Optimal Search, Learning and Implementation*

Alex Gershkov and Benny Moldovanu †

November 28, 2011

Abstract

We characterize the incentive compatible, constrained efficient policy ("second-best") in a dynamic matching environment, where impatient, privately informed agents arrive over time, and where the designer gradually learns about the distribution of agents' values. We also derive conditions on the learning process ensuring that the complete-information, dynamically efficient allocation of resources ("first-best") is incentive compatible. Our analysis reveals and exploits close, formal relations between the problem of ensuring implementable allocation rules in our dynamic allocation problems with incomplete information and learning, and between the classical problem, posed by Rothschild *[21],* of finding optimal stopping policies for search that are characterized by a *reservation price property.*

Keywards: dynamic mechanism design, learning, optimal stopping.

JEL: D61, D78, D83.

*The comments of an associate editor and an anonymous referee greatly improved the quality of the exposition. We wish to thank Sergiu Hart, Philippe Jehiel, Alessandro Pavan, Xianwen Shi, Phil Reny and Asher Wolinsky for helpful remarks. Participants at the workshop "Information and Dynamic Mechanism Design" June 2009, Bonn, made very fruitful comments. We are grateful to the German Science Foundation and to the European Research Council for financial support.

†Gershkov: Department of Economics and Center for the Study of Rationality, The Hebrew University of Jerusalem, Mt. Scopus, 91905 Jerusalem, Israel. Tel. 97225883252, Email: alexg@huji.ac.il; Moldovanu: Department of Economics, University of Bonn, Lennestr. 37, 53113 Bonn, Germany, Tel. 49228739242, Email: mold@uni-bonn.de.

1 Introduction

We characterize the incentive compatible, constrained efficient policy ("second-best") in a dynamic allocation environment with several heterogenous objects, where impatient, privately informed agents arrive over time, and where the designer gradually learns about the distribution of agents' values. The second-best policy is characterized by a suitable generalization to heterogenous objects of the so-called *reservation price property*, first discussed in Rothschild's [21] classical paper. We also derive conditions on the learning process ensuring that the complete-information, dynamically efficient allocation of resources ("first-best") is incentive compatible, in which case it also displays a reservation price property.

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 [9] (DLR hereafter). In the DLR model, a finite set of possibly heterogenous, commonly ranked objects needs to be assigned to a set of heterogeneous 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. In the static counterpart of this problem all agents are present at the same point in time, and the optimal matching is assortative: the agent with the highest type should get the object with the highest quality, and so on (see Becker [4])[1].

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]. Gershkov and Moldovanu [10] (GM) added incomplete information to Albright's learning model and

[1] In another classical static model, Mussa and Rosen studied screening under incomplete information in a framework where agents need to be matched to commonly ranked qualities.

showed, via an example, 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 statistical dependence between the types of different agents is ruled out (see for example Parkes and Singh, [19], and Bergemann and Valimäki [5]). In this paper, the types of early agents become, via the learning process, informative about the types of later agents and independence fails.

If all payments can be delayed until a time in the future when no new arrivals occur, the efficient allocation can always be implemented since payments can be then conditioned on the actual allocation in each instance[2]. But such uncoupling of the physical and monetary parts is not always realistic in applications, and we will abstract from it here [3].

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 informational externalities[4]. Segal [23] analyzed revenue maximization in a static environment with an unknown distribution of the agents' values, and also observed that agents have an informational effect on others. But, the type of problems highlighted in our present paper do not occur in Segal's static model since a standard VCG mechanism always leads there to the efficient outcome.

In our model, a necessary condition for extracting truthful information about values is the monotonicity of the (possible random) allocation rule, i.e., agents with higher values should not be worse-off than contemporaneous agents with lower values. 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 revealed information induces it- leading to

[2] See also Athey and Segal [2].

[3] Our main results can also be seen as a measure of the cost of having online payments.

[4] Dasgupta and Maskin [8] and Jehiel and Moldovanu [11] have analyzed efficient implementation in static models with direct informational externalities. Kittsteiner and Moldovanu [12] used these insights in a dynamic model with direct externalities and without learning.

a failure of truthful revelation- if the designer decides in response to deny present resources in order to keep them for the "sunnier" future. GM [10] derived an implicit condition on the structure of the allocation policy (and thus on **endogenous** variables) ensuring that efficient implementation is possible. They showed that monotonicity holds if the impact of currently revealed information on today's values is higher than the impact on option values. This observation translates to the dynamic framework with learning the single-crossing idea appearing in the theory of static efficient implementation with interdependent values. But, the resulting set of conditions was unsatisfactory since it is not formulated in terms of the primitives of the learning model.

Two natural research directions are suggested by the above insights:

1. Since the complete information, dynamically efficient policy is likely to be implementable only under restrictive conditions, it is of interest to characterize the optimal policy respecting the incentive constraints (second-best). We are able to offer here a complete characterization by using several concepts that were developed in the context of *majorization* theory. The crucial insight is that the second-best policy is deterministic, i.e. it allocates to each type of agent, at each point in time, a well defined available quality instead of a lottery over several feasible qualities.

2. It is of interest to derive direct conditions - that can be checked in applications - on the **exogenous** parameters of the allocation cum learning environment that allow the implementation of the first best (e.g., conditions on the initial beliefs about the environment and on the learning protocol). We offer here two such sets of sufficient conditions ensuring that the first-best is indeed implementable.

Our analysis of the above questions 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*. In particular, and letting aside for a while the mechanism design/dynamic efficiency interpretation, our results about the above second question can also be seen as offering conditions ensuring that the optimal search policy without recall for highest prices for several (possibly heterogenous) objects exhibits the relevant generalization of the reservation price property[5].

It is important to note that in the classical search model, price quotations are non-strategic, and the monotonicity requirement behind the reservation price property is only a convenient, intuitive feature, facilitating the use 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 particular, our characterization of the second-best mechanism (question 1 above) has no counterpart in the classical search literature.

The paper is organized as follows: In Section 2 we present the dynamic allocation and learning model and we recall a result, due to Albright [1], that characterizes the efficient dynamic allocation policy under complete information about the arriving agents' values. In Section 3 we add incomplete information about the agents' value (while keeping the assumption that the designer gradually learns about the distribution of values). We first characterize incentive compatible allocations in terms of a monotonicity property. An example shows that the efficient allocation, as described by Albright, need not be implementable. We then characterize the incentive-compatible dynamic policy that maximizes expected welfare while respecting incentive compatibility (second best), using mathematical ideas from *majorization theory.* In particular, we show that the second-best policy is always deterministic, and that it satisfies a generalized form of a reservation price property appearing in classical search models. In Section 4 we offer two sets of sufficient conditions under which the second-best policy coincides with the first-best (in other words, we offer conditions under which the complete information dynamically efficient policy characterized by Albright is incentive compatible). A 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 puts a 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 4, the bound becomes tighter in

[5] Our results can also be easily adapted for the version where a buyer who wants to buy several units searches for the lowest prices offered by sellers who each possess one unit.

early search stages, while in the second the bound becomes tighter in later periods. 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. Section 5 concludes. All proofs are relegated to an Appendix.

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 $0 \leq q_1 \leq q_2 ... \leq q_m$ 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.

We assume that there are one or more unknown parameters of the distribution F from which agents' types are sampled. The beliefs about these parameters are originally given by a prior distribution which is then sequentially updated via Bayes' rule as additional information is observed. Denote by Φ_n the designer's prior over possible distribution functions, and by $\Phi_k(x_n, ..., x_{k+1})$ his beliefs about the distribution function F after observing types $x_n, ..., x_{k+1}$. Given such beliefs, let $\widetilde{F}_k(x|x_n, ..., x_{k+1})$ denote the distribution of the next type x_k, conditional on observing $x_n, ..., x_{k+1}$, while

$\widetilde{f}_k(x|x_n, ..., x_{k+1})$ denote the corresponding density. We assume that the distribution $\widetilde{F}_k(x|x_n, ..., x_{k+1})$ is symmetric with respect to observed signals - a feature satisfied by the standard Bayesian learning model used here.

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

We start by characterizing the dynamically efficient allocation under **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 $n, ..., k+1$, and of allocations to those agents[6]. Let $\mathcal{H}_k$ be the set of all histories at period k. Denote by χ_k the ordered set of signals reported by the agents that arrived at periods $n, ..., k+1$. Finally, denote by Π_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 Π_n and a history H_k completely determine the set Π_k.

The result below characterizes, at each period, the dynamically efficient policy in terms of cutoffs 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 (see Becker [4]).

Theorem 1 *(Albright, 1977)*

1. *Assume that types $x_n, .., x_{k+1}$ have been observed, and consider the arrival of an agent with type x_k in period $k \geq 1$. There exist functions $0 = a_{0,k}(\chi_k, x_k) \leq a_{1,k}(\chi_k, x_k) \leq a_{2,k}(\chi_k, x_k) ... \leq a_{k,k}(\chi_k, x_k) = \infty$ such that the efficient dynamic policy - which maximizes the expected value of the total reward - assigns the item with the $i-th$ smallest type*

[6] Since we allow for random mechanisms, the history needs to include the results of the previous randomizations. But, a mechanism that depends on these can be replicated by another mechanism that only depends on the result of the current randomization. For notational simplicity we shall therefore exclude the results of the previous randomizations from the specification of histories.

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. *These functions are related to each other by the following recursive formulae:*

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

where[7]

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

These cutoffs have very natural interpretation: for each object i and period k the cutoff $a_{i,k}(\chi_k, x_k)$ equals the expected value of the agent's type to which the item with $i - th$ smallest type is assigned in a problem with $k - 1$ periods before the period $k - 1$ signal is observed.

3 The Incentive Efficient (Second-Best) Policy

We now focus on the additional constraints imposed by incentive compatibility in the model with **incomplete information and learning.** We therefore assume below that the agents' types are private information. Without loss of generality, we can restrict attention to direct mechanisms where every agent, upon arrival, reports his type and where the mechanism specifies which item the agent gets (if any), and a payment.[8]

[7] We set $+\infty \cdot 0 = -\infty \cdot 0 = 0$.

[8] Since agents observe the history, they are better informed after directly observing types. Yet, the argument holds because: (i) if a policy is implementable by a general mechanism then, with private values, it is also implementable via an augmented mechanism where, in addition, agents report types, but the designer does not use this information; (ii) this augmented mechanism can be replicated by a direct one.

In this Section we characterize the incentive compatible, optimal solution (second best). The second-best allocation - that maximizes expect welfare under the incentive constraints - turns out to be deterministic: it uses cutoffs that at each period partition the set of types into disjoint intervals associated with available qualities such that higher types obtain a higher quality. This is the appropriate generalization of the reservation price property.

We first need to characterize incentive compatible allocations in our model: an allocation policy (which may be random) is implementable under incomplete information if and only if, in each period and for every history of events at preceding periods, the expected quality allocated to the current agent is non-decreasing in the agent's reported type.

Proposition 1 *For a fixed allocation policy, denote by $Q_k(H_k, x)$ the expected quality allocated to an agent arriving at period k after history H_k, and reporting signal x. An allocation policy is implementable if and only if for any k and for any H_k the expected quality $Q_k(H_k, x)$ is non-decreasing in x.*

Proof. See Appendix A. ■

3.1 An Illustration

In this subsection we illustrate some of our insights in a simple example where the designer is uncertain about a shape parameter affecting the distribution of values.

Example 1 *There is one object of quality $q = 1$, and two periods (this corresponds to the general case of $q_2 = 1$ and $q_1 = 0$). The agents' valuations distribute on $[0, 1]$: with probability α the distribution is $F(x) = x$, while with probability $1 - \alpha$ the distribution is $F(x) = x^\theta$ where $\theta > 0$.*

After observing the type of the agent who arrives at period 2 (the first period !) the designer's belief about the next period's type given the current observation is given by

$$\widetilde{f}_1(x|x_2) = \frac{\alpha + (1-\alpha)\theta^2 x_2^{\theta-1} x^{\theta-1}}{\alpha + (1-\alpha)\theta x_2^{\theta-1}}.$$

Therefore, the expected type at the next period is given by:

$$E[x|x_2] = \frac{\frac{\alpha}{2} + (1-\alpha)\frac{\theta^2}{\theta+1} x_2^{\theta-1}}{\alpha + (1-\alpha)\theta x_2^{\theta-1}}.$$

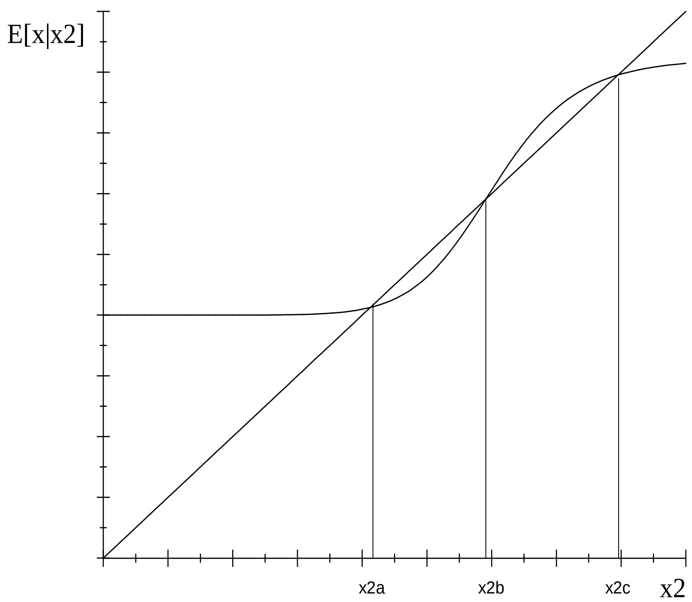

Figure 1: The efficient allocation is not monotone.

The figure below shows the function $E[x|x_2]$ *for values* $\alpha = 0.2$ *and* $\theta = 12$. *The designer would like to allocate the object in period 2 if and only if his value exceeds the expected value of the agent that arrives at the next period. Therefore, the designer wants to allocate the object at period 2 if and only if* $x_2 \in (x_{2a}, x_{2b}] \cup (x_{2c}, 1]$ *where* $x_{2a} \approx 0.513$, $x_{2b} \approx 0.691$, *and* $x_{2c} \approx 0.895$.*The efficient (first-best) allocation is not monotone, and hence, by Proposition 1, it is not implementable. The reason for this non-monotonicity is the following: when* x_2 *is relatively low, the designer is almost certain that the distribution is uniform (since the density of the second distribution is very low at the bottom of the type distribution); but, at some point, the density of the second distribution becomes very steep and, for low* α, *this causes the designer to adjust his beliefs also very steeply: he becomes more and more optimistic about the future, which is detrimental to a current agent with relatively high types.*

We shall show below that the second-best mechanism - which needs to induce a monotonic allocation - is characterized by a fixed cutoff above which the object is allocated in the current period, and below which the object is car-

ried to the next period.[9] *When choosing this cutoff, the goal of the designer is to minimize the disutility from misallocating the object relative to the efficient allocation. For example, if the designer chooses a cutoff* $b \in (x_{2a}, x_{2b})$, *misallocation occurs if the signal of the first agent if* $x_2 \in (x_{2a}, b] \cup (x_{2b}, x_{2c})$. *It is straightforward to see that a cutoff at* x_{2a} *dominates such an allocation since then the object will be misallocated only if* $x_2 \in (x_{2b}, x_{2c})$. *Moreover, any cutoff below* x_{2a} *is dominated by the cutoff of* x_{2a}. *Similarly, the cutoff of* x_{2c} *dominates the allocation induced by any cutoff in the interval* $[x_{2b}, x_{2c})$, *and by any cutoff above* x_{2c}. *Therefore, the optimal cutoff is either* x_{2a} *or* x_{2c}. *In other words, the optimal cutoff is one of the types that makes the designer indifferent between allocating the object and keeping it.*

Which of these two cutoffs is optimal? A cutoff at x_{2a} *causes misallocation in case the signal of the first agent belongs to the interval* (x_{2b}, x_{2c}) *since then the efficient allocation prescribes to keep the object until the next period. The disutility from the misallocation is*

$$\int_{x_{2b}}^{x_{2c}} \left(E\left[x|x_2\right] - x_2\right) \widetilde{f}_2(x_2) dx_2.$$

Setting the cutoff at x_{2c}, *causes misallocation in case the signal of the first agent belongs to the interval* (x_{2a}, x_{2b}) *since then efficient allocation prescribes to allocate the object to that agent. The disutility from the misallocation is given by*

$$\int_{x_{2a}}^{x_{2b}} \left(x_2 - E\left[x|x_2\right]\right) \widetilde{f}_2(x_2) dx_2.$$

Because $\int_{x_{2a}}^{x_{2b}} \left(x_2 - E\left[x|x_2\right]\right) \widetilde{f}_2(x_2) dx_2 < \int_{x_{2b}}^{x_{2c}} \left(E\left[x|x_2\right] - x_2\right) \widetilde{f}_2(x_2) dx_2$, *the cutoff* x_{2c} *is here optimal. Note how the density of next arrival's type, its expected value, and all types that make the designer indifferent between allocating today and tomorrow are involved in this calculation. Finally, note that backward induction can be used for cases with more periods in order to show that second-best mechanisms have a cut-off structure.*

Remark 1 *Chen and Wang [7] study the learning process of a revenue-maximizing seller who employs posted prices in a related model with one object, infinite horizon, and time discounting. In their setting there are two possible true distributions of values. Assuming that these two distributions*

[9] For this simple case with one object, this reduces to an argument originally due to Riley and Zeckhauser [22].

are ordered in the hazard rate order, and have an increasing hazard rate (this holds in the above example for any $\theta \geq 1$), they show that, after each rejection, more and more weight is put on the stochastically worse distribution, and therefore prices decline over time. It is important to note that these authors did not consider the optimal mechanism which would constitute the revenue analog of what we do here[10] .

3.2 The General Case

We now characterize the policy that maximizes expected welfare over the entire class of incentive compatible policies. The main mathematical idea used to prove that this policy is deterministic relies on several concepts from majorization theory (see Lemma 1 in Appendix A). Majorization is a measure of dispersion for vectors, akin to the second-order stochastic dominance relation among distributions. We first need the following definitions and a well-known result:

Definition 1 *1. For any $n-$tuple $\gamma = (\gamma_1, \gamma_2, .., \gamma_n)$ let $\gamma_{(j)}$ denote the jth largest coordinate (so that $\gamma_{(n)} \leq \gamma_{(n-1)} \leq ... \leq \gamma_{(1)}$). Let $\alpha = (\alpha_1, \alpha_2, .., \alpha_n)$ and $\beta = (\beta_1, \beta_2, .., \beta_n)$ be two $n-$tuples. We say that α is* majorized *by β and we write $\alpha \prec \beta$ if the following system of $n-1$ inequalities and one equality is satisfied:*

$$\begin{aligned} \alpha_{(1)} &\leq \beta_{(1)} \\ \alpha_{(1)} + \alpha_{(2)} &\leq \beta_{(1)} + \beta_{(2)} \\ ... &\leq ... \\ \alpha_{(1)} + \alpha_{(2)} + ..\alpha_{(n-1)} &\leq \beta_{(1)} + \beta_{(2)} + \beta_{(n-1)} \\ \alpha_{(1)} + \alpha_{(2)} + .. + \alpha_{(n)} &= \beta_{(1)} + \beta_{(2)} + .. + \beta_{(n)} \end{aligned}$$

2. A function $\Psi : \Re^n \to \Re$ is called Schur-convex *if $\Psi(\alpha) \leq \Psi(\beta)$ for any α, β such that $\alpha \prec \beta$.*

Theorem 2 *(see Marshall and Olkin [15]) Assume that $\Psi : \mathbb{R}^n \to \mathbb{R}$ is symmetric and has continuous partial derivatives. Then Ψ is Schur-convex*

[10]Babaioff et al. [3] offer several bounds on the ratio between the revenue that can be obtained by employing posted prices in models where the distribution is unknown and the revenue obtained when the distribution is known.

if and only if for all $(y_1, ..., y_n) \in \mathbb{R}^n$ *and all* $i, j \in \{1, .., n\}$ *it holds that*

$$\left(\frac{\partial \Psi (y_1, ..., y_n)}{\partial y_i} - \frac{\partial \Psi (y_1, ..., y_n)}{\partial y_j} \right) (y_i - y_j) \geq 0.$$

Theorem 3 *1. At each period* k*, expected welfare (calculated before the arrival of the period* k *agent) is a Schur-convex, linear function of the available qualities at that period.*

2. The incentive compatible, optimal mechanism (second-best) is deterministic. That is, for every history at period k*,* H_k*, and for every type* x *of the agent that arrives at that period, there exists a quality* q *that is allocated to that agent with probability* 1 *.*

3. At each period, the optimal mechanism partitions the type set of the arriving agent into a collection of disjoint intervals such that all types in a given interval obtain the same quality with probability 1*, and such that higher types obtain a higher quality.*

Proof. See Appendix A. ■

The proof of the Theorem proceeds by backward induction. The decision at the last period is trivial. At the last but one period there are at most two objects left, and the question is whether to allocate the higher quality to the current agent. We adapt a *no-haggling* argument due to Riley and Zeckhauser [22] in order to show that the optimal mechanism uses a cutoff-rule (see also the Illustration above). This argument is also analogous to the well known one showing that the revenue-maximizing mechanism for a seller facing one buyer - whose virtual value need not be increasing ! - is always a take-it-or-leave-it offer at a certain reserve price. The analogy follows because in all these cases the decision variable is binary.

The main challenge is then to prove that the optimal allocation is deterministic also at earlier periods, where several objects with different qualities are available. Assume then that the optimal mechanism is deterministic at all later stages k, $k-1, ..1$, and consider stage $k+1$. Given the observed signals so far, let $b_{i,k+1}$ denote here the expected value of the agent's type to which the item with $i-th$ smallest type is assigned in the future (from stage k on, before the period k signal is observed). Then, because the optimal mechanism from period k on is deterministic, and because cutoffs are chosen optimally, expected welfare is given by $\sum_{i=1}^{k} q_{(i)} b_{i,k+1}$ where $q_{(i)}$ denotes here $i-$th lowest quality among the items available for allocation at period k. By monotonicity of implementable rules, by the definition of the terms $b_{i,k+1}$, and by Theorem 2, expected welfare is a Schur-convex function of the available qualities at stage k. Since this function takes higher values for more dispersed vectors of qualities - in the sense of majorization - the designer prefers to leave for the future the most disperse set of feasible qualities (among all feasible and incentive compatible allocations that are welfare equivalent at period $k+1$). Turning this argument on its head, the allocation at period $k+1$ should be the **least dispersed** possible among all welfare equivalent and incentive compatible allocations. This implies that, for any agent arriving at $k+1$, the allocation should be either deterministic, or should randomize among at most two neighboring qualities.[11] Finally, randomization among two neighboring qualities cannot be optimal - this follows by the same argument used above for binary decisions.

4 When the Second Best Coincides with the First Best

After having characterized the second-best policy, we now look for conditions on the model's primitives under which the first-best, complete information policy is implementable.

The next result, related to a result in GM [10], displays an **implicit** sufficient condition on the cutoffs of the efficient, complete information allocation, characterized in Theorem 1 above.

Proposition 2 *Assume that for any k, χ_k, $i \in \{0,..,k\}$, 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. See Appendix B. ■

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

[11] In other words, if the designer seeks to allocate the object of expected quality $\tilde{q}$ to the current agent, he may use different lotteries between the available qualities. However, different lotteries leave different vectors of the qualities available for the future allocations. The lottery that randomizes between two neighboring qualities leaves for future allocation the most dispersed vector of expected ordered qualities of the objects.

result requires the effect of the current information on the current value to be stronger than the effect on the option value, similarly to the well known *single-crossing* condition that appears in the theory of efficient design with interdependent valuations. Under such conditions, the second best policy is also first best, i.e., the incentive constrained optimal policy coincides with the optimal policy under complete information.

Theorem 4 *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:*

suff1 $\widetilde{F}_k(x|\chi_k) \succsim_{FOSD} \widetilde{F}_k(x|\chi'_k)$

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

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

Proof. See Appendix B. ■

The first condition (stochastic dominance) above says that higher observations should lead to optimism about future observations[12], 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 tighter (i.e., the bound on the optimism associated to higher observation gradually decreases) in the early search periods. But learning models typically have a relatively high gradient in early learning periods, implying that our second condition is likely to be violated if there are many search periods (Example 3 in Appendix B illustrates this phenomenon).

In order to obtain sufficient 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 since in many

[12] The stochastic dominance condition is, for example, a simple consequence of a standard setting found in the literature (Milgrom [16]): 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 θ, and by $H(\theta)$ the corresponding probability distribution - the prior belief which gets then updated after each observation. If $f(x|\theta)$ has the Monotone Likelihood Ratio (MLR) property, then $\widetilde{F}_k(x|\chi_k) \succsim_{FOSD} \widetilde{F}_k(x|\chi'_k)$.

learning models (in particular in those where beliefs converge, say, to the true distribution) the impact of later observations 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. 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 2.

Theorem 5 *Assume that, for all k, all x, and all $n-k \geq i \geq 1$, the conditional distribution function $\widetilde{F}_k\left(x|x_n,..,x_{k+1}\right)$ and the density $\widetilde{f}_k\left(x|x_n,\cdots,x_{k+1}\right)$ are continuously differentiable with respect to x_{k+i}. If for all x, χ_k, and all $n-k \geq i \geq 1$, it holds that*

$$0 \geq \frac{\partial \widetilde{F}_k\left(x|\chi_k\right)}{\partial x_{k+i}} \geq -\frac{1}{n-k}\frac{\partial \widetilde{F}_k\left(x|\chi_k\right)}{\partial x} \tag{2}$$

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

Proof. See Appendix B. ■

Example 2 below further illustrates these conditions on a simple two-periods setting.

Remark 2 *While the left hand inequality in condition (2) is just another way to express the stochastic dominance condition also employed in Theorem 4, it is worth to deeper explore the right hand side.*

1. *Putting aside differentiability, this condition is equivalent to requiring that the function $\widetilde{F}_k\left(x+\frac{z}{n-k}|x_{k+1},.,x_{k+i}+z,x_{k+i+1},.,x_n\right)$ is non-decreasing in z. In other words, after having made $n-k$ observations, a small shift to the right - which moves the value of the distribution upwards - is enough to compensate the downward effect on the distribution's value 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).*

2. *Alternatively, denote by $x^*(u)$ the $u-th$ percentile of next type's conditional distribution function $\widetilde{F}_k\left(x|\chi_k\right)$, that is, $\widetilde{F}_k\left(x^*(u)|\chi_k\right) = u$.*

Condition (2) implies that

$$\frac{\partial x^*(u)}{\partial x_{k+i}} = - \left. \frac{\frac{\partial \widetilde{F}_k(x|\chi_k)}{\partial x_{k+i}}}{\frac{\partial \widetilde{F}_k(x|\chi_k)}{\partial x}} \right|_{x=x^*(u)} \leq \frac{1}{n-k}.$$

In other words, the effect of an increase in a previous observation on ***any*** *percentile of the distribution governing the next observation is bounded by* $\frac{1}{n-k}$*, where* $n-k$ *is the number of observations already made.*

3. *Note that our condition guarantees that* $\forall i,k,n, \frac{\partial a_{i,k}(\chi_k,x_k)}{\partial x_k} \leq \frac{1}{n-k}$ *although* $\frac{\partial a_{i,k}(\chi_k,x_k)}{\partial x_k} \leq 1$ *seems sufficient for the implementation of the efficient allocation. Nevertheless, the long-term effect of each non-terminal observation makes it impossible to obtain tighter conditions that apply generally. To see that, recall equation 12 which shows that each cutoff is given by the expectation of the second highest value among the type observed next period, and two adjacent next-period cutoffs. In particular, the current observation affects today's cutoffs via: 1. an impact on next period cutoffs, and 2. a shift of the relevant distribution of the second highest order-statistic. The second effect is bounded by*

$$\frac{\partial E\left(x_{k-1}|x_k,\chi_k\right)}{\partial x_k} = \frac{\partial}{\partial x_k}\int_0^\infty \left(1-\widetilde{F}_{k-1}\left(x|x_k,\chi_k\right)\right) \leq \frac{1}{n-k+1}$$

With a bound of 1 *instead of* $\frac{1}{n-k}$*, the first effect would be bounded by*

$$1-\left[\widetilde{F}_{k-1}\left(a_{i,k-1}|x_k,\chi_k\right)-\widetilde{F}_{k-1}\left(a_{i-1,k-1}|x_k,\chi_k\right)\right]$$

where $a_{i,k-1}$ *are tomorrow's optimal cutoffs. Since for any period* $k-1$ *there exists an* i *such that* $\widetilde{F}_{k-1}\left(a_{i,k-1}|x_k,\chi_k\right)-\widetilde{F}_{k-1}\left(a_{i-1,k-1}|x_k,\chi_k\right) \leq \frac{1}{k-1}$*, we would obtain*

$$1-\left[\widetilde{F}_{k-1}\left(a_{i,k-1}|x_k,\chi_k\right)-\widetilde{F}_{k-1}\left(a_{i-1,k-1}|x_k,\chi_k\right)\right] \geq \frac{k-2}{k-1}$$

which is arbitrarily close to 1 *if the number of the remaining objects is high. Thus, the combined effect may bigger than* 1 *for any* $k \geq \frac{n}{2}+1$ *which would violate the single-crossing condition.*

4. *Albright [1] displays a list of families of distributions/conjugate priors for which, at each period, the sets of types that get assigned to various qualities form an ordered sequence of intervals (see Theorem 1). In those cases, the efficient dynamic allocation (first best) is implementable. For example, consider a normal distribution of values* $\widetilde{x} \sim N(\mu, 1)$ *with unknown mean* μ*, and prior beliefs about* μ *of the form* $\widetilde{\mu} \sim N(\mu_0, 1/\tau)$ *where* $\tau > 0$*. We now show that the conditions of Theorem 5 are indeed satisfied.*

   After observing $x_n, ..x_{k+1}$ *the posterior on* $\widetilde{\mu}$ *is given by* $N(\overline{\mu}, 1/(\tau + n - k))$ *where*

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

   This yields

$$\widetilde{F}_k(x|x_n, ..., x_{k+1}) = N(\overline{\mu}, 1 + 1/(\tau + n - k)).$$

   Note that

$$\widetilde{F}_k(x + \frac{z}{\tau + (n-k)}|x_n, .., x_i + z, .., x_{k+1}) = \widetilde{F}_k(x|x_n, .., x_i, .., x_{k+1}) \quad (3)$$

   so that the stochastic dominance condition necessarily holds. By differentiating with respect to z *both sides of the identity (3), and by letting* z *go to zero, we obtain that*

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

   as desired.

Example 2 *There are two periods. The agents' valuations distribute on* $[0, 1]$*: with probability* α *the distribution is* $F(x) = x$*, while with probability* $1 - \alpha$ *the distribution is* $F(x) = x^\theta$ *where* $\theta > 0$*. Recall that the designer's belief about the next period's type given the current observation is given by*

$$\widetilde{f}_1 (x|x_2) = \frac{\alpha + (1 - \alpha) \theta^2 x_2^{\theta - 1} x^{\theta - 1}}{\alpha + (1 - \alpha) \theta x_2^{\theta - 1}}.$$

Condition (2) requires here that

$$
\begin{aligned}
0 &\geq \frac{\partial \widetilde{F}_1(x|x_2)}{\partial x_2} \geq -\frac{\partial \widetilde{F}_1(x|x_2)}{\partial x} \Longleftrightarrow \\
0 &\geq \frac{\alpha(1-\alpha)\theta(\theta-1)x_2^{\theta-2}x\left(x^{\theta-1}-1\right)}{\left(\alpha+(1-\alpha)\theta x_2^{\theta-1}\right)} \geq -\alpha-(1-\alpha)\theta^2 x_2^{\theta-1}x^{\theta-1}
\end{aligned}
$$

The left-hand-side inequality - stochastic dominance - holds for any $\theta \geq 1$. *It is also easy to see that the right-hand-side inequality holds for any* α *whenever* θ *is close enough to* 1. *Hence the efficient allocation is implementable for any* α *if* θ *is relatively close to* 1. *This is very intuitive since in those cases the two possible distributions are very close to each other and hence, for any* x_2, *the designer's belief about the agent that arrives next will not be significantly updated. When* θ *is larger, the condition holds for a smaller range of values of the parameter* α, *where the designer is relatively more confident that the uniform distribution is the true one. For example, when* $\theta = 2$ *the sufficient condition holds for any* $\alpha \geq 1/3$.

As already mentioned above, our condition is not necessary for implementability. While it guarantees that for any $x_2 \in [0,1]$, $\frac{\partial E[x|x_2]}{\partial x_2} \leq 1$, *a necessary and sufficient condition for the monotonicity of the efficient allocation in this example is that this inequality holds at the indifference points* $E[x|x_2] = x_2$. *The next figure, plotted for* $\alpha = 0.8$ *and* $\theta = 18$ *reveals that* $\frac{\partial E[x|x_2]}{\partial x_2} > 1$ *for relatively high* x_2. *Nevertheless, the efficient allocation is here monotone and hence implementable.*

4.1 Search for the Lowest Price and the Reservation Price Property

In a famous paper, Rothschild [21] 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 the searcher accepts (or stops search) at any price less than or equal to R, and rejects (or continues to search) 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

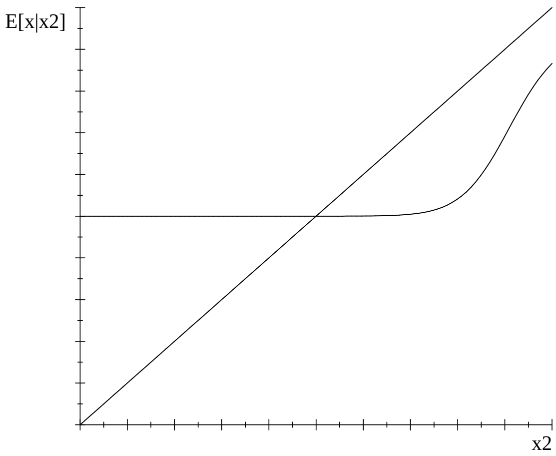

Figure 2: Sufficiency of Condition (2).

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 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. The optimal Bayesian search rule need not generally have this property (see Rothschild [21] and Kohn an Shavel [13] for examples). Rothschild showed that the reservation price property holds for a searcher equipped with a Dirichlet prior about the parameters of a multinomial distribution governing the price quotations[13]. 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

[13]The Dirichlet is the conjugate prior of the multinomial distribution, so the posterior is also Dirichlet in this case.

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. For the one-object case studied by Rothschild, various answers to this problem were offered by Rosenfield and Shapiro [20], Morgan [17], Seierstad [24] and Bickchandani and Sharma [6].

The conditions derived in our paper are more stringent than those obtained in the search literature, 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.[14]

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 [20]. In order to understand the relation between our results and theirs, recall first our condition from Theorem 5: For all x , χ_k ,and all $n-k \geq i \geq 1$

$$0 \geq \frac{\partial \widetilde{F}_k\left(x|\chi_k\right)}{\partial x_{k+i}} \geq -\frac{1}{n-k}\frac{\partial \widetilde{F}_k\left(x|\chi_k\right)}{\partial x}. \tag{4}$$

The first requirement in the paper by Rosenfield and Shapiro is identical to our stochastic dominance condition (the left hand side of condition (4)), 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, χ_k and all $n-k \geq i \geq 1$

$$\int_x^\infty \frac{\partial \widetilde{F}_k\left(y|\chi_k\right)dy}{\partial x_{k+i}} \geq -\frac{1}{n-k}(1-\widetilde{F}_k(x|\chi_k). \tag{5}$$

In other words, theirs is simply the "average" version of the right hand side side of our condition (4), and hence it is obviously implied by it.

Seierstad [24] offers another variant. Besides stochastic dominance, his condition reads (again in the differentiable case): For all x, k and χ_k

$$\sum_{i=1}^{n-k} \frac{\partial \widetilde{F}_k\left(x|\chi_k\right)}{\partial x_{k+i}} \geq -\widetilde{f}_k\left(x|\chi_k\right) \tag{6}$$

which is also clearly implied by our condition (4). The reason why we need stronger conditions than both Rosenfield and Shapiro's and Seierstad's is

[14] These more stringent conditions are needed even if all objects are homogenous.

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 Conclusion

We have derived conditions on the primitives of the learning environment that allow efficient dynamic implementation, and we have characterized the constrained efficient policy in terms of a generalized reservation price property. In yet another interpretation, our results can be seen as delineating the loss entailed by requiring "online" payments in dynamic allocation problems.

An interesting alternative approach for analyzing learning in our dynamic mechanism design environment would be to restrict attention to some simple class of indirect mechanisms that may be appealing for applications, e.g., a menu of prices at each period. It is important though to point out that such mechanisms entail some sub-optimality because the designer is not able then to elicit precise information about the agents' types. Thus she will learn less than in a direct mechanism, and each particular specification of prices also determines how much is being learned.

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 [25]). 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). But, 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, similarly to the phenomena illustrated here.

Our model can be easily generalized to allow for random arrival of agents (e.g., arrivals governed by a stochastic process). In such a framework an interesting extension is to also allow learning about arrival rates. We shall analyze such settings in future research.

Finally, one can perform exercises analogous to the present one also for other (non-Bayesian) learning models.

6 Appendix A: The Incentive Efficient (Second-Best) Policy

Proof of Proposition 1. $\Rightarrow$ For a given implementable allocation policy, assume by contradiction that there exist a period k, a history H_k, and two signals of the current agent $x' > x''$ such that $Q_k(H_k, x') < Q_k(H_k, x'')$. Denote by $P_k(H_k, x)$ the expected payment of the agent that arrives at period k after history H_k and reports x. The incentive constraint for type x'' implies:

$$x'' Q_k(H_k, x'') - P_k(H_k, x'') \geq x'' Q_k(H_k, x') - P_k(H_k, x'). \tag{7}$$

Since by assumption $x' > x''$ and $Q_k(H_k, x') < Q_k(H_k, x'')$, the above inequality implies that

$$\begin{aligned} x'(Q_k(H_k, x'') - Q_k(H_k, x')) &> x''(Q_k(H_k, x'') - Q_k(H_k, x')) \\ &\geq P_k(H_k, x'') - P_k(H_k, x'') \end{aligned}$$

which further implies that

$$x'(Q_k(H_k, x'') - Q_k(H_k, x')) > P_k(H_k, x'') - P_k(H_k, x')$$

The above inequality contradicts the incentive compatibility constraint for type x'.

$\Leftarrow$ We prove this part by constructing a payment scheme that implements a given monotonic allocation. Consider the following payment scheme:

$$P_k(H_k, x) = x Q_k(H_k, x) - \int_0^x Q_k(H_k, y)\, dy.$$

The expected utility of an agent with type x that arrives at period k after history H_k and reports truthfully is given by $\int_0^x Q_k(H_k, y)\, dy$. If he reports $x' \neq x$, his expected utility is given by

$$(x - x') Q_k(H_k, x') + \int_0^{x'} Q_k(H_k, y)\, dy.$$

We need to show that for any k , H_k, x, x' we have

$$(x - x')\, Q_k(H_k, x') + \int_0^{x'} Q_k(H_k, y)\, dy \le \int_0^{x} Q_k(H_k, y)\, dy.$$

The last inequality can be written as

$$(x - x')\, Q_k(H_k, x') \le \int_{x'}^{x} Q_k(H_k, y)\, dy.$$

which is true by the monotonicity of $Q_k(H_k, x)$. ■

For the proof of the Theorem 3, we need first the following Lemma:

Lemma 1 *Consider a set of* m *numbers* $q_m \ge q_{m-1} \ge ...q_1 \ge 0$, *and assume that* q_j *is deleted from the set with probability* p_j *,where* $0 \le p_j \le 1$ *and* $\sum_j p_j = 1$[15]. *Let* $Q = \sum_{j=1}^{m} p_j q_j$ *denote the expectation of the deleted term. Denote by* $\widetilde{q}_{(m-1)}$ *the expectation of the highest order statistic out of the* $(m-1)$ *remaining terms, by* $\widetilde{q}_{(m-2)}$ *the second highest order statistic, and so on, until* $\widetilde{q}_{(1)}$.

1. *If there exists* i *such that* $Q = q_i$, *then the* $(m-1)$ *dimensional vector* $(\widetilde{q}_{(m-1)}, \widetilde{q}_{(m-2)}, ...\widetilde{q}_{(1)})$ *is majorized by the* $(m-1)$ *dimensional vector* $(q_m, q_{m-1}, ..q_{i+1}, q_{i-1}, ..q_1)$ *obtained by deleting* q_i *with probability* 1.

2. *If there exist no* i *such that* $Q = q_i$, *let* l *and* $\alpha \in (0,1)$ *be such that* $Q = \sum_{j=1}^{m} p_j q_j = \alpha q_l + (1-\alpha) q_{l+1}$ [16]. *Delete* q_l *with probability* α, *and* q_{l+1} *with probability* $(1-\alpha)$, *and denote by* $\widetilde{\widetilde{q}}_{(m-1)}, \widetilde{\widetilde{q}}_{(m-2)}, .., \widetilde{\widetilde{q}}_{(1)}$ *the expectations of the order statistics out of the* $(m-1)$ *remaining terms. Then the* $(m-1)$ *dimensional vector* $(\widetilde{q}_{(m-1)}, \widetilde{q}_{(m-2)}, ...\widetilde{q}_{(1)})$ *is majorized by the* $(m-1)$ *dimensional vector* $(\widetilde{\widetilde{q}}_{(m-1)}, \widetilde{\widetilde{q}}_{(m-2)}, ...\widetilde{\widetilde{q}}_{(1)})$.

Proof. 1. If there exists j with $p_j = 1$, the claim is obvious. Assume therefore that at least two probabilities p_j are strictly positive (in particular this implies $q_1 < Q < q_m$). We need to show that the following holds:

$$\begin{aligned} \sum_{l=1}^{m-1} \widetilde{q}_{(l)} &= \sum_{l=1}^{m} q_j - q_i \\ \forall k &= 1, .., m-1, \sum_{l=k}^{m-1} \widetilde{q}_{(l)} \le q_m + ... + q_{i+1} + q_{i-1} + ... + q_k \end{aligned}$$

[15] In our application below, the deleted element is the quality of the object allocated today, while the remaining qualities stay for future allocation.

[16] Note that such an l and α must exist and are unique.

The first equality is clear by the the definition of $\widetilde{q}_{(l)}$. Note also that , for each $l = 1, ..m-1$, we have $q_l \leq \widetilde{q}_{(l)} \leq q_{l+1}$. For the inequalities there are two cases:

i. $k \leq i$. Then $q_l \leq \widetilde{q}_{(l)}$ implies:

$$\sum_{l=1}^{k-1} \widetilde{q}_{(l)} \geq \sum_{l=1}^{k-1} q_l \Leftrightarrow \sum_{l=1}^{m-1} \widetilde{q}_{(l)} - \sum_{l=1}^{k-1} \widetilde{q}_{(l)} \leq (\sum_{l=1}^{m} q_j - q_i) - \sum_{l=1}^{k-1} q_l$$

$$\Leftrightarrow \sum_{l=k}^{m-1} \widetilde{q}_{(l)} \leq q_m + ... + q_{i+1} + q_{i-1} + ... + q_k$$

ii. $k > i$. Then $\widetilde{q}_{(l)} \leq q_{l+1}$ implies

$$\sum_{l=k}^{m-1} \widetilde{q}_{(l)} \leq \sum_{l=k+1}^{m} q_l$$

2. The proof is very similar to the above one, and is omitted here. ■

Proof of Theorem 3. We prove the statement by backward induction, and we divide the proof in several steps. The argument for period 1 (last period) is obvious.

1. The argument for the last but one period. Consider period $k = 2$, and define $b_{1,2}(\chi_2, x_2) = \int_0^\infty x_1 d\widetilde{F}_1(x_1|\chi_2, x_2)$. This is the expectation of the agent's type who arrives at period 1, as a function of the observed history. Denote by $q_{(2:\Pi_2)} \geq q_{(1:\Pi_2)} \geq 0$ the two highest remaining qualities - only these are relevant here for welfare maximizing allocations - , and by $p(x_2)$ the probability that the period 2 agent gets the object with the higher quality. It is easy to see that incentive compatibility is equivalent here to p being monotonically increasing. The designer's problem is given by:

$$\max_p \int_0^\infty (p(x_2)[q_{(2:\Pi_2)}\ x_2 + q_{(1:\Pi_2)}\ b_{1,2}(\chi_2, x_2)]+$$
$$(1 - p(x_2))[q_{(1:\Pi_2)}\ x_2 + q_{(2:\Pi_2)}\ b_{1,2}(\chi_2, x_2)])d\widetilde{F}_2(x_2|\chi_2)$$

over increasing functions p with range in $[0, 1]$.

After simple manipulations the above reduces to

$$\max_p \left(q_{(2:\Pi_2)} - q_{(1:\Pi_2)}\right) \int_0^\infty p(x_2)[x_2 - b_{1,2}(\chi_2, x_2)]d\widetilde{F}_2(x_2|\chi_2)$$

over increasing functions p with range in $[0, 1]$.

Note that the solution of the above maximization problem does not depend on the values of $q_{(2:\Pi_2)} \geq q_{(1:\Pi_2)} \geq 0$. The problem is completely analogous to the classical problem faced by a revenue-maximizing seller who wants to allocate an indivisible object to a unique buyer whose virtual valuation (which need not be increasing) is given by the function $x_2 - b_{1,2}(\chi_2, x_2)$. By an argument originally due to Riley and Zeckhauser (1984), the solution is deterministic, and is given by

$$p(x_2) = \begin{cases} 1, \text{for } x_2 \geq b^*_{1,2}(\chi_2) \\ 0, \text{ otherwise} \end{cases}$$

where $b^*_{1,2}(\chi_2) = \arg\max_{x \in [0,\infty)} \int_x^\infty [x_2 - b_{1,2}(\chi_2, x_2)] d\widetilde{F}_2(x_2|\chi_2)$. In particular, note that in order to be a maximizer $b^*_{1,2}(\chi_2)$ must necessarily belong to the set

$$\beta_{1,2}(\chi_2) = \left\{ \begin{array}{c} x : \exists \varepsilon > 0 \text{ such that } \forall y \in (x, x+\varepsilon) \text{ it holds that } b_{1,2}(\chi_2, y) \leq y \\ \text{and } \forall z \in (x-\varepsilon, x) \text{ it holds that } b_{1,2}(\chi_2, y) \geq y \end{array} \right\}$$

In other words, if $b_{1,2}(\chi_2, x_2)$ is continuous in x_2, then $b^*_{1,2}(\chi_2)$ is one of the solutions of $x_2 = b_{1,2}(\chi_2, x_2)$.

Finally, note that the expected welfare after the allocation of period 2 has been made, but before the period 1 agent arrives, is given by $q_{(1:\Pi_1)} b_{1,2}(\chi_2, x_2)$.

2. The formula for expected welfare. Assume that the allocation at stages $1, 2, ..k$ is deterministic and uses cutoffs $b^*_{i-1,j}(\chi_j)$, $j = 1, 2, ..k$. Let $b_{1,2}(\chi_2, x_2) = \int_0^\infty x_1 d\widetilde{F}_1(x_1|\chi_2, x_2)$ and define inductively

$$\begin{aligned} b_{i,k+1}(\chi_{k+1}, x_{k+1}) &= \int_{B_{i,k}} x_k d\widetilde{F}_k(x_k|\chi_{k+1}, x_{k+1}) \\ &+ \int_{\underline{B}_{i,k}} b_{i-1,k}(\chi_k, x_k) d\widetilde{F}_k(x_k|\chi_{k+1}, x_{k+1}) \\ &+ \int_{\overline{B}_{i,k}} b_{i,k}(\chi_k, x_k) d\widetilde{F}_k(x_k|\chi_{k+1}, x_{k+1}) \end{aligned} \tag{8}$$

where

$$\begin{aligned} \underline{B}_{i,k} &= \left\{ x_k : x_k \leq b^*_{i-1,k}(\chi_k) \right\} \\ B_{i,k} &= \left\{ x_k : b^*_{i-1,k}(\chi_k) < x_k \leq b^*_{i,k}(\chi_k) \right\} \\ \overline{B}_{i,k} &= \left\{ x_k : x_k > b^*_{i,k}(\chi_k) \right\} \end{aligned}$$

That is, $b_{i,k+1}(\chi_{k+1}, x_{k+1})$ 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) in a non-random mechanism that uses cutoffs $b^*_{i,k}(\chi_k)$. Note that for any k, χ_k, x_k and i, we have $b_{i,k}(\chi_k, x_k) \geq b_{i-1,k}(\chi_k, x_k)$.

Denote by $q_{(i:\Pi_k)}$ the $i-$th lowest quality among the items available for allocation at period k, Π_k. We now show that the expected utility after the allocation at period $k+1$ has been completed is given by

$$\sum_{i=1}^{k} q_{(i:\Pi_k)} b_{i,k+1}(\chi_{k+1}, x_{k+1}), \tag{9}$$

By Theorem 2, the above function is a Schur-convex function of the available qualities at stage k.

The statement holds for period 2 (see point 1 above), and assume by induction that the expected utility from allocating the object of quality q in period k to an agent with type x_k is $qx_k + \sum_{i=1}^{k-1} q_{(i:\Pi_k \backslash q)} b_{i,k}(\chi_k, x_k)$ where $\Pi_k \backslash q$ is the set of the available objects after the allocation of the object of quality q at period k. Taking the expectation over x_k and using the inductive formulae 8, we obtain:

$$
\begin{aligned}
&\sum_{j=1}^{k}\int_{B_{j,k}}\left[q_{(j:\Pi_k)}x_k+\sum_{i=1}^{k-1}q_{(i:\Pi_k\setminus q_j)}b_{i,k}(\chi_k,x_k)\right]d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
=\;&\sum_{j=1}^{k}q_{(j:\Pi_k)}\int_{B_{j,k}}x_k d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
&+\sum_{j=1}^{k}\sum_{i=1}^{k-1}q_{(i:\Pi_k\setminus q_j)}\int_{B_{j,k}}b_{i,k}(\chi_k,x_k)d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
=\;&\sum_{j=1}^{k}q_{(j:\Pi_k)}\int_{B_{j,k}}x_k d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
&+\sum_{j=1}^{k}\sum_{i=1}^{j-1}q_{(i:\Pi_k)}\int_{B_{j,k}}b_{i,k}(\chi_k,x_k)d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
&+\sum_{j=1}^{k}\sum_{i=j+1}^{k}q_{(i:\Pi_k)}\int_{B_{j,k}}b_{i-1,k}(\chi_k,x_k)d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
=\;&\sum_{j=1}^{k}q_{(j:\Pi_k)}\int_{B_{j,k}}x_k d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
&+\sum_{i=1}^{k-1}q_{(i:\Pi_k)}\sum_{j=i+1}^{k}\int_{B_{j,k}}b_{i,k}(\chi_k,x_k)d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
&+\sum_{i=2}^{k}q_{(i:\Pi_k)}\sum_{j=1}^{i-1}\int_{B_{j,k}}b_{i-1,k}(\chi_k,x_k)d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
=\;&\sum_{j=1}^{k}q_{(j:\Pi_k)}\int_{B_{j,k}}x_k d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
&+\sum_{i=1}^{k-1}q_{(i:\Pi_k)}\int_{\overline{B}_{i,k}}b_{i,k}(\chi_k,x_k)d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})\\
&+\sum_{i=2}^{k}q_{(i:\Pi_k)}\int_{\underline{B}_{i,k}}b_{i-1,k}(\chi_k,x_k)d\widetilde{F}_k(x_k|\chi_{k+1},x_{k+1})=\sum_{j=1}^{k}q_{(j:\Pi_k)}b_{j,k+1}(\chi_{k+1},x_{k+1}).
\end{aligned}
$$

The the third equality is obtained by changing the order of summation, and the fourth equality follows from the definition of sets $\overline{B}_{j,k}$ and $\underline{B}_{j,k}$.

3. The optimal mechanism is deterministic. We now show that the optimal allocation at period $k+1$ is non-random. Take any incentive compatible mechanism such that the expected quality assigned to type x after history H_{k+1} is given by $Q_{k+1}(H_{k+1}, x)$. Recall that expected utility from stage k on is a Schur-convex function of the remaining qualities. In particular, this function is monotonically increasing in the majorization order. Apply now Lemma 1 to Π_{k+1}, the set of items available for allocation at period $k+1$, with "deletion of quality q_j " taken to mean "allocate quality q_j at period $k+1$". This application yields another incentive compatible mechanism that generates at least the same welfare as the original one, and that uses for any history H_{k+1} and arriving type x at stage $k+1$ either a non-random allocation rule, or a random rule that assigns positive probability only to two neighboring qualities[17].

Together with the necessary monotonicity of the expected quality, the above argument implies that, without loss of generality, we can restrict attention to allocation rules that divide the type space of the arriving agent into intervals $[0, \overline{x}^1), [\overline{x}^1, \underline{x}^2), [\underline{x}^2, \overline{x}^2), ..., [\overline{x}^k, \underline{x}^{k+1}), [\underline{x}^{k+1}, \infty)$ where $[\underline{x}^l, \overline{x}^l)$ is the interval of the types that get assigned to object $q_{(l:\Pi_{k+1})}$ with probability 1, while $[\overline{x}^{l-1}, \underline{x}^l)$ is the interval of randomization between $q_{(l-1:\Pi_{k+1})}$ and $q_{(l:\Pi_{k+1})}$. To complete the proof we need to show that for each potential interval of randomization $[\overline{x}^{l-1}, \underline{x}^l]$ there exists a cutoff $x_*^{l-1} \in [\overline{x}^{l-1}, \underline{x}^l]$ such that the designer can increase expected welfare by using (instead of randomization) a deterministic policy that allocates the object of quality q_{l-1} if $x \in [\overline{x}^{l-1}, x_*^{l-1})$ and allocate the object of quality q_l if $x \in [x_*^{l-1}, \underline{x}^l)$. Since this argument involves only two adjacent qualities, the proof is identical to the one used at point 1 above, and we omit it here.

4. The determination of optimal cutoffs. We finally show that the optimal cutoffs at period $k+1$, $b^*_{i,k+1}(\chi_{k+1})$ must belong to the set

$$\beta_{i,k+1}(\chi_{k+1}) = \left\{ \begin{array}{c} x \ / \ \exists \varepsilon > 0 \text{ such that: } \forall y \in (x, x+\varepsilon) \text{ holds } b_{i,k+1}(\chi_{k+1}, y) \leq y \\ \text{and } \forall z \in (x-\varepsilon, x) \text{ holds } b_{i,k+1}(\chi_{k+1}, y) \geq y \end{array} \right\}. \tag{10}$$

If such set is empty, that is, if $b_{i,k+1}(\chi_{k+1}, x_{k+1}) > x_{k+1}$ for any x_{k+1}, object i is never allocated to agent that arrives at period k after history χ_{k+1}, and we set $b^*_{i,k+1}(\chi_{k+1}) = \infty$ in this case.

By the derivation at point 3 above, the expected welfare at period $k+1$ if an object of quality q is allocated to the agent arriving at $k+1$ and if future

[17] The new mechanism is incentive compatible since, by the constructions used in Lemma 1, each type gets the same expected quality as in the original mechanism.

allocations are governed by the optimal policy is given by

$$qx_{k+1} + \sum_{j=1}^{k} q_{\left(j:\Pi_{k+1\backslash q}\right)} b_{i,k+1}(\chi_{k+1}, x_{k+1}). \tag{11}$$

Assume by contradiction that there exists a cutoff $x_i^* \left(\chi_{k+1}\right)$ which is used for allocating an object with quality $q_{(i:\Pi_{k+1})}$ such that $x_i^* \left(\chi_{k+1}\right) \notin \beta_{i,k+1} \left(\chi_{k+1}\right)$. Then either there exists $\varepsilon > 0$ such that for any $y \in \left(x_i^* \left(\chi_{k+1}\right), x_i^* \left(\chi_{k+1}\right) + \varepsilon\right)$ we have $b_{i,k+1}(\chi_{k+1}, y) > y$, or there exists $\varepsilon > 0$ such that for any $y \in \left(x_i^* \left(\chi_{k+1}\right) - \varepsilon, x_i^* \left(\chi_{k+1}\right)\right)$ we have $b_{i,k+1}(\chi_{k+1}, y) < y$. Take the first case (the second is analogous), and change the cutoff from $x_i^* \left(\chi_{k+1}\right)$ to $x_i^* \left(\chi_{k+1}\right) + \epsilon$ where $0 < \epsilon \leq \varepsilon$. Such a change is possible only if the cutoff for the adjacent higher quality is above $x_i^* \left(\chi_{k+1}\right) + \epsilon$. Then the change increases expected welfare since it has an impact only if $x_{k+1} \in \left(x_i^* \left(\chi_{k+1}\right), x_i^* \left(\chi_{k+1}\right) + \epsilon\right)$ in which case the current agent gets object $q_{(i-1:\Pi_{k+1})}$ instead of $q_{(i:\Pi_{k+1})}$. The effect of the increase on (11) is given by

$$\left(b_{i,k+1}(\chi_{k+1}, x_{k+1}) - x_{k+1}\right) \left(q_{(i:\Pi_{k+1})} - q_{(i-1:\Pi_{k+1})}\right) > 0.$$

If the cutoff for the adjacent higher quality object $q_{(i+1:\Pi_{k+1})}$ is also equal to $x_i^* \left(\chi_{k+1}\right)$, then adjust both cutoffs upwards by ϵ. In this case, the increase has an impact only if $x_{k+1} \in \left(x_i^* \left(\chi_{k+1}\right), x_i^* \left(\chi_{k+1}\right) + \epsilon\right)$, and the effect on welfare is

$$\left(b_{i+1,k+1}(\chi_{k+1}, x_{k+1}) - x_{k+1}\right) \left(q_{(i+1:\Pi_{k+1})} - q_{(i-1:\Pi_{k+1})}\right) > 0$$

To complete the proof we need to show that the selection of cutoffs from the set $\beta_{i,k+1} \left(\chi_{k+1}\right)$ is independent of the qualities of the available objects. This however, follows from the linearity of expected welfare in the available qualities. ■

7 Appendix B: Coincidence of Second Best and First Best

Proof of Proposition 2. GM [10] showed that the efficient allocation is implementable if and only if for any k, $i \leq k$ and χ_k the set of types that is matched with a given quality $\{x : a_{i,k}(\chi_k, x) > x \geq a_{i-1,k}(\chi_k, x)\}$ is convex. The characterization of the complete information efficient allocation provided by Albright states that for any k, $i \leq k$, x and χ_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, χ_k and $i \in \{0, .., k\}$, 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. ■

Before proving the main results about the implementability of the complete information efficient allocation (first best), we prove two useful structural results about the efficient cutoffs. First, we show that the average of all but the extreme cutoffs equals the expectation about the next type. Note that, by Theorem 1 we can write

$$a_{i,k+1}(\chi_{k+1}, x_{k+1}) = E_{x_k|x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \tag{12}$$

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

$$\begin{cases} a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k) & \text{if} \quad x_k \leq a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k) \\ x_k & \text{if} \quad a_{i-1,k}(\chi_{k+1}, x_{k+1}, x_k) < x_k \leq a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) \\ a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) & \text{if} \quad x_k > a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) \end{cases} . \tag{13}$$

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 $\widetilde{F}_k(x_k|\chi_{k+1}, x_{k+1})$ is symmetric with respect to the observed signals, then $a_{i,k+1}(\chi_{k+1}, x_{k+1})$ is symmetric as well.

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

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

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

$$\sum_{i=1}^{k} \left[a_{i-1,k}(\chi_k, x_k) \mathbf{1}_{\underline{A}_{i,k}} + a_{i,k}(\chi_k, x_k) \mathbf{1}_{\overline{A}_{i,k}} \right] = \sum_{i=1}^{k-1} a_{i,k}(\chi_{k+1}, x_{k+1}, x_k) \tag{14}$$

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

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

where the first equality follows from (14) , 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 $\widetilde{F}_k(x|\chi_k) \succsim_{FOSD} \widetilde{F}_k(x|\chi'_k)$. Then for any $i \in \{1, ..., k-1\}$ 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

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

Stochastic dominance immediately implies that the cutoffs are non-decreasing in x_2. We now apply the induction argument, and assume that, for any χ_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. ■

Proof of Theorem 4. Lemma 2 and the second condition in the Theorem's statement imply that

$$\sum_{i=1}^{k-1} (a_{i,k}(\chi_k, x_k) - a_{i,k}(\chi_k, x'_k)) = (k-1) \left(E_{x_{k-1}|\chi_k, x_k} x_{k-1} - E_{x_{k-1}|\chi_k, x'_k} x_{k-1} \right)$$
$$\leq \frac{k-1}{k-1} (x_k - x'_k) . \quad (15)$$

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 15 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 2, the efficient dynamic policy is then implementable. ■

Example 3 *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 $f_2(x) = 1 - \frac{b_2}{2} + b_2 x$, where $b_1, b_2 \in [-2, 2)$. Note that*

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

Using Bayesian updating we get that

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

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 χ_k and χ'_k be two sequences of observed signals that differ only in one coordinate, with $\chi_k \geq \chi'_k$. Then by simple calculations we obtain

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

Finally, if

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

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. To see that the second condition of Theorem 4 will not hold for sufficiently high number of periods, note that at the first period $E\left(x|\chi_{n-1}\right) - E\left(x|\chi_{n-1}'\right) = \frac{b_1-b_2}{12}\left[\Pr\left(b_i = b_1|x_n\right) - \Pr\left(b_i = b_1|x_n'\right)\right]$ *is independent of the number of future observations* n*. Therefore, there exist number of periods* n *and observations* x_n *and* x_n' *such that* $E\left(x|\chi_{n-1}\right) - E\left(x|\chi_{n-1}'\right) > \frac{x_n - x_n'}{n-2}$.

Proof of Theorem 5. Note first that

$$\begin{aligned}\frac{\partial E\,(x|\chi_k)}{\partial x_{k+i}} &= \frac{\partial}{\partial x_{k+i}}\int_0^\infty \left(1-\widetilde{F}_k\,(x|\chi_k)\right)dx_k \\ &\leq \frac{1}{n-k}\int_0^\infty \frac{\partial \widetilde{F}_k\,(x|\chi_k)}{\partial x}dx_k = \frac{1}{n-k} \qquad (16)\end{aligned}$$

where the inequality follows from the condition of the theorem. By Proposition 2, it is sufficient to show that for any k, any history of reports χ_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 χ_{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 (16) 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 $\widetilde{f}_1\left(x|x_n, ..., x_2\right)$ 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 $i \in \{1, ..., k-1\}$ 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

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

Since $a_{i,k}(\chi_{k+1}, x_{k+1}, x_k)$ is continuously differentiable in x_{k+1} for any $i \in \{1, ..., k-1\}$ by the induction argument, and since $\widetilde{f}_k\left(x_k|\chi_{k+1}, x_{k+1}\right)$ 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

$$\frac{\partial}{\partial x_{k+1}} \int_0^{\infty} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) \widetilde{f}_k \left(x_k | x_{k+1}, \chi_{k+1}\right) dx_k$$

$$= \int_0^{\infty} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \widetilde{f}_k \left(x_k | x_{k+1}, \chi_{k+1}\right) dx_k \tag{17}$$

$$+ \int_0^{\infty} \frac{\partial \widetilde{f}_k \left(x_k | x_{k+1}, \chi_{k+1}\right)}{\partial x_{k+1}} G_{i,k}(x_k, x_{k+1}, \chi_{k+1}) dx_k. \tag{18}$$

Consider first the term in the sum above (17):

$$\int_0^{\infty} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \widetilde{f}_k \left(x_k | x_{k+1}, \chi_{k+1}\right) dx_k$$

$$= \int_0^{a^*_{i-1,k}\left(x_{k+1},\chi_{k+1}\right)} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \widetilde{f}_k \left(x_k | x_{k+1}, \chi_{k+1}\right) dx_k$$

$$+ \int_{a^*_{i-1,k}\left(x_{k+1},\chi_{k+1}\right)}^{a^*_{i,k}\left(x_{k+1},\chi_{k+1}\right)} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \widetilde{f}_k \left(x_k | x_{k+1}, \chi_{k+1}\right) dx_k$$

$$+ \int_{a^*_{i,k}\left(x_{k+1},\chi_{k+1}\right)}^{\infty} \frac{\partial G_{i,k}(x_k, x_{k+1}, \chi_{k+1})}{\partial x_{k+1}} \widetilde{f}_k \left(x_k | x_{k+1}, \chi_{k+1}\right) dx_k$$

$$\leq \frac{1}{n-k+1} - \frac{1}{n-k+1} \left[\widetilde{F}_k \left(a^*_{i,k}\left(x_{k+1}, \chi_{k+1}\right) | x_{k+1}, \chi_{k+1}\right) - \widetilde{F}_k \left(a^*_{i-1,k}\left(x_{k+1}, \chi_{k+1}\right) | x_{k+1}, \chi_{k+1}\right)\right]$$

where the existence of the fixed points $a^*_{i,k}\left(x_{k+1}, \chi_{k+1}\right)$ and $a^*_{i-1,k}\left(x_{k+1}, \chi_{k+1}\right)$ 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 $x_k \in \left[a^*_{i-1,k}\left(x_{k+1}, \chi_{k+1}\right), a^*_{i,k}\left(x_{k+1}, \chi_{k+1}\right)\right]$.

Consider now the second term in the sum (18):

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

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

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

Combining now the two terms 17 and 18 we obtain

$$
\begin{aligned}
&\frac{\partial}{\partial x_{k+1}}\int_0^\infty G_{i,k}(x_k,x_{k+1},\chi_{k+1})\widetilde{f}_k\left(x_k|x_{k+1},\chi_{k+1}\right)dx_k \\
\leq\ & \frac{1}{n-k+1}- \\
& \frac{1}{n-k+1}\left[\widetilde{F}_k\left(a^*_{i,k}\left(x_{k+1},\chi_{k+1}\right)|x_{k+1},\chi_{k+1}\right)-\widetilde{F}_k\left(a^*_{i-1,k}\left(x_{k+1},\chi_{k+1}\right)|x_{k+1},\chi_{k+1}\right)\right] \\
& +\frac{1}{n-k+1}\frac{1}{n-k}-\frac{n-k}{n-k+1}\int_{a^*_{i-1,k}\left(x_{k+1},\chi_{k+1}\right)}^{a^*_{i,k}\left(x_{k+1},\chi_{k+1}\right)}\frac{\partial\widetilde{F}_k\left(x_k|x_{k+1},\chi_{k+1}\right)}{\partial x_{k+1}}dx_k \qquad (19)
\end{aligned}
$$

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

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

Integrating with respect to x both sides of the assumed inequality

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

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

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] Babaioff, M., Blumrosen, L., Dughmi, S. and Singer. Y.: "Posting Prices with Unknown Distributions," *Innovations in Computer Science* (ICS 2011).

[4] Becker, G.S. (1973): "A Theory of Marriage: Part I.", *Journal of Political Economy* **81**, 813-846.

[5] Bergemann, D. and Välimäki, J. 2007: "Dynamic Marginal Contribution Mechanisms." Cowles Foundation Discussion Paper 1616.

[6] Bickhchandani, S. and S. Sharma, (1996): "Optimal Search with Learning." *Journal of Economic Dynamics and Control* **20**, 333-359.

[7] Chen, Y. and Wang, R. (1999): "Learning Buyers' Valuation Distribution in Posted-Price Selling" , *Economic Theory* **14**, 417–428.

[8] Dasgupta, P. and Maskin, E. (2000): "Efficient Auctions." *Quarterly Journal of Economics,* **115**(2): 341-388.

[9] Derman, C., Lieberman, G, and Ross, S. (1972): "A Sequential Stochastic Assignment Problem." *Management Science*, **18**(7): 349-355.

[10] Gershkov, A. and Moldovanu, B. (2009): "Learning About the Future and Dynamic Efficiency." *American Economic Review* **99**(4): 1576-1587.

[11] Jehiel, P. and Moldovanu, B. (2001): "Efficient Design with Interdependent Valuations." *Econometrica,* **69**(5): 1237-1259.

[12] Kittsteiner, T. and Moldovanu, B. (2005): "Priority Auctions and Queue Disciplines that Depend on Processing Time." *Management Science,* **51**(2), 236-248.

[13] Kohn, M. and Shavell, S. (1974): "The theory of search." *Journal of Economic Theory,* **9,** 123

[14] 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.

[15] Marshall, A. and Olkin, I. (1979): "Inequalities: The Theory of Majorization and its Applications", Mathematics in Science and Engineering, Vol. 143, Academic Press, New-York.

[16] Milgrom, P. (1981): "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* **12**(2), 280-391.

[17] Morgan, P.B. (1985): "Distributions of the Duration and Value of Job Search with Learning." *Econometrica* **53**, 1199-1232.

[18] Mussa, M. and Rosen, S. (1978): "Monopoly and Product Quality", *Journal of Economic Theory* **18**, 301-317

[19] 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)

[20] Rosenfield, D. and Shapiro, R. (1981): "Optimal Adaptive Price Search." *Journal of Economic Theory,* **25**, 1-20.

[21] Rothschild, M. (1974): "Searching for the Lowest Price When the Distribution of Prices Is Unknown." *Journal of Political Economy*, **82**(4), 689-711.

[22] Riley, J. and Zeckhauser,R. (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm." *Quarterly Journal of Economics*, **98** (2), 267-289.

[23] Segal, I. (2003): "Optimal Pricing Mechanisms with Unknown Demand" *American Economic Review* **93** (3), 509-529

[24] Seierstad, A. (1992): "Reservation Prices in Optimal Stopping." *Operations Research*, **40**, 409-414

[25] Talluri, K. and van Ryzin, G. (2004): *The Theory and Practice of Revenue Management,* Springer: New York.