Economics Letters 118 (2013) 121–125

Contents lists available at SciVerse ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Non-Bayesian optimal search and dynamic implementation

Alex Gershkov [a,b,*], Benny Moldovanu [c]

[a] Department of Economics, Hebrew University of Jerusalem, Israel
[b] School of Economics, University of Surrey, United Kingdom
[c] Department of Economics, University of Bonn, Germany

ARTICLE INFO

Article history:
Received 2 August 2012
Received in revised form
11 September 2012
Accepted 21 September 2012
Available online 5 October 2012

JEL classification:
D4
D8

Keywords:
Dynamic mechanism design
Optimal stopping
Learning

ABSTRACT

We show that a non-Bayesian learning procedure leads to very permissive implementation results concerning the efficient allocation of resources in a dynamic environment where impatient, privately informed agents arrive over time, and where the designer gradually learns about the distribution of agents' values. This contrasts the rather restrictive results that have been obtained for Bayesian learning in the same environment, and highlights the role of the learning procedure in dynamic mechanism design problems.

1. Introduction

We analyze the implementation of the efficient dynamic policy in a model where impatient, privately informed agents arrive over time, and where the designer gradually learns about the distribution of agents' values using a non-Bayesian updating procedure.

The complexities of Bayesian updating make Bayesian updating rules impossible to implement in practical applications: agents often rely on simpler, non-Bayesian heuristics for updating their beliefs. We show that a simple, non-Bayesian updating procedure that was used in the classical search literature leads to very permissive implementation results, contrasting the rather restrictive results that have been obtained for Bayesian learning in the same mechanism design environment.

Our study highlights the role of learning in dynamic mechanism design problems, and adds a new dimension that is mostly lacking both in the classical (static) mechanism design theory, and in the more recent literature on dynamic mechanism design. In particular, we show that under-reaction to new information has important consequences in frameworks where the information is endogenously generated by strategic agents (see Epstein et al. (2008) for a study about the role of under-reaction in a dynamic setting where new information is exogenously generated).

The present allocation model is based on a classical model due to Derman et al. (1972) (DLR hereafter): a finite set of heterogenous and commonly ranked objects is assigned to a set of agents who arrive one at a time. After each arrival, the designer decides which object (if any) to assign to the present agent. Both the attribute of the present agent (that determines his value for the various available objects) and the future distribution of attributes are known to the designer in the DLR analysis.

Learning about future values in the complete-information DLR model has been analyzed by Albright (1977). Gershkov and Moldovanu (2009) added incomplete information to Albright's model, and derived an implicit condition ensuring that efficient implementation is possible. Roughly speaking, implementation is possible if the impact of currently revealed information on today's values is higher than the impact on option values. This insight replaces in the dynamic framework with learning the *single-crossing* idea appearing in the theory of static efficient implementation with interdependent values.[1] Gershkov and Moldovanu (2012) offered conditions on the exogenous parameters of the model – the prior beliefs about the agents' values – that allow efficient implementation. Since these conditions are relatively restrictive, they

* Corresponding author at: Department of Economics, Hebrew University of Jerusalem, Israel.
E-mail addresses: alexg@huji.ac.il (A. Gershkov), mold@uni-bonn.de (B. Moldovanu).

[1] See for example Dasgupta and Maskin (2000) and Jehiel and Moldovanu (2001) who analyzed static models with direct informational externalities.

0165-1765/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2012.09.026

also characterized the incentive-efficient, second-best mechanism based on Bayesian learning.

In the present paper we study an adaptive, non-Bayesian learning process that have been used in a search framework (see Rothschild, 1974) by Bickhchandani and Sharma (1996). This process constructs a posterior that is a convex combination of a prior and the empirical distribution, with more and more weight given to the empirical distribution. The process is consistent in the sense that it uniformly converges to the true distribution as the number of observations goes to infinity. This is a consequence of the Glivenko–Cantelli Theorem.

We prove that, with this process, the efficient allocation is always implementable since new information is incorporated in option values at a slow rate, so that the impact of new information on present values is always higher. We keep all other assumptions of traditional dynamic maximization. In particular, our designer maximizes expected utility given current beliefs and is aware of future biases due to learning. Given Bayes' Theorem, this means that dynamic inconsistencies may appear along the way—this is a common feature of models that relax some aspect of Bayesian learning (see also Epstein et al., 2008).

A word of caution is needed here: Our result does not imply that the non-Bayesian procedure is "better" than Bayesian updating! Since expectations depend on the learning process, a non-Bayesian designer will generally prefer a different policy than a Bayesian one. Thus, our result just says that the complete information efficient allocation – whose calculation proceeds *given* an assumed learning procedure – can always be implemented for the particular adaptive process studied here. Examples will illustrate this issue.

The paper is organized as follows: In Section 2 we present the dynamic allocation and learning model. In Section 3 we first recall two results: 1. The characterization of the efficient allocation policy under complete information due to Albright (1977); 2. An implicit condition on the structure of the efficient policy ensuring that this policy can be implemented also under incomplete information, due to Gershkov and Moldovanu (2009). In Section 4 we focus on the non-Bayesian learning model. Theorem 2 shows that, given the learning model, the implicit condition is always satisfied, and hence the corresponding efficient allocation policy is always implementable. Section 5 concludes. The proof of the theorem is relegated to an Appendix.

2. The model

There are m items and n agents. Each item i is characterized by a "quality" q_i, and each agent j is characterized by a "type" x_j. If an item with quality $q_i \geq 0$ is assigned to an agent with type x_j and this agent is asked to pay p, then this agent enjoys a utility given by $q_i x_j - p$. Getting no item generates a utility of zero. The goal is to find an assignment that maximizes total welfare.

Agents arrive sequentially, one agent per period of time, and each agent can transact (in both physical and monetary terms) only upon arrival.

Note that 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 (*assortative matching)*.

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

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

The function F is not known to the designer nor to the agents. At the beginning of the allocation process the designer has a prior Φ_n over possible distribution functions, and he updates his beliefs after each additional observation. Denote by $\Phi_k\,(x_n, \ldots, x_{k+1})$ the designer's beliefs about the distribution function F after observing types $x_n, \ldots, x_{k+1}$. Given such beliefs, let $\widetilde{F}_k(x|x_n, \ldots, x_{k+1})$ denote the distribution of the next type x_k, conditional on observing $x_n, \ldots, x_{k+1}$. Finally, we assume that each agent, upon arrival observes the whole history of the previous play.

3. The dynamic efficient allocation

Albright (1977) derived the efficient dynamic policy under complete information, i.e., when the agent's type is revealed to the designer upon the agent's arrival. The efficient allocation maximizes, at each decision period, the sum of the expected utilities of all agents, given all the information available at that period. That is, the designer solves the following recursive maximization problem: if at period k the set of objects still available for allocation is Π_k, the designer solves

$$\max_{q_i \in \Pi_k} [q_i \cdot x_k + V_{k-1}(\Pi_k \setminus \{q_i\} \mid x_n, \ldots, x_k)] \tag{1}$$

where $V_{k-1}(\Pi_k \setminus \{q_i\} \mid x_n, \ldots, x_k)$ denotes the expected utility from the optimal future allocation of the remaining inventory $\Pi_k \setminus \{q_i\}$ given that the designer has already observed types $x_n, \ldots, x_k$.

It is important to note that, due to the presence of learning about the uncertain environment, the expectation V_{k-1} is determined by the *prior beliefs*, by the agents' *types observed so far*, and by the *belief updating process*. Thus, a non-Bayesian designer will generally prefer a different policy than a Bayesian one.

Gershkov and Moldovanu (2009) displayed an implicit sufficient condition on these cutoffs ensuring that the efficient allocation is implementable also under incomplete information. These observations are gathered in the next theorem.

Let the history at period k, H_k, be the ordered set of all signals reported by the agents that arrived at periods $n, \ldots, k+1$, and of allocations to those agents. Let $\mathcal{H}_k$ be the set of all histories at period k. Denote by χ_k the ordered set of signals reported by the agents that arrived at periods $n, \ldots, k+1$.

Theorem 1. 1. *Albright (1977) Assume that types $x_n, \ldots, x_{k+1}$ have been observed, and consider the arrival of an agent with type x_k in period $k \geq 1$. There exist functions $0 = a_{0,k}(\chi_k, x_k) \leq a_{1,k}(\chi_k, x_k) \leq a_{2,k}(\chi_k, x_k) \cdots \leq a_{k,k}(\chi_k, x_k) = \infty$ such that the efficient dynamic policy – which maximizes the expected value of the total reward – assigns the item with the i-th smallest type if $x_k \in (a_{i-1,k}(\chi_k, x_k), a_{i,k}(\chi_k, x_k)]$. The functions $a_{i,k}(\chi_k, x_k)$ do not depend on the q's.*

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

$$\begin{aligned} a_{i,k+1}(\chi_{k+1}, x_{k+1}) &= \int_{A_{i,k}} x_k d\widetilde{F}_k(x_k|\chi_{k+1}, x_{k+1}) \\ &\quad + \int_{\underline{A}_{i,k}} a_{i-1,k}(\chi_k, x_k) d\widetilde{F}_k(x_k|\chi_{k+1}, x_{k+1}) \\ &\quad + \int_{\overline{A}_{i,k}} a_{i,k}(\chi_k, x_k) d\widetilde{F}_k(x_k|\chi_{k+1}, x_{k+1}) \end{aligned} \tag{2}$$

where[2]

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

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