J. Math. Anal. Appl. 419 (2014) 1322–1332

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Robust Markov perfect equilibria

Anna Jaśkiewicz [a], Andrzej S. Nowak [b,*]

[a] Institute of Mathematics and Computer Science, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
[b] Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Podgórna 50, 65-246 Zielona Góra, Poland

A R T I C L E   I N F O

Article history:
Received 10 December 2013
Available online 27 May 2014
Submitted by J.A. Filar

Keywords:
Markov decision model
Quasi-hyperbolic discounting
Non-randomised Markov perfect equilibrium
Robust equilibrium

A B S T R A C T

In this paper we study a Markov decision model with quasi-hyperbolic discounting and transition probability function depending on an unknown parameter. Assuming that the set of parameters is finite, the sets of states and actions are Borel and the transition probabilities satisfy some additivity conditions and are atomless, we prove the existence of a non-randomised robust Markov perfect equilibrium.

1. Introduction and the model

The quasi-hyperbolic discounting concept in a dynamic choice model reveals time-consistency in making decisions by a person whose utility changes over time, see the seminal papers by Strotz [26] and Phelps and Pollak [25]. As suggested in [25] finding a time-consistent rational (optimal in some sense) policy can be done by considering a non-cooperative intergenerational game played by short lived generations. When each generation has a countably many descendants one can think of a game between countably many *selves* representing the same person whose utility changes over discrete time, see [4,15] and Subsection 4.3 in [20] for general accounts and discussion. The Markov perfect equilibrium concept supported by a discussion in [21] seems to be appropriate solution for the study of dynamic games under the quasi-hyperbolic discounting. More precisely, a Markov perfect equilibrium is a fixed point of some operator in a suitably chosen strategy function space. This fact emphasises the difference between the approaches used in game theory and in dynamic programming. The first paper on quasi-hyperbolic discounting is the work of Alj and Haurie [1], where the state and action spaces are finite. Further literature on this type of problems (within deterministic and stochastic framework) the reader may find in [2,4,15,20] and the references cited therein.

* Corresponding author.
E-mail addresses: anna.jaskiewicz@pwr.wroc.pl (A. Jaśkiewicz), a.nowak@wmie.uz.zgora.pl (A.S. Nowak).

http://dx.doi.org/10.1016/j.jmaa.2014.05.061
0022-247X/

All the aforementioned papers except [20] deal with very specific transition law and one-dimensional state and action spaces. The proofs are based on certain lattice programming techniques exploiting monotonicity and concavity assumptions. The models with Borel state and action spaces are treated only in [20], but the main focus is put on the risk-sensitive generations that employ the exponential utility aggregator function. Assumptions made on the transition probability function in [20] are relatively weak when compared to previous papers and, moreover, they allow to apply the Dvoretzky–Wald–Wolfowitz theorem [10] on elimination of randomisation to obtain non-randomised equilibria in the atomless case. Such equilibria are extremely desirable in economic applications, see for instance, [16,17]. All the papers discussed above are concerned with the transition probability functions that are perfectly known to all players. In [3] this assumption is weakened by saying that the transitions are dependent on some parameter controlled by Nature. Then, the concept of a Markov perfect equilibrium is slightly modified by the idea of maxmin optimisation. The case is also considered in this paper. An equilibrium we are interested in is now called *robust* and is shown to exist for a large class of Markov decision models with Borel state and action spaces. Firstly, we prove the equilibrium in the class of randomised strategies (Theorem 1) and then in the class of non-randomised strategies in the atomless case. The latter result (Theorem 2) is more important when one thinks of applications in economics. Our proof makes use of the Dvoretzky–Wald–Wolfowitz theorem and is new compared with the approach taken in [3]. Specifically, the authors in [3] deal with special consumption/investment models, the assumptions are stronger and the state and action spaces are subsets in the real line. The only exception is that the set of parameters in [3] may be infinite. In this paper, this set is finite but Nature can apply infinitely many probability distributions for selecting a parameter. For a further discussion of our results the reader is referred to Remarks 1 and 2.

Let R be the set of all real numbers and $R_+ = [0, \infty)$, $T = N = \{1, 2, ...\}$. By a Borel space Y we mean a non-empty Borel subset of a complete separable metric space endowed with the Borel σ-algebra $\mathcal{B}(Y)$. Let $P(Y)$ denote the space of all probability measures on Y endowed with the weak topology and the Borel σ-algebra, see Chapter 7 in [7]. Let S and A be Borel spaces. Assume that $A(s)$ is a non-empty compact subset of A for each $s \in S$ and the set

$$C := \{(s, a) : s \in S,\ a \in A(s)\} \subset S \times A$$

is Borel. Let F be the set of all Borel measurable selectors of the set-valued mapping $s \mapsto A(s)$. By Corollary 1 in [8], $F \neq \emptyset$. By Φ we denote the set of all transition probabilities ϕ from S to A such that $\phi(A(s)|s) = 1$ for each $s \in S$. Clearly, F can be seen as a subset of Φ, so $\Phi \neq \emptyset$.

We consider a *Markov decision model* with *quasi-hyperbolic preferences* and unknown transition probability function in which the decision maker is viewed as a sequence of autonomous temporal *selves*. The selves are indexed by the respective periods $t \in T$ in which they choose their actions. Let S and A, considered above, be the *state space* and *the action space*, respectively. Then $A(s)$ is the set of all *actions available* in state $s \in S$. Further, assume that Θ is a *finite set* of some parameters. At the beginning of t-th period, self t observes the state $s_t \in S$ and chooses (possibly at random) $a_t \in A(s_t)$. Then, $q(\cdot|s_t, a_t, \theta)$ is the probability distribution of the next state. Note that q is a transition probability from $S \times A \times \Theta$ to S. Self t's satisfaction is reflected by an *instantaneous utility function* $u : C \mapsto R_+$ that remains unchanged over all periods. It is assumed that θ is chosen according to a certain probability measure $\gamma_t \in \mathcal{P}$, where $\mathcal{P}$ denotes the action set of *Nature* and it is assumed to be a closed subset of $P(\Theta)$. Thus, $\mathcal{P}$ can be viewed as a compact subset of an Euclidean space.

We impose the following assumptions on the model.

(A1) There exist probability measures $\mu_1^\theta, \ldots, \mu_l^\theta$ on S and Borel measurable functions $g_1, \ldots, g_l : C \mapsto [0, 1]$ such that