The Trade-off Between Fast Learning and Dynamic Efficiency

Alex Gershkov and Benny Moldovanu*

01.04.2008

Abstract

In both static and dynamic, independent private values setups, the efficient allocation is implementable if the distribution of agents' values is known. Lack of knowledge about the distribution is inconsequential in the static case. But, if distribution of agents' values is not known in a dynamic framework, and if the designer gradually learns about it by observing present values, endogenously arising informational externalities may prevent the implementation of the efficient allocation if present observations have a large impact on expectations about the future. We provide necessary and sufficient conditions for the efficient allocation to be implementable, and we draw a parallel to situations with direct informational externalities.

1 Introduction

In this paper we analyze the conditions under which an efficient allocation of resources is achievable in a dynamic private values setting where the designer gradually learns about the distribution of agents' values.

As an application, consider the sequential allocation of scarce spectrum to emerging technologies. Whenever a new technology appears, a benevolent designer is faced with a problem: should he allocate a valuable chunk of spectrum right now while forfeiting (at least for a long period of time) the option of allocating it to a possibly superior, future technology? It makes sense to assume here that current firms have a better knowledge of their technological prowess than the designer. Moreover, it is reasonable to assume that the assessment of

*We are grateful for financial support from the German Science Foundation and from the Max Planck Society. Gershkov, Moldovanu: Department of Economics, University of Bonn, Lennestr. 37, 53113 Bonn. Emails: alex.gershkov@uni-bonn.de, mold@uni-bonn.de.

future technological possibilities, can be improved by observing the current technology. As we shall show below, if current technology owners are strategic, the learning process about the future interacts with, and disturbs in a subtle way the process by which the present information necessary for an efficient decision can be extracted. Efficient implementation becomes possible only if learning is not "too fast".

Another good example is a hiring process in an organization who is filling several positions at different hierarchy levels over time (an university department, say). While all candidates prefer higher level positions, some aspect of their ability to perform the needed tasks at each level of the hierarchy is privately known to the candidates. If contracts with current employees are difficult to adjust, the designer must carefully compare the quality of present candidates - which can serve as a signal about the "market" - to the expected quality of candidates that may become available in the future.

The above examples are typical for a large class of situations where "investors" faced with a fixed amount of resources need to distribute them over time to a sequence of randomly arriving opportunities. As above, a transaction typically consists of a physical and a monetary part. An important assumption in our model is that both parts cannot be made contingent on information revealed in the future by other strategic agents. This assumption, which is discussed in a more detailed way in Section 7.1, is natural in contexts where new information arrives with a significant delay.

The theoretical study of the efficient allocation of resources to privately informed, strategic agents revolves around the seminal analysis due to Vickrey, Clarke and Groves (Vickrey [25], Clarke [7], Groves [13]). The VCG mechanisms work by aligning private and social interests via individual payments that correspond to the externality imposed by an agent's presence on others. Given these payments, all agents have the same goal - to maximize social welfare. An important requirement behind the VCG construction is that the externality imposed by an agent on others does not depend on that agent's information. This requirement is fulfilled in the standard *private values* framework where agents are completely informed about their own values for the various outcomes. Moreover, since the VCG mechanism implements the efficient allocation in dominant strategies, it can be executed even if the joint distribution of values is unknown. Thus, informational assumptions about this distribution are immaterial, and learning about it is not required from a welfare point of view.

The classical VCG model is static - all agents live simultaneously, and an allocation decision is taken only once. But, a generalization of the marginal externality idea allows the implementation of the sequentially efficient allocation also in dynamic settings with private values as long as the distribution of values is known to the designer (see Parkes and Singh, 2003). Knowledge of this distribution is now crucial since, in order to compute the correct monetary transfers that align private and social interests, the designer needs to calculate the correct expected externality imposed by an agent over time. Athey and Segal [3] and Bergemann and Valimäki [4] generalize the VCG analysis to a framework where agents obtain private information over time, focusing on budget balancedness, and on individual rationality, respectively.

The assumption, used in most studies of dynamic mechanism design, that the designer knows the (future) distribution of agents' values is unlikely to be fulfilled in most applications. Whereas this assumption is completely inconsequential in the static case, we show here that dropping it and replacing it with the assumption that the designer continually updates his beliefs about this distribution may lead to an impossibility of implementing the efficient allocation.

In a nut-shell, a necessary condition for extracting truthful information about values is the monotonicity of the allocation rule: agents with higher values should not be worse-off than contemporaneous agents with lower values. This requirement, however, may not coincide with what a dynamic efficient allocation stipulates: after observing a high value of a present agent, a "sufficiently optimistic" designer wants to keep, say, a high quality object for future allocation since he may expect even higher values in the future. In contrast, he may want to allocate the object right now if the present agent's value is lower, in which case the designer becomes "sufficiently pessimistic" about the future. When it happens, this lack of monotonicity due to sufficiently fast learning destroys incentive compatibility - the ability to extract truthful information from the present agent - and the efficient assignment cannot be implemented. Implementation is possible only if, roughly speaking, the belief updating process is not too fast, i.e., if the designer's beliefs do not change dramatically after each new observation.

We want to stress here that it is the combination of dynamic decision making and learning that may prevent the implementation of the efficient allocation. This happens even if all valuations are private because the externality imposed by a present agent on others becomes, via the designer's observation cum learning process, indirectly dependent on that agent's information. In other words, learning generates *indirect informational externalities.*

The above basic insight is both simple and general. In order not to obscure its simplicity, we chose to present it in the lean and elegant dynamic assignment framework due to Derman, Lieberman and Ross [9] (DLR hereafter). Our insight continues to play a role in much more general models of dynamic mechanism design.

In the DLR model, a finite set of heterogenous, commonly ranked objects needs to be assigned to a set of agents who arrive one at a time. After each arrival the designer needs to decide which object (if any) to assign to the present agent. Using dynamic programming, DLR have characterized the dynamically efficient policy in terms of cutoffs that, compared to the attribute of the present agent, determine at each point in time which object this agent should get. In their model both the attribute of the present agent (which determine his value for the various available objects) and the future distribution of attributes are known to the designer. We first show that the efficient policy characterized by these authors continues to be implementable even if arriving agents have private information about their attributes. (Bayesian) learning in the complete-information DLR model has been analyzed by Albright [2] who showed that DLR's qualitative results continue to hold also for the case where the social planner uses current information in order to learn about the distribution of forthcoming agents's types. In both these papers, however, the agents do not act strategically, and hence the problem of implementability does not arise. Our main results are presented in a model that adds both private information about attributes and strategic behavior to Albright's model.

In the DLR model with either complete information, or with private values and a known distribution of types, the cutoffs (or reservation values) determining the dynamic efficient policy turn out to be *Markovian*, i.e., they do not depend on the previous history in previous stages, nor on the (reported) type of the currently arriving agent. In contrast, Albright's model with learning about the distribution of values exhibits history-dependent cutoffs. Their dependence on the type of the agent who arrives at the current stage creates the implementation problem discussed above, whenever these types are private information. If learning is not "too fast", it is possible though to describe the efficient cutoffs in terms of fixed points that do not depend on the type of the arriving agent, and efficient implementation becomes possible.

The extensive literature on search considered the optimal stopping problem faced by a decision maker who is confronted with a stream of price quotations generated from some unknown, exogenous distribution[1]. In his seminal paper, Rothschild [28] derived sufficient conditions for the optimal policy to satisfy *the reservation price property*[2]. A search policy satisfies the reservation-price property if stopping search at any price would imply stopping the search also at all more favorable prices. At each point in time, this reservation-price may depend on the history of all observed offers, as in our model. The structure of Rothschild's model (and of many models in the large literature that followed) is different from ours in one main respect: the price quotations are not generated by strategic agents, and thus there are no information revelation problems, nor any incentive constraints[3]. Thus the monotonicity requirement behind the reservation price property is not due to an explicit economic constraint, whereas monotonicity is tightly related to incentive compatibility in our model.

In contrast to DLR, who focused on dynamic welfare maximization, there is an extensive literature on dynamic revenue maximization in the field of Management and Operation Research[4]. The standard assumption in this strand of literature is that the distribution of agents's attributes (i.e., the demand function) is known. Riley and Zeckhauser [26] considered a single object revenue-maximizing procedure where there is learning about the distribution of the agents' values. In their model, the optimal mechanism is a sequence of take-it-or-leave-it offers. While this sequence can be either increasing or decreasing over time, Chen and Wang [6] showed that the optimal sequence of the offers will decline over time if the possible distributions are ranked in the hazard rate order. This happens because then each rejection shifts the seller's beliefs towards the lower distribution, i.e., the seller becomes more pessimistic.

As discussed above, our private values model exhibits indirect informational externalities

[1]For an extensive survey of the classical search literature, see Lippman and McCall [20].

[2]Rothschild's characterization is performed under the assumption that price quotations are derived from some discrete distribution. Rosenfield and Shapiro [27] offer a generalization to continuous distributions.

See also Kohn and Shavel [19] who showed that the optimal decision whether to continue the search is always characterized by the cutoff level of the best available offer.

[3]Other differences are: 1) Usually, search stops after one acceptance, which corresponds to the single-object case in our paper. 2) There is an explicit cost of search.

[4]See the surveys by Bitran and Caldentey [5] and Elmaghraby and Keskinocak [10], and to the book by Talluri and Van Ryzin [30]. Gershkov and Moldovanu [12] explore the relationship between the revenue maximizing and welfare maximizing policies.

due to learning. Dasgupta and Maskin (2000) and Jehiel and Moldovanu (2001) have examined both the possible extensions and limitations of the VCG analysis to static frameworks with *direct informational externalities*, where the values of agents depend on information a-priori held by others (theirs is a so-called framework with *interdependent values*). In particular, efficient implementation - via generalized VCG mechanisms - is possible only if the private information is one-dimensional and if a single-crossing condition holds. Even with a single privately informed agent, efficient implementation is generically impossible if the private information is multi-dimensional. Kittsteiner and Moldovanu [18] used these insights in a dynamic auction model for queueing environments where agents have private information about own processing times.

We will show that the introduction of learning about the distribution of values in the dynamic model with private values yields qualitative insights that parallel the case of direct informational externalities in a dynamic model where the distribution of values is known. In particular, the efficient policy can always be implemented if the designer is able to delay the necessary monetary transfers after all observations about types have been made, analogously to insights developed by Mezzetti [21] for the static case with interdependent values, and by Athey and Segal [3] for a dynamic model where private information arrives over time.

Segal [29] analyzes revenue maximization in a static environment (i.e., all agents live simultaneously and there is one allocation decision to be made) with an unknown distribution of the agents' values. Although this paper deals with a private values environment, Segal observes that, analogously to what happens in our model, each agent has an informational effect on others, since his/her disclosed types reveals also some information about the types of his opponents. As a consequence, in the revenue maximizing procedure, the terms of trade for a given agent will be affected by information conveyed by others. Since in Segal's static private values environment utilization of a standard VCG mechanism does not require any knowledge about the joint distribution of values and therefore always leads to the efficient outcome, the type of problems highlighted in our dynamic environment cannot occur there.

Finally, we would like to mention several other strands of the literature have emphasized the trade-off between the maximization of current versus future profits in the presence of less than perfect information. For example, optimal experimentation by a monopolist facing an unknown demand was studied in Mirman, Samuelson and Urbano [22] in a model with two periods, and by Keller and Rady [17] in the infinite horizon case. Since myopically maximizing the current profit may not be the most informative strategy, in their models there is a trade-off between immediate profit maximization and more intensive learning that may increase the future profits[5]. For another, quite different example, consider the so-called *ratchet effect* analyzed by Freixas, Guesnerie and Tirole [11] and Weitzman [31] among others: in a repeated principal-agent interactions, leakage of private information in some period affects the optimal reward scheme at future periods. Thus, the agent will strategically choose his present action in order to manipulate the principal's beliefs, allowing for future information rents.

The paper is organized as follows. In Section 2 we present a non-technical illustration of our main idea. Section 3 describes the sequential assignment model and the various informational scenarios that will be treated. In Section 4 we present the main result of Derman, Lieberman, Ross [9] concerning the efficient dynamic assignment under complete information about presently arriving agents and under certainty about the distribution of types. In Section 4 we show that the DLR solution can be implemented, via appropriate payments, also in a private values framework with incomplete information about the types of arriving agents (while keeping the assumption that the designer knows the distribution of types). In Section 5 we present the main result of Albright [2] about the efficient dynamic assignment in a framework where arriving types are observable (complete information), but where the designer does not know the distribution of types. After each observation, the designer's belief about that distribution is updated. In Section 6, the main section, we show that the first best dynamic assignment need not be implementable if there is both incomplete information about arriving types, and learning about the distribution of types. We also offer several sufficient conditions under which efficient implementation is possible, and illustrate these with examples using conjugate prior families and Bayesian updating[6]. In Section 6.1 we show that the dynamic efficient assignment can always be implemented if all payments to the designer can be delayed until the end of the entire allocation process. In Section 7 we draw a parallel between the difficulties illustrated in the previous sections and those arising in a framework without learning but where arriving agents have interdependent values. Section

[5] The problem of the single decision maker that has to make a sequence of decisions in an unknown environment is also considered in Aghion et al. [1]. Their main result displays conditions under which, eventually, the decision maker adopts the correct decision.

[6] Our general treatment is not confined to Bayesian learning.

8 concludes.

2 An Illustration

One object can be allocated to one of two agents that arrive sequentially, one per period. Each agent derives utility from the object only if he gets it in the period of arrival. Valuations x_1 and x_2 are private and independently drawn from the interval $[0, 2]$, according to a uniform distribution. We call the "complete information case" the situation where the value of agents becomes known to the designer upon arrival, while in the "incomplete information case" the designer must first extract this information. For implementing allocations in the later case, we assume that both physical and monetary transfers to an agent must be made "online" upon arrival (i.e., they cannot be delayed and made contingent on information that may arrive at a later point in time).

Assume first, as commonly done in the auction/mechanism design literature, that the designer knows the distribution of values. The complete information dynamic efficient allocation is very simple: allocate the object to the first arriving agent if $x_1 \geq 1$, and allocate the object to the second agent otherwise (note that 1 is the expectation of the value for the second agent). This allocation can be obviously implemented also in the incomplete information case by posting a price of 1 in the first period, and a price of 0 in the second.

Assume now that the designer does not precisely know the distribution of values: with probability 0.5 he believes that the distribution is uniform on the interval $[0, 1]$, while with probability 0.5 he believes that it is uniform on $[1, 2]$. Let us first compute the complete information efficient dynamic allocation:

1. If the $x_1 \leq 1$, the designer updates his belief about x_2 , and the posterior is that x_2 is uniformly distributed on $[0, 1]$. Thus, the efficient allocation calls for assigning the object to the first agent if $x_1 \geq 0.5$, and to the second agent otherwise.

2. If the $x_1 > 1$, the posterior is that x_2 is uniformly distributed on $[1, 2]$. Thus, the efficient allocation calls for assigning the object to the first agent if $x_1 \geq 1.5$, and to the second agent otherwise.

3. If the object has not been allocated at the first stage, it is always assigned to the second agent.

4. To conclude, the first agent should get the object if $x_1 \in [0.5, 1] \cup [1.5, 2]$, and the second agent should get the object otherwise. In particular, the set of types for which agent 1 should get the object is not convex.

Unfortunately, the complete information efficient allocation cannot be implemented in the incomplete information case. To see that, assume that the designer makes a (possibly negative) transfer of t_1 to the first agent if this agent announces $\widehat{x}_1 \in [0.5, 1]$ and a transfer of τ_1 if this agent announces $\widehat{x}_1 \in [1, 1.5]$. Implementability of the efficient allocation requires that:

$$\begin{aligned} x_1 + t_1 &\geq \tau_1 \text{ for any } x_1 \in [0.5, 1] \\ \tau_1 &\geq x_1 + t_1 \text{ for any } x_1 \in [1, 1.5] \end{aligned}$$

While the first inequality implies $t_1 \geq \tau_1 - 0.5$, the second one implies $t_1 \leq \tau_1 - 1.5$. Together, these inequalities yield $0.5 + t_1 \geq 1.5 + t_1 \Leftrightarrow 0.5 \geq 1.5,$ a contradiction.

The problem is that the designer "learns too fast", i.e., the report of the first agent has a very significant impact on the updating process of the designer's belief. The information gained through learning may be used to the detriment of the first agent who becomes then reluctant to report truthfully.

Observe that the complete information efficient allocation can be implemented in the incomplete information case if the transfer to the first agent can be delayed until after the second period. The following scheme does the job:

1. If $\widehat{x}_1 \in [0.5, 1]$ the first agent gets the object and pays either 0.5 if $\widehat{x}_2 \leq 1$ or 100 if $\widehat{x}_2 > 1$

2. If $\widehat{x}_1 \in [1.5, 2]$ the first agent gets the object and pays either 1.5 if $\widehat{x}_2 \geq 1$ or 100 if $\widehat{x}_2 < 1$

3. If $\widehat{x}_1 \in [0, 0.5] \cup [1, 1.5]$ the first agent does not get the object and pays nothing.

4. If the object has not been allocated in the first stage it is assigned to the second agent without any further transfers[7].

[7] In this example, the second agent is always indifferent between lying and telling the truth, and we assume that he tells the truth.

In the above scheme, the second agent's report serves as a check on the first agent's report. Since the first agent's payment (which does not directly depend the agent's own report) can be delayed, it can be used as a punishment if a lie is subsequently detected.

3 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 type $q_i \geq 0$ is assigned to an agent with type x_j then this agent enjoys a utility given by $q_i x_j$. Getting no item generates utility of zero. The goal is to find an assignment that maximizes total welfare. In a static problem where all types are given, it is well known (see Hardy, Littlewood and Polya, [15]) that total welfare is maximized by assigning the item with the highest type to the agent with the highest type, the item with the second highest type to the agent with the second highest type, and so on... This assignment rule is called "assortative matching".

The main feature that makes the present problem interesting is the assumption that agents arrive sequentially, one agent per period of time, and that each can only be served upon arrival, i.e. there is no recall. Let period n denote the first period, period $n-1$ denote the second period, ...period 1 denote the last period. After an item is assigned, it cannot be reallocated in the future. 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 w.l.o.g. that $m = n$.

While the items' properties $0 \leq q_1 \leq q_2 ... \leq q_m$ are assumed to be known constants, the agents' properties are assumed to be independent and identically distributed random variables X_i on $[0, +\infty)$ with common c.d.f. F.

We consider below four distinct versions of the model. These differ in the respective underlying informational assumptions:

1. **Complete Information:** Arriving agents know their type, and this type is also observable by the designer at the time of arrival. There is uncertainty about the future (represented by the known distribution of types F).

2. **Incomplete Information:** Arriving agents know their type, but this is not observable by the designer. Thus, he/she may need to spend some resources in order to extract

information about types (i.e, some information rents are left to the agents). This yields a standard private values models with independent types.

3. **Complete Information + Learning:** Agents know their type and the type is also observable by the designer at arrival times. But, the distribution from which the types are sampled, F, is unknown. At stage n, the belief about F is originally described by a prior distribution $\Phi = \Phi_n$, which is then updated as successive types are observed. The posterior belief about the distribution of types after observing types $x_n, ..., x_m$ is given by $\Phi_{m-1}(x_n, ..., x_m)$. For any possible belief, we assume that the implied marginal distribution of types has a finite mean.

4. **Incomplete Information + Learning:** Agents know their type, but the type is not observable by the designer, and resources are needed to extract information about these. Moreover, the distribution from which the types are sampled is unknown. The beliefs about this distribution evolve as described above.

4 Dynamic Efficiency under Complete Information

Derman, Lieberman and Ross [9] have characterized the allocation policy that maximizes the total expected welfare in the complete-information model. Their main results are summarized in the Theorem below:

Theorem 1 *(DLR, 1972)*

1. *Consider the arrival of an agent with type x in period $k \geq 1$. There exist constants $0 = a_{0,k} \leq a_{1,k} \leq a_{2,k} ... \leq a_{k,k} = \infty$ such that the efficient sequential policy - that maximizes the expected value of the total reward - assigns the item with the $i-th$ smallest type if $x \in (a_{i-1,k},\, a_{i,k}]$. The constants $a_{i,k}$ depend on the distribution F but not on the $q's$.*[8]

2. *Each $a_{i,k}$ equals the expected value of the agent's type to whom the item with $i-th$ smallest type is assigned in a problem with $k-1$ periods. These constants can be*

[8] The basic form of the optimal policy remains the same for any supermodular reward function. But then the constants $a_{i,k}$ depend on the q's.

calculated by the following recursive formulae:

$$a_{i,k+1} = \int_{a_{i-1,k}}^{a_{i,k}} x dF(x) + a_{i-1,k} F(a_{i-1,k}) + a_{i,k}[1 - F(a_{i,k})] \tag{1}$$

where we set $+\infty \cdot 0 = -\infty \cdot 0 = 0$.

3. *In a problem with* n *periods and with items types* $0 \le q_1 \le q_2 \le ... \le q_n$*, total expected welfare is given by* $\sum_{i=1}^{n} q_i a_{i,n+1}$.

Note that the complete information efficient dynamic policy - called *first best* - is history independent: the efficient cutoffs at any stage k are not affected by previous decisions nor by previous reports.

5 Dynamic Efficiency under Incomplete Information

This section analyses the case where the designer cannot observe arriving types, but knows the distribution F from which types are drawn. We start with the characterization of the set of implementable allocations.

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. Both physical allocation and payment at period k may, in principle, depend on the signals previously observed. Let the history at period k, $H_k \in [0, \infty)^{n-k-1}$, be the ordered set of all signals reported by the agents that arrived at periods $n, ..., k+1$.[9] An allocation policy is called *deterministic* if, at any period k, and for any possible type of the agent that arrives at k it applies a non-random allocation rule.

Denote by $\phi_k : H_k \times [0, \infty) \times \Pi_k \to \Pi_k$ a deterministic allocation policy for period k, where Π_k is the set of available objects at k, and denote by $P_k : H_k \times [0, \infty) \times \Pi_k \to \mathbb{R}$ the associated payment rule. Note that the cardinality of set Π_k is k.

The next Proposition shows that a deterministic allocation policy is implementable if and only if, at each stage, it is based on a partition of the agents' type space.

[9] The history should also include also the set of the past decisions. Since the set of the present available objects contains all the relevant information, we omit this part for notational simplicity.

Proposition 1 *Assume that* $q_j \neq q_l$ *for any* $q_j, q_l \in \Pi_k$, $j \neq l$. *A deterministic policy* ϕ_k *is implementable if and only if there exist* $k+1$ *functions* $0 = y_{0,\Pi_k}(H_k) \leq y_{1,\Pi_k}(H_k) \leq y_{2,\Pi_k}(H_k) \leq \cdots \leq y_{k,\Pi_k}(H_k) = \infty$, *such that* $x_i \in (y_{j-1,\Pi_k}(H_k), y_{j,\Pi_k}(H_k)) \Rightarrow \phi_k(H_k, x_i, \Pi_k) = q_{(j)}$ *where* $q_{(j)}$ *denotes the j'th lowest element of the set* Π_k.[10] *Moreover, the associated payment scheme must satisfy* $P_k(H_k, x_t, \Pi_k) = P_k(H_k, \widetilde{x_k}, \Pi_k)$ *if* $\phi_k(H_k, x_k, \Pi_k) = \phi_k(H_k, \widetilde{x_k}, \Pi_k)$.

Proof. $\Longrightarrow$ If two reports lead to the same physical allocation, then, in any incentive compatible mechanism, the associated payments should be the same as well. A direct mechanism is equivalent here to a mechanism where the arriving agent at period k chooses an object and a payment from a menu $(q_j, P_j)_{j=1}^k$. For any history H_k, if some type x_k prefers the pair (q_m, P_m) over any other pair (q_l, P_l) with $q_m > q_l$, then any type $\widetilde{x_k} > x_k$ also prefers (q_m, P_m) over (q_l, P_l). This implies that $\phi_k(H_k, \widetilde{x}, \Pi_t) \geq \phi_k(H_k, x, \Pi_t)$ for any k, H_k and Π_t. Therefore an agent who arrives at period k gets object $q_{(j)}$ if he reports a type contained in the interval $(y_{j-1,\Pi_k}(H_k), y_{j,\Pi_k}(H_k))$. A similar argument shows that $\phi_t(H_k, y_{i,\Pi_k}(H_{k+1}), \Pi_t) \in \{q_i, q_{i+1}\}$ for $i \in \{1, 2, ..., k\}$.

$\Longleftarrow$ The proof is constructive. Given a partition-based allocation policy, we design a payment scheme P_j that, for any $j \in \{1, ..., k\}$, will induce type $x_k \in (y_{j-1,\Pi_k}(H_k), y_{j,\Pi_k}(H_k)]$ to choose the object with quality $q_{(j)}$. Without loss of generality, we assume that an agent who is indifferent between two best price-quality pairs in the menu $(q_j, P_j)_{j=1}^k$ chooses the object with lower type among the two. Consider then the following payment scheme

$$P_j(H_k, x_k, \Pi_k) = \sum_{i=2}^{j} (q_{(i)} - q_{(i-1)}) y_{i-1,\Pi_k}(H_k). \tag{2}$$

for $x_k \in (y_{j-1,\Pi_k}(H_k), y_{j,\Pi_k}(H_k)]$ and $P_1(H_k, x_k, \Pi_k) = 0$. Note that type $x_k = y_{j,\Pi_k}(H_k)$ is indifferent between $\left(q_{(j)}, P_j\right)$ and $\left(q_{(j+1)}, P_{j+1}\right)$. Moreover, any type above $y_{j,\Pi_k}(H_k)$ prefers $\left(q_{(j+1)}, P_{j+1}\right)$ over $\left(q_{(j)}, P_j\right)$, while any type below, prefers $\left(q_{(j)}, P_j\right)$ over $\left(q_{(j+1)}, P_{j+1}\right)$. Therefore, any type $x_k \in (y_{j,\Pi_k}(H_k), y_{j+1,\Pi_k}(H_k)]$ prefers $\left(q_{(j+1)}, P_{j+1}\right)$ over any other pairs in the menu.[11] ■

[10] Types at the boundary between two intervals can be assigned to either one of the neighboring elements of the partition. That is $\phi_t(H_k, y_{i,\Pi_k}(H_k), \Pi_t) \in \{q_i, q_{i+1}\}$, $i = 1, 2, .., k-1$.

[11] The payment given in (2) is not the only one implementing the partition $0 \leq y_{1,\Pi_k}(H_k) \leq y_{2,\Pi_k}(H_k) \leq \cdots \leq y_{k,\Pi_k}(H_k) = \infty$. Adding to the payment any function that does not depend on the reported type of the agent will not change the implemented partition. However, in any *individually rational* mechanism q_1 should be non-positive.

We assumed above that the inventory available at k contains only objects with distinct types. If there are some identical objects, there exist other implementable policies that do not take the form of partitions. But, for each such policy, there exists another implementable policy that is based on a partition, and that generates the same expected utility for all agents and for the designer.

Together with the partition-based characterization of Theorem 1, Proposition 1 shows that incomplete information per-se is not an obstacle towards achieving first-best dynamic efficiency.

Corollary 1 *The first-best policy is implementable also under incomplete information.*

6 Dynamic Efficiency under Complete Information and Learning

In this section we address the dynamic welfare maximization problem where the types of arriving agents are observable, but where the distribution of types is unknown. In this framework, each observation of an agent's type reveals some information about the distribution of future agents' types. This information can be used to improve the forecast of types arriving at later stages, and thus to improve the decision making in the present period where one needs to assess and compare present and option values.

The next result extends the complete information insight of Derman, Lieberman and Ross [9] to the setup with learning:

Theorem 2 *(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 $0 = a_{0,k}\left(\Phi_{k-1}(H_k, x_k)\right) \leq a_{1,k}\left(\Phi_{k-1}(H_k, x_k)\right) \leq a_{2,k}\left(\Phi_{k-1}(H_k, x_k)\right) ... \leq a_{k,k}\left(\Phi_{k-1}(H_k, x_k)\right) = \infty$ such that the efficient dynamic policy - which maximizes the expected value of the total reward - assigns the item with the $i-th$ smallest quality if $x_k \in \left(a_{i-1,k}\left(\Phi_{k-1}(H_k, x_k)\right),\ a_{i,k}\left(\Phi_{k-1}(H_k, x_k)\right)\right]$. The constants $a_{i,k}\left(\Phi_{k-1}(H_k, x_k)\right)$ do not depend on the $q's$.*

2. *Each $a_{i,k+1}\left(\Phi_k(H_{k+1}, x_{k+1})\right)$ 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:

$$a_{i,k+1}\left(\Phi_k(H_{k+1},x_{k+1})\right)=\int_{A_{i,k}} xdG(x) \tag{3}$$

$$+\int_{\underline{A}_{i,k}} a_{i-1,k}\left(\Phi_{k-1}(H_k,x_k)\right)dG(x)+\int_{\overline{A}_{i,k}} a_{i,k}\left(\Phi_{k-1}(H_k,x_k)\right)dG(x)$$

where G is the distribution of the X's which is derived from the current posterior $\Phi_k(H_k,x_{k+1})$, and where $\underline{A}_{i,k}=\{x:x\leq a_{i-1,k}\left(\Phi_{k-1}(H_k,x_k)\right)\}$,

$A_{i,k}=\{x:a_{i-1,k}\left(\Phi_{k-1}(H_k,x_k)\right)<x\leq a_{i,k}\left(\Phi_{k-1}(H_k,x_k)\right)\}$ and

$\overline{A}_{i,k}=\{x:x>a_{i,k}\left(\Phi_{k-1}(H_k,x_k)\right)\}$.[12]

In contrast to the case where the distribution of types is known, the first-best policy is now history dependent because every past observation affects the designer's beliefs about the characteristics of forthcoming agents.

As was noted by Albright, the actual calculation of the cutoffs is generally quite difficult. The next example illustrates the Theorem's result for a distribution of types that is uniform on the interval $[0,W]$, where the upper bound W is unknown. We assume that the designer's prior over W is a Pareto distribution. Note that the family of Pareto distributions is the *conjugate prior* family to the uniform distribution with an unknown upper bound.[13]

Example 1 *The agents' types distribute uniformly on the interval $[0,W]$, but W is unknown. The designer is endowed with a prior belief about the parameter W that is given by a Pareto distribution, $P(\alpha,R)$ with parameters α, R, where $\alpha>1$. That is, the density of the prior is given by:*

$$\begin{cases}\frac{\alpha R^{\alpha}}{W^{\alpha+1}} & \text{if } W>R\\ 0 & \text{otherwise}\end{cases}.$$

By the defining property of conjugate families, after observing agents' types $x_n,..,x_m$, the designer's beliefs about W will still be described by a Pareto distribution, but with updated parameters $\alpha\left(x_n,..,x_m\right)$ and $R\left(x_n,..,x_m\right)$.

Assume that at the penultimate period, $k=2$, the parameters of the Pareto distribution are $\alpha\left(x_n,..,x_3\right)=\alpha_2$ and $R\left(x_n,..,x_3\right)=R_2$. If at that period an agent with type x_2 arrives,

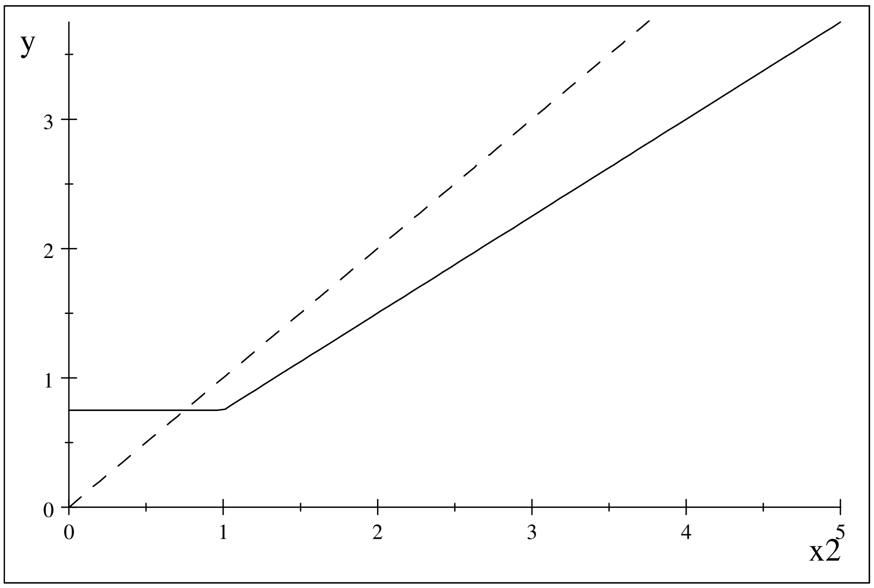
[12] We set $+\infty\cdot 0=-\infty\cdot 0=0$.

[13] Given Bayesian updating of beliefs, a class of prior probabilities is said to be conjugate if it is closed under sampling.

and if there are two available objects with types, $p_1 \leq p_2$, the efficient allocation is to assign the object with type p_2 whenever $x_2 > E[X_1|X_2 = x_2]$. The expected value of the type at period 1 (which coincides with the efficient cutoff at period 2) is given by

$$a_{1,2}(x_2) := E[X_1|X_2 = x_2] = \frac{\alpha_2 + 1}{2\alpha_2} \max\{R_2, x_2\}$$

The next figure plots the efficient cutoff $a_{1,2}(x_2)$ (solid line) and the type of the agent arriving in period 2 (dashed line) for values $R_2 = 1$ and $\alpha_2 = 2$.



Note that $\alpha_2 > 1$ implies $\frac{\alpha_2+1}{2\alpha_2} < 1$, and thus the unique solution to the equation $\frac{\alpha_2+1}{2\alpha_2} \max\{R_2, x\} = x$ is $x^* = \frac{\alpha_2+1}{2\alpha_2} R_2$. Therefore, the efficient allocation rule can be described as follows: if the agent's type is above $\frac{\alpha_2+1}{2\alpha_2} R_2$, he gets the object p_2; otherwise, he gets p_1.

We now proceed to stage $k = 3$. Assume that the agent that arrives at this stage has a type of x_3. Similar computations to the above yield the following efficient allocation rule:

1. *The agents gets the object with characteristic p_3 if and only if*

$$x_3 \geq a_{2,3}(x_3) = \max\{R_3, x_3\} \left(\frac{\alpha_3 + 2}{8\alpha_3 + 8} + \frac{1}{2} + \frac{1}{2\alpha_3} \right)$$

2. *He gets the object with characteristic p_1 if and only if*

$$x \leq a_{1,3}(x_3) = \max\{R_3, x_3\} \left(\frac{3\alpha_3 + 2}{8\alpha_3 + 8} + \frac{1}{2\alpha_3} \right)$$

3. He gets the object p_2 otherwise.

Finally, the updating rule is $\alpha_2 = \alpha_3 + 1$ and $R_2 = \max\{R_3, x_3\}$.

7 The Trade-off between Dynamic Efficiency and Learning under Incomplete Information

In an exact parallel to our analysis of complete information first-best policy characterized by Derman, Lieberman and Ross (see Theorem 1 and Corollary1) we now ask whether it is possible to implement the first best policy with learning, described in Theorem 2, when we add incomplete information.

Although history dependent allocation rules were not necessary for the implementation of the first best policy without learning, recall that the general implementation result, Proposition 1, allowed for such rules. Moreover, recall that its proof used only the incentive compatibility constraints of the currently arriving agent. Therefore, an analogous result holds even if the distribution of agents' types is unknown, and if learning takes place:

Proposition 2 *Assume that $q_j \neq q_l$ for any $q_j, q_l \in \Pi_k$, $j \neq l$. A deterministic policy ϕ_k is implementable if and only if there exist $k+1$ functions $0 = y_{0,\Pi_k}(H_k) \leq y_{1,\Pi_k}(H_k) \leq y_{2,\Pi_k}(H_k) \leq \cdots \leq y_{k,\Pi_k}(H_k) = \infty$, such that $x_i \in (y_{j-1,\Pi_k}(H_k), y_{j,\Pi_k}(H_k)) \Rightarrow \phi_k(H_k, x_i, \Pi_k) = q_{(j)}$ where $q_{(j)}$ denotes the j'th lowest element of the set Π_k.[14] Moreover, the associated payment scheme must satisfy $P_k(H_k, x_t, \Pi_k) = P_k(H_k, \widetilde{x_k}, \Pi_k)$ if $\phi_k(H_k, x_k, \Pi_k) = \phi_k(H_k, \widetilde{x_k}, \Pi_k)$.*

Proof. The proof is similar to the proof of Proposition 1 and it is omitted here[15]. ∎

Roughly speaking, an implementable policy may use the current information revealed by the period k agent only to determine the allocation at all future stages. But, in marked contrast to the case with a known distribution, Theorem 2 shows that, at stage k, the efficient policy is described in terms of cutoffs that generally depend on the type of the agent arriving at k. As we shall see below, it is this feature that generates problems when trying to extract information from this agent, and hence when trying to implement the first-best policy. The crucial question is whether we can alternatively describe the efficient policy at any stage

[14] Types at the boundary between two intervals can be assigned to either one of the neighboring elements of the partition. That is $\phi_k(H_k, y_{i,\Pi_k}(H_k), \Pi_t) \in \{q_i, q_{i+1}\}$, $i = 0, 2, .., k-1$.

[15] For the case of several identical objects see the remark after Proposition 1.

k in terms of cutoffs that are independent of the type of the agent arriving at k- such a description will involve the existence of certain fixed points (recall the construction of the optimal policy for stage $k = 2$ in Example 1).

It is interesting to note here that Morgan [23] focused on the existence of fixed points that do not depend on current information for the determination of the reservation value function in an optimal search model with learning. His analysis was conducted under the assumption of complete information, and the goal of his exercise was to find simpler ex-ante calculable formulas for probabilities of search duration and for the expected value of search.

Out next Theorem exhibits a necessary and sufficient condition for the implementability of the first-best policy in the incomplete-information + learning framework, and an additional sufficient condition that can be easily checked in applications:

Theorem 3 *1) Consider the first-best policy (which takes into account learning), characterized in Theorem 2. This policy is implementable if and only if for any period k, for any object q_i, and for any belief $\Phi_{k-1}(H_k, x_k)$, the set*

$A_{i,k} = \{x_k : a_{i-1,k}\left(\Phi_{k-1}H_k, x_k)\right) < x_k \leq a_{i,k}\left(\Phi_{k-1}(H_k, x_k)\right)\}$ *is convex (i.e., it is an interval)*

2) Assume that for any k , and for any $i \in \{0, .., k\}$, the cutoff $a_{i,k}\left(\Phi_{k_{-1}}(H_k, x_k)\right)$ is differentiable with respect to the signal of the agent arriving at k, and assume that for any x_k it holds that

$$\frac{\partial}{\partial x_k} a_{i,k}\left(\Phi_{k-1}(H_k, x_k)\right) < 1,$$

Then, there exist constants $0 = a^*_{0,k}\left(H_k\right) \leq a^*_{1,k}\left(H_k\right) \leq a^*_{2,k}\left(H_k\right) ... \leq a^*_{k,k}\left(H_k\right) = \infty$ that do not depend on x_k such that for any $i \in \{1, ..., k\}$

$$x_k \in \left(a_{i-1,k}\left(\Phi_{k-1}(H_k, x_k)\right), a_{i,k}\left(\Phi_{k-1}(H_k, x_k)\right)\right] \Longrightarrow x_k \in \left(a^*_{i-1,k}\left(H_k\right), a^*_{i,k}\left(H_k\right)\right].$$

In particular, the first-best policy can be implemented also under incomplete information.

Proof. Recall from Theorem 2 that the current agent k should get the object with characteristic q_i if and only if his type belongs to the set

$$A_{i,k} = \{x_k : a_{i-1,k}\left(\Phi_{k-1}(H_k, x_k)\right) < x_k \leq a_{i,k}\left(\Phi_{k-1}(H_k, x_k)\right)\}.$$

The first claim follows then immediately from Proposition 2.

For the second part, recall that $a_{i,k}(\Phi_{k-1}(H_k,x_k))$ equals the expected value of the agent's type to which the item with i-th smallest type is assigned. This yields $a_{i,k}(\Phi_{k-1}(H_k,x_k)) \geq 0$ for any i, k and any posterior Φ_{k-1}. The inequality $\frac{\partial}{\partial x_k} a_{i,k}(\Phi_{k-1}(H_k,x_k)) < 1$ implies then that the equation

$$x = a_{i,k}(\Phi_{k-1}(H_k, x))$$

has a unique solution, which we denote by $a^*_{i,k}(H_k)$. From Theorem 2 we know that $a_{i,k}(\Phi_{k-1}(H_k,x_k)) \geq a_{i-1,k}(\Phi_{k-1}(H_k,x_k))$, which, in turn, implies that $a^*_{i,k}(H_k) \geq a^*_{i-1,k}(H_k)$. Therefore, the set $\left\{a^*_{i,k}(H_k)\right\}_{i=0}^{k}$ partitions agent's k type space. Moreover, by definition, the cutoffs in this set are independent of agent's k type. Finally, if $x_k > a_{i-1,k}(\Phi_{k-1}(H_k,x_k))$, the definition of $a^*_{i-1,k}(H_k)$ implies that $x_k > a^*_{i-1,k}(H_k)$, and $x_k \leq a_{i,k}(\Phi_{k-1}(H_k,x_k))$ implies $x_k \leq a^*_{i,k}(H_k)$. Therefore,

$$x_k \in (a_{i-1,k}(\Phi_{k-1}(H_k,x_k)), a_{i,k}(\Phi_{k-1}(H_k,x_k))] \Longrightarrow x_k \in (a^*_{i-1,k}(H_k), a^*_{i,k}(H_k)].$$

■

If the cutoffs $a_{i,k}(\Phi_{k-1}(H_k,x_k))$ are only slightly affected by the information conveyed by the signal x_k (which means that the current optimal allocation doesn't significantly react to new information, and that learning is a gradual) then the efficient allocation can be implemented. Recall that in the limiting case of no learning - where we know that the efficient allocation is implementable - we have $\frac{\partial}{\partial x_k} a_{i,k}(\Phi_{k-1}) = 0$. An example in the Appendix shows that differentiability and $\frac{\partial}{\partial x_k} a_{i,k}(\Phi_{k-1}) < 1$ are only the sufficient but not necessary conditions.

We present below two examples of distributions where the resulting sets $A_{i,k}$ are convex. Hence the above Theorem implies that the first-best policy is implementable in both cases.

Example 2 *(Albright 1977)*

1. *The agents' types distribute uniformly with support $[w,W]$, but the bounds w and W are unknown. The conjugate prior family is the Bivariate Pareto distribution with parameters α, r, R, and with density*

$$\begin{cases} \frac{\alpha(\alpha+1)(R-r)^{\alpha}}{(W-w)^{\alpha+1}} & \text{if } W > R \text{ and } w < r \\ 0 & \text{otherwise} \end{cases}.$$

The efficient cutoffs for period k are given by $a_{i,k}(\Phi_{k-1}) = \frac{r_k+R_k}{2} + \frac{R_k-r_k}{2} b_{i,n}(\alpha_k)$ *where* $b_{i,n}(\alpha_k) = a_{i,k}(\Phi_{k-1})$ *with* $r_k = -1$, $R_k = 1$ *and* $\alpha = \alpha_k$. *The updating rule for the hyperparameters of the distribution of* (w, W) *is as follows:*

$$\begin{aligned} \alpha_k &= \alpha_{k+1} + 1 \\ r_k &= \min\{x, r_{k+1}\} \\ R_k &= \max\{x, R_{k+1}\} \end{aligned}$$

2. *The agents' types distribute according to a Gamma distribution* $Gamma(\alpha, \beta)$ *with unknown inverse scale (or rate) parameter* β. *The conjugate family for for this case is of the form* $Gamma(\gamma, \delta)$. *The efficient cutoffs for period k are given by* $a_{i,k}(\Phi_{k-1}) = \delta_k b_{i,n}(\gamma_k, \alpha_k)$, *where* $b_{i,n}(\alpha_k) = a_{i,k}(\Phi_{k-1})$ *for* $\delta = 1$, *and* $\gamma = \gamma_k$. *The updating rule for the hyperparameters for the distribution of* β *is*

$$\begin{aligned} \gamma_k &= \gamma_{k+1} + \alpha \\ \delta_k &= \delta_{k+1} + x. \end{aligned}$$

Our next result considers a simple but structurally important case where the first-best policy has a very special structure, and therefore can be always implemented:

Proposition 3 *Assume that* $q_1 \leq q_2 \leq ..q_n$, *and assume that, for any sequence of arriving types, the sequence of beliefs* $\Phi = \Phi_n, \Phi_{n-1}, .., \Phi_1$ *induce successive marginal distributions of types that form a submartingale (supermartingale). Then, the complete information, dynamically efficient policy is to allocate the items successively in ascending (descending) order of types:* $q_1,\ q_2, ..q_n \quad (q_n,\ q_{n-1}, .., q_1)$. *These policies are implementable using type-independent transfers*[16].

Proof. The simple form of the efficient policies follows by Theorem 2 of Derman, Lieberman and Ross (1972). The second claim follows trivially since no information needs to be extracted from the agents in order to implement these policies. ■

The entire above analysis has assumed that the agents' types are independently distributed. Qualitatively similar dynamic implementation problems occur if types are correlated, even if the distribution of types is known - see also Section 8 below.

[16] If the induced sequence is a martingale, then any allocation policy is efficient and implementable.

7.1 Delayed Mechanisms

As we showed in the illustration (see Section 2), the ability to delay payments until all relevant information has been revealed allows much more flexibility. This contrasts the insight for independent private values frameworks without learning where delaying payments does not improve the ability to achieve efficient outcomes (see Parkes and Singh [24]).

Assume now that the designer still needs to allocates the objects upon arrival, but that he is able to delay payments until the end of stage one (which is the last stage). Consider the following mechanism: upon arrival, agents report their types, and objects are assigned according to the first-best allocation policy described in Theorem 2. If agent k gets object $q_{(j)}$, $j \geq 2$ he pays $\sum_{i=2}^{j}(q_{(i)} - q_{(i-1)})x_{(i-1)}$ where $x_{(1)}, .., x_{(n)}$ represent the order set of the revealed types, which is fully known at the end of stage one, and if gets object $p_{(1)}$ he pays zero. It is easy to see that reporting truthfully is an equilibrium. Thus, we have proved[17]:

Proposition 4 *If payments can be delayed until no more arrivals occur, it is always possible to implement the first-best allocation policy.*

Although theoretically appealing, a delayed scheme such as the one above may be problematic in real-life situations where arrivals are separated by significant periods of time because:

1. The first arriving agents may not even "exist" anymore when the later ones arrive.

2. The last agent's signal has no effect on his own allocation (both physical and monetary), but it does influence the payments of all preceding agents. It is not clear what are the incentives of this agent to report truthfully[18]. Moreover, this feature is conducive to collusive agreements, and more significantly so than in static frameworks where the colluders need to jointly solve both a physical and a monetary allocation problem under two-sided incomplete information.

3. It seems difficult both to write contracts that cover numerous contigencies that may occur in a distant, uncertain future, and to execute monetary payments based on events that may not be verifiable any more since they lie in the distant past.

[17] An analogous result for a much more general model has been proven by Athey and Segal [3] - see their Proposition 1.

[18] See also Mezzetti [21].

8 Sequential Assignment with Interdependent Values

In this Section we very briefly describe a simple variation of Derman-Lieberman-Ross basic model (without any learning) where agents have interdependent types. Our goal is to show that this model exhibits, both conceptually and technically, the same problems/possible solutions as those encountered in the incomplete information + learning framework analyzed above.

We will show how interdependence in types leads to history dependent optimal partitions, and we exhibit the necessary mechanism adjustments needed to implement the welfare maximizing allocation in the sequential arrival model with incomplete information. As above, not all incentive-compatible mechanisms can be replicated by "online" mechanisms where monetary transfers need to be simultaneous to physical allocations.

A general formulation is to assume that the type (or value) of the agent arriving in period k, is a certain function of the types of all other agents. The types of agents arriving at later stages cannot be observed at stage k, and they do not influence the analysis: one computes the expected value over all these signals, and uses that value for allocation decisions at stage k. Thus, only signals of agents arriving before period k are of interest. We choose below the simplest formulation for our purposes: if the agent who arrives at stage k has type x_k and if he gets an object with type p_i, then his utility is given by $q_i v_k$ where $v_k = x_k + \alpha x_{k+1}$, where x_{k+1} is the type of the agent that arrived at the previous stage $k+1$, and where $\alpha > 0$. Thus, the agent arriving in period k cares only about his own signal and the signal of the agent who arrived one period earlier.

Proposition 5 *1. Assume that the values of the agents in period $n, n-1, .., k+1$ were $v_n, v_{n-1}, .., v_{k+1}$, and consider the arrival of an agent with value v_k (or type $x_k = v_k - \alpha x_{k+1}$) in period $k \geq 1$. There exist functions $-\infty = a_{0,k}(v_k) \leq a_{1,k}(v_k) \leq a_{2,k}(v_k) ... \leq a_{k,k}(v_k) = \infty$ such that the efficient sequential policy assigns the item with the $i-th$ remaining smallest type if $v_k \in (a_{i-1,k}(v_k), a_{i,k}(v_k)]$. These functions depend on the distribution of agents' types F, but not on the $q's$.*

2. Denote by $c_{i,k}(v_k)$ the i-th lowest term in the set $\left\{ \{a_{i,k}(v_k)\}_{i=1}^{k-1}, v_k \right\}$. Then the expected utility from stage k onward is given by

$$\sum_{i=1}^{k} q_{(i)} c_{i,k}(v_k)$$

where $q_{(i)}$ denotes the $i-th$ lowest available object.

3. *Assume that $\alpha < 1$. Then, for each i, and for each $v_n, v_{n-1}, .., v_{k+1}$, the equation $a_{i,k}(v_k) = v_k$ has a unique solution, denoted by $a^*_{i,k} = a^*_{i,k}(v_n, v_{n-1}, .., v_{k+1})$.*

4. *The efficient assignment is implementable if $\alpha < 1$.*

Proof. 1-2. The proof is by induction. The claims are correct for $k = 1$. Suppose that the claims are correct if there are still m stages to go. Then, at the previous stage $(m+1)$, the expected utility from allocating the item with type $p_{(j)}$ to the current agent is given by:

$$\begin{aligned} & q_{(j)}v_{m+1} + E_{x_m}\sum_{i=1}^{j-1} q_{(i)}c_{i,m}(v_m) + E_{x_m}\sum_{i=j+1}^{m+1} q_{(i)}c_{i-1,m}(v_m) \\ = \; & q_{(j)}v_{m+1} + \sum_{i=1}^{j-1} q_{(i)}E_{x_m}c_{i,m}(v_m) + \sum_{i=j+1}^{m+1} q_{(i)}E_{x_m}c_{i-1,m}(v_m) \end{aligned}$$

Letting $a_{i,m+1}(v_{m+1}) = E_{x_m}c_{i,m}(v_m)$ and applying the well known theorem (due to Hardy) about the optimality of assortative matching completes the proof, as in the standard private values problem (see Theorem 1).

3-4. By the above claims, $a_{i,k}(v_k)$ represents the stage k value expectation for the agent that will be assigned in the future to the remaining item with the $i-$lowest quality. We now calculate the derivative $\frac{da_{i,k}(v_k)}{dv_k}$. If the remaining item with the $i-$lowest quality will be allocated in stages $k-2, k-3, ..1$, the value of the respective agent does not depend on v_k, and the contribution of these terms to the derivative is thus zero. The remaining item with the $i-$lowest type will be allocated in stage $k-1$ with probability $[F(a_{i,k-1}(v_{k-1})) - F(a_{i-1,k-1}(v_{k-1}))]$. Note that $v_{k-1} = x_{k-1} + \alpha x_k = x_{k-1} + \alpha v_k - \alpha^2 x_{k+1}$, and hence $\frac{dv_{k-1}}{dv_k} = \alpha$. Together, these observations yield:

$$\frac{da_{i,k}(v_k)}{dv_k} = \alpha[F(a_{i,k-1}(v_{k-1})) - F(a_{i-1,k-1}(v_{k-1}))] < 1$$

We obtain that the equation $a_{i,k}(v_k) = v_k$ has a unique solution. The rest of the proof follows as in Theorem 3**.** In particular, consider for each period k the following system of prices, one for each the remaining k objects:

$$\begin{aligned}
t^k_{(1)} &= 0 \\
t^k_{(2)} &= (q^k_{(2)} - q^k_{(1)})a^*_{1,k} \\
&... \\
t^k_{(j)} &= \textstyle\sum_{i=1}^{j-1}(q^k_{(i+1)} - q^k_{(i)})a^*_{i,k} \\
&.... \\
t^k_{(k)} &= \textstyle\sum_{i=1}^{k-1}(q^k_{(i+1)} - q^k_{(i)})a^*_{i,k}
\end{aligned}$$

Faced with the above menu of prices, each arriving agent truthfully reveals his type, and the complete information efficient sequential assignment is implemented. ■

The above construction extends the logic behind dynamic VCG mechanisms to models with interdependent values. The insight behind the extension parallels the one used for the construction of generalized VCG mechanisms for static environments with one-dimensional signals and interdependent values (see for example Dasgupta and Maskin[8] and Jehiel and Moldovanu[16])

9 Conclusion

We have analyzed the conditions under which an efficient allocation of resources is implementable in a dynamic private values setting where the designer gradually learns about the distribution of agents' values. Learning by the designer generates indirect informational externalities which may destroy the incentives for truthful revelation. Efficient implementation is possible only if the belief updating process is not too fast, i.e.,the designer's beliefs do not dramatically change after each new observation. Analogous phenomena occur in a model where the designer is completely informed about the distribution of the agents' values, but these are interdependent. In particular, we can use an insight from that theory (see Jehiel and Moldovanu [16]) to conclude that dynamic efficient implementation is generically impossible in the model with learning if values are private and if agents' signals are multidimensional. Thus, inefficiency is the rule in the general dynamic model with arbitrarily heterogenous objects unless, for each object, the decision over its assignment can be separated from the assignment decisions over other objects.

Theoretically, it is always possible to implement the efficient policy if the designer is able to delay the necessary monetary transfers after all allocation decisions have been made[19]. Thus, for applications, it is important to single out the frameworks where such schemes are feasible and robust.

10 Appendix

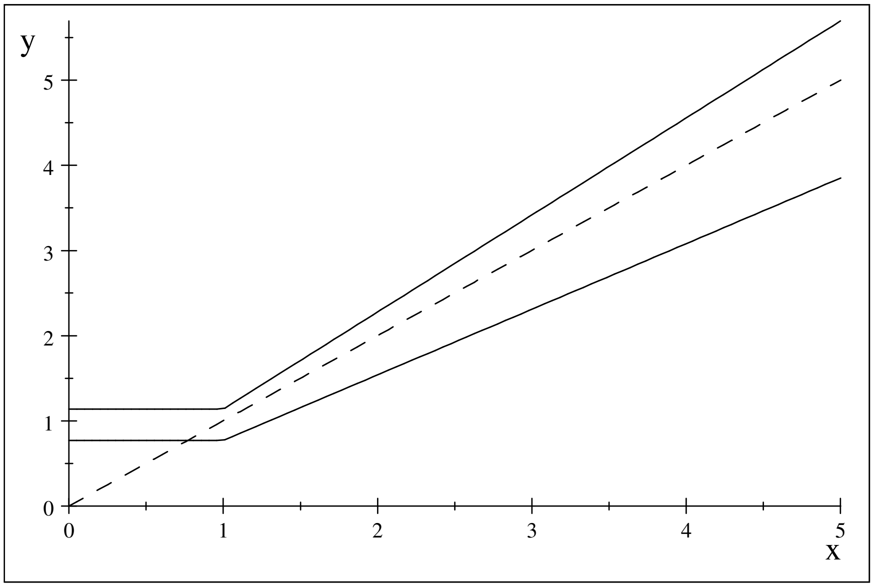
The example below shows that the sufficient condition on derivatives in Theorem 3 is not necessary.

Example 3 (Example 1 continued) *The agents' types distribute uniformly on* $[0, W]$, *with* W *unknown. The designer has a prior belief about* W *given by a Pareto distribution,* $P(\alpha, R)$ *with parameters* α, R , *where* $\alpha > 1$. *Let* $k = 3$, *and assume that the current hyper-parameters of the Pareto distribution describing the current belief about the upper bound* W *are* $R_3 = 1$ *and* $\alpha_3 = 1.1$. *The efficient allocation is described by the following type-dependent cutoffs*

$$\begin{aligned} a_{2,3}(\Phi_2) &= \max\{R_3, x_3\}\left(\frac{\alpha_3+2}{8\alpha_3+8}+\frac{1}{2}+\frac{1}{2\alpha_3}\right) = 1.1391 \max\{1, x_3\} \\ a_{1,3}(\Phi_2) &= \max\{R_3, x_3\}\left(\frac{3\alpha_3+2}{8\alpha_3+8}+\frac{1}{2\alpha_3}\right) = 0.77 \max\{1, x_3\} \end{aligned}$$

The above cutoffs are not everywhere differentiable in x_3, *and* $\frac{\partial}{\partial x_3} a_{2,3}(\Phi_2) > 1$ *for* $x_3 > 1$. *Nevertheless, as the next figure illustrates, there do exist type-independent cutoffs that induce the same allocation (hence allowing first-best implementation).*

[19] The computer science literature distinguishes between "online" and "offline" mechanisms. Delaying payments is possible in the latter form.



In this case, the agent that arrives at period $k = 3$ should never gets the object with quality q_3, because a high type of that agent induces the belief that subsequent agents will have even higher types. The agent gets object q_2 if his type is above 0.77, and gets object q_1 otherwise. The payment scheme that induces truthful revelation is

$$q_3 = \begin{cases} 0 & \text{if } x \in [0, 0.77] \\ 0.77\,(q_2 - q_1) & \text{if } x > 0.77 \end{cases}.$$

References

[1] Aghion, P., Bolton, P., Harris, C. and Jullien, B. (1991): "Optimal Learning by Experimentation", *Review of Economic Studies*, **58**, 621-654.

[2] Albright, S.C. (1977): "A Bayesian Approach to a Generalized House Selling Problem," *Management Science*, **24**(4), 432-440.

[3] Athey, S. and Segal, Y (2007): "An Efficient Dynamic Mechanism", discussion paper, Stanford University.

[4] Bergemann, D. and Välimäki, J. (2007): "Dynamic Marginal Contribution Mechanisms", discussion paper, Yale University.

[5] Bitran, G.R. and Caldentey, R. (2003): "An Overview of Pricing Models for Revenue Management", *Manufacture & Services Operations Management* **5**(3), 203-229.

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

[7] Clarke E. (1971): "Multipart Pricing of Public Goods", *Public Choice* **8**, 19-33.

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

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

[10] Elmaghraby, W. and Keskinocak, P. (2003): "Dynamic Pricing in the Presence of Inventory Considerations: Research Overview, Current Practices, and Future Directions", *Management Science* **49** (10), 1287-1309.

[11] Freixas, X., Guesnerie, R. and Tirole, J. (1985): "Planning under Incomplete Information and the Ratchet Effect", *Review of Economic Studies* **52**, 173-191.

[12] Gershkov, A. and Moldovanu, B. (2007): "The Dynamic Assignment of Heterogenous Objects: A Mechanism Design Approach", *Working paper*, University of Bonn.

[13] Groves T. (1973): "Incentives in Teams", *Econometrica* **41**, 617-631.

[14] Hansen R. (1985): "Auctions with Contingent Payments", *American Economic Review* **75**, 862-865.

[15] Hardy, G., Littlewood, J.E., and Polya, G. (1934): *Inequalities*, Cambridge University Press: Cambridge.

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

[17] Keller, G. and Rady, S. (1999): "Optimal Experimentation in a Changing Environment", *Review of Economic Studies,* **66**, 475-507.

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

[19] Kohn, M. and Shavel, S. (1974): "The Theory of Search", *Journal of Economic Theory,* **9**, 93-123.

[20] Lippman, S. and McCall, J. (1976): "The Economics of Job Search: A Survey", *Economic Inquiry* **14**(2), 155-189.

[21] Mezzetti, C, (2004): "Mechanism Design with Interdependent Valuations: Efficiency", *Econometrica* 72(5), 1617-1627

[22] Mirman, L., Samuelson, L. and Urbano, A. (1993): "Monopoly Experimentation", *International Economic Review,* **34**(3), 549-563.

[23] Morgan, P. (1985): "Distributions of the Duration and Value of Job Search with Learning", *Econometrica,* **53**(5), 1199-1232.

[24] 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)

[25] Vickrey, W. (1961): "Counterspeculation, Auctions and Competitive Sealed Tenders", *Journal of Finance* **16**, 8-37.

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

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

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

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

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

[31] Weitzman, M. (1980): "The "Ratchet Principle" and Performance Incentives", *Bell Journal of Economics* **11**, 302-308.