

# Learning About The Future and Dynamic Efficiency

by

Alex Gershkov and Benny Moldovanu<sup>1</sup>

21.7.2008

Abstract: We study an allocation problem where a set of heterogeneous objects needs to be allocated to agents arriving over time. The basic model is of the private, independent values type. The dynamically efficient allocation is implementable if the distribution of agents' values is known. Whereas lack of knowledge about the distribution is inconsequential in the static case, endogenous informational externalities arise if the designer gradually learns about the distribution by observing present values. These externalities may prevent the implementation of the dynamically 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 point out the role of delayed payments.

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, 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, and that the designer's assessment of future technological possibilities is improved by observing the current technology. If technology owners are strategic, we show below that

---

<sup>1</sup>We wish to thank three anonymous referees for their very helpful comments. We are grateful for financial support from the German Science Foundation via SFB - TR 15, and from the Max Planck Research Prize. Gershkov, Moldovanu: Department of Economics, University of Bonn, Lennestr. 37, 53113 Bonn. Emails: alex.gershkov@uni-bonn.de, mold@uni-bonn.de.

the learning process about the future disturbs the ability to extract the information necessary for efficient decision making. Another good example is hiring by an organization who fills positions at different hierarchy levels over time. While all candidates prefer higher level positions, some aspect of their ability to perform the needed tasks at each level 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. Finally, consider a firm allocating franchises in a country where it did not operate before, and where contractual clauses or other impediments prevent reallocation during a significant period of time. (Local) potential franchisees may have some superior information regarding future profitability in the respective market, and hence the elicitation of private information is crucial for a franchisor that is supposed to get a share of the future profits.

The above examples are typical for a large class of situations where a decision maker faced with a fixed amount of resources needs to distribute them over time to a sequence of randomly arriving opportunities<sup>2</sup>. 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 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, static analysis due to William Vickrey, Edward Clarke and Theodor Groves (Vickrey (1961), Clarke (1971), Groves (1973)). The VCG mechanisms align private and social interests via individual payments that correspond to the externality imposed by an agent's presence on others. For the VCG

---

<sup>2</sup>More generally, the phenomena analyzed here occur whenever a decision maker needs to learn about the future from current, strategic agents while taking irreversible actions that also affect the set of future, feasible decisions. Numerous aspects of public policy and regulation (e.g., environmental decisions) as well as commercial decisions taken under conditions of unknown demand (e.g., by airline, hotel or freight companies) fall within this class.

construction to work, the externality imposed by an agent on others cannot depend on that agent's information. This requirement is fulfilled in the standard *private values* framework.

A generalization of the marginal externality idea allows the implementation of the sequentially efficient allocation also in dynamic, private values settings, as long as the distribution of values is known to the designer (see David C. Parkes and Satinder Singh (2003)). Knowledge of this distribution is crucial since, in order to compute the monetary transfers that align private and social interests, the designer needs to calculate the correct expected externality imposed by an agent over time. Susan Athey and Ilya Segal (2007) and Dirk Bergemann and Juuso Valimäki (2007) 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, that the designer knows the (future) distribution of agents' values is unlikely to be fulfilled in most applications. This assumption is inconsequential in the static case since the VCG mechanism implements the efficient allocation in dominant strategies. When learning about the distribution becomes necessary, the information revealed by an agent plays a dual role: it affects both the value of the current allocation, and the option value of any future allocation. As a consequence, the expected externality imposed by a present agent on others - which is important in order to implement the efficient allocation - becomes dependent on that agent's information. In other words, learning generates *indirect informational externalities*<sup>3</sup>. 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 be satisfied by the dynamic efficient allocation: after observing a present high value, 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

---

<sup>3</sup>The implementation of any given allocation function - not necessarily the efficient one - where the option values depends on the current information will raise similar issues.

which case the designer becomes "sufficiently pessimistic" about the future. Efficient implementation will be possible only if, roughly speaking, the designer's beliefs do not change dramatically after each new observation. This insight translates to a private values dynamic framework with learning an observation made by Eric Maskin (1992) in a static framework with *direct informational externalities* (also called *interdependent values*).

Partha Dasgupta and Eric Maskin (2000) and Philippe Jehiel and Benny Moldovanu (2001) have examined both the possible extensions and limitations of the VCG analysis to static frameworks where the values of agents depend on information held by others. 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<sup>4</sup>. Single crossing translates here into the requirement that beliefs change only gradually.

An important difference between our setting with private values and learning, and a standard setting with interdependent values is that in the former "interdependence" may or may not arise depending on the social choice function to be implemented. With learning, the designer uses today's signal in order to adjust the option value attached to the future. This creates problems for dynamic efficiency because externality payments (needed for VCG mechanisms) involve precisely these option values. This problem may not arise for other social choice functions.

The efficient policy is always implementable if the designer is able to delay monetary transfers till after all observations about types have been made, analogously to insights developed by Claudio Mezzetti (2004) for the static case with interdependent values, and by Athey and Segal (2007) for a dynamic model where private information arrives over time.

The above basic insights are both simple and general. In order not to obscure their simplicity, we present then in the lean and elegant dynamic assignment framework due to Cyrus Derman, Gerald Lieberman and Sheldon Ross (1972) (DLR hereafter). In

---

<sup>4</sup>Thomas Kittsteiner and Benny Moldovanu (2005) used these insights in a dynamic auction model for queueing environments where agents have private information about own processing times.

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 decides which object to assign to the present agent. In their model 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. We first show that the efficient policy characterized by DLR continues to be implementable even if arriving agents have private information about their attributes.<sup>5</sup> (Bayesian) learning in the complete-information DLR model has been analyzed by Christian Albright (1977). In both these papers, the agents do not act strategically, and hence the issues of implementability/monetary payments do not arise. Our main results are presented in a model that adds private information about attributes, and strategic behavior to Albright's learning model.

Ilya Segal (2003) analyzes revenue maximization in a static environment with an unknown distribution of the agents' values. Although he deals with a private values environment, Segal observes that, analogously to what happens in our model, each agent has an informational effect on others. As a consequence, in the revenue maximizing procedure, the terms of trade for a given agent will be affected by information conveyed by others. The type of problems highlighted here cannot occur in Segal's static model since a standard VCG mechanism always leads there to the efficient outcome.

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.<sup>6</sup> In his seminal paper, Michael Rothschild (1974) derived sufficient conditions for the optimal policy to satisfy *the reservation price property, i.e.*, stopping search at any price should imply stopping also at all

---

<sup>5</sup>In contrast to DLR, who focused on dynamic welfare maximization, there is an extensive literature on dynamic revenue maximization (See Talluri and Van Ryzin (2004)). Gershkov and Moldovanu (2008) explore the relationship between revenue maximizing and welfare maximizing policies. Riley and Zeckhauser (1983) considered a single object revenue-maximizing procedure where there is learning about the distribution of the agents' values.

<sup>6</sup>For an extensive survey of the classical search literature, see Steven Lippman and John J. McCall (1976).

more favorable prices. In Rothschild's model, price quotations are not generated by strategic agents, and thus there are no information revelation problems, nor incentive constraints. Thus the monotonicity requirement behind the reservation price property is not due there to an explicit economic constraint, whereas monotonicity is tightly related to incentive compatibility in our model.

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, Leonard J. Mirman et al., (1993) and Godfrey Keller and Sven Rady (1999) analyze optimal experimentation by a monopolist facing an unknown demand<sup>7</sup>, while Xavier Freixas et al., (1985), Martin L. Weitzman (1980) focus on the so-called *ratchet effect*: in a repeated principal-agent interaction, leakage of private information in some period affects the optimal reward scheme at future periods, and the agent choose his present action in order to manipulate the principal's beliefs.<sup>8</sup>

The paper is organized as follows. In Section I we present a non-technical illustration of our main idea. Section II describes the model. In Section III we show that the DLR efficient dynamic allocation can be implemented, via appropriate payments, also in a private values framework with incomplete information about the types of arriving agents (while assuming that the designer knows the distribution of types). In Section IV, the main section, we show that the complete information, efficient dynamic assignment need not be implementable if the designer, who does not know the distribution of types, updates his belief about this distribution after each observation. We also offer several sufficient conditions under which efficient implementation is possible. Finally, we show that the dynamic efficient assignment can always be implemented if all payments can be delayed until the end of the entire allocation process. Section V concludes.

---

<sup>7</sup>The problem of the single decision maker that has to make a sequence of decisions in an unknown environment is also considered in Philippe Aghion et al. (1991). Their main result displays conditions under which, eventually, the decision maker adopts the correct decision.

<sup>8</sup>This is also the subject of a small literature on dynamic regulation where it is shown that *regulatory lags* - the practice whereby information revealed by the firm is not immediately utilized - can be beneficial (see Jean-Jacques Laffont and Jean Tirole (1993))

# I. An Illustration

An indivisible object can be allocated in one of two periods. Agent  $i$ ,  $i = 1, 2$  is present in period  $i$  only and has a valuation  $x_i$  for the object. Both physical and monetary transfers to an agent must be made "online", in the period where they are present (i.e., they cannot be delayed and made contingent on information that may arrive at a later point in time).

Assume first that valuations are private and independently drawn from the interval  $[0, 2]$ , according to a uniform distribution. The complete information dynamic efficient allocation is to 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 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 compute first the complete information efficient dynamic allocation: If the  $x_1 < (>)1$ , the posterior is that  $x_2$  is uniformly distributed on  $[0, 1]$  ( $[1, 2]$ ). This yields

$$E[x_2/x_1] = \begin{cases} 0.5 & \text{if } x_1 < 1 \\ 1 & \text{if } x_1 = 1 \\ 1.5 & \text{if } x_1 > 1 \end{cases}$$

Thus, the first agent should get the object if  $x_1 \in [0.5, 1] \cup [1.5, 2]$ , and the set of types for which agent 1 should get the object in the efficient allocation is not convex. To see that this allocation cannot be implemented, assume that the designer makes a (possibly negative) transfer of  $t_1$  to the first agent if this agent announces  $\hat{x}_1 \in [0.5, 1]$  and a transfer of  $\tau_1$  if this agent announces  $\hat{x}_1 \in [1, 1.5]$ . 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 report of the first agent has a significant impact on the updating process of the designer's belief. The information gained through learning is used to the detriment of the first agent who becomes then reluctant to report truthfully.<sup>9</sup>

The complete information efficient allocation can be implemented if it is possible to delay the transfer to the first agent until after the second period. The following scheme does the job:

1. If  $\hat{x}_1 \in [0.5, 1] \cup [1.5, 2]$  the first agent gets the object and pays  $\hat{x}_2$ .
2. If  $\hat{x}_1 \in [0, 0.5] \cup [1, 1.5]$  the first agent does not get the object and pays nothing.
3. If the object has not been allocated in the first stage it is assigned to the second agent without any further transfers<sup>10</sup>.

In the above scheme, the first agent pays the realized externality, which is independent of his own report. This corrects his incentives to truthfully report his private information.

## II. 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  $x_j$  then this agent enjoys a utility given by  $q_i x_j$ . Getting no

---

<sup>9</sup>The knowledgeable reader will notice the similarity of this example to the original example provided by Maskin (1992) for the static setting without learning, but with interdependent values.

<sup>10</sup>In this example, the second agent is always indifferent between lying and telling the truth, and we assume that he tells the truth.



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 w.l.o.g. that  $m = n$ . While the items' properties  $0 \leq q_1 \leq q_2 \dots \leq q_m$  are assumed to be known, the agents' types are assumed to be independent and identically distributed random variables  $X_i$  on  $[0, +\infty)$  with common c.d.f.  $F$ . We consider two distinct versions of this model:

1. **Incomplete Information:** Arriving agents privately know their type, but the designer only knows the distribution  $F$  from which types are sampled. This yields a standard private values models with independent types.
2. **Incomplete Information + Learning:** Arriving agents privately know their type, and the distribution from which types are sampled,  $F$ , is unknown to the designer. At stage  $n$ , the belief about  $F$  is originally described by a prior distribution  $\Phi = \Phi_n$ , which is updated as successive types are observed. The posterior belief about the distribution of types after observing types  $x_n, \dots, x_m$  is given by  $\Phi_{m-1}(x_n, \dots, x_m)$ . For any possible belief, we assume that the implied marginal distribution of types has a finite mean.

In both models we assume that the entire history is publicly observable.

### III. Dynamic Efficiency under Incomplete Information

Here we analyze 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 allocation policies: without loss of generality, we can restrict attention to direct mechanisms where every agent, upon arrival reports his characteristic  $x_i$  and where the mechanism specifies which item the agent gets, and a payment.<sup>11</sup> Both physical allocation and payment at period  $k$  may, in principle, depend on the signals previously observed.

Let the history at period  $k$ ,  $H_k$ , be the ordered set of all signals reported by the agents that arrived at periods  $n, \dots, k + 1$  and the ordered set of objects allocated to those agents. Let  $\mathcal{H}_k$  be the set of all histories at period  $k$ . 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  $\phi_k : \mathcal{H}_k \times [0, \infty) \rightarrow \Pi_k$  a deterministic allocation policy for period  $k$ , where  $\Pi_k$  is the set of available objects at  $k$ , and denote by  $P_k : \mathcal{H}_k \times [0, \infty) \rightarrow \mathbb{R}$  the associated payment rule. Note that the cardinality of set  $\Pi_k$  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.<sup>12</sup>

**Proposition 1** *Assume that  $q_j \neq q_l$  for any  $q_j, q_l \in \Pi_k, j \neq l$ . A deterministic policy  $\phi_k$  is implementable if and only if there exist  $k + 1$  functions  $0 = y_0(H_k) \leq y_1(H_k) \leq y_2(H_k) \leq \dots \leq y_k(H_k) = \infty$ , such that  $x \in (y_{j-1}(H_k), y_j(H_k)) \Rightarrow \phi_k(H_k, x) = q_{(j)}$*

---

<sup>11</sup>Since agents observe the history, they are better informed after directly observing types. Yet, the argument holds because: 1. 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. 2. This augmented mechanism can be replicated by a direct one.

<sup>12</sup>The used notion of (Bayes-Nash) implementability does not require uniqueness of the equilibrium implementing the given social choice function. But note that in the schemes constructed below agents have dominant strategies.

where  $q_{(j)}$  denotes the  $j$ 'th lowest element of the set  $\Pi_k$ .<sup>13</sup>

**Proof.**  $\implies$  If two reports lead to the same physical allocation, then, in any incentive compatible mechanism, the associated payments should be the same as well. Therefore, a direct mechanism is equivalent here to a scheme 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  $x$  prefers the pair  $(q_m, P_m)$  over any other pair  $(q_l, P_l)$  with  $q_m > q_l$ , then any type  $\tilde{x} > x$  also prefers  $(q_m, P_m)$  over  $(q_l, P_l)$ . This implies that  $\phi_k(H_k, \tilde{x}) \geq \phi_k(H_k, x)$  for any  $k$  and  $H_k$ . Therefore an agent arriving at period  $k$  gets object  $q_{(j)}$  if he reports a type contained in the interval  $(y_{j-1}(H_k), y_j(H_k))$ . A similar argument shows that  $\phi_k(H_k, y_i(H_{k+1})) \in \{q_i, q_{i+1}\}$  for  $i \in \{1, 2, \dots, k\}$ .

$\Leftarrow$  The proof is constructive. Given a partition-based allocation policy, we design a payment scheme  $P_j$  that, for any  $j \in \{1, \dots, k\}$ , will induce type  $x \in (y_{j-1}(H_k), y_j(H_k)]$  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

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

for  $x \in (y_{j-1}(H_k), y_j(H_k)]$  and  $P_1(H_k, x_k) = 0$ . Note that type  $x = y_j(H_k)$  is indifferent between  $(q_{(j)}, P_j)$  and  $(q_{(j+1)}, P_{j+1})$ . Moreover, any type above  $y_j(H_k)$  prefers  $(q_{(j+1)}, P_{j+1})$  over  $(q_{(j)}, P_j)$ , while any type below, prefers  $(q_{(j)}, P_j)$  over  $(q_{(j+1)}, P_{j+1})$ . Therefore, any type  $x \in (y_j(H_k), y_{j+1}(H_k)]$  prefers  $(q_{(j+1)}, P_{j+1})$  over any other pairs in the menu.<sup>14</sup> ■

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

<sup>13</sup>Types at the boundary between two intervals can be assigned to either one of the neighboring elements of the partition. That is  $\phi_t(H_k, y_i(H_k)) \in \{q_i, q_{i+1}\}$ ,  $i = 1, 2, \dots, k-1$ .

<sup>14</sup>The payment given in (1) is not the only one implementing the partition  $0 \leq y_1(H_k) \leq y_2(H_k) \leq \dots \leq y_k(H_k) = \infty$ . 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 the payment for the worst object  $P_1$  should be non-positive.

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.

Derman, Lieberman and Ross (1972) have characterized the allocation policy that maximizes total expected welfare in a complete-information model where the designer gets to observe the type of each arriving agent (monetary incentives for information revelation are not needed then). Using the DLR characterization, and Proposition 1, we obtain:

**Theorem 2**    1. (DLR, 1972) *The complete information, welfare maximizing policy has the following form: Consider the arrival of an agent with type  $x$  in period  $k \geq 1$ ; There exist constants  $0 = a_{0,k} \leq a_{1,k} \leq a_{2,k} \dots \leq a_{k,k} = \infty$  that depend on the distribution  $F$  but not on the  $q$ 's, such that the policy assigns the item with the  $i$  – th smallest quality if  $x \in (a_{i-1,k}, a_{i,k}]$ .*<sup>15</sup>

2. *The first-best policy is implementable under incomplete information.*

The second part follows from Proposition 1 since the welfare maximizing policy identified in the first part is obviously partition-based. Thus, incomplete information *per-se* is not an obstacle towards achieving first-best dynamic efficiency. Note also that the above described policy is history independent: the efficient cutoffs at any stage  $k$  are not affected by previous decisions, nor by previous reports.

## IV. Dynamic Efficiency under Incomplete Information and Learning

In this section we investigate the dynamic welfare maximization problem under the more realistic assumption that the distribution of types is unknown. Each inference

---

<sup>15</sup>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. The optimal policy has a similar form for any supermodular reward function, but then the constants  $a_{i,k}$  may also depend on the  $q$ 's.

about an agent's type reveals now 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 process where present and option values are compared.

Although history dependent allocation rules were not necessary for the implementation of the first best policy without learning, 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, Proposition 1 holds even if the distribution of agents' types is unknown, and if learning takes place. 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. The next result extends the DLR insight to the setup with learning:

**Theorem 3** (*Albright, 1977*)

1. *The complete information, welfare maximizing policy has the following form: Assume that types  $x_n, \dots, x_{k+1}$  have been observed, and consider the arrival of an agent with type  $x_k$  in period  $k \geq 1$ . There exist functions  $0 = a_{0,k}(\Phi_{k-1}(H_k, x_k)) \leq a_{1,k}(\Phi_{k-1}(H_k, x_k)) \leq a_{2,k}(\Phi_{k-1}(H_k, x_k)) \dots \leq a_{k,k}(\Phi_{k-1}(H_k, x_k)) = \infty$ , that do not depend on the  $q$ 's, such that the policy assigns the item with the  $i$ -th smallest quality if  $x_k \in (a_{i-1,k}(\Phi_{k-1}(H_k, x_k)), a_{i,k}(\Phi_{k-1}(H_k, x_k))]$ .*
2. *Each  $a_{i,k+1}(\Phi_k(H_{k+1}, x_{k+1}))$  equals the expected value of the agent's type to which the item with  $i$ -th smallest type is assigned in a problem with  $k$  periods before the period  $k$  signal is observed. These functions are related to each other by the following recursive formulae:*

$$\begin{aligned}
 a_{i,k+1}(\Phi_k(H_{k+1}, x_{k+1})) &= \int_{A_{i,k}} x dG(x) + \int_{\underline{A}_{i,k}} a_{i-1,k}(\Phi_{k-1}(H_k, x_k)) dG(x) \\
 &\quad + \int_{\bar{A}_{i,k}} a_{i,k}(\Phi_{k-1}(H_k, x_k)) dG(x) \tag{2}
 \end{aligned}$$

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

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

$$\bar{A}_{i,k} = \{x_k : x_k > a_{i,k}(\Phi_{k-1}(H_k, x_k))\}.$$

In marked contrast to the case with a known distribution, the efficient policy at stage  $k$  is described in terms of cutoffs that generally depend on the type of the agent arriving at  $k$  because every observation affects the designer's beliefs about the characteristics of forthcoming agents. This feature creates the implementation problems, and the crucial question is whether we can alternatively describe the efficient policy at any stage  $k$  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.<sup>16</sup>

Our 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 4** *Consider the first-best policy, characterized in Theorem 3.*

1) *Assume that  $q_i \neq q_j$  for all  $i, j$  such that  $i \neq j$ <sup>17</sup>. The first-best 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  $A_{i,k} = \{x_k : a_{i-1,k}(\Phi_{k-1}(H_k, x_k)) < x_k \leq a_{i,k}(\Phi_{k-1}(H_k, x_k))\}$  is convex (i.e., it is an interval).*

2) *If for any  $k$ ,  $H_k$  and for any  $i \in \{0, \dots, k\}$ , the cutoff  $a_{i,k}(\Phi_{k-1}(H_k, x_k))$  is differentiable with respect to the signal of the agent arriving at  $k$ , and  $\frac{\partial}{\partial x_k} a_{i,k}(\Phi_{k-1}(H_k, x_k)) < 1$  for any  $x_k$  and  $H_k$ , then, the first-best policy can be implemented also under incomplete information.<sup>18</sup>*

**Proof.** By Theorem 3, the current agent  $k$  gets the object with quality  $q_i$  if and only if his type belongs to  $A_{i,k} = \{x_k : a_{i-1,k}(\Phi_{k-1}(H_k, x_k)) < x_k \leq a_{i,k}(\Phi_{k-1}(H_k, x_k))\}$ .

---

<sup>16</sup>Morgan (1985) focused on the existence of fixed points that do not depend on current information for the determination of the reservation value function in a complete information, optimal search model with learning. The goal was to find simpler ex-ante calculable formulas for probabilities of search duration and for the expected value of search

<sup>17</sup>This assumption excludes here cases where the efficient allocation from some stage on does not depend on the revealed information and can therefore be trivially implemented.

<sup>18</sup>These are only sufficient but not necessary conditions for implementability.

The first claim follows then immediately from Proposition 1: by Theorem 3, if the sets  $A_{i,k}$  are convex for any  $i$ , then for any  $m, l \in \{1, \dots, k\}$  with  $m > l$  it must hold that  $\inf\{x : x \in A_{m,k}\} \geq \sup\{x : x \in A_{l,k}\}$ . In other words, the intervals  $A_{i,k}$  are ordered.

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  $\Phi_{k-1}$ . The inequality  $\frac{\partial}{\partial x_k} a_{i,k}(\Phi_{k-1}(H_k, x_k)) < 1$  implies that the equation  $x = a_{i,k}(\Phi_{k-1}(H_k, x))$  has at most one solution. Let  $a_{i,k}^*(H_k)$  denote this solution whenever it exists, and let  $a_{i,k}^*(H_k) = \infty$  otherwise. By Theorem 3 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  $\{a_{i,k}^*(H_k)\}_{i=0}^k$  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

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

The result follows from the first part. ■

Recall that in the limiting case of no learning - where we know that the efficient allocation is implementable - we have  $\frac{\partial}{\partial x_k} a_{i,k}(\Phi_{k-1}) = 0$ . Similarly, if the cutoffs  $a_{i,k}(\Phi_{k-1}(H_k, x_k))$  are only slightly affected by the information conveyed by the signal  $x_k$  (i.e., the current optimal allocation doesn't significantly react to new information) then the efficient allocation can be implemented.<sup>19</sup> Examples where the resulting sets  $A_{i,k}$  are convex include: 1. Agents' types are uniformly distributed on the interval  $[0, W]$ , and the designers' prior about  $W$  is a Pareto distribution  $P(\alpha, R)$  with  $\alpha > 1$ ; 2. Agents' types distribute according to a Gamma distribution  $Gamma(\alpha, \beta)$  with unknown rate parameter  $\beta$ , and the designer's beliefs about  $\beta$  are given by another Gamma distribution  $Gamma(\gamma, \delta)$ .<sup>20</sup>

<sup>19</sup>The first-best policy can be always implemented (using type-independent transfers) if the sequence of beliefs  $\Phi = \Phi_n, \Phi_{n-1}, \dots, \Phi_1$  induces successive marginal distributions of types that form a sub(super)-martingale. Then, the efficient policy is to allocate the items successively in ascending (descending) order of types.

<sup>20</sup>Both these examples assume *conjugate priors* given Bayesian learning. For more details, see

The above analysis has assumed that agents' types are independently distributed. Even if the distribution of types is known, qualitatively similar dynamic implementation problems occur if types are correlated. In a static environment with correlated types - where the efficient allocation can always be implemented - Jacques Cremer and Richard McLean (1985), (1988) show how the designer can extract the entire surplus by using payments contingent on the information revealed by other agents.<sup>21</sup> As shown in Section I, the ability to delay payments confers more flexibility precisely because it enables the use of contingent payments. This contrasts the insight for dynamic independent private values frameworks without learning where delaying payments does not improve the ability to achieve the dynamically efficient outcome (see Parkes and Singh (2003)).

Assume then that the designer still needs to allocate the objects upon arrival, but that he is able to delay payments until the end of stage one, 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 3. 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  $x_{(1)}, \dots, x_{(n)}$  represent the ordered set of the *revealed* types, which is fully known at the end of stage one; if he gets object  $q_{(1)}$ , he pays zero. It is easy to see that reporting truthfully is an equilibrium. Thus, we obtain<sup>22</sup>:

**Proposition 5** *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 above may be problematic in real-life situations where arrivals are separated by significant periods of

---

Albright (1977).

<sup>21</sup>Note though that in our framework with independent types, full surplus extraction is impossible even if payments are delayed. For independent types, the role of posterior information and contingent payments towards increasing revenue from sales has been pointed out by Robert G. Hansen (1985). Mezzetti (2007) studies surplus extraction with interdependent values.

<sup>22</sup>An analogous result for a much more general model has been proven by Athey and Segal (2007) - see their Proposition 1.



time because: 1) Early agents may not even "exist" when later ones arrive; 2) It may be difficult to write contracts that cover numerous contingencies covering a distant, uncertain future, and to execute monetary payments based on events that lie in the distant past; 3) The last agent's information 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<sup>23</sup>. 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.

## V. 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 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 Jehiel and Moldovanu (2001) 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 heterogeneous objects unless, for each object, the decision over its assignment can be separated from the assignment decisions over other objects. Theoretically, it is possible to implement the efficient policy if the designer is able to delay the necessary monetary transfers after all allocation decisions have been made. Thus, for applications, it is important to single out the frameworks where such schemes are feasible and robust. Finally, recall that we assumed the entire history to be publicly observable. Even if

---

<sup>23</sup>See also Mezzetti (2004).

we allow the designer to withhold information about the history (which may be advantageous), inefficiencies continue to occur if he needs to learn about the distribution of values.

## References

- Aghion, Philippe, Patrick Bolton, Christopher Harris and Bruno Jullien. 1991: "Optimal Learning by Experimentation." *Review of Economic Studies*, 58(4): 621-654.
- Albright, Christian. 1977: "A Bayesian Approach to a Generalized House Selling Problem." *Management Science*, 24(4): 432-440.
- Athey, Susan, and Ilya Segal. 2007: "An Efficient Dynamic Mechanism." discussion paper, Stanford University.
- Bergemann, Dirk, and Juuso Välimäki. 2007: "Dynamic Marginal Contribution Mechanisms." Cowles Foundation Discussion Paper 1616.
- Clarke, Edward. 1971: "Multipart Pricing of Public Goods." *Public Choice*, 11(1): 19-33.
- Cremer, Jacques, and Richard P. McLean. 1985: "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent." *Econometrica*, 53(2): 345-362
- Cremer, Jacques, and Richard P. McLean. 1988: "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." *Econometrica*, 56(6): 1247-1257.
- Dasgupta, Partha, and Eric Maskin. 2000: "Efficient Auctions." *Quarterly Journal of Economics*, 115(2): 341-388.
- Derman, Cyrus, Gerald J. Lieberman, and Sheldon M. Ross. 1972: "A Sequential Stochastic Assignment Problem." *Management Science*, 18(7): 349-355.

- Freixas, Xavier, Roger Guesnerie and Jean Tirole. 1985: "Planning under Incomplete Information and the Ratchet Effect", *Review of Economic Studies*, 52(2): 173-191.
- Gershkov, Alex, and Benny Moldovanu. 2008: "Dynamic Revenue Maximization with Heterogenous Objects: A Mechanism Design Approach", *Working paper*, University of Bonn.
- Groves, Theodore. 1973: "Incentives in Teams." *Econometrica*, 41(4): 617-631.
- Hansen, Robert G. 1985: "Auctions with Contingent Payments." *American Economic Review*, 75(4): 862-865.
- Jehiel, Philippe, and Benny Moldovanu. 2001: "Efficient Design with Interdependent Valuations." *Econometrica*, 69(5): 1237-1259.
- Keller, Godfrey, and Sven Rady. 1999: "Optimal Experimentation in a Changing Environment." *Review of Economic Studies*, 66(3): 475-507.
- Kittsteiner, Thomas, and Benny Moldovanu. 2005: "Priority Auctions and Queue Disciplines that Depend on Processing Time", *Management Science*, 51(2), 236-248.
- Laffont, Jean-Jacques, and Jean Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*, Cambridge: MIT Press.
- Lippman, Steven, and John J. McCall. 1976: "The Economics of Job Search: A Survey." *Economic Inquiry*, 14(2): 155-189.
- Maskin, Eric. 1992. "Auctions and Privatization." in *Privatization*, ed. Horst Siebert, 115-136. J.C.B. Mohr Publisher.
- Mezzetti, Claudio. 2004: "Mechanism Design with Interdependent Valuations: Efficiency." *Econometrica*, 72(5): 1617-1627.
- Mezzetti, Claudio. 2007: "Mechanism Design with Interdependent Valuations: Surplus Extraction." *Economic Theory* 31(3): 473-488.
- Mirman, Leonard J., Larry Samuelson and Amparo Urbano. 1993: "Monopoly Experimentation." *International Economic Review*, 34(3): 549-563.

- Morgan, Peter B 1985: "Distributions of the Duration and Value of Job Search with Learning." *Econometrica*, 53(5): 1199-1232.
- Parkes, David C., and Satinder Singh. 2003: "An MDP-Based Approach to Online Mechanism Design." *Proceedings of 17th Annual Conference on Neural Information Processing Systems (NIPS 03)*
- Riley, John, and Richard Zeckhauser. 1983: "Optimal Selling Strategies: When to Haggle, When to Hold Firm." *The Quarterly Journal of Economics*, 98(2): 267-289.
- Rothschild, Michael. 1974: "Searching for the Lowest Price When the Distribution of Prices Is Unknown." *Journal of Political Economy*, 82(4): 689-711.
- Segal, Ilya. 2003: "Optimal Pricing Mechanisms with Unknown Demand." *American Economic Review*, 93(3), 509-529.
- Talluri, Kalyan T., and Garrett J. van Ryzin. 2004: *The Theory and Practice of Revenue Management*, Springer: New York.
- Vickrey, William. 1961: "Counterspeculation, Auctions and Competitive Sealed Tenders." *Journal of Finance* 16(1). 8-37.
- Weitzman, Martin L. 1980: "The "Ratchet Principle" and Performance Incentives." *Bell Journal of Economics*, 11(1), 302-308.