Games and Economic Behavior 104 (2017) 329–349

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Optimal dynamic information provision ☆

Jérôme Renault [a], Eilon Solan [b,*], Nicolas Vieille [c]

[a] TSE (GREMAQ, Université Toulouse 1), 21 allée de Brienne, 31000 Toulouse, France
[b] School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997800, Israel
[c] Department Economics and Decision Sciences, HEC Paris, 1, rue de la Libération, 78 351 Jouy-en-Josas, France

ARTICLE INFO

Article history:
Received 20 February 2015
Available online 8 May 2017

JEL classification:
C73
C72

Keywords:
Dynamic information provision
Optimal strategy
Greedy algorithm
Commitment

ABSTRACT

We study a dynamic model of information provision. A state of nature evolves according to a Markov chain. An advisor with commitment power decides how much information to provide to an uninformed decision maker, so as to influence his short-term decisions. We deal with a stylized class of situations, in which the decision maker has a risky action and a safe action, and the payoff to the advisor only depends on the action chosen by the decision maker. The greedy disclosure policy is the policy which, at each stage, minimizes the amount of information being disclosed in that stage, under the constraint that it maximizes the current payoff of the advisor. We prove that the greedy policy is optimal in many cases – but not always.

1. Introduction

In this paper, we study a dynamic model of information provision, with the goal of identifying the optimal information provision policy. The participants are (a) an advisor and (b) a sequence of short-lived "investors", who each decide whether to take a risky action, such as a short-run investment. The payoff from the risky action is contingent on the state of nature, which is unknown to the investor and evolves according to a Markov chain. The investor may get information on the current state through the advisor, who receives a fixed fee whenever the risky action is chosen.

How much information to disclose is the choice of the advisor. We assume that the advisor has commitment power. To be specific, the advisor *ex ante* commits to an information provision rule, which maps histories into distributions over signals.

Since the investors are short lived, they behave myopically, hence they take the risky action if and only if their expected payoff is above that of an outside option, normalized to zero. It follows that the problem faced by the advisor reduces to a Markov decision problem (MDP) in which the state space is the set of posterior beliefs of the investors, and the action space is the set of information provision rules. In that MDP, the advisor chooses the quality of information to provide so as to maximize the (expected) discounted frequency of stages in which the risky action is taken. Advising is thus both honest, in that realized signals cannot be manipulated, and strategic, in that the information content of the signal is strategic.

There are two (mutually exclusive) interpretations of our game. In the first one, the advisor does observe the successive states of nature but does commit to a dynamic information *disclosure* policy, prior to the beginning of the game. An alter-

☆ The research of Renault and Vieille was supported by Agence Nationale de la Recherche (Grant ANR-10-BLAN 0112). Solan acknowledges the support of the Israel Science Foundation, Grants #212/09 and #323/13. The authors thank Omri Solan for numerics that led to the counterexample presented in Section 5 and the referees and Associate Editor for their insightful comments.

* Corresponding author.
E-mail addresses: jerome.renault@tse-fr.eu (J. Renault), eilons@post.tau.ac.il (E. Solan), vieille@hec.fr (N. Vieille).

http://dx.doi.org/10.1016/j.geb.2017.04.010
0899-8256/

native interpretation is to assume instead that the advisor does *not* observe the states of nature. Choosing in a given stage a distribution of signals, contingent on the state, then amounts to choosing a statistical experiment *à la* Blackwell, whose outcome is public. Under this latter interpretation, the advisor chooses which information will be *publicly* obtained.

As in any dynamic optimization problem, the advisor faces a trade-off between maximizing the probability of investment in the current stage, and preserving his continuation value. By disclosing information at a given date, the advisor may increase his current payoff, but then gives up part of his information advantage in later dates (except if successive states are independent).

Writing the dynamic programming equation that characterizes the value of the optimization problem is rather straightforward. *Solving* it, that is, identifying an optimal policy of the advisor, is not. Our main contribution is to prove that often, but not always, the trade-off faced by the advisor is solved in a very simple, myopic way. We define the greedy policy as the one that, at any given date, *minimizes* the amount of information being disclosed, subject to the current payoff of the advisor being *maximized*. We prove that this policy is optimal when there are only two states of nature. We then focus on a class of Markov chains, described by a renewal property. Within this class, we prove that the greedy policy is optimal, whenever the distribution of the initial state is close enough to the invariant distribution of the Markov chain of states. We also prove that the greedy policy is also eventually optimal, irrespective of the distribution of the initial state. That is, any optimal policy eventually coincides with the greedy policy. However, as we show by means of a counterexample, the greedy policy may fail to be optimal for some distributions of the initial state.

Our paper belongs to the literature on information design, and on so-called persuasion mechanisms. In the static case, Kamenica and Gentzkow (2011) characterized optimal persuasion mechanisms, using insights from Aumann and Maschler (1995). A dynamic model of persuasion was studied in Ely (2017), where the underlying state of nature is fixed. The sequence of posterior beliefs of the uninformed agent then follows a martingale, and the goal of the informed agent is to maximize suspense over time, defined as the L^1-variation of the sequence of posteriors. Independently of us, Ely (2017) deals with the same model as ours, fully solves one specific example, and analyzes a number of variants, including the case of two uninformed agents, and of more-than-two actions. Relatedly, Honryo (2011) analyzes a model of dynamic persuasion, where the sender has several pieces of information over the (fixed) quality of a proposal. These pieces, both negative and positive, may be sent sequentially. In Hörner and Skrzypacz's (2016), a seller may transmit gradually (verifiable) information, in exchange for payments.

Che and Hörner (2015) and Halac et al. (forthcoming) address the issue of the optimal design of information disclosure, in the context of social learning/experimentation and of contests respectively.

Our paper also relates to the literature on dynamic Bayesian games, see, e.g., Mailath and Samuelson (2001), or Athey and Bagwell (2008), Escobar and Toikka (2013), Renault (2006), Hörner et al. (2010, 2015), or Renault et al. (2013), who study dynamic sender–receiver games.

2. Model and preliminary observations

2.1. Model

We consider a stylized class of two-player dynamic games between an *advisor* and short-lived *investors*. Over time, the advisor privately observes the successive realizations $(\omega_n)_{n\in\mathbf{N}}$ of a changing state of nature, with values in a finite set Ω. In each stage n, the advisor chooses which information to disclose to the current investor, who decides whether to invest or not. The game then moves to the next stage.[1]

The advisor receives a fee whenever investment takes place and discounts future payoffs according to the discount factor $\delta < 1$. The investor's utility from investing is $r(\omega_n)$, with $r : \Omega \to \mathbf{R}$. When the investor's belief is $p \in \Delta(\Omega)$, the expected payoff from investing (net of the fee to the investor) is therefore given by the scalar product $\langle p, r\rangle = \sum_{\omega\in\Omega} p(\omega)r(\omega)$.

We assume that the sequence $(\omega_n)_{n\in\mathbf{N}}$ is an irreducible Markov chain with transitions $(\pi(\omega' \mid \omega))_{\omega,\omega'\in\Omega}$. While investors know the distribution of the sequence $(\omega_n)_{n\in\mathbf{N}}$, the only additional information received along the play comes from the advisor. That is, previous investment outcomes are not observed by the investor.[2] The investor chooses to invest if and only if the expected (net) payoff from investing is nonnegative.[3]

Accordingly, the *investment region* is $I := \{p \in \Delta(\Omega), \langle p, r\rangle \geq 0\}$ and the *investment frontier* is $\mathcal{F} := \{p \in \Delta(\Omega), \langle p, r\rangle = 0\}$. We also denote by $J := \Delta(\Omega) \setminus I$ the *noninvestment* region.

The game reduces to a stochastic optimization problem, in which the advisor manipulates the posterior beliefs of the investors, so as to maximize the expected discounted frequency of stages in which investment takes place. An *information disclosure policy* for the advisor specifies for each stage, the probability distribution of the message to be sent in that stage, as a function of previous messages and of the information privately available to the advisor, that is, past and current states.

As mentioned in the Introduction, an equivalent and alternative interpretation would be to assume that the advisor does *not* observe the sequence $(\omega_n)_{n\in\mathbf{N}}$ and chooses instead in each stage a statistical experiment *à la* Blackwell. Such an

[1] We are not explicit about the message set. It will be convenient to first assume that it is rich enough, e.g., equal to Ω. We will show that w.l.o.g. two messages suffice for our results.

[2] This assumption is discussed at the end of this section.

[3] In particular, we assume that the investor invests whenever indifferent.