Learning in Bayesian Regulation

Semih Koray and Ismail Saglam

Bogazici University

April 2005

Learning in Bayesian Regulation*

Semih Koray† and Ismail Saglam‡

† Department of Economics, Bilkent University, Bilkent 06800, Ankara, Turkey.
E-mail: ksemih@bilkent.edu.tr

‡ Corresponding Author. Department of Economics, Bogazici University, Bebek 34342, Istanbul, Turkey. E-mail: saglamis@boun.edu.tr

*The authors thank Leonid Hurwicz, Murat Sertel, and seminar participants at Bilkent University and Bogazici University. Saglam also acknowledges the support of Turkish Academy of Sciences, in the framework of 'Distinguished Young Scientist Award Program' (TUBA-GEBIP-2004). The usual disclaimer applies.

Learning in Bayesian Regulation

Semih Koray and Ismail Saglam

We examine the issue of learning in a generalized principal-agent model with incomplete information. We show that there are situations in which the agent prefers a Bayesian regulator to have more information about his private type. Moreover, the outcome of the Bayesian mechanism regulating the agent is path-dependent; i.e. the convergence of the regulator's belief to the truth does not always yield the complete information outcome.

Keywords: Learning, Principle-Agent Model, Bayesian Regulation, Incomplete Information Learning.

1 Introduction

The issue of learning has occupied an important place in the recent literature of game theory while most of the pioneering studies have focused on learning in repeated games with incomplete information. For example, Jordan (1991) considers a noncooperative normal form game where each player is endowed with full Bayesian rationality and has prior beliefs about his opponents' privately known payoffs. The Bayesian Nash equilibrium of this game needs not coincide with the Nash equilibrium of the complete information (true) game. However, Jordan shows that under certain restrictions on beliefs the players in a repeated play of the described normal form game can learn to play the Nash equilibrium of the complete information game even though they will not necessarily attain complete information. Kalai and Lehrer (1993) and Blume and Easley (1995) obtain a similar convergence result for infinitely repeated games that involve non-myopic players. The empirical evaluations of the Jordan's Bayesian learning model was recently evaluated in Cox, Shachat and Walker (2001), which shows that when the true game had a unique pure strategy equilibrium, the experimental subjects' play converged to the equilibrium, while this was not the case if the true game had multiple equilibria.

In the existing literature, learning occurs while each player maximizes his infinite horizon expected utility and updates his prior beliefs using the Bayes rule. Here, in this paper we examine the issue of Bayesian learning as a direct goal of (one of the) players in a static decision problem and ask the following questions: in a principal-agent model of regulation with incomplete information, (i) how is 'more information' described in a situation of 'incomplete' learning where the belief of the regulator about the regulated agent does not convergence to the truth? (ii) is 'more information' about the regulated agent always desirable for the regulator and the principal or, conversely, undesirable for the regulated agent? (iii) can all modes of convergence of beliefs yield the (equilibrium) outcome of the complete information game?

We consider a generalized principal-agent model that borrows from Guesnerie and Laffont (1984). The model involves three players: a principal, an agent, and a regulator. The agent has a private type parameter which is unobservable to the other players. However, this parameter is commonly known to lie in a finite interval of reals. Both the principal and the agent derive a transferrable utility from a decision of the agent. The environment is economic in the sense that the most preferred decision differ across the principal and the agent, hence the regulation of the economic activity from the viewpoint of social equity is inevitable.[1]

Under the described incomplete information structure the regulator without loss of generality restricts himself, by the Revelation Principle (Gibbard (1973) and Myerson (1979)), to direct revelation mechanisms which induce the agent to truthfully report his private information. The optimal regulatory mechanism involves a type-dependent subsidy from the principal to the agent. The objective function of the regulator is the social welfare defined as a weighted sum of the agent's and the principal's utilities net of the subsidy taken/given, with the weight assigned to the welfare of the agent not exceeding that of the principal. The aim of the regulator is to choose the optimal decision for each possible type of the agent conditional upon that the mechanism must be incentive-compatible and individually rational for the agent. At the optimal solution there exists an inavoidable (ex-post) welfare loss owing to the positive informational rents that should be guaranteed to the agent, whose net utility is disfavoured by the social welfare function. Since this welfare loss depends on the regulator's prior belief as well as its support, there exists an incentive for the regulator to learn about the private information of the agent.

Leaving aside the epistemological questions searching for the source of the 'new information', we consider learning situations in which the regulator is able to confine the true type of the agent confidently to a smaller support prior to regulation. That is, learning in our model involves a new belief, obtained after finite or infinite number of revisions (stages) and defined on a narrower support that contains the truth. To explicitly define complete learning, we first describe the complete information belief together with the social welfare it induces. Since this extreme belief, which is indeed a dirac function at the truth, is not admissible in our incomplete information setup, we assume that learning is incomplete in a finite-stage setup. So we examine learning by distinguishing two benchmark cases: single-stage incomplete learning and infinite-stage complete learning.

For single-stage learning, we first answer what valuable or more informa-

[1]Note that our characterization of the principal-agent problem, in general, differs from a classical one, where the principal coincides with the regulator whose goal is to maximize the principal's welfare.

tion means. We show that a regulator can ascertain that he has obtained more information only if he has updated his beliefs over the learned finer support using the Bayes rule. We then examine two cases in which the support of a Bayesian regulator's belief shrinks from the left or the right. In the former case, more information about the regulated agent is interestingly found to increase his own welfare. This gives rise to the truthful signalling of the agent about his type space before the implementation of the mechanism. In the latter case where the original support shrinks from the right, it is the social welfare that is found to increase. For intermediate cases in which the support of the regulator's belief shrinks from both sides, the effect of learning on the social welfare is ambiguous. In that case, we may assume that the Bayesian regulator or the society does not have a clear incentive to do learning.[2]

Finally, our paper considers infinite-stage complete learning with an infinite sequence of revisions of the prior belief over their respective supports converging to the complete information belief. Here we define two modes of convergence: absolute (strong form) convergence and pseudo (weak form) convergence. An interesting result we establish is that the outcome of the Bayesian regulatory mechanism is path-dependent. That is, a sequence of updated beliefs approaching to the truth ensures complete information outcome only if the mode of convergence is sufficiently strong.

The organization of the paper is as follows: Section 2 introduces the Bayesian regulation model. Section 3 introduces some preliminaries on learning and presents our results. Finally, Section 4 concludes.

2 Model

Consider two players with quasi-linear utility functions

$$u_p(x, t, \theta) = V_p(x, \theta) - t, \tag{1}$$

$$u_a(x, t, \theta) = V_a(x, \theta) + t, \tag{2}$$

[2]The potential for non-learning has also been shown by Kiefer and Nyarko (1989) in a setting of optimal control of an unknown linear process and by Feldman and McLennan (1989) in a repeated statistical decision problem with normal disturbances.

where V_p and V_a (u_p and u_a) stand for the utilities (net utilities) of the principal and the agent, respectively. Here, θ is the agent's private information about his utility function, x is called a decision and t is the total monetary transfer from the principal to the agent.

The private type parameter θ of the agent is commonly known to lie in some closed real continuum Θ. Define $\theta_1 = \max(\Theta)$ and $\theta_0 = \min(\Theta)$.

The assumptions about the utility functions V_p and V_a are:

A0. $\text{argmax}_x V_p(x,\theta) \neq \text{argmax}_x V_a(x,\theta)$

A1. $\partial(V_p + V_a)/\partial x > 0$

A2. $\partial^2(V_p + V_a)/\partial x^2 < 0$

A3. $\partial^2 V_p/\partial x \partial\theta \leq 0$

A4. $\partial^2 V_a/\partial x \partial\theta \leq 0$

A5. $\partial V_a/\partial\theta < 0$

A6. $\partial^3 V_a/\partial x \partial\theta^2 \leq 0$

A7. $\partial^3 V_a/\partial x^2 \partial\theta \leq 0$

The regulator announces a contract between the principal and the agent. The instruments of the contract are the control of the decision x and the transfer t to the agent. By the Revelation Principle, the regulator can restrict himself to direct revelation mechanisms which ask the agent to report his private information and which give to the agent no incentive to lie.

The optimal regulatory policy is designed to satisfy two conditions first of which is that the agent must never expect a greater net utility by misreporting than he could by truthfully reporting his private information:

$$\text{(IC)} \quad u_a(x(\theta), t(\theta), \theta) \geq u_a(x(\hat{\theta}), t(\hat{\theta}), \theta), \text{ for all } \theta, \hat{\theta} \in \Theta \tag{3}$$

The second condition is that the regulator can never force the agent to be regulated unless the agent obtains a nonnegative net utility:

$$\text{(IR)} \quad u_a(x(\theta), t(\theta), \theta) \geq 0, \text{ for all } \theta \in \Theta \tag{4}$$

Let $U_a(\theta, \hat{\theta})$ denote the net utility of the agent when he reports his private parameter as $\hat{\theta}$ while θ is the actual parameter. Condition (IC) implies that $U_a(\theta, \theta) = U_a(\theta)$ satisfies

$$U_a(\theta) = \max_{\hat{\theta}\in\Theta} \; u_a(x(\hat{\theta}), t(\hat{\theta}), \theta) = u_a(x(\theta), t(\theta), \theta) \tag{5}$$

for all $\theta \in \Theta$. From the envelope theorem, we obtain

$$\frac{dU_a}{d\theta} = \frac{\partial u_a}{\partial \theta} = \frac{\partial V_a}{\partial \theta}. \tag{6}$$

Similarly, denote by $U_p(\theta)$ the net utility of the principal when the agent truthfully reports his private parameter as θ.

The social welfare $W(\theta)$ is defined as the sum of the principal's net utility and a fraction of the agent's net utility:

$$W(\theta) = U_p(\theta) + \alpha U_a(\theta), \tag{7}$$

where $\alpha \in [0, 1]$ is the relative weight assigned to the net utility of the agent. Integrating (6) using the assumption (A5) yields

$$U_a(\theta) = -\int_{\theta}^{\theta_1} \frac{\partial}{\partial \tilde{\theta}} V_a(x(\tilde{\theta}), \tilde{\theta}) d\tilde{\theta}. \tag{8}$$

Inserting $U_p(\theta) = V_p(x(\theta), \theta) - t(\theta)$ and $t(\theta) = U_a(\theta) - V_a(\theta)$ into (7), the actual social welfare becomes:

$$W(\theta) = V_p(x(\theta), \theta) + V_a(x(\theta), \theta) + (1 - \alpha) \int_{\theta}^{\theta_1} \frac{\partial}{\partial \tilde{\theta}} V_a(x(\tilde{\theta}), \tilde{\theta}) d\tilde{\theta} \tag{9}$$

Assumptions 6 and 7 are sufficent for the optimal decision $x(.)$, if exists, to be nonincreasing and implemented by the described subsidy mechanism. However, it is known that there exists no feasible solution $x(.)$ that maximizes (9) unless the two players' welfares are equally weighted in the social welfare function or that the utility of the agent is seperable in its two arguments. The common remedy is simply to have a Bayesian regulator.

We consider a Borel field $\mathcal{T}^{\Theta}$ on the type space Θ and regard the subset $\mathcal{A}^{\Theta}$ of probability measures on $\mathcal{T}^{\Theta}$ with densities that are strictly positive at each element of Θ as the set of admissible prior beliefs for the regulator.

Let $f \in \mathcal{A}^\Theta$ be the prior belief of the regulator and F be the respective cumulative distribution function. We assume that f becomes common knowledge before the regulator asks the agent to report his type.

Let the pair (f, Θ) denote the information structure that is commonly known by all parties in the society.

The objective function of the regulator under the information structure (f, Θ) is the expected social welfare:

$$\int_{\theta_0}^{\theta_1} \left(V_p(x(\theta), \theta) + V_a(x(\theta), \theta) + (1-\alpha) \int_{\theta}^{\theta_1} \frac{\partial}{\partial \tilde{\theta}} V_a(x(\tilde{\theta}), \tilde{\theta}) d\tilde{\theta} \right) f(\theta) d\theta \tag{10}$$

Modifying (10), we obtain the problem of the Bayesian regulator as:

$$\max_{x(.)} \int_{\theta_0}^{\theta_1} \Bigg(V_p(x(\theta), \theta) + V_a(x(\theta), \theta)$$

$$+ (1-\alpha) \frac{F(\theta)}{f(\theta)} \frac{\partial}{\partial \theta} V_a(x(\theta), \theta) \Bigg) f(\theta) d\theta \tag{11}$$

s.t. (IC) and (IR)

To simplify the solution and the analysis of the Bayesian regulatory mechanism, we henceforth assume that for all $\Theta \subset I\!\!R$ and $f \in \mathcal{A}^\Theta$:

A8. $F(\theta)/f(\theta)$ is nondecreasing in θ

Proposition 1. *The solution to Bayesian regulation problem (2) satisfies*

$$\frac{\partial V_p}{\partial x} + \frac{\partial V_a}{\partial x} = -(1-\alpha) \frac{F(\theta)}{f(\theta)} \frac{\partial^2 V_a}{\partial x \partial \theta}. \tag{12}$$

Proof. See Appendix.

We henceforth assume $\alpha \in [0, 1)$ and $\partial^2 V_a / \partial x \partial \theta < 0$ so as to be in the Bayesian framework where the beliefs of the regulator affects the optimal program (12) through the term $F(\theta)/f(\theta)$, so called "the inverse of the reverse hazard rate".

Let $\bar{x}^f$ denote the solution to (12), and let $\bar{U}_p^f(\theta), \bar{V}_p^f(\bar{x}^f(\theta), \theta), \bar{U}_a^f(\theta)$, $\bar{V}_a^f(\bar{x}^f(\theta), \theta)$, $\bar{t}^f(\theta)$, and $\bar{W}^f(\theta)$ respectively denote the net and gross utilities of the principal and the agent, the subsidy and the social welfare evaluated at the report $\theta \in \Theta$ under the belief $f(.)$.

3 Learning the Unknown: Preliminaries

3.1 Dominance of Beliefs

We first define a dominance relation over beliefs to compare the regulatory outcomes that these beliefs lead to.

Definition 1. Let $f_1 \in \mathcal{A}^{\Theta_1}$ and $f_2 \in \mathcal{A}^{\Theta_2}$, where $\Theta_1, \Theta_2 \subset \Theta$. The belief f_1 stochastically dominates (in inverse of the reverse hazard rate) the belief f_2 on $\Theta_1 \cap \Theta_2$ if $F_1(\theta)/f_1(\theta) \leq F_2(\theta)/f_2(\theta)$ for all $\theta \in \Theta_1 \cap \Theta_2$.

Lemma 1. *Let $f_1 \in \mathcal{A}^{\Theta_1}$ and $f_2 \in \mathcal{A}^{\Theta_2}$, where $\Theta_1, \Theta_2 \subset \Theta$, be such that f_1 stochastically dominates the belief f_2 on $\Theta_1 \cap \Theta_2$. Then*

$$\bar{x}^{f_1}(\theta) > \bar{x}^{f_2}(\theta) \quad \text{and} \quad \bar{U}_a^{f_1}(\theta) > \bar{U}_a^{f_2}(\theta) \tag{13}$$

for all $\theta \in \Theta_1 \cap \Theta_2$.

Proof. See Appendix.

The finding that the optimal decision $\bar{x}^f$ is decreasing in the rate F/f will be the crux of many welfare results and discussion in the paper. Lemma 1 implies that using the described dominance concept the agent can rank some admissible beliefs if they have the same support, while a similar preference relation over the beliefs is not available for the society (or the principal). In other words, on a given support of positive length there exists no belief of the regulator which is desired most by the whole society. However, this negative result needs not take anything away from the potential and incentive for learning about the true type of the agent. Indeed, as the rest of this paper will make it clear, there are situations where the social welfare is very sensitive to the support of beliefs that are believed to contain the searched type parameter.

Below, we first describe the complete information belief which denotes the extreme end of learning.

3.2 Complete Information Belief

Hereafter, we fix and denote by θ^T the private type parameter of the agent.

Definition 2. Given the type space Θ, the regulator has complete information about $\theta^T \in \Theta$ if his belief is represented by the Dirac delta (unit impulse) function $\delta(\theta - \theta^T) \in \mathcal{T}^\Theta$, which is defined as

$$\delta(\theta - \theta^T) = 0 \;\; if \;\theta \neq \theta^T, \;\; and \quad \int_{\theta_0}^{\theta_1} \delta(\theta - \theta^T) d\theta = 1. \tag{14}$$

Denote the belief $\delta(\theta - \theta^T)$ by f^T. Note that $f^T \notin \mathcal{A}^\Theta$, i.e. f^T is not an admissible belief in the model. Nevertheless, we are interested in it since in the continuum of admissible beliefs there exist some that are arbitrarily 'close' to f^T. Moreover, the belief f^T points to the desired direction of learning, as clearly implied by the following result.

Proposition 2. *Actual social welfare at the optimal regulatory solution attains its supremum for the belief* f^T*:*

$$\bar{W}^{f^T}(\theta^T) = \sup_{f \in \mathcal{A}^\Theta} \bar{W}^f(\theta^T). \tag{15}$$

Proof. See Appendix.

The above result is not suprising since the complete information eliminates any welfare loss that would otherwise arise when $\alpha \neq 1$ due to the informational rents the agent gets under an incentive compatible contract.

3.3 Complete and Incomplete Learning

Learning can roughly be described as a change of the current information structure towards the state of the complete information structure $I^T = (f^T, \Theta^T)$, where $\Theta^T = \{\theta^T\}$. We say that learning is *complete* if the information structure resulting from learning is equal to I^T, and it is *incomplete* otherwise. Since I^T is not an admissible structure in our incomplete information model, we can speak of complete learning only in the limit of an infinite-stage process. Thus, learning achieved in a finite number of stages is incomplete by assumption.

In the next section, we first consider single-stage incomplete learning that can straightforwardly be generalized to the finite-stage case. There, we define 'more information' resulting from learning and examine its effect on the regulatory outcome. We then consider infinite-stage complete learning. For both situations, we simply assume that the learning of the regulator is exogenous; and the underlying learning technology is such that it always pays to spend on learning from the viewpoint of the society.

4 Results

4.1 Single-Stage Incomplete Learning

Consider a single-stage incomplete learning process, prior to regulation, which changes the current information structure (f^0, Θ^0) to (f^1, Θ^1) where $f^i \in \mathcal{A}^{\Theta^i}$ and $\Theta^1 \subset \Theta^0$ with $\Theta_1 \notin \{\Theta^0, \Theta^T\}$.

Suppose further that the regulator has not acquired any additional information about the distribution of the types in the finer support Θ^1. Then the posterior belief f^1 on Θ^1 should be obtained by some (pre-announced) update rule from the prior f^0 on Θ^0. In the following definition we state the minimal restriction on f^1 to ensure that the information structure (f^1, Θ^1) is superior to (f^0, Θ^0).

Definition 3. The structure (f^1, Θ^1) contains *valuable* (or more) information about θ^T than the structure (f^0, Θ^0) if $\Theta^1 \subset \Theta^0$ and $f^1(\theta^T)/f^0(\theta^T) \geq f^1(\theta)/f^0(\theta)$ for all $\theta \in \Theta^1$.

In the single-stage learning we consider the information about θ^T is incomplete. Thus, more information resulting from learning does not necessarily imply that the regulator, and the society, are aware of its presence. Indeed, one can naturally ask the following question: can the regulator be ever *certain* that he has "more information" under some incomplete learning? Note that the regulator can simply check whether Θ^1 is a subset of Θ^0. So, the above question boils down to whether the regulator can certify that $f^1(\theta^T)/f^0(\theta^T) \geq f^1(\theta)/f^0(\theta)$ for all $\theta \in \Theta^1$ without actually knowing what the value of θ^T is. Apparently, the answer is 'yes' only if $f^1(\theta)/f^0(\theta)$ is constant over Θ^1. This brings us to define the following.

Definition 4. The belief f^1 on Θ^1 is the Bayesian update of f^0 on Θ^0 where $\Theta^1 \subset \Theta^0$ if $f^1(\theta) = f^0(\theta)(1+\gamma)$ for all $\theta \in \Theta^1$, where $\gamma = [\int_{\Theta^1} f(\theta)d\theta]^{-1} - 1$.

The following result shows that the regulator can convince the society that he knows more about the regulated agent only if the regulator is a Bayesian learner.

Proposition 3. *The regulator knows that the structure* (f^1, Θ^1) *contains more information about* θ^T *than the structure* (f^0, Θ^0) *only if* f^1 *is the Bayesian update of* f^0*.*

In sequel, we point to situations in which the agent prefers the Bayesian regulator to have more information about his private type.

Proposition 4. *Suppose the regulator knows that the learned structure* (f^1, Θ^1) *contains more information than the prior structure* (f^0, Θ^0)*, where* $\min(\Theta^1) > \min(\Theta^0)$ *and* $\max(\Theta^1) = \max(\Theta^0)$*. Then the welfare of the regulated agent is higher under the learned structure, i.e.* $\bar{U}_a^{f^1}(\theta) > \bar{U}_a^{f^0}(\theta)$ *for all* $\theta \in \Theta^1$*.*

Proof. See Appendix.

With Bayesian learning that shrinks the type space from the left, the regulator's belief under the learned information structure stochastically dominates his original belief. Then the welfare of the agent increases by Lemma 1, whereas the changes in the welfare of the principal and the society are ambiguous. Proposition 4 definitely points to the possibility of honest signalling of the agent about his type space before the implementation of the regulatory mechanism.

Corollary 1. *Let* (f^0, Θ^0) *be the current information structure and the regulator be known to use Bayes rule in updating his beliefs. Then the agent finds it profitable to signal that his type parameter cannot be in the interval*

$[\min(\Theta^0), \theta^T)$.

The following proposition symmetrically examines learning with right-sided contraction of the type space.

Proposition 5. *Suppose the regulator knows that the learned structure* (f^1, Θ^1) *contains more information than the prior structure* (f^0, Θ^0)*, where* $\min(\Theta^1) = \min(\Theta^0)$ *and* $\max(\Theta^1) < \max(\Theta^0)$*. Then the welfare of the regulated agent is lower whereas the welfare of the principal and the society are both higher under the learned structure, i.e.* $\bar{U}_a^{f^1}(\theta) < \bar{U}_a^{f^0}(\theta)$*,* $\bar{U}_p^{f^1}(\theta) > \bar{U}_p^{f^0}(\theta)$ *and* $\bar{W}^{f_1}(\theta) > \bar{W}^{f^0}(\theta)$ *for all* $\theta \in \Theta^1$*.*

Proof. See Appendix.

Note that Bayesian learning that shrinks the type space only from the right leaves the inverse of the reverse hazard rate, hence the optimal decision variable, unchanged. Nevertheless, the informational rents of the agent become reduced as the upper bound of the integral expression in (8) becomes smaller under the new information structure. With lower informational rents, the social welfare in (9) becomes higher independently from the weight α of the agent's welfare. It follows that the welfare of the principal, which coincides with the social welfare when $\alpha = 0$, is higher, too. Obviously, the regulator must keep on this kind of learning until a point where the expected gain of getting more information is balanced by the cost of learning.

For the general case in which the support of the regulator's initial belief shrinks from both sides under Bayesian learning, we can draw no conclusions as to how the actual welfare of the two agents will change. However, there are readily available numerical simulations in which learning with the two-sided contraction of the support leads to a loss in the actual social welfare. What we can conclude is that more information does not always result in more social welfare.

4.2 Infinite-Stage Complete Learning

We assume that the regulator initially face the information structure (f^0, Θ^0) and he learns, before the implementation of the mechanism, an infinite sequence of structures $\{(f^i, \Theta^i)\}_{i=1}^{\infty}$.

A direct extension of Propositions 4 and 5 for single-stage learning to infinite-stage learning is possible by respectively assuming left-sided and right-sided sequential learning of the type space. Note, however, that left-sided learning that shrinks the type space uniformly from the left cannot yield a finer support than $\Theta^{\infty} = [\theta^T, \max(\Theta^0)]$, whereas right-sided learning that contracts the type space uniformly from the right can at best produce $\Theta^{\infty} = [\min(\Theta^0), \theta^T]$. So, single-sided uniform learning of the type space cannot be considered as an ultimate mode of learning leading to the complete information (provided that θ^T is in the interior of Θ^0).

Nevertheless, for complete learning one may consider the hybrid of left- and right-sided contraction as follows: Let $\{(f^i, \Theta^i)\}_{i=0}^{\infty}$ be a sequence such that Θ^{i+1} is either left-sided or right-sided contraction of Θ^i, the belief f^{i+1} is a Bayesian update of f^i, and $\Theta^{\infty} = \Theta^T$. Then, we know by Proposition 4 that the welfare of the agent increases at each left-sided contraction of the type space over the learning sequence. On the other hand, Proposition 5 implies that the welfare of the agent decreases while the welfare of the principal and the society increase at each right-sided contraction of the type space. But, superimposing these two partially opposite results is not conclusive to guess the level (or even the direction in) which the welfare of the agents and the society will converge to (change) over the infinite sequence of complete learning.

Without any hope of extending the welfare results obtained under single-stage incomplete learning to infinite-stage complete learning, we shall consider an infinite sequence of social welfare induced by a respective sequence of learned information structures and study its limit behavior by appealing to the following convergence definitions.

Definition 5. Consider $\{(f^i, \Theta^i)\}_{i=0}^{\infty}$, where $\Theta^i \subset \Theta^0$ and $f^i \in \mathcal{A}^{\Theta^i}$. We say that the sequence f^i:

i) pseudo converges to f^T if for any given $\epsilon > 0$ and $\theta \in \Theta^0 \backslash \{\theta^T\}$ there is an integer $N(\theta, \epsilon)$ such that $|f^i(\theta) - f^T(\theta)| < \epsilon$ for every $i \geq N$;

ii) absolutely converges to f^T if it converges in pseudo sense and if for any $\epsilon > 0$ there is an integer $N(\epsilon)$ such that $f^i(\theta^T) > 1/\epsilon$ for all $i > N$.

The inequality $|f^i(\theta) - f^T(\theta)| < \epsilon$ that must be checked for pseudo convergence boils down to $f^i(\theta) < \epsilon$ by the definitions of f^i and f^T. Moreover, for $\theta \in \Theta^0 \backslash \Theta^i$ this condition is trivially satisfied, as f^i is zero everywhere outside its support Θ^i. Note also that the above notion of absolute convergence is the well-known pointwise convergence of a real-valued sequence in probability theory.

Below we first show that the absolute convergence of the regulator's belief to the truth eliminates any welfare loss due to informational asymmetry and yields complete information outcome.

Proposition 6. *Let $\{(f^i, \Theta^i)\}_{i=0}^{\infty}$ where $\Theta^i \subset \Theta^0$ and $f^i \in \mathcal{A}^{\Theta^i}$ be such that f^i absolutely converges to f^T. Then*

$$\lim_{i\to\infty} \bar{W}^{f^i}(\theta^T) = \bar{W}^{f^T}(\theta^T). \tag{16}$$

Proof. See Appendix.

Notice first that the notion of absolute convergence is too strong (that is why we call it absolute) in the sense that the regulator reduces the density assigned to all types other than the true type θ^T down to zero while he increases the density of θ^T unboundedly. Thus, the rate $F^i(\theta^T)/f^i(\theta^T)$ goes to zero, making the social welfare approach to its supremum level. However, with such learning it seems as if the regulator has already distinguished θ^T from the other types. It is true that Bayesian updating of beliefs which would uniquely allow the regulator to verify the availability of more information in each stage of learning implies absolute convergence of beliefs. But, in the infinite-stage 'complete' learning model Bayesian learning is no longer indispensable as we assume that the regulator, representing the society, is only interested in the limit social welfare. Obviosuly, the stronger the notion of convergence of beliefs, the more expensive will be the cost of learning in reality. Thus, it is meaningful to examine whether our notion of convergence can be weakened without relaxing the requirement that it yields the full-information social welfare in the limit. The following result shows that pseudo convergence, our weaker definition of convergence, may not yield the most desired outcome.

Proposition 7. *There exists a sequence $\{(f^i, \Theta^i)\}_{i=0}^{\infty}$ where $\Theta^i \subset \Theta^0$, $f^i \in$*

$\mathcal{A}^{\Theta^i}$, *f^i pseudo converges to f^T, such that*

$$\lim_{i\to\infty} \bar{W}^{f^i}(\theta^T) \neq \bar{W}^{f^T}(\theta^T). \tag{17}$$

Proof. See Appendix.

The example provided in the proof of Proposition 7 is a sequence of information structures with $\Theta^0 = [0, 2]$ and $\theta^T = 1$. The densities at each stage of learning are decreasing over their supports before θ^T while they are constant at the level 1/4 everywhere after θ^T. So, initially inverse of the reverse hazard rate at θ^T is $F^0(1)/f^0(1) = (3/4)/(1/4) = 3$. At each stage of learning, the support shrinks to half of its size around θ^T as we have $\min(\Theta^i) = 1 - 2^{-i}$ and $\max(\Theta^i) = 1 + 2^{-i}$. Thus, the sequence f^i is pseudo converging to f^T. Moreover, the density after θ^T remains to stay at 1/4, which implies that in learning stage $i \geq 1$ half of the probability mass after θ^T, which is $(1 - F^{i-1}(1))/2$, is redistributed over the interval before θ^T. Thus, in stage i the probability mass over the interval $[\min(\Theta^i), \theta^T]$ becomes $F^i(1) = F^{i-1}(1) + (1 - F^{i-1}(1))/2$. In the limit of this non-Bayesian learning, $F^i(1)$ approaches 1, while $f^i(1)$ remains at (1/4). Since, $F^i(1)/f^i(1)$ does not converge to zero, the sequence does not yield the complete information outcome at the report $\theta^T = 1$. We should note that the constancy of f^i at or above θ^T in the proposed example is immaterial for the proof of Proposition 7. In fact, we can trivially modify the described example so as to make the density f^i increase monotonically (but boundedly) after θ^T but get the same result.

5 Conclusions

In a generalized principle-agent model, we have examined a Bayesian regulator's learning about the private information of the regulated agent. We have specified what "more information" means and demonstrated that more information about the informed agent needs not be undesirable for him. We have also characterized situtations in which the principal and the society benefit from the regulator's learning. The final result we have obtained is that outcomes of Bayesian regulatory mechanisms are, in general, path-dependent.

The convergence of the regulator's belief to the truth needs not lead to the complete information outcome.

Our findings support the view that one should be careful in determining what to expect from Bayesian mechanisms with their existing specifications. It has long been noticed that the subjective nature of beliefs may cast some doubts on the implementability of the Bayesian mechanisms. Crew and Kleindorfer (1986), Vogelsang (1988), Koray and Sertel (1990) criticized the Bayesian approach in regulation on the grounds of unaccountability and manipulability of the regulator's subjective prior beliefs. In a very recent study, Koray and Saglam (2005) examine the same issue in the Baron and Myerson (1982) model of monopoly regulation. They show that all interest groups in the society are extremely sensitive to the prior belief of the regulator. There exist beliefs yielding values arbitrarily close to the supremum of actual welfare and expected welfare of the regulated agent (monopolist) and the principal (consumers), respectively. Moreover, under some other beliefs one can come as close to the infimum of actual welfare of both parties as possible. When the belief of the regulator is unverifiable by the public, the existence of such critical beliefs leads to a bargaining game over the beliefs between a corrupt regulator and the interest groups in the society, which distorts the regulatory outcome predicted by Baron and Myerson.

What we add to the previous results is that Bayesian mechanisms may yield unpredictable and sometimes undesirable outcomes even in the presence of a benevolent and sincere regulator if the socially efficient type of learning is not completely specified as part of the regulatory mechanism.

References

Baron D. and R.B. Myerson (1982): "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50, 911-930.

Blume, L. and D. Easley (1995): "What has the Rational Learning Literature Taught Us? in Alan Kirman and Mark Salmon, eds. Learning and Rationality in Economics. Oxford: Blackwell, pp. 12-39.

Cox J.C., Shachat J. and M. Walker (2001): "An Experimental Test of Bayesian Learning in Games," *Games and Economic Behavior*, 34, 11-33.

Crew M.A. and P.R. Kleindorfer (1986): *The Economics of Public Utility Regulation.* Cambridge, MA: MIT Press.

Feldman, M. and A. McLennan (1989): "Learning in a Repeated Statistical Decision Problem with Normal Disturbances," Manuscript, University of Illinois, Urbana-Champaign.

Gibbard A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587-602.

Guesnerie R. and J.J. Laffont (1984): "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," *Journal of Public Economics*, 25, 329-369.

Jordan J.S. (1991): "Bayesian Learning in Normal Form Games," *Games and Economic Behavior*, 3, 60-81.

Kalai E. and E. Lehrer (1993): "Rational Learning Leads to Nash Equilibria," *Econometrica*, 61, 1019-1045.

Kiefer, N. and Y. Nyarko (1989): "Control of an Unknown Linear Process with Learning," *International Economic Review*, 30 (3), 571-86.

Koray S. and M.R. Sertel (1990): "Pretend-but-Perform Regulation and Limit Pricing," *European Journal of Political Economy*, 6, 451-472.

Koray S. and I. Saglam (2005): "The Need for Regulating a Bayesian Regulator," *Journal of Regulatory Economics*, forthcoming.

Myerson R.B. (1979): "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 47, 61-74.

Vogelsang, I. (1988): "A Little Paradox in the Design of Regulatory Mechanisms," *International Economic Review*, 29, 467-476.

Appendix

Proof of Proposition 1. The integrand in the objective function of (2) is differentiated with respect to $x(\theta)$ to obtain the optimality condition (12). Using the asssumptions (A2) and (A7), it is easy to check that the same integrand is concave in x.

To show that the solution to (12) satisfies the incentive-compatibility constraint (IC), we will first prove that the optimal solution $\bar{x}$ is nonincreasing in θ. Total differentiation of (12) with respect to θ yields

$$\left(\frac{\partial^2 V_p}{\partial x^2} + \frac{\partial^2 V_a}{\partial x^2} + (1-\alpha)\frac{F(\theta)}{f(\theta)}\frac{\partial^3 V_a}{\partial^2 x^2 \partial\theta}\right)\frac{d\bar{x}}{d\theta} =$$

$$\left(-(1-\alpha)\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) - 1\right)\frac{\partial^2 V_a}{\partial x \partial\theta} - \frac{\partial^2 V_p}{\partial x \partial\theta} - (1-\alpha)\frac{F(\theta)}{f(\theta)}\frac{\partial^3 V_a}{\partial x \partial\theta^2}.$$

Using the assumptions (A2), (A3), (A4), (A6) and (A7) together with the assumption that $F(\theta)/f(\theta)$ is nondecreasing in θ, we conclude that $d\bar{x}/d\theta$ is nonpositive.

The net utility of the agent when he truthfully reports his type as θ is

$$U_a(\theta) = -\int_\theta^{\theta_1} \frac{\partial}{\partial\tilde{\theta}} V_a(\bar{x}(\tilde{\theta}), \tilde{\theta}) d\tilde{\theta}$$

by (8). The net utility of the agent when he misreports its unknown parameter as $\hat{\theta}$ while θ is the true parameter is

$$U_a(\theta, \hat{\theta}) = V_a(\bar{x}(\hat{\theta}), \theta) + U_a(\hat{\theta}) - V_a(\bar{x}(\hat{\theta}), \hat{\theta}). \tag{18}$$

Subtracting $U_a(\theta)$ from (18) we get

$$\begin{aligned} U_a(\theta, \hat{\theta}) - U_a(\theta) &= -\int_{\hat{\theta}}^{\theta} \frac{\partial}{\partial\tilde{\theta}} V_a(\bar{x}(\tilde{\theta}), \tilde{\theta}) d\tilde{\theta} + V_a(\bar{x}(\hat{\theta}), \theta) - V_a(\bar{x}(\hat{\theta}), \hat{\theta}) \\ &= -\int_{\hat{\theta}}^{\theta} \frac{\partial}{\partial\tilde{\theta}}\Big(V_a(\bar{x}(\tilde{\theta}), \tilde{\theta}) - V_a(\bar{x}(\hat{\theta}), \tilde{\theta})\Big) d\tilde{\theta} \le 0 \end{aligned}$$

from (A4) and $d\bar{x}(\theta)/d\theta \le 0$. Thus, the optimal program (12) is incentive-compatible.

Finally to check condition (IR), i.e. $U_a(\theta) \geq 0$ at the optimal solution $\bar{x}$, is straightforward from (8) thanks to assumption (A5). Q.E.D.

Proof of Lemma 1. Total differentiation of (12) at the optimal decision $\bar{x}^f$ with respect to $F(\theta)/f(\theta)$ yields

$$\left(\frac{\partial^2 V_p}{\partial x^2} + \frac{\partial^2 V_a}{\partial x^2} + (1-\alpha)\frac{F(\theta)}{f(\theta)}\frac{\partial^3 V_a}{\partial^2 x^2 \partial\theta}\right)\frac{d\bar{x}^f}{d[F(\theta)/f(\theta)]} = -(1-\alpha)\frac{\partial^2 V_a}{\partial x \partial\theta}.$$

From assumptions (A2), (A4) with strict inequality and (A7) it follows that $\bar{x}^f$ is decreasing in $F(\theta)/f(\theta)$. Considering equation (8), using assumptions (A4) and (A5) and $F_1(\theta)/f_1(\theta) < F_2(\theta)/f_2(\theta)$, we conclude that $\bar{U}_a^{f_1}(\theta) > \bar{U}_a^{f_2}(\theta)$ for all $\theta \in \Theta$. Q.E.D.

Proof of Proposition 2. Actual social welfare at θ^T is

$$\begin{aligned} W(\theta^T) &= U_p(\theta^T) + \alpha U_a(\theta^T) \\ &= V_p(x(\theta^T), \theta^T) - t(\theta^T) + \alpha(V_a(x(\theta^T), \theta^T) + t(\theta^T)) \\ &= V_p(x(\theta^T), \theta^T) + V_a(x(\theta^T), \theta^T) - (1-\alpha)U_a(\theta^T). \end{aligned}$$

When $\alpha = 1$, we have

$$W(\theta^T) = V_p(x(\theta^T), \theta^T) + V_a(x(\theta^T), \theta^T)$$

irrespective of the information structure. Note that the above equation remains to hold true even when $\alpha \neq 1$ if there is complete information about the private type of the agent since optimality and (IR) condition together leads to $t = -V_a$ or $U_a = 0$, i.e. no informational rents for the agent. The optimal solution $\bar{x}(\theta^T)$ then satisfies

$$\frac{\partial}{\partial x}V_p(x(\theta^T), \theta^T) + \frac{\partial}{\partial x}V_a(x(\theta^T), \theta^T) = 0,$$

yielding $\bar{W}^{f^T}(\theta^T)$.

Noticing that for all $f \in \mathcal{A}^\Theta$ we have $W^f = V_p^f + V_a^f - (1-\alpha)U_a^f \leq V_p^f + V_a^f$ by the condition (IR) reading $U_a^f \geq 0$, and using $\partial(V_p^f + V_a^f)/\partial x > 0$, $\partial\bar{x}/\partial(F/f) < 0$ when $\partial^2 V_a/\partial x \partial\theta < 0$ and $F^T(\theta^T)/f^T(\theta^T) = 0$, we establish $\bar{W}^{f^T}(\theta^T) = \sup_{f \in \mathcal{A}^\Theta} \bar{W}^f(\theta^T)$. Q.E.D.

Proof of Proposition 4. Since f^1 is a Bayesian update of f^0 on a finer support, $f^1(\theta) > f^0(\theta)$ and hence $F^1(\theta) < F^0(\theta)$ for all $\theta \in [\min(\Theta^1), \max(\Theta^0))$ while $F^1(\max(\Theta^0)) = F^0(\max(\Theta^0)) = 1$. This implies that $F^1(\theta)/f^1(\theta) < F^0(\theta)/f^0(\theta)$ for all $\theta \in \Theta^1$. Then from Lemma 1, $\bar{x}^{f^1}(\theta) > \bar{x}^{f^0}(\theta)$ and $\bar{U}_a^{f^1}(\theta) > \bar{U}_a^{f^0}(\theta)$ for all $\theta \in \Theta^1$. Q.E.D.

Proof of Proposition 5. Since f^1 is a Bayesian update of f^0, $f^1(\theta) = f^0(\theta)(1+\gamma)$ for all $\theta \in \Theta^1$, where $\gamma = [F(\max(\Theta_1))]^{-1} - 1$. Note that $F^1(\theta)/f^1(\theta) = F^0(\theta)/f^0(\theta)$ and therefore $x^{f^1}(\theta) = x^{f^0}(\theta)$ for all $\theta \in \Theta^1$. Then from (8) we obtain $\bar{U}_a^{f^1}(\theta) < \bar{U}_a^{f^0}(\theta)$, since $\max \Theta^1 < \max \Theta^0$.

We have $\bar{W}^{f^1}(\theta) > \bar{W}^{f^0}(\theta)$ since $\bar{W}^{f^1}(\theta) = \bar{V}_p^{f^1} + \bar{V}_a^{f^1} - (1-\alpha)\bar{U}_a^{f^1} = \bar{W}^{f^0}(\theta) + (1-\alpha)(\bar{U}_a^{f^0} - \bar{U}_a^{f^1})$. Finally, $\bar{U}_p^{f^1}(\theta) > \bar{U}_p^{f^0}(\theta)$ follows from the fact that $\bar{W}^{f^0}(\theta) = \bar{U}_p^{f^0}(\theta)$ when $\alpha = 0$. Q.E.D.

Proof of Proposition 6. By the definition of absolute convergence $\lim_{i\to\infty} f^i(\theta^T) = \infty$, so $\lim_{i\to\infty} F^i(\theta^T)/f^i(\theta^T) = 0$. It then follows from the program (12) using Lemma 1 and assumption (A1) that $\lim_{i\to\infty} \bar{W}^{f^i}(\theta^T) = \bar{W}^{f^T}(\theta^T)$. Q.E.D.

Proof of Proposition 7. We will prove that the below sequence of beliefs satisfies the conditions stated in the proposition.

$$\Theta^0 = [0,2], \quad \theta^T = 1 \tag{19}$$

$$f^0(\theta) = \begin{cases} 5/4 - \theta & \text{if } \theta \in [0,1) \\ 1/4 & \text{if } \theta \in [1,2] \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

$i \geq 1:$

$$\Theta^i = [1 - 2^{-i}\,,\, 1 + 2^{-i}] \tag{21}$$

$$F^{i-1}(1) + (1/2)(1 - F^{i-1}(1)) = \frac{1}{2}\left(h_i + \frac{1}{4}\right)(1 - \min(\Theta^i)) \tag{22}$$

$$f^i(\theta) = \begin{cases} h^i - \frac{h^i-(1/4)}{1-\min(\Theta^i)}(\theta - \min(\Theta^i)) & \text{if}\, \theta \in [\min(\Theta^i), 1) \\ 1/4 & \text{if}\, \theta \in [1, \max(\Theta^i)] \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

First note from (21) that $\lim_{i\to\infty} \Theta^i = \{1\}$. Thus, $f^i(\theta) = 0$ is zero in the limit for all $\theta \neq 1$ implying that the sequence pseudo converges to f_T. Moreover, we have

$$\frac{F^i(1)}{f^i(1)} = \frac{F^{i-1}(1) + (1 - F^{i-1}(1))/2}{f^{i-1}(1)} > \frac{F^0(1)}{f^0(1)}$$

for all $i \geq 1$. So, $\lim_{i\to\infty} F^i(1)/f^i(1) \neq 0$. We then conclude from the program (12) using Lemma 1 and assumption (A1) that $\lim_{i\to\infty} \bar{W}^{f^i}(1) \neq \bar{W}^{f^T}(1)$. Q.E.D.