at the following *quantile preserving* updating procedure<sup>7</sup>: Prior to any observation, the designer estimates the unknown distribution by the uniform distribution. Suppose that  $m$  observations were observed, and order them in increasing order  $\{x_{(1)}, \dots, x_{(m)}\}$ . Let  $x_{(0)} = 0$  and  $x_{(m+1)} = 1$ . Then, the type of the next arrival is estimated according to the density

$$f_k(x|x_n, \dots, x_{n-m+1}) = \sum_{i=1}^{m+1} \frac{\mathbf{1}_{[x_{(i-1)}, x_{(i)})}(x)}{(m+1)(x_{(i)} - x_{(i-1)})}.$$

In other words, each interval of the form  $[x_{(i-1)}, x_{(i)})$  gets assigned a probability  $p_i = \frac{1}{m+1}$ , and the density within the interval is constant. The rationale behind the equal weights of  $\frac{1}{m+1}$  for each interval becomes apparent by recalling that, for  $m$  large,

$$E[X_{i,m}] \approx F^{-1}\left(\frac{i}{m+1}\right) \text{ and } F(E[X_{i,m}]) - F(E[X_{i-1,m}]) \approx \frac{1}{m+1}$$

where the  $X_{i,m}$  is the  $i$ -th highest order statistic,  $i = 1, \dots, m$ , of a random variable  $X$  distributed according to distribution  $F$ . As above, the Glivenko-Cantelli theorem implies that the above estimated distribution uniformly converges to the true distribution. Our last result shows that the efficient allocation associated with this estimation procedure is always implementable.

**Theorem 13** *Assume that the designer uses the maximum entropy/quantile preserving learning procedure. Then the efficient dynamic policy can always be implemented under incomplete information.*

**Proof.** Analogously to our proof of Theorem 8 we will show that for any  $i, k, \chi_k$  and  $x_k$ , the cutoff  $a_{i,k}(\chi_k, x_k)$  is continuously differentiable in  $x_k$ , and that  $\frac{\partial}{\partial x_k} a_{i,k}(\chi_k, x_k) \leq \frac{1}{n-k+2}$ . We prove this result by induction on  $k$ , the number of remaining periods. Note first that Lemma 3 yields the monotonicity of  $a_{i,k+1}(\chi_{k+1}, x_{k+1})$  in  $x_{k+1}$ .

We denote by  $x_{(i)}$ , the  $i$ -th lowest observation among the  $n - k + 1$  observations made up to an including period  $k$ , with  $x_{(0)} = 0$  and  $x_{(n-k+2)} =$

---

<sup>7</sup>They studied search with recall and did not look at the reservation price property for search without recall.

1. For  $k = 2$ , we have:

$$\begin{aligned}
a_{1,2}(x_n, \dots, x_2) &= \sum_{i=1}^n \int_{x_{(i-1)}}^{x_{(i)}} \frac{x}{n(x_{(i)} - x_{(i-1)})} dx \\
&= \frac{1 + 2 \sum_{i=1}^{n-1} x_{(i)}}{2n} = \frac{1 + 2 \sum_{i=2}^n x_i}{2n} \Rightarrow \\
\frac{\partial a_{1,2}(x_n, \dots, x_2)}{\partial x_2} &= \frac{1}{n}.
\end{aligned}$$

Assume now that the statement holds for all periods up to  $k$ . This implies that there exists at most one solution to the equation  $a_{i,k}(\chi_k, x_k) = x_k$ , denoted by  $a_{i,k}^*(\chi_k)$ . Let  $l = \max \{j : x_{(j)} \leq a_{i-1,k}^*(\chi_k)\}$  and  $m = \max \{j : x_{(j)} \leq a_{i,k}^*(\chi_k)\}$ , and assume, for simplicity, that  $m > l$  (the case  $m = l$  is analogous). Using the definition of  $a_{i,k+1}(\chi_{k+1}, x_{k+1})$  we obtain

$$\begin{aligned}
&a_{i,k+1}(\chi_{k+1}, x_{k+1}) \\
&= \sum_{j=1}^l \frac{\int_{x_{(j-1)}}^{x_{(j)}} a_{i-1,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})} + \frac{\int_{x_{(l)}}^{a_{i-1,k}^*(\chi_k)} a_{i-1,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(l+1)} - x_{(l)})} \\
&\quad + \frac{\int_{x_{(l+1)}}^{x_{(l+1)}} x_k dx_k}{(n-k+1)(x_{(l+1)} - x_{(l)})} + \sum_{j=l+2}^m \frac{\int_{x_{(j-1)}}^{x_{(j)}} x_k dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})} \\
&\quad + \frac{\int_{x_{(m)}}^{a_{i,k}^*(\chi_k)} x_k dx_k}{(n-k+1)(x_{(m+1)} - x_{(m)})} + \frac{\int_{a_{i,k}^*(\chi_k)}^{x_{(m+1)}} a_{i,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(m+1)} - x_{(m)})} \\
&\quad + \sum_{j=m+1}^{n-k+1} \frac{\int_{x_{(j-1)}}^{x_{(j)}} a_{i,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})}.
\end{aligned}$$

Let  $j$  the index satisfying  $x_{(j)} = x_{k+1}$ . There are three different cases: 1.  $x_{k+1} \leq x_{(l)}$ ; 2.  $x_{(m)} \geq x_{k+1} > x_{(l)}$ ; 3.  $x_{k+1} > x_{(m)}$ . We prove the result for the first case; the proofs of the other two cases are very similar, and we omit

them here. We obtain:

$$\begin{aligned}
& \frac{\partial a_{i,k+1}(\chi_{k+1}, x_{k+1})}{\partial x_{k+1}} \\
= & \frac{a_{i-1,k}(\chi_k, x_{(j)})}{(n-k+1)(x_{(j)} - x_{(j-1)})} - \frac{a_{i-1,k}(\chi_k, x_{(j)})}{(n-k+1)(x_{(j+1)} - x_{(j)})} \\
& + \sum_{j=1}^l \frac{\int_{x_{(j-1)}}^{x_{(j)}} \frac{\partial a_{i-1,k}(\chi_k, x_k)}{\partial x_{k+1}} dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})} + \frac{a_{i-1,k}^*(\chi_k)}{x_{(l)}} \frac{\int_{x_{(l)}} \frac{\partial a_{i-1,k}(\chi_k, x_k)}{\partial x_{k+1}} dx_k}{(n-k+1)(x_{(l+1)} - x_{(l)})} \\
& + \frac{a_{i,k}^*(\chi_k)}{(n-k+1)(x_{(m+1)} - x_{(m)})} + \sum_{j=m+1}^{n-k+1} \frac{\int_{x_{(j-1)}}^{x_{(j)}} \frac{\partial a_{i,k}(\chi_k, x_k)}{\partial x_{k+1}} dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})} \\
& - \frac{\int_{x_{(j-1)}}^{x_{(j)}} a_{i-1,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})^2} + \frac{\int_{x_{(j)}}^{x_{(j+1)}} a_{i-1,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(j+1)} - x_{(j)})^2}
\end{aligned}$$

Note that

$$\begin{aligned}
& \sum_{j=1}^l \frac{\int_{x_{(j-1)}}^{x_{(j)}} \frac{\partial a_{i-1,k}(\chi_k, x_k)}{\partial x_{k+1}} dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})} + \frac{a_{i-1,k}^*(\chi_k)}{x_{(l)}} \frac{\int_{x_{(l)}} \frac{\partial a_{i-1,k}(\chi_k, x_k)}{\partial x_{k+1}} dx_k}{(n-k+1)(x_{(l+1)} - x_{(l)})} \quad (16) \\
& + \frac{a_{i,k}^*(\chi_k)}{(n-k+1)(x_{(m+1)} - x_{(m)})} + \sum_{j=m+1}^{n-k+1} \frac{\int_{x_{(j-1)}}^{x_{(j)}} \frac{\partial a_{i,k}(\chi_k, x_k)}{\partial x_{k+1}} dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})} \\
\leq & \frac{1}{(n-k+1)} \frac{n-k-m+l+1}{n-k+2} \leq \frac{1}{(n-k+1)} \frac{n-k}{n-k+2}
\end{aligned}$$

where the first inequality follows from the inductive assumption (  $\frac{\partial a_{i-1,k}(\chi_k, x_k)}{\partial x_{k+1}} \leq \frac{1}{n-k+2}$  ) while the second inequality follows because  $m > l$ . In addition,

$$\begin{aligned}
& \frac{a_{i-1,k}(\chi_k, x_{(j)})}{(n-k+1)(x_{(j)} - x_{(j-1)})} - \frac{a_{i-1,k}(\chi_k, x_{(j)})}{(n-k+1)(x_{(j+1)} - x_{(j)})} \\
& - \frac{\int_{x_{(j-1)}}^{x_{(j)}} a_{i-1,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(j)} - x_{(j-1)})^2} + \frac{\int_{x_{(j)}}^{x_{(j+1)}} a_{i-1,k}(\chi_k, x_k) dx_k}{(n-k+1)(x_{(j+1)} - x_{(j)})^2} \\
\leq & \frac{a_{i-1,k}(\chi_k, x_{(j)})}{(n-k+1)(x_{(j)} - x_{(j-1)})} - \frac{a_{i-1,k}(\chi_k, x_{(j)})}{(n-k+1)(x_{(j+1)} - x_{(j)})} \\
& - \frac{a_{i-1,k}(\chi_k, x_{(j-1)})}{(n-k+1)(x_{(j)} - x_{(j-1)})} + \frac{a_{i-1,k}(\chi_k, x_{(j+1)})}{(n-k+1)(x_{(j+1)} - x_{(j)})} \\
= & \frac{a_{i-1,k}(\chi_k, x_{(j)}) - a_{i-1,k}(\chi_k, x_{(j-1)})}{(n-k+1)(x_{(j)} - x_{(j-1)})} + \frac{a_{i-1,k}(\chi_k, x_{(j+1)}) - a_{i-1,k}(\chi_k, x_{(j)})}{(n-k+1)(x_{(j+1)} - x_{(j)})} \\
= & \frac{1}{n-k+1} \left[ \frac{\partial}{\partial x_{k+1}} a_{i-1,k}(\chi_k, x'_k) + \frac{\partial}{\partial x_{k+1}} a_{i-1,k}(\chi_k, x''_k) \right] \\
\leq & \frac{1}{n-k+1} \frac{2}{n-k+2}
\end{aligned} \tag{17}$$

where  $x'_k \in [x_{(j-1)}, x_{(j)}]$  and  $x''_k \in [x_{(j)}, x_{(j+1)}]$ . The first inequality follows from the monotonicity of  $a_{i-1,k}(\chi_k, x_k)$ , and the last inequality follows from the induction argument. Combining (16) and (17) we obtain

$$\frac{\partial a_{i,k+1}(\chi_{k+1}, x_{k+1})}{\partial x_{k+1}} \leq \frac{1}{n-k+1}.$$

as desired. ■

## 6 Conclusion

We have derived conditions on the primitives of the learning environment that allow efficient dynamic implementation. The analysis has used insights from static mechanism design with interdependent values, and has revealed close connections to the problem of ensuring that optimal search policies display a reservation price property.

In contrast to our focus on dynamic welfare maximization, there is an extensive literature on dynamic revenue maximization in the field of yield or revenue management (see the book of Talluri and Van Ryzin [19]). Roughly

speaking, this literature considers intuitive pricing schemes, and does not focus on implementation issues (since in most considered settings this is not an issue). As soon as learning about the environment takes place simultaneously with allocation decisions, one has to be more careful: not all ad-hoc pricing schemes will be generally implementable, and the revenue maximization exercise must take this fact into account, as first illustrated by Riley and Zeckhauser [17] in their "haggling" model. Their argument can be adapted to our own framework in order to show that the second-best optimal policy (in situations where the first best cannot be implemented) is also deterministic, and hence has the form of cutoffs.

## References

- [1] Albright, S.C. (1977): "A Bayesian Approach to a Generalized House Selling Problem." *Management Science*, **24**(4), 432-440.
- [2] Athey, S. and Segal, I. (2007): "An Efficient Dynamic Mechanism." discussion paper, Stanford University.
- [3] Bergemann, D. and Välimäki, J. 2007: "Dynamic Marginal Contribution Mechanisms." Cowles Foundation Discussion Paper 1616.
- [4] Bikhchandani, S. and S. Sharma, (1996): "Optimal Search with Learning." *Journal of Economic Dynamics and Control* **20**, 333-359.
- [5] Chou, C.F, and Talmain, G. (1993): "Nonparametric Search." *Journal of Economic Dynamics and Control*, **17**: 771-784.
- [6] Dasgupta, P. and Maskin, E. (2000): "Efficient Auctions." *Quarterly Journal of Economics*, **115**(2): 341-388.
- [7] Derman, C., Lieberman, G, and Ross, S. (1972): "A Sequential Stochastic Assignment Problem." *Management Science*, **18**(7): 349-355.
- [8] Gershkov, A. and Moldovanu, B. (2009): "Learning About the Future and Dynamic Efficiency." forthcoming *American Economic Review*.
- [9] Jehiel, P. and Moldovanu, B. (2001): "Efficient Design with Interdependent Valuations." *Econometrica*, **69**(5): 1237-1259.
- [10] Kittsteiner, T. and Moldovanu, B. (2005): "Priority Auctions and Queue Disciplines that Depend on Processing Time." *Management Science*, **51**(2), 236-248.

- [11] Lippman, S.A. and McCall, J. (1981): "The Economics of Uncertainty: Selected Topics and Probabilistic Methods." in *Handbook of Mathematical Economics*, Ed. K.J. Arrow and M.D. Intriligator, Volume I, Elsevier, 211-284.
- [12] Milgrom, P. (1981): "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* **12**(2), 280-391.
- [13] Morgan, P.B. (1985): "Distributions of the Duration and Value of Job Search with Learning." *Econometrica* **53**, 1199-1232.
- [14] Parkes, D.C., and Singh, S. (2003): "An MDP-Based Approach to Online Mechanism Design." *Proceedings of 17th Annual Conference on Neural Information Processing Systems (NIPS 03)*
- [15] Rosenfield, D. and Shapiro, R. (1981): "Optimal Adaptive Price Search." *Journal of Economic Theory*, **25**, 1-20.
- [16] Rothschild, M. (1974): "Searching for the Lowest Price When the Distribution of Prices Is Unknown." *Journal of Political Economy*, **82**(4), 689-711.
- [17] Riley, J. and Zeckhauser (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm." *Quarterly Journal of Economics*, **98** (2), 267-289.
- [18] Seierstad, A. (1992): "Reservation Prices in Optimal Stopping." *Operations Research*, **40**, 409-414
- [19] Talluri, K. and van Ryzin, G. (2004): *The Theory and Practice of Revenue Management*, Springer: New York.