A RISK MINIMIZATION PROBLEM FOR FINITE HORIZON SEMI-MARKOV DECISION PROCESSES WITH LOSS RATES

QIULI LIU*
School of Mathematical Sciences
South China Normal University
Guangzhou 510631, China

XIAOLONG ZOU
School of Economics and Statistics
Guangzhou University
Guangzhou 510006, China

(Communicated by Onésimo Hernández-Lerma)

Abstract. This paper deals with the risk probability for finite horizon semi-Markov decision processes with loss rates. The criterion to be minimized is the risk probability that the total loss incurred during a finite horizon exceed a loss level. For such an optimality problem, we first establish the optimality equation, and prove that the optimal value function is a unique solution to the optimality equation. We then show the existence of an optimal policy, and develop a value iteration algorithm for computing the value function and optimal policies. We also derive the approximation of the value function and the rules of iteration. Finally, a numerical example is given to illustrate our results.

1. Introduction. In the field of Markov decision processes (MDPs), many criteria have been proposed to study stochastic optimal control problems such as expected discount criteria, expected average criteria, risk-sensitive criteria, the first passage criteria, and risk probability criteria and so on [4, 5, 6, 7, 8, 10, 11, 12, 15, 3, 20, 23]. Among all the criteria, risk probability criteria have received much attention as they have rich applications in many areas such as equipment maintenance, queuing systems, reliability engineering, risk analysis and so on[1, 2, 14, 17, 16, 21].

Risk probability criteria have been widely studied in the literature for MDPs by many authors via different methods; see, for instance, [10, 11, 12, 20, 18] and their extensive references. According to the characterization of the optimality problems, the existing works on risk probability criteria can be roughly classified into two groups: (i) minimizing the risk probability that the total reward is not greater than (or less than) a given initial threshold value, (ii) minimizing the risk probability that the total loss exceeds a given initial threshold value. We next briefly describe some

2010 Mathematics Subject Classification. Primary: 90C40; Secondary: 93E20.

Key words and phrases. Semi-Markov decision processes, loss rate, risk probability, optimal policy.

Research supported by Natural Science Foundation of Guangdong Province (Grant No.2014A030313438), Zhujiang New Star (Grant No. 201506010050) and Guangdong Province outstanding young teacher training plan (Grant No. YQ2015050).

* Corresponding author. Emails: liuql@m.scnu.edu.cn; zouxiao@gzhu.edu.cn.
existing works on the two groups respectively, and then show our motivation and main results of this paper on the group (ii). In the group (i), the risk probability is defined as $P^\pi(Z_r \leq \lambda)$, where $Z_r$ denotes the nonnegative total reward. Several authors minimize the risk probability that a random total reward is less than or equal to a given initial threshold value in discrete time Markov decision processes [15, 22, 25]. For example, Ohtsubo and Sakaguchi [22] consider the risk probability that the total reward is not greater (or less) than a given initial threshold value. In addition, a few works have been devoted to risk probability criteria in semi-Markov decision processes (SMDPs). More precisely, Huang, Guo and Song study the risk probability that the system reaches a prescribed reward level during a first passage time to some target set for SMDPs [10]. Recently, Huang and Guo investigate the risk probability criteria for finite horizon SMDPs and introduce a class of horizon-relevant policies which depend on the usual states and the planning horizons [12].

In the group (ii), the risk probability is defined as $P^\pi(Z_l > \lambda)$, where $Z_l$ denotes the nonnegative total loss. To the best of the authors’ knowledge, there are only few studies dealing with the loss case; see, for instance, Ohtsubo considers the optimality problem of minimizing the risk probability that total loss exceeds a threshold value [18]. Obviously, from the fact $P^\pi(Z_r \leq \lambda) = 1 - P^\pi(Z_r > \lambda)$, we see that the problem of minimizing $P^\pi(Z_r > \lambda)$ is not equivalent to that of minimizing $P^\pi(Z_r \leq \lambda)$. A review of existing references show that all of the related works for risk probability criteria are mainly on the case of reward for SMDPs. However, the case of loss for SMDPs has not been explored yet. In the present paper, we devote ourself to the risk probability criteria for SMDPs with loss rates.

As is well known, the business cycle is a series of peaks and troughs [16, 21]. A business cycle is basically defined in terms of periods of expansion or recession. During recessions, the economy is contracting, as measured by decreases in indicators like employment, industrial production, sales and personal incomes[16, 21]. The decision maker is often concerned with the total loss incurred during the recessions. Hence, a natural optimality question is how to choose optimal policies such that the risk of the system that the total loss incurred during a finite time is minimized. That is the main motivation of the current paper. Therefore, the optimality problem of minimizing the risk probability that the total loss incurred is meaningful and novel, it will be further studied in the present paper.

Compared with Huang, Guo and Li [12] for finite horizon semi-Markov decision processes with risk probability, this paper:

(1) deals with the risk probability criteria $P^\pi(Z_l > \lambda)$ in the group (ii) for finite horizon SMDPs with loss rates other than the reward case [12];

(2) gives the approximation value of the value function and the exact number of iteration steps for the stop of the iterations (See Theorem 3.2 below for details).

In this paper, we further study finite horizon SMDPs, and our goal is to find an optimal policy which minimizes the risk probability that the total loss incurred by a system during a finite horizon exceeds a loss level. The main contributions of this paper are as follows. Firstly, we characterize the risk probability (Lemma 3.1). Secondly, we establish the optimal equation via a successive approximation technique, and further, we provide an iteration algorithm for computing the value function (Theorem 3.1 (a))and prove that the value function is its unique solution (Theorem 3.1 (b)). Thirdly, we show the existence of a deterministic stationary ($\epsilon$-optimal) optimal policy (Theorem 3.1 (c) and Theorem 3.2 (c)). Fourthly, we draw a conclusion that the error of the approximation of the value function may be
sufficiently small within finite iteration steps (Theorem 3.2 (a),(b)). These results are new for risk probability criteria in finite horizon SMDPs. The remainder of this paper is organized as follows. We describe the mathematical model and introduce the terminology in Section 2. We then prove that an optimal value function is a unique solution to an optimality equation and there exist optimal (ε-optimal) policies in Section 3. An example illustrating possible applications of the obtained results is given in Section 4.

2. The model and optimality criterion. We shall give a brief review of the main concepts of SMDPs. The model of SMDPs we are concerned with is the five-tuple as below:

\[ (E, (A(i) \subset A, i \in E), Q(\cdot, |i, a), c(i, a)) \]

where \( E \) and \( A \) are denumerable state and action spaces, respectively; \( A(i) \) denotes the set of admissible actions at state \( i \in E \), which is assumed to be finite. The transition mechanism of the SMDPs is described by the semi-Markov kernel \( Q(\cdot, |i, a) \) on \( \mathbb{R}_+ \times E \) given \( K \), where \( \mathbb{R}_+ := [0, +\infty) \) and \( K := \{(i, a) | i \in E, a \in A(i)\} \) is the set of all feasible state-action pairs. It is assumed that (i) \( Q(\cdot, |i, a) \), for any fixed \( j \in E \), and \( (i, a) \in K \), is a nondecreasing, right continuous real function on \( \mathbb{R}_+ \) such that \( Q(0, j|i, a) = 0 \); (ii) \( Q(t, |i, a) \), for each fixed \( t \in \mathbb{R}_+ \), is a sub-stochastic kernel on \( E \) given \( K \); and (iii) \( P(|i, a) := Q(\infty, |i, a) \) is a stochastic kernel on \( E \) given \( K \). If an action \( a \in A(i) \) is chosen in state \( i \), then \( Q(t, j|i, a) \) is the joint probability that the sojourn time in state \( i \) is not greater than \( t \in \mathbb{R}_+ \) and the next state is \( j \in E \). Finally, the real function \( c : K \to \mathbb{R}_+ \) denotes the loss rate.

**Remark 1.** Compared with the risk probability criteria in SMDPs [10, 12, 20], in our model (1) a loss rate is considered, which results in a different definition of the risk probability in (7) below.

A finite horizon SMDP with risk probability criteria evolves as follows: at the initial decision epoch \( s_0 = 0 \), the system occupies state \( i_0 \), and the decision maker has a loss level \( \lambda_0 \) over a planning horizon \( t_0 \) in mind (that is, he/she should try to control the loss incurred \( \lambda_0 \) within the planning horizon \( t_0 \)), then he/she chooses an action \( a_0 \in A(i_0) \) according to the state \( i_0 \), the planning horizon \( t_0 \) and the loss level \( \lambda_0 \). As a consequence of this action selection, the system remains in \( i_0 \) until time \( s_1 \), at which point the system jumps to \( i_1 \) and then the next decision epoch occurs. At time \( s_1 \), a loss \( c(i_0, a_0)(s_1 - s_0) \) is incurred, leading to a remaining loss level \( \lambda_1 := \lambda_0 - c(i_0, a_0)(s_1 - s_0) \) over a remaining planning horizon \( t_1 := [t_0 - (s_1 - s_0)]^+ \) for the decision maker, where \( [x]^+ := \max\{x, 0\} \). According to the current state \( i_1 \), the current planning horizon \( t_1 \) and the current loss level \( \lambda_1 \) as well as the previous state \( i_0 \) and the previous loss level \( \lambda_0 \), the decision maker adopts an action \( a_1 \in A(i_1) \) and the same sequence of events occur. The decision process evolves in this way, and so we obtain an admissible horizon-relevant (h-r in short) history \( h_n \) of the SMDPs up to the \( n \)th decision epoch, i.e.,

\[
h_n = (s_0, i_0, t_0, \lambda_0, a_0, \ldots, s_{n-1}, i_{n-1}, t_{n-1}, \lambda_{n-1}, a_{n-1}, s_n, i_n, t_n, \lambda_n),
\]

where \( s_0 = 0, s_{m+1} \geq s_m, \ (i_m, a_m) \in K, \ t_0 \in \mathbb{R}_+, t_{m+1} := [t_m - (s_{m+1} - s_m)]^+, \ \lambda_0 \in \mathbb{R} := (-\infty, +\infty), \ \lambda_{m+1} := \lambda_m - c(i_m, a_m)(s_{m+1} - s_m) \) for \( m = 0, 1, \ldots, n-1 \), and \( i_n \in E \). Let \( H_n \) denote the set of all admissible h-r histories \( h_n \) of the system up to the \( n \)th decision epoch. Assume that \( H_n \) is endowed with a Borel \( \sigma \)-algebra.
Remark 2. (a) The h-r history $h_n$ here is similar to those in [11, 12]. For more details, readers are referred to Remark 2.2 in [11, 12].

(b) The case $\lambda_n < 0$ is allowed here, which means that the planning horizon $t_n$ maybe equal to 0 in some decision epoch, while in [12] the case $\lambda_n < 0$ is thought to be risk-free on behalf of the decision maker at the $n$th decision epoch.

To state the finite horizon SMDPs with risk probability criteria we are concerned with, we need to introduce the classes of policies.

Definition 2.1. An h-r randomized historic policy, or simply an h-r policy, is a sequence $\pi = \{\pi_n, n \geq 0\}$ of stochastic kernels $\pi_n$ on $A$ given $H_n$ such that

$$\pi_n(A(i_n)|h_n) = 1 \quad \forall h_n \in H_n, n = 0, 1, \ldots.$$  

Remark 3. Compared with the policies in infinite horizon discounted and expected criteria [4, 7, 8, 23], the h-r policy here depends on the loss levels, the decision epochs, the planning horizon, the states and actions.

The set of all h-r policies is denoted by $\Pi$.

Notation. Let $\Phi$ denote the set of all stochastic kernels $\varphi$ on $A$ given $E \times \mathbb{R}_+ \times \mathbb{R}$ satisfying $\varphi(A(i)|i, t, \lambda) = 1$ for all $(i, t, \lambda) \in E \times \mathbb{R}_+ \times \mathbb{R}$, and $\mathcal{F}$ the set of measurable functions $f$ from $E \times \mathbb{R}_+ \times \mathbb{R}$ to $A$ such that $f(i, t, \lambda) \in A(i)$ for all $(i, t, \lambda) \in E \times \mathbb{R}_+ \times \mathbb{R}$.

Let us now introduce several subclasses of policies.

Definition 2.2. (a) An h-r policy $\pi = \{\pi_n\}$ is said to be h-r randomized Markov if there is a sequence $\{\varphi_n\}$ of stochastic kernels $\varphi_n \in \Phi$ such that $\pi_n(\cdot|h_n) = \varphi_n(\cdot|i_n, t_n, \lambda_n)$ for every $h_n \in H_n$ and $n \geq 0$. We write such a policy as $\pi = \{\varphi_n\}$.

(b) An h-r randomized Markov policy $\pi = \{\varphi_n\}$ is said to be h-r randomized stationary if $\varphi_n$ are independent of $n$. We write $\pi = \{\varphi, \varphi, \ldots\}$ as $\varphi$ for brevity.

(c) An h-r randomized Markov policy $\pi = \{\varphi_n\}$ is said to be h-r deterministic Markov if there is a sequence $\{f_n\}$ of measurable functions $f_n \in \mathcal{F}$ such that $\varphi_n(\cdot|i, t, \lambda)$ is the Dirac measure at $f_n(i, t, \lambda)$ for every $(i, t, \lambda) \in E \times \mathbb{R}_+ \times \mathbb{R}$ and $n \geq 0$. We write such a policy as $\pi = \{f_n\}$.

(d) An h-r deterministic Markov policy $\pi = \{f_n\}$ is said to be h-r deterministic stationary if $f_n$ are independent of $n$. We write $\pi = \{f, f, \ldots\}$ for simplicity.

For convenience of description, we denote by $\Pi_{RM}, \Pi_{RS}, \Pi_{DM}$ and $\Pi_{DS}$ the sets of all h-r randomized Markov, h-r randomized stationary, h-r deterministic Markov and h-r stationary policies, respectively. Obviously, $\mathcal{F} = \Pi_{DS} \subset \Pi_{DM} \subset \Pi$ and $\Phi = \Pi_{RS} \subset \Pi_{RM} \subset \Pi$.

Let $(\Omega, \mathcal{F})$ be the measurable space consisting of the sample space $\Omega$ and the corresponding product $\sigma$-algebra $\mathcal{F}$.

For each $\omega \in \Omega$ represented by

$$\omega = (s_0, i_0, t_0, \lambda_0, a_0, \ldots, s_n, i_n, \lambda_n, a_n, \ldots),$$

we define the coordinate variables $S_n$, $J_n$, $T_n$, $\lambda_n$ and $A_n$ $(n = 0, 1, \ldots)$ on $(\Omega, \mathcal{F})$ as follows:

$$S_n(\omega) = s_n, \quad J_n(\omega) = i_n, \quad T_n(\omega) = t_n, \quad \lambda_n(\omega) = \lambda_n, \quad A_n(\omega) = a_n,$$

where $S_n$ denotes the $n$th decision epoch, while $J_n$, $T_n := [T_{n-1} - (S_n - S_{n-1})]^+$, $\lambda_n := \lambda_{n-1} - c(J_{n-1}, A_{n-1})(S_n - S_{n-1})$ and $A_n$ represent the system state, the planning horizon, the loss level and the action chosen at the $n$th decision epoch.
respectively. Also, we denote by $X_0 := 0, X_n := S_n - S_{n-1} (n \geq 1)$ the sojourn times between two successive decision epochs. Given $(i, t, \lambda) \in E \times \mathbb{R}_+ \times \mathbb{R}$ and $\pi \in \Pi$, by the well-known Tulcea’s theorem [7], there exist a unique probability measure on $(\Omega, \mathcal{F})$ satisfying

$$P_{(i,t,\lambda)}^\pi(S_0 = 0, J_0 = i, T_0 = t, \lambda_0 = \lambda) = 1,$$

$$P_{(i,t,\lambda)}^\pi(A_n = a|h_n) = \pi_n(a|h_n),$$

$$P_{(i,t,\lambda)}^\pi(S_{n+1} - S_n \leq s, J_{n+1} = j|h_n, a_n) = Q(s, j|i_n, a_n),$$

$$P_{(i,t,\lambda)}^\pi(T_{n+1} = [t_n - (s_{n+1} - s_n)]^+|h_n, a_n, s_{n+1}) = 1,$$

$$P_{(i,t,\lambda)}^\pi(\lambda_{n+1} = \lambda_n - c(i_n, a_n)(s_{n+1} - s_n)|h_n, a_n, s_{n+1}) = 1. \quad (6)$$

Here, we define an underlying continuous-time state-action process \{Z(t), W(t), t \in \mathbb{R}_+\} corresponding to the stochastic process \{S_n, J_n, A_n, n \geq 0\}, by

$$Z(t) = \begin{cases} J_n, & \text{for } S_n \leq t < S_{n+1}, n = 0, 1, \ldots, \\ \partial_E, & \text{for } t \geq S_\infty, \end{cases}$$

$$W(t) = \begin{cases} A_n, & \text{for } S_n \leq t < S_{n+1}, n = 0, 1, \ldots, \\ \partial_A, & \text{for } t \geq S_\infty, \end{cases}$$

where $\partial_E$ and $\partial_A$ are the extra state and action joined to $E$ and $A$, respectively, and $S_\infty$ is the accumulation point of the sequence \{S_n\}, i.e., $S_\infty := \lim_{n \to \infty} S_n$.

**Definition 2.3.** The stochastic process \{Z(t), W(t), t \in \mathbb{R}_+\} is called a semi-Markov decision process.

To deal with the finite horizon optimization problem, we fix an arbitrary T-horizon (with $T \in \mathbb{R}_+$). To ensure the existence of an optimal policy, we impose the following assumption.

**Assumption 2.1.** $P_{(i,t,\lambda)}^\pi(\{S_\infty > T\}) = 1$, for every $(i, t, \lambda) \in E \times [0, T] \times \mathbb{R}$ and $\pi \in \Pi$.

**Remark 4.** Note that Assumption 2.1 above is trivially fulfilled in DTMDPs (short for discrete-time Markov Decision Processes, which are special SMDPs with sojourn time fixed) [19, 22, 24, 25, 26] with $S_\infty = \infty$.

To verify Assumption 2.1 above, we need the following condition.

**Condition 2.1.** There exist constants $\delta > 0$ and $\xi > 0$ such that

$$D(\delta|i, a) := Q(\delta, E|i, a) \leq 1 - \xi \quad \forall (i, a) \in K.$$

**Remark 5.** For the verification of Assumption 2.1, the reader is referred to the proof of Proposition 2.1 in [11]. Moreover, Condition 2.1 can be more easily verified since it is imposed on the *primitive* data of the model (1).

For each initial state $(i, t, \lambda) \in E \times [0, T] \times \mathbb{R}$, we define the risk probability $F^\pi$ of the SMDP \{Z(t), W(t), t \in \mathbb{R}_+\} under a policy $\pi \in \Pi$ by

$$F^\pi(i, t, \lambda) := P_{(i,t,\lambda)}^\pi \left( \int_0^t c(Z(s), W(s)) \, ds > \lambda \right), \quad (7)$$

which measures the risk of the system that the total loss incurred during the interval $[0, t]$ exceeds $\lambda$ when using the policy $\pi$. Thus, the optimization problem we are
interested in is to minimize the system’s risk \( F^\pi(i, t, \lambda) \) over the set \( \Pi \). That is, our goal is to find a policy \( \pi^* \in \Pi \) such that
\[
F^\pi(i, t, \lambda) = \inf_{\pi \in \Pi} F^\pi(i, t, \lambda) =: F^*(i, t, \lambda) \quad \forall (i, t, \lambda) \in E \times [0, T] \times \mathbb{R},
\]
where \( F^*(i, t, \lambda) \) is the optimal value function.

**Remark 6.** (a) Different from the previous study on the probability criteria in SMDPs with reward rates \([9, 10, 12]\), this paper deals with the loss rate case.
(b) Obviously, we have
\[
(b) \text{ Obviously, we have } F^*(i, t, \lambda) = 1 - P^\pi_{(i, t, \lambda)}(\int_0^t c(Z(s), W(s))ds \leq \lambda).
\]
Our optimality problem can be equivalently transformed to the problem of maximizing \( \hat{F}^*(i, t, \lambda) \), while it is not equivalent to the one in \([12]\) since Huang \([12]\] considers the problem of minimizing \( \hat{F}^*(i, t, \lambda) \). Moreover, the reward rate \( r \) in \([12]\] is replaced by the loss rate \( c \) in this paper.

Using arguments as in the proof of Proposition 2.2 in \([11]\], we can also prove that \( \Pi_{RM} \) is a sufficient set of policies for our optimality problem, i.e.,
\[
F^*(i, t, \lambda) = \inf_{\pi \in \Pi_{RM}} F^\pi(i, t, \lambda) \quad \forall (i, t, \lambda) \in E \times [0, T] \times \mathbb{R},
\]
which indicates that it suffices to find optimal policies in the class of randomized Markov policies.

**Definition 2.4.** For each \( \epsilon > 0 \), a policy \( \pi' \in \Pi \) is called \( \epsilon \)-optimal if
\[
F^{\pi'}(i, t, \lambda) \leq F^*(i, t, \lambda) + \epsilon, \quad \forall (i, t, \lambda) \in E \times [0, T] \times \mathbb{R},
\]
A 0-optimal is simply called an optimal policy.

### 3. On the optimal value function and optimality equation

In this section, we show our main results. That is, we prove that the value function is a unique solution to the optimality equation, and there exists an optimal (or \( \epsilon \)-optimal) policy. In addition, we obtain an algorithm for computing \( \epsilon \)-optimal policies.

We define the following sets of functions. Let \( \mathcal{F}_m \) be the set of functions \( F : E \times [0, T] \times \mathbb{R} \rightarrow [0, 1] \) such that \( F(\cdot, \cdot, \cdot) \) is Borel measurable on \( E \times [0, T] \times \mathbb{R} \) and \( F(i, t, \lambda) = 1 \) if \( \lambda < 0 \) for each \((i, t) \in E \times [0, T]\); and \( \mathcal{F}_m \) the set of functions \( F \in \mathcal{F}_m \) such that \( F(i, t, \cdot) \) is monotone nonincreasing and right continuous on \( \mathbb{R} \) for each \((i, t) \in E \times [0, T]\), while \( F(i, \cdot, \lambda) \) is monotone nondecreasing and right continuous on \([0, T]\) for each \((i, \lambda) \in E \times \mathbb{R}\). Also, we define operators \( H^\varphi, H \) on \( \mathcal{F}_m \) as follows: for \( F \in \mathcal{F}_m \), \( a \in A(i), \ (i, t) \in E \times [0, T] \) and \( \varphi \in \Phi \), if \( \lambda \geq 0 \),
\[
H^\varphi F(i, t, \lambda) := \sum_{a \in A(i)} \varphi(a|i, t, \lambda) H^a F(i, t, \lambda),
\]
\[
H F(i, t, \lambda) := \min_{a \in A(i)} H^a F(i, t, \lambda),
\]
and \( H^a F(i, t, \lambda) = H^\varphi F(i, t, \lambda) = H F(i, t, \lambda) := 1 \) if \( \lambda < 0 \).
Remark 7. (a) The operators $H^\varphi$ and $H$ above are different from those for first passage SMDPs with risk probability criteria [12] and those for finite DTMDPs with risk probability criteria [18]. Also, these operators are important in characterizing the value function and optimal policies for first passage SMDPs with risk probability criteria and loss rates; see Theorem 3.5, Theorem 3.6 below. (b) In fact, the operators $H^\varphi$ and $H$ are monotone, that is, $H^\varphi F \geq H^\varphi G$ and $HF \geq HG$ if $F, G \in \mathcal{F}_m$, and $F \geq G$.

For each $(i, t, \lambda) \in E \times [0, T] \times \mathbb{R}$ and $\pi \in \Pi$, by the definition of $F_\pi(i, t, \lambda)$, we have

$$F_\pi(i, t, \lambda) = P_{(i, t, \lambda)}^\pi \left( \int_0^t c(Z(s), W(s))ds > \lambda \right)$$

$$= P_{(i, t, \lambda)}^\pi \left( \sum_{m=0}^\infty \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda \right)$$

$$= P_{(i, t, \lambda)}^\pi \left( \bigcup_{n=1}^\infty \left( \sum_{m=0}^n \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda \right) \right)$$

$$= \lim_{n \to \infty} P_{(i, t, \lambda)}^\pi \left( \sum_{m=0}^n \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda \right),$$

where the second equality follows from Assumption 2.1, and the last equality is due to the nonnegativity of the loss rate and the continuity of probability measures. Based on (8), we define $F_{\pi-1}(i, t, \lambda) := I_{(-\infty, 0)}(\lambda)$, and

$$F_n(i, t, \lambda) := \left\{ \begin{array}{ll} P_{(i, t, \lambda)}^\pi \left( \sum_{m=0}^n \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda \right), & \lambda \geq 0; \\ 1, & \text{otherwise} \end{array} \right.$$ 

for every $n \geq 0$ and $(i, t) \in E \times [0, T]$. Obviously, $F_n^\pi \leq F_{n+1}^\pi$ for all $n \geq -1$ and $\lim_{n \to \infty} F_n^\pi = F^\pi$.

The following lemma is fundamental to our main results.

Lemma 3.1. Let $\pi = \{\varphi_0, \varphi_1, \ldots\} \in \Pi_{RM}$ be arbitrary.

(a) For each $n \geq -1$, $F_n^\pi \in \mathcal{F}_m$ and $F^\pi \in \mathcal{F}_m$.

(b) For each $n \geq -1$, $F_{n+1}^\pi = H^{\varphi_0} F_n^\pi$ and $F^\pi = H^{\varphi_0} F^\pi$, where $^{(1)}\pi = \{\varphi_1, \varphi_2, \ldots\} \in \Pi_{RM}$. In particular, $F^f = H^f F^f$ for every $f \in \mathbb{F}$.

Proof. (a) To show that $F_n^\pi \in \mathcal{F}_m$, it suffices to prove that $F_n^\pi(i, \cdot, \cdot)$ is Borel-measurable on $[0, T] \times \mathbb{R}$ for each $i \in E$. We show this by induction. It is obviously true for $n = -1$. Suppose that $F_n^\pi(i, \cdot, \cdot)$ is Borel-measurable for some $n \geq -1$ and all $\pi \in \Pi_{RM}$. It then follows that for any $\pi = \{\varphi_0, \varphi_1, \ldots\} \in \Pi_{RM}$, we have

$$H^{\varphi_0} F_n^\pi(i, t, \lambda) = \sum_{a \in A(i)} \varphi_0(a|i, t, \lambda) \left[ I_{(\lambda, +\infty)}(c(i, a)t)(1 - D(t|i, a)) \right]$$

$$+ \sum_{j \in E} \int_0^t F_n^\pi(j, t - u, \lambda - c(i, a)u)Q(du, j|i, a)$$

is well-defined and measurable in $(t, \lambda)$ for each $i \in E$, where $^{(1)}\pi = \{\varphi_1, \varphi_2, \ldots\} \in \Pi_{RM}$. By the property of conditional expectation, the Markov property, (2)–(6)
and a direct calculation, we have

\[
F_{n+1}^\pi(i, t, \lambda) = P_{(i,t,\lambda)}^\pi \left( \sum_{m=0}^{n+1} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda \right) = P_{(i,t,\lambda)}^\pi \left( \sum_{m=0}^{n+1} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda, S_1 > t \right) \\
+ P_{(i,t,\lambda)}^\pi \left( \sum_{m=0}^{n+1} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda, S_1 \leq t \right) \\
= P_{(i,t,\lambda)}^\pi \left( \int_0^t c(J_0, A_0)ds > \lambda, S_1 > t \right) \\
+ P_{(i,t,\lambda)}^\pi \left( \sum_{m=1}^{n+1} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda - c(J_0, A_0)X_1, S_1 \leq t \right) \\
= E_{(i,t,\lambda)}^\pi \left[ P_{(i,t,\lambda)}^\pi \left( c(J_0, A_0)t > \lambda, S_1 > t|S_0, J_0, T_0, \lambda_0, A_0 \right) \right] \\
+ E_{(i,t,\lambda)}^\pi \left[ \sum_{m=1}^{n+1} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda - c(J_0, A_0)X_1, S_1 \leq t \right| \left[ S_0, J_0, T_0, \lambda_0, A_0, S_1, J_1, T_1, \lambda_1 \right] \\
= E_{(i,t,\lambda)}^\pi \left[ I_{(\lambda, +\infty)}(c(J_0, A_0)t) P_{(i,t,\lambda)}^\pi \left( S_1 > t|S_0, J_0, T_0, \lambda_0, A_0 \right) \right] \\
+ E_{(i,t,\lambda)}^\pi \left[ I_{\{S_1 \leq t\}} P_{(i,t,\lambda)}^\pi \left( \sum_{m=1}^{n+1} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda - c(J_0, A_0)X_1 \right| S_0, J_0, T_0, \lambda_0, A_0, S_1, J_1, T_1, \lambda_1 \right] \\
= \sum_{a \in A(i)} \varphi_0(a|i, t, \lambda) \left[ I_{(\lambda, +\infty)}(c(i, a)t)(1 - D(t|i, a)) \right] + \sum_{a \in A(i)} \varphi_0(a|i, t, \lambda) \sum_{j \in E} \int_0^t Q(du, j|i, a) P_{(i,t,\lambda)}^\pi \left( \sum_{m=1}^{n+1} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda - c(J_0, A_0)X_1|S_0 = 0, J_0 = i, T_0 = t, \lambda_0 = \lambda, A_0 = a, S_1 = u, J_1 = j, T_1 = [t - u]^+, \lambda_1 = \lambda_0 - c(i, a)u \right) \\
= \sum_{a \in A(i)} \varphi_0(a|i, t, \lambda) \left[ I_{(\lambda, +\infty)}(c(i, a)t)(1 - D(t|i, a)) + \sum_{j \in E} \int_0^t Q(du, j|i, a) \right. \\
\times P_{(i,t-u,\lambda_0-c(i,a)u)}^{(1)} \left[ \sum_{m=0}^{n} \int_{S_m \wedge t}^{S_{m+1} \wedge t} c(Z(s), W(s))ds > \lambda - c(i, a)u \right] \\
= \sum_{a \in A(i)} \varphi_0(a|i, t, \lambda) \left[ I_{(\lambda, +\infty)}(c(i, a)t)(1 - D(t|i, a)) + \sum_{j \in E} \int_0^t Q(du, j|i, a) \right. \\
\times F_n^{(1)}(j, t-u, \lambda_0 - c(i, a)u) \\
\left. \times F_n^{(1)}(j, t-u, \lambda_0 - c(i, a)u) \right]
\]

where the fourth equality is due to the property of conditional expectation, the sixth equality follows from the properties (2)–(6), and the seventh equality is due
to the Markov property. Also, note that \( F_{n+1}^\pi(i, t, \lambda) = H_iF_n^{(1)\pi}(i, t, \lambda) = 1 \) for \( \lambda < 0 \) by definitions. Hence,

\[
F_{n+1}^\pi(i, t, \lambda) = H_iF_n^{(1)\pi}(i, t, \lambda) \quad \forall (i, t, \lambda) \in E \times [0, T] \times \mathbb{R},
\]

and thus \( F_{n+1}^\pi(i, \cdot, \cdot) \) is measurable in \((t, \lambda)\) for each \( i \in E \). Then, by the induction hypothesis, \( F_n^\pi(i, \cdot, \cdot) \) is measurable for all \( n \geq 1 \). Moreover, since a (pointwise) limit of measurable functions is still measurable, we have \( F_\pi = \lim_{n \to \infty} F_n^\pi \in \mathcal{F}_m \).

(b) From the proof of part (a), we see \( F_{n+1}^\pi = H_iF_n^{(1)\pi} \). Letting \( n \to \infty \) and using the dominated convergence theorem, we obtain \( F_\pi = H_iF^{(1)\pi} \), and so the last statement is obvious.

To give the proof of the following Lemma, we need to introduce the following notation. Let \( \mathcal{F}_m \) be the set of functions \( F : E \times [0, T] \times \mathbb{R} \to [-1, 1] \) such that \( F(i, \cdot, \cdot) \) is measurable on \([0, T] \times \mathbb{R}\) for each \( i \in E \). Define an operator from \( \mathcal{F}_m \) into itself as follows: for each \( f \in \mathcal{F} \), \( F \in \mathcal{F}_m \) and \((i, t) \in E \times [0, T] \), if \( \lambda \geq 0 \),

\[
\mathcal{P}^f F(i, t, \lambda) := \sum_{j \in E} \int_0^t F(j, t-u, \lambda - c(i, f(i, t, \lambda))u)Q(du, j|i, f(i, t, \lambda)),
\]

and \( \mathcal{H}^f F(i, t, \lambda) := 1 \) for \( \lambda < 0 \).

The following lemma is key to the existence of optimal policies.

**Lemma 3.2.** Suppose that Condition 2.1 holds. Then, the following statements hold.

(a) If \( F, G \in \mathcal{F}_m \) and \( F - G \leq \mathcal{P}^f (F - G) \) for some policy \( f \in F \), then \( F \leq G \).

(b) For every \( f \in \mathcal{F} \), \( \mathcal{P}^f \) is the unique solution in \( \mathcal{F}_m \) to the equation \( F = H_iF \).

**Proof.** (a) To show the desired result, we first introduce a new notation. Define

\[
F_\delta(t) := \begin{cases} 
1, & t \geq \delta \\
1 - \xi, & 0 \leq t < \delta \\
0, & t < 0,
\end{cases}
\]

where \( \xi, \delta \) are as in Condition 2.1. Let \( F_\delta^{(n)} \) denote the \( n \)-fold convolution of \( F_\delta \), defined by

\[
F_\delta^{(n)}(t) = \int_0^t F_\delta^{(n-1)}(t-s)dF_\delta(s), \quad t \in R_+.
\]

It follows from Condition 2.1 and the argument of the proof of Theorem 1 in [17] and Lemma 3.1 in [13] that for any \( n > k \) and \( t > 0 \),

\[
F_\delta^{(n)}(t) \leq (1 - \xi^k)^{|n/k|},
\]

where \( k \) is a nonnegative integer satisfying \( k > T/\delta \), and \(|n/k|\) denotes the largest integer not larger than \( n/k \). Below, under Condition 2.1, we use an induction argument to show that

\[
(\mathcal{P}^f)^n F(i, t, \lambda) - (\mathcal{P}^f)^n G(i, t, \lambda) \leq F_\delta^{(n)}(t), \quad (9)
\]

for all \((i, t, \lambda) \in E \times [0, T] \times \mathbb{R}, n \geq 1 \).

In fact, when \( n = 1 \), noting that \( F - G \leq 1 \), we get

\[
\mathcal{P}^f (F - G)(i, t, \lambda) \leq D(t|i, f(i)) \leq F_\delta(t),
\]

where the last inequality is a direct result of Condition 2.1 and the definition of \( F_\delta(t) \).
Suppose (9) holds for \( l = n \). For the case of \( l = n + 1 \),
\[
(\overline{F}^l)^{n+1}(F - G)(i, t, \lambda) = \overline{F}^l((\overline{F}^l)^n(F - G))(i, t, \lambda) \leq \overline{F}^l(F^{(n)}_\delta)(i, t, \lambda) = \sum_{j \in E} \int_0^t F^{(n)}_\delta(t - u)Q(du, j|i, f(i, t, \lambda)).
\]

Suppose \( t \in [m_\delta, (m_\delta + 1)\delta) \) for some nonnegative integer \( m_\delta \), then since \( F^{(n)}_\delta(t) \) is a step function, we may rewrite the above integral as
\[
(\overline{F}^l)^{n+1}(F - G)(i, t, \lambda) \leq pF^{(n)}_\delta(t)Q([0, t - m_\delta\delta), E|i, f(i, t, \lambda)) + F^{(n)}_\delta(t - \delta)Q([t - (m_\delta - 1)\delta, t - m_\delta\delta), E|i, f(i, t, \lambda))
\]
\[
+ F^{(n)}_\delta(t - 2\delta)Q([t - (m_\delta - 2)\delta, t - (m_\delta - 1)\delta), E|i, f(i, t, \lambda)) + \cdots + F^{(n)}_\delta(0)Q([t - t, t), E|i, f(i, t, \lambda)).
\]

For sake of notation, we denote \( p = Q([0, t - m_\delta\delta), E|i, f(i, t, \lambda)) \). According to Condition 2.1, we have \( p \leq 1 - \xi \). Thus there exists a non-negative real number \( q \), such that \( p + q = 1 - \xi \).

\[
(\overline{F}^l)^{n+1}(F - G)(i, t, \lambda) \leq pF^{(n)}_\delta(t) + (1 - p)F^{(n)}_\delta(t - \delta)
\]
\[
= (p + q - q)F^{(n)}_\delta(t) + (1 - p)F^{(n)}_\delta(t - \delta)
\]
\[
\leq (p + q)F^{(n)}_\delta(t) - qF^{(n)}_\delta(t - \delta) + (1 - p)F^{(n)}_\delta(t - \delta)
\]
\[
= (1 - \xi)F^{(n)}_\delta(t) + \xi F^{(n)}_\delta(t - \delta)
\]
\[
= F^{(n+1)}_\delta(t),
\]
where the last inequality is due to the non-decreasing property of \( F^{(n)}_\delta(t) \).

By the induction hypothesis, if \( F - G \leq \overline{F}^l(F - G) \), then we have
\[
(F - G)(i, t, \lambda) \leq (\overline{F}^l)^n(F - G)(i, t, \lambda) \leq F^{(n)}_\delta(t) \leq (1 - \xi^k)^{n/k}
\]
for every \( i, t, \lambda \in E \times [0, T] \times \mathbb{R} \) and all \( n \geq 1 \). Therefore, noting that \( 0 < 1 - \xi^k < 1 \) and letting \( n \to \infty \) in (10), it is clear that \( F(i, t, \lambda) \leq G(i, t, \lambda) \) for every \( i, t, \lambda \in E \times [0, T] \times \mathbb{R} \).

(b) It follows from Lemma 3.1(b), \( F^f \in \mathcal{F}_m \) is a solution to the equation \( F = H^f F \).

Let \( G \) be another solution in \( \mathcal{F}_m \) to \( F = H^f F \). Then, we have \( F^f - G = \overline{F}^l(F^f - G) \), which gives \( F^f = G \) by part (a), and so the uniqueness follows.

\[\square\]

Lemma 3.3. Let \( \{F_n, n \geq -1\} \) be a nondecreasing sequence in \( \mathcal{F}_m \). Then, \( \lim_{n \to \infty} HF_n(i, t, \lambda) = H \lim_{n \to \infty} F_n(i, t, \lambda) \) for every \( (i, t, \lambda) \in E \times [0, T] \times \mathbb{R} \).

Proof. Since \( F_n \uparrow F \), the limit \( F = \lim_{n \to \infty} F_n \exists \) exists. It follows from the monotonicity of the operator \( H \) that \( HF_n \leq HF \) for each \( n \geq -1 \). Therefore, \( \lim_{n \to \infty} HF_n \leq HF \). We shall show that the reverse inequality is also true if \( A(\cdot) \) is finite. Fix an arbitrary \( (i, t, \lambda) \in E \times [0, T] \times \mathbb{R} \), we consider the sets
Lemma 3.4. Suppose that Assumption 2.1 holds. By the finiteness of $A(i)$, the monotonicity of $H$, and the assumption that $F_n \uparrow F$, we see that each of these sets is nonempty and $A_n \downarrow A^*$ for each $n \geq -1$. Let $a_n$ be such that $H^{a_n} F_n(i, t, \lambda) = H F_n(i, t, \lambda)$ and $a_n \in A_n$. By the finiteness of $A(i)$ and the fact $A_n \downarrow A^*$, there exists $a^* \in A^*$ and a subsequence $\{a_{n_k}\}$ of $a_n$ satisfying $a_{n_k} = a^*$ for any $n_k \geq m$. Since $H$ is a monotone operator and $F_n \uparrow F$, we have $\lim_{k \to \infty} H F_n(i, t, \lambda)$ exists and

$$
\lim_{n \to \infty} H F_n(i, t, \lambda) = \lim_{k \to \infty} H^{a_{n_k}} F_{n_k}(i, t, \lambda) = \lim_{k \to \infty} H^{a^*} F_m(i, t, \lambda).
$$

(13)

Letting $m \to \infty$ in the above expression gives that $\lim_{n \to \infty} H F_n(i, t, \lambda) \geq H F(i, t, \lambda)$. The proof is completed. \hfill \Box

In fact, Lemma 3.3 gives a condition of interchanging limits and minima.

Lemma 3.4. Suppose that Assumption 2.1 holds.

(a) If $F \in \mathcal{F}_m$, then $HF \in \mathcal{F}_m$, and there exists an $f \in \mathcal{F}$ such that $HF = H^f F$.

(b) If $F \in \mathcal{F}_r$, then both $H^a F$ and $HF$ are in $\mathcal{F}_r$ for any $a \in A(i)$.

(c) If $F_n \in \mathcal{F}_r$ and $F_n \leq F_{n+1}$ for each $n \geq 0$, then $\lim_{n \to \infty} F_n \in \mathcal{F}_r$.

Proof. (a) Under Assumption 2.1, the measurable selection theorem (see Proposition D.5 in [7]) ensures the existence of an $f \in \mathcal{F}$ such that

$$
H^f F(i, t, \lambda) = \min_{a \in A(i)} H^a F(i, t, \lambda) = HF(i, t, \lambda)
$$

for every $(i, t, \lambda) \in E \times [0, T] \times \mathbb{R}$, and thus (a) follows.

(b) It follows from the definition of $H^a$ and $F \in \mathcal{F}_r$ that $H^a F \in \mathcal{F}_m$, and furthermore, $H^a F(i, t, \cdot)$ is monotone nonincreasing and right continuous on $\mathbb{R}$ for each $(i, t) \in E \times [0, T]$, and on the other hand, $H^a F(i, \cdot, \lambda)$ is right continuous on $[0, T]$ for each $(i, \lambda) \in E \times \mathbb{R}$ and $a \in A(i)$. To prove that $H^a F \in \mathcal{F}_r$, we need only show that $H^a F(i, \cdot, \lambda)$ is monotone nondecreasing on $[0, T]$ for each $(i, \lambda) \in E \times \mathbb{R}$. Indeed, for fixed $(i, \lambda) \in E \times \mathbb{R}$ and $a \in A(i)$, if $t_2 > t_1 > \frac{\lambda}{c(i, a)}$, by a direct calculation, we see that

$$
H^a F(i, t_2, \lambda) - H^a F(i, t_1, \lambda)
$$

$$
= [1 - D(t_2 | i, a)] + \sum_{j \in E} \int_{\frac{\lambda}{c(j, a)}}^{\frac{\lambda}{c(i, a)}} Q(du, j | i, a) F(j, t_2 - u, \lambda - c(i, a) u)
$$

$$
+ \sum_{j \in E} \int_{\frac{\lambda}{c(i, a)}}^{t_2} Q(du, j | i, a) \times [1 - D(t_1 | i, a)] - \sum_{j \in E} \int_{0}^{\frac{\lambda}{c(j, a)}} Q(du, j | i, a)
$$

$$
\times F(j, t_1 - u, \lambda - c(i, a) u) - \sum_{j \in E} \int_{\frac{\lambda}{c(j, a)}}^{t_1} Q(du, j | i, a) \times 1
$$

$$
= \sum_{j \in E} \int_{0}^{\frac{\lambda}{c(j, a)}} Q(du, j | i, a) F(j, t_2 - u, \lambda - c(i, a) u)
$$
which implies that
\[ H^aF(i, t_2, \lambda) - H^aF(i, t_1, \lambda) \]

\[ = \sum_{j \in E} \int_0^{t_2} Q(du, j|i, a)F(j, t_2 - u, \lambda - c(i, a)u) \]

\[ - \sum_{j \in E} \int_0^{t_1} Q(du, j|i, a)F(j, t_1 - u, \lambda - c(i, a)u) \]

\[ = \sum_{j \in E} \int_0^{t_1} Q(du, j|i, a)F(j, t_2 - u, \lambda - c(i, a)u) \]

\[ + \sum_{j \in E} \int_{t_1}^{t_2} Q(du, j|i, a)F(j, t_2 - u, \lambda - c(i, a)u) \]

\[ - \sum_{j \in E} \int_0^{t_1} Q(du, j|i, a)F(j, t_1 - u, \lambda - c(i, a)u) \]

\[ = \sum_{j \in E} \int_0^{t_1} Q(du, j|i, a)\left[F(j, t_2 - u, \lambda - c(i, a)u) - F(j, t_1 - u, \lambda - c(i, a)u)\right] \]

\[ + \sum_{j \in E} \int_{t_1}^{t_2} Q(du, j|i, a)F(j, t_2 - u, \lambda - c(i, a)u) \]

\[ \geq 0 \]

which implies that \( H^aF(i, \cdot, \lambda) \) is monotone nondecreasing on \([0, T]\) for each \((i, \lambda) \in E \times \mathbb{R}\).

We now turn to proving \( HF \in \mathcal{F}_r \). By part (a), we easily see that \( HF \in \mathcal{F}_m \). By the monotonicity of \( H \), \( HF(i, t, \cdot) \) is nonincreasing on \( \mathbb{R} \) and \( HF(i, \cdot, \lambda) \) is nondecreasing on \([0, T]\). Hence, to prove that \( HF \in \mathcal{F}_r \), it need only to show that \( HF(i, t, \cdot) \) is right continuous on \( \mathbb{R} \) for each \((i, t) \in E \times [0, T]\) and \( HF(i, \cdot, \lambda) \) is right continuous on \([0, T]\) for each \((i, \lambda) \in E \times \mathbb{R}\). Indeed, for each \( \lambda \in \mathbb{R} \), let \( \{\lambda_k\} \) be an arbitrary sequence such that \( \lambda_k \downarrow \lambda \). Using the monotonicity of \( HF(i, t, \cdot) \) on \( \mathbb{R} \), we have

\[ HF(i, t, \lambda_k) \leq HF(i, t, \lambda). \]  \hspace{1cm} (14)

which implies that

\[ \lim_{\lambda_k \downarrow \lambda} HF(i, t, \lambda_k) \leq HF(i, t, \lambda). \]  \hspace{1cm} (15)

On the other hand, since \( A(i) \) is finite, there exists a subsequence \( \{\lambda_{k_m}\} \) and \( a^* \in A(i) \) such that \( HF(i, t, \lambda_{k_m}) = H^{a^*}F(i, t, \lambda_{k_m}) \). Then, by the right continuity of \( H^{a^*}F \) for any \( a \in A(i) \), it follows that

\[ \lim_{\lambda_k \downarrow \lambda} HF(i, t, \lambda_k) = \lim_{\lambda_{k_m} \downarrow \lambda} H^{a^*}F(i, t, \lambda_{k_m}) = H^{a^*}F(i, t, \lambda) \geq HF(i, t, \lambda). \]  \hspace{1cm} (16)

Therefore, combining (15) with (16) gives the desired result that \( \lim_{\lambda_k \downarrow \lambda} HF(i, t, \lambda_k) = HF(i, t, \lambda) \). We next show that \( HF(i, \cdot, \lambda) \) is right continuous on \([0, T]\) for each
Theorem 3.5. Under Condition 2.1, the following assertions hold.

(a) From Lemma 3.4 (b), we have \( F_n^\ast \in \mathcal{F}_n \). Hence, \( F_n^\ast \) is in \( \mathcal{F}_m \), and thus \( HF_n^\ast \) is well defined for every \( n \geq -1 \). Furthermore, it is easy to see that \( F_n^\ast \leq F_{n+1}^\ast \) for all \( n \geq -1 \), and hence \( \lim_{n \to -\infty} F_n^\ast =: \bar{F} \) exists. By Lemma 3.4 (c), \( \bar{F} \) is in \( \mathcal{F}_r \).

To complete the proof, it suffices to prove that \( \bar{F} \leq F^\ast \) and \( \bar{F} \geq F^\ast \).

To show \( \bar{F} \leq F^\ast \), it suffices to show that \( F_n^\ast \leq F_n^\pi \) for all \( \pi \in \Pi_{RM} \) and \( n \geq -1 \). We now prove this fact by induction. Indeed, it is obviously true for \( n = -1 \). Suppose that \( F_n^\ast \leq F_n^\pi \) for all \( \pi \in \Pi_{RM} \) and some \( n \geq -1 \). Then, for any \( \eta = \{\eta_0, \eta_1, \ldots\} \in \Pi_{RM} \), we have

\[
F_{n+1}^\ast = HF_n^\ast \leq HF_n^{(1)} \eta \leq H^{\eta_0} F_n^{(1)} \eta = F_{n+1}^{\eta_0}.
\]

where the first inequality is due to the induction hypothesis, and the last equality follows from Lemma 3.1 (b). Therefore, by the induction hypothesis, we obtain \( F_n^\ast \leq F_n^\pi \) for all \( \pi \in \Pi_{RM} \) and \( n \geq -1 \). Hence, \( \bar{F} \leq F^\pi \) for all \( \pi \in \Pi_{RM} \), which together with the arbitrariness of \( \pi \) yields \( \bar{F} \leq F^\ast \).
Now it remains to prove that $\tilde{F} \geq F^*$. By Lemma 3.4 (a), we obtain some decision rule $f \in \Pi_{DM}$ such that $\tilde{F} = F^f$. Indeed, it follows from Lemma 3.3 that $\tilde{F} = \lim_{n \to \infty} HF_n^* = H \lim_{n \to \infty} F_n^* = H F \tilde{F}$. On the other hand, by Lemma 3.4 (a), there exists an $f \in \Pi_{DM}$ such that $HF = H^f \tilde{F}$. Therefore, we have $\tilde{F} = H^f \tilde{F}$. It follows from Lemma 3.2 that $\tilde{F} = F^f$, so that $\tilde{F} \geq F^*$. We already know $\tilde{F} \leq F^*$. This completes the proof.

(b) It follows from the proof of Theorem 3.1 (a) that $F^*$ satisfies the optimality equation $F^* = HF^*$. Suppose that $G$ is a solution to $F = HF$. By Lemma 3.4 (a), there exist decision rules $g$, $f$ such that $G = H\eta G$ and $F^* = H^f F^*$. Hence, $G = H\eta G \leq H^f$ and $F^* = H^f F^* \leq H^f G^*$ on $E \times [0, T] \times \mathbb{R}$. Thus, $F^* - G \leq H^f F^* - G^*$ and $G - F^* \leq H^f (G - F^*)$ on $E \times [0, T] \times \mathbb{R}$. By Lemma 3.2, this yields that $G = F^*$.

(c) Combining part (b) and Lemma 3.4 yields that there exists $f \in F$ such that $F^* = H^f F^*$. The optimality of $f \in F$ is an immediate result of Lemma 3.2 (b).

We now state another main result concerning the approximation value of $F^*$ and the rules of iteration.

**Theorem 3.6.** Suppose that Condition 2.1 holds.

(a) For any $n \geq 1$, we have

$$\sup_{i,t,\lambda} \left| F^*_{nk}(i, t, \lambda) - F^*_{nk-1}(i, t, \lambda) \right| \leq (1 - \xi^k) \sup_{i,t,\lambda} \left| F^*_{(n-1)k}(i, t, \lambda) - F^*_{(n-1)k-1}(i, t, \lambda) \right|, \tag{18}$$

where $k := \left\lceil \frac{2}{\xi} \right\rceil + 1$ as in Lemma 3.2. Here and below, we write $\sup_{i,t,\lambda}$ instead of $\sup_{(i,t,\lambda) \in E \times [0, T] \times \mathbb{R}}$ for sake of convenience.

(b) Given any sufficiently small $\epsilon > 0$, let $n_0 := \left\lceil \frac{\ln(1/\epsilon)}{\ln(1-\xi^k)} \right\rceil$. Then, we have

$$0 \leq F^*(i, t, \lambda) - F^*_{n_0k}(i, t, \lambda) < \epsilon. \tag{19}$$

(c) There exists a policy $f^* \in F$ satisfying $F^*_{(n_0+1)k} = H^{f^*} F^*_{(n_0+1)k-1}$, and such a policy is $\epsilon$-optimal, with $n_0$ as in (b) above.

**Proof.** (a) For fixed $n \geq 1$,

$$F^*_{nk}(i, t, \lambda) - F^*_{nk-1}(i, t, \lambda) = HF^*_{nk-1}(i, t, \lambda) - HF^*_{nk-2}(i, t, \lambda) \leq H^{f_i} F^*_{nk-1}(i, t, \lambda) - H^{f_i} F^*_{nk-2}(i, t, \lambda) = \overline{H}^{f_i} (HF^*_{nk-2}(i, t, \lambda) - HF^*_{nk-3}(i, t, \lambda)) \leq \overline{H}^{f_i} (\overline{H}^{f_2} F^*_{nk-2}(i, t, \lambda) - \overline{H}^{f_2} F^*_{nk-3}(i, t, \lambda)) \leq ... \leq \overline{H}^{f_i} (\overline{H}^{f_2} \cdots (\overline{H}^{f_k} (F^*_{(n-1)k}(i, t, \lambda) - F^*_{(n-1)k-1}(i, t, \lambda)) \leq (1 - \xi^k) \sup_{i,t,\lambda} \left| F^*_{(n-1)k}(i, t, \lambda) - F^*_{(n-1)k-1}(i, t, \lambda) \right|$$

where $f_l \in F, l = 1, 2, ..., k$ satisfy $HF^*_{nk-l-1} = H^{f_l} F^*_{nk-l-1}$, and the last inequality can be proved in the same way as (9).
(b) We claim that
\[ 0 \leq F_{(n_0+1)k}^* - F_{(n_0+1)k-1}^* \leq \xi^k(1 - \xi^k)^{l-1}, \quad \forall \, l \geq 0. \] 
Indeed, for all \( l \geq 0 \), it is clear that \( F_{(n_0+1)k}^* - F_{(n_0+1)k-1}^* \geq 0 \). Next we prove by induction that \( F_{(n_0+1)k}^* - F_{(n_0+1)k-1}^* \leq \xi^k(1 - \xi^k)^{l-1} \). By (18), we see that
\[
F_{n_0k}^* - F_{n_0k-1}^* \\
\leq (1 - \xi^{n_0}) \sup_{i,t,\lambda} \left[ F_0^*(i,t,\lambda) - F_1^*(i,t,\lambda) \right] \\
\leq (1 - \xi^{n_0}) = (1 - \xi^{k})^{\sup_{i,t,\lambda} \left[ F_1^*(i,t,\lambda) - F_0^*(i,t,\lambda) \right]} < (1 - \xi^{k})^{\sup_{i,t,\lambda} \left[ F_1^*(i,t,\lambda) - F_0^*(i,t,\lambda) \right]} = \frac{\xi^k}{k}(1 - \xi^k)^{l-1},
\]
which shows (21) for \( l = 0 \). Suppose that \( F_{(n_0+l)k}^* - F_{(n_0+l)k-1}^* \leq \frac{\xi^k}{k}(1 - \xi^k)^{l-1} \) for some \( l \geq 0 \). By the induction hypothesis, we find that
\[
F_{(n_0+l+1)k}^*(i,t,\lambda) - F_{(n_0+l+1)k-1}^*(i,t,\lambda) \\
\leq (1 - \xi^k) \sup_{i,t,\lambda} \left[ F_{(n_0+l)k}^*(i,t,\lambda) - F_{(n_0+l)k-1}^*(i,t,\lambda) \right] \\
< \frac{\xi^k}{k}(1 - \xi^k)^{l+1}.
\]
Thus (21) holds for all \( k \geq 0 \). It follows that
\[
\sup_{i,t,\lambda} \left[ F_1^*(i,t,\lambda) - F_{n_0k-1}^*(i,t,\lambda) \right] \\
= \sup_{i,t,\lambda} \left[ \sum_{n=1}^{\infty} \sum_{m=1}^{k} (F_{(n_0+l-1)k+m-1}^* - F_{(n_0+l-1)k+m-2}^*)(i,t,\lambda) \right] \\
\leq \sum_{n=1}^{\infty} \sum_{m=1}^{k} \sup_{i,t,\lambda} \left[ F_{(n_0+l-1)k+m-1}^* - F_{(n_0+l-1)k+m-2}^*(i,t,\lambda) \right] \\
< \sum_{n=1}^{\infty} \sum_{m=1}^{k} \frac{\xi^k}{k}(1 - \xi^k)^{l-1} \\
= \epsilon,
\]
where the second inequality is due to (21) and the nonincreasing property of
\[
\left\{ \sup_{i,t,\lambda} \left[ (F_{n+1}^* - F_n^*)(i,t,\lambda) \right], n \geq -1 \right\}.
\]
In fact, for all \((i,t,\lambda) \in E \times [0,T] \times \mathbb{R}, \)
\[
F_{n+1}^*(i,t,\lambda) - F_n^*(i,t,\lambda) \\
= H F_n^*(i,t,\lambda) - H F_{n-1}^*(i,t,\lambda) \\
\leq H f_1 F_n^*(i,t,\lambda) - H f_1 F_{n-1}^*(i,t,\lambda) \\
= \sum_{j \in E} \int_0^t Q(du,j|i,f(i))(F_n^* - F_{n-1}^*)(j,t - u,\lambda - c(i,f(i))u) \\
\leq \sup_{i,t,\lambda} (F_n^* - F_{n-1}^*)(i,t,\lambda),
\]
where \( f_1 \in F \) satisfies \( HF_{n-1}^* = H f_1 F_{n-1}^* \). Thus the nonincreasing property follows.
Since \( F_{n_0k-1} \leq F_{n_0k} \leq F_n^* \), part (b) follows.
(c) The existence of $f'$ is ensured by Theorem 3.5 (a) and Lemma 3.4 (a). Using the argument as in (20) as well as Lemma 3.4 (b), we have
\[
\sup_{i,t,\lambda} \left[ F'(i, t, \lambda) - F^*(n_0+1)_K(i, t, \lambda) \right] \\
= \sup_{i,t,\lambda} \left[ H F'(i, t, \lambda) - H F^*_{n_0, k+1}(i, t, \lambda) \right] \\
\leq (1 - \xi^k) \sup_{i,t,\lambda} \left[ F'(i, t, \lambda) - F^*_{n_0, k+1}(i, t, \lambda) \right] \\
\leq (1 - \xi^k) \left\{ \sup_{i,t,\lambda} \left[ F'(i, t, \lambda) - F^*_{n_0+1}(i, t, \lambda) \right] + \sup_{i,t,\lambda} \left[ F^*_{n_0+1}(i, t, \lambda) - F^*_{n_0+1, k+1}(i, t, \lambda) \right] + \cdots + \sup_{i,t,\lambda} \left[ F^*_{n_0+1}(i, t, \lambda) - F^*_{n_0+1, k+1}(i, t, \lambda) \right] \right\},
\]
which together with (18), gives
\[
\sup_{i,t,\lambda} \left[ F'(i, t, \lambda) - F^*(n_0+1)_K(i, t, \lambda) \right] \\
\leq k \frac{1 - \xi^k}{\xi^k} \sup_{i,t,\lambda} \left[ F^*_{n_0+1}(i, t, \lambda) - F^*_{n_0+1, k+1}(i, t, \lambda) \right] \\
\leq k \frac{1 - \xi^k}{\xi^k} \epsilon \xi^k (1 - \xi^k)^{-1} \\
= \epsilon.
\]
Therefore, the fact $F^*_{n_0+1, k} \geq F^*$ yields that $F' - F^* \leq F' - F^*_{n_0+1, k} \leq \epsilon$, which means that $f'$ is an $\epsilon$-optimal policy. \hfill \Box

By Theorem 3.6, we develop an effective way of evaluating the value function and optimal policies.

**The iteration algorithm for $\epsilon$-optimal policies** includes the following two steps:

**Step 1.** For any sufficiently small $\epsilon > 0$, compute $F^*_{n_0+1, k-1}$ and $F^*_{n_0+1, k}$ using the iteration algorithm in Theorem 3.5(a), with $n_0 = \left\lfloor \frac{\ln(1 - \epsilon^k)}{\ln(1 - \xi^k)} \right\rfloor$.

**Step 2.** Seek a policy $f'$ such that $F^*_{n_0+1, k} = H f' F^*_{n_0+1, k-1}$. Then, by Theorem 3.6 (c), the policy $f'$ is $\epsilon$-optimal.

4. **An application to financial market.** In this section, we apply our results to the business cycle and demonstrate how to compute the value function and an $\epsilon$ optimal policy based on the value iteration algorithm.

**Example 4.1.** (A business cycle) Consider a business cycle with three states, say 1, 2 and 3, which represent the depression, the recovery and the peak ones, respectively. During recessions, the company decision maker try to take the corresponding investment strategy to adapt their company to fluctuation in the business cycle. Suppose that in state 1 the decision maker may take either an action $a_{11}$ with a loss rate $c(1, a_{11})$ or an action $a_{12}$ with a loss rate $c(1, a_{12})$, while in state 2 he may choose another action $a_{21}$ with a loss rate $c(2, a_{21})$ or another action $a_{22}$ with a loss rate $c(2, a_{22})$. However, the system in state 3 incurs the loss rate $c(3, a_{31}) = 0$ with non-action denoted by $a_{31}$. The transition mechanism of the model is given by the
that the distribution function $\mathcal{G}$ is of the following particular form:

$$G \leq 0$$

A finite horizon $[0, T]$ is chosen when the system is in state 2, the system jumps to state 3 with probability one after an exponential-distributed random time with parameter $\mu(2, a)$ after 0. Moreover, we assume that when it falls in state 3, the system transits to state 3 with probability one after an exponential-distributed random time with parameter $\mu(3, a)$. For such a business cycle, the company manager wishes to find an $\epsilon$ optimal policy with the minimum risk probability over a finite horizon $[0, T]$ with $T > 0$.

We now formulate the system as an SMDP model with the data as follows. The state space $E = \{1, 2, 3\}$; the action sets $A(1) = \{a_{11}, a_{12}\}$, $A(2) = \{a_{21}, a_{22}\}$ and $A(3) = \{a_{31}\}$; the horizon $T$ is assumed to be 15; the semi-Markov kernel $Q(\cdot, \cdot | i, a)$ is of the following particular form: $Q(t, j | i, a) = G(t | i, a)p(j | i, a)$ for every $t \in \mathbb{R}_+$, $i \in E$, $j \in E$ and $a \in A(i)$, in which $G(t | i, a)$ and $p(j | i, a)$ denote the distribution functions of the sojourn time and the transition probabilities, respectively. Suppose that the distribution function $G(t | i, a)$ are given by

$$G(t | 1, a_{11}) = \begin{cases} \frac{t}{\mu(1, a_{11})}, & 0 \leq t \leq \mu(1, a_{11}) \\ 1, & t > \mu(1, a_{11}) \end{cases}$$

$$G(t | 1, a_{12}) = \begin{cases} \frac{t}{\mu(1, a_{12})}, & 0 \leq t \leq \mu(1, a_{12}) \\ 1, & t > \mu(1, a_{12}) \end{cases}$$

$$G(t | 2, a_{21}) = 1 - e^{-\mu(2, a_{21})t}, \quad t \in \mathbb{R}_+$$

$$G(t | 2, a_{22}) = 1 - e^{-\mu(2, a_{22})t}, \quad t \in \mathbb{R}_+$$

$$G(t | 3, a_{31}) = 1 - e^{-\mu(3, a_{31})t}, \quad t \in \mathbb{R}_+$$

with the parameters

$$\mu(1, a_{11}) = 30, \quad \mu(1, a_{12}) = 20, \quad \mu(2, a_{21}) = 0.05,$$

$$\mu(2, a_{22}) = 0.025 \quad \mu(3, a_{31}) = 0.3. \quad (22)$$

The transition probabilities of states, $p(j | i, a)$, are defined by

$$p(1 | 1, a_{11}) = 0, \quad p(2 | 1, a_{11}) = 0.9, \quad p(3 | 1, a_{11}) = 0.1;$$

$$p(1 | 1, a_{12}) = 0, \quad p(2 | 1, a_{12}) = 0.7, \quad p(3 | 1, a_{12}) = 0.3;$$

$$p(1 | 2, a_{21}) = 0, \quad p(2 | 2, a_{21}) = 0.6, \quad p(3 | 2, a_{21}) = 0.4;$$

$$p(1 | 2, a_{22}) = 0, \quad p(2 | 2, a_{22}) = 0.2, \quad p(3 | 2, a_{22}) = 0.8;$$

$$p(1 | 3, a_{31}) = 0, \quad p(2 | 3, a_{31}) = 0, \quad p(3 | 3, a_{31}) = 1; \quad (23)$$

and the loss rates, $c(i, a)$, are given as

$$c(1, a_{11}) = 5, \quad c(1, a_{12}) = 6, \quad c(2, a_{21}) = 4, \quad c(2, a_{22}) = 3. \quad (24)$$

From the data above, Condition 2.1 holds with $\delta = 1$ and $\xi = \min\{29/30, 19/20, e^{-0.05}, e^{-0.025}, e^{-0.3}\}$, and thus Assumption 2.1 is fulfilled. Therefore, the value iteration is valid and the existence of an $\epsilon$ optimal policy is ensured by Theorems 3.1 (a) and 3.2. The constant $k$ is assumed to be 16 ($k > \frac{\xi}{\mathcal{F}}$). By Theorem 3.2, the iterations are stopped if $n = 2023 =: n_0$ (where $\epsilon \equiv 10^{-4}$), and the condition $0 \leq F^* - F^{n_0+1} k < 10^{-4}$ is satisfied.
Note that, since \( c(3, a_{31}) = 0 \) and state 3 is absorbing, \( F^*(3, t, \lambda) = I_{(-\infty, 0]}(\lambda) \) for all \((t, \lambda) \in [0, 15] \times \mathbb{R}\). Hence, we only compute the functions \( F^*_{(n_{0}+1)k}(1, t, \lambda) \) and \( F^*_{(n_{0}+1)k}(2, t, \lambda) \). After applying the value iteration algorithm described in Theorem 3.1 (a), we obtain some numerical results shown in Figure 1 and Figure 2. Figure 1 and Figure 2 show the optimal value functions \( F^*_{(n_{0}+1)k}(i, t, \lambda) \) for each \((i, t) \in \{1, 2\} \times [0, 15]\), while \( F^*_{(n_{0}+1)k}(i, t, \lambda) \) is monotone nondecreasing in \( t \) for each \((i, \lambda) \in \{1, 2\} \times [0, 90]\).

![Figure 1. The function \( F^*_{(n_{0}+1)k}(1, t, \lambda) \)](image)

Figure 3 plots the function \( H^a F^*_{(n_{0}+1)k-1}(i, 10, \lambda) \) for the planning horizon \( t = 10 \). It is clear to see that the risk with \( a_{12} \) is much lower than that with \( a_{11} \) for the case \( \lambda \in (0, 49.9) \), while the risk with \( a_{12} \) is much higher than that with \( a_{11} \) for the case \( \lambda \in (49.9, 120) \). Figure 4 have the similar conclusions for the planning horizon \( t = 15 \). For better illustrating \( \epsilon \)-optimal policy, we also plot the function \( \lambda^*(i, t) \) below, which depends on the states and planning horizons.

In view of Figure 5 and the analysis above, we can define a policy \( f^* \) by

\[
f^*(1, t, \lambda) = \begin{cases} a_{12}, & \lambda < \lambda^*(1, t), \\ a_{11}, & \lambda \geq \lambda^*(1, t) 
\end{cases} \quad f^*(2, t, \lambda) = \begin{cases} a_{22}, & \lambda < \lambda^*(2, t), \\ a_{21}, & \lambda \geq \lambda^*(2, t) 
\end{cases}
\]

such that \( F^*(i, t, \lambda) = H^f F^*(i, t, \lambda) \) for every \((i, t, \lambda) \in \{1, 2\} \times [0, 15] \times \mathbb{R}\); for example, for the planning horizon \( t = 15 \), we have

\[
f^*(1, 15, \lambda) = \begin{cases} a_{12}, & \lambda < 74.95, \\ a_{11}, & \lambda \geq 74.95 
\end{cases} \quad f^*(2, 15, \lambda) = \begin{cases} a_{22}, & \lambda < 44.95, \\ a_{21}, & \lambda \geq 44.95 
\end{cases}
\]

and for the case \( t = 10 \), we have

\[
f^*(1, 10, \lambda) = \begin{cases} a_{12}, & \lambda < 49.9, \\ a_{11}, & \lambda \geq 49.9 
\end{cases} \quad f^*(2, 10, \lambda) = \begin{cases} a_{22}, & \lambda < 29.95, \\ a_{21}, & \lambda \geq 29.95 
\end{cases}
\]
By Theorem 3.6, such a policy $f^*$ is $\epsilon$-optimal with the minimum risk probability.

REFERENCES

[1] N. Bauerle and U. Rieder, *Markov Decision Processes with Application to Finance*, Universitext, Springer, Heidelberg, 2011.
Figure 4. The function \( H^a F^*_{(n_0+1)k-1}(i, 15, \lambda) \)

Figure 5. The function \( \lambda^*(i, t) \)

[2] M. Bouakiz and Y. Kebir, Target-level criterion in Markov decision processes, *Journal of Optimization Theory and Applications*, 86 (1995), 1–15.

[3] M. K. Ghosh and S. Subhamay, Non-stationary semi-Markov decision processes on a finite horizon, *Stochastic Analysis and Applications*, 31 (2013), 183–190.
[4] X. P. Guo and O. Hernández-Lerma, *Continuous-Time Markov Decision Processes: Theory and Applications*, Springer-Verlag, Berlin, 2009.

[5] X. P. Guo and J. Yang, A new condition and approach for zero-sum stochastic games with average payoffs, *Stochastic Analysis and Applications*, 26 (2008), 537–561.

[6] X. P. Guo, P. Shi and W. P. Zhu, Strong average optimality for controlled nonhomogeneous Markov chains, *Stochastic Analysis and Applications*, 19 (2001), 115–134.

[7] O. Hernández-Lerma and J. B. Lasserre, *Discrete-time Markov Control Processes*, Basic optimality criteria, Springer-Verlag, New York, 1996.

[8] O. Hernández-Lerma and J. B. Lasserre, *Further Topics on Discrete-Time Markov Control Processes*, Springer-Verlag, New York, 1999.

[9] Y. H. Huang and X. P. Guo, Optimal risk probability for first passage models in semi-Markov decision processes, *Journal of Mathematical Analysis and Applications*, 359 (2009), 404–420.

[10] Y. H. Huang, X. P. Guo and X. Y. Song, Performance analysis for controlled semi-Markov systems with application to maintenance, *Journal of Optimization Theory and Applications*, 150 (2011), 395–415.

[11] Y. H. Huang and X. P. Guo, Finite horizon semi-Markov decision processes with application to maintenance systems, *European Journal of Operational Research*, 212 (2011), 131–140.

[12] Y. H. Huang, X. P. Guo and Z. F. Li, Minimum risk probability for finite horizon semi-Markov decision processes, *Journal of Mathematical Analysis and Applications*, 402 (2013), 378–391.

[13] Y. H. Huang and X. P. Guo, Mean-variance problems for finite horizon semi-Markov decision processes, *Applications of Mathematical Optimization*, 72 (2015), 233–259.

[14] N. Limnios and G. Oprisan, *Semi-Markov Processes and Reliability*, Birkhäuser Boston, Inc., Boston, MA, 2001.

[15] J. Y. Liu and S. M. Huang, Markov decision processes with distribution function criterion of first-passage time, *Applications of Mathematical Optimization*, 43 (2001), 187–201.

[16] P. M. Madhani, Rebalancing fixed and variable pay in a sales organization: A business cycle perspective, *Compensation Benefits Review*, 42 (2010), 179–189.

[17] J. W. Mamer, Successive approximations for finite horizon semi-Markov decision processes with application to asset liquidation, *Oper. Res.*, 34 (1986), 638–644.

[18] Y. Ohtsubo, Minimizing risk models in stochastic shortest path problems, *Mathematical Methods of Operations Research*, 57 (2003), 79–88.

[19] Y. Ohtsubo, Optimal threshold probability in undiscounted Markov decision processes with a target set, *Applied Mathematical Optimization*, 57 (2004), 519–532.

[20] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, Inc., New York, 1994.

[21] C. Ruhm, Are recessions good for your health?, *Quarterly Journal of Economics*, 115 (2000), 617–650.

[22] M. Sakaguchi and Y. Ohtsubo, Markov decision processes associated with two threshold probability criteria, *Journal of Control Theory and Applications*, 11 (2013), 548–557.

[23] Q. D. Wei and X. P. Guo, New average optimality conditions for semi-Markov decision processes in Borel spaces, *Journal of Optimization Theory and Applications*, 153 (2012), 709–732.

[24] D. J. White, Minimising a threshold probability in discounted Markov decision processes, *J. Math. Anal. Appl.*, 173 (1993), 634–646.

[25] Y. H. Wu, Bounds for the ruin probability under a Markovian modulated risk model, *Communications in Statistics: Stochastic Models*, 15 (1999), 125–136.

[26] S. X. Yu, Y. L. Lin and P. F. Yan, Optimization models for the first arrival target distribution function in discrete time, *J. Math. Anal. Appl.*, 225 (1998), 193–223.

Received June 2017; revised July 2017.

E-mail address: liuql@m.scnu.edu.cn
E-mail address: zouxiaol@gzhu.edu.cn