

**SOLUTION AND FORECAST HORIZONS FOR INFINITE HORIZON  
NONHOMOGENEOUS MARKOV DECISION PROCESSES**

**Torpong Cheevaprawatdomrong**

JONG STIT CO., LTD.  
BANGKOK, THAILAND  
TONYCHEE@YAHOO.COM

**Irwin E. Schochetman**

MATHEMATICS AND STATISTICS  
OAKLAND UNIVERSITY  
ROCHESTER, MICHIGAN 48309  
SCHOCHET@OAKLAND.EDU

**Robert L. Smith**

INDUSTRIAL AND OPERATIONS ENGINEERING  
THE UNIVERSITY OF MICHIGAN  
ANN ARBOR, MICHIGAN 48109  
RLSMITH@UMICH.EDU

**Alfredo Garcia**

SYSTEMS AND INFORMATION ENGINEERING  
UNIVERSITY OF VIRGINIA  
CHARLOTTESVILLE, VIRGINIA 22901  
AGARCIA@VIRGINIA.EDU

November 19, 2005

ABSTRACT.

We consider the problem of solving a nonhomogeneous infinite horizon Markov Decision Process (MDP) problem in the general case of potentially *multiple* optimal first period policies. More precisely, we seek an algorithm that, given a finite subset of the problem's potentially infinite data set, delivers an optimal first period policy. Such an algorithm can thus recursively generate, within a rolling horizon procedure, an infinite horizon optimal solution to the original infinite horizon problem. However it can happen that for a given problem no such algorithm exists. In this case, it is impossible to solve the problem with a finite amount of data. We say such problems fail to be *well-posed*. Under the assumption of increasing marginal returns in actions (with respect to states) and stochastically increasing states into which the system transitions (with respect to actions), we provide an algorithm that is guaranteed to solve the corresponding nonhomogeneous MDP whenever the problem is well-posed. The algorithm proceeds by discovering, in finite time, a forecast horizon for which an optimal solution delivers an optimal first period policy to the infinite horizon problem. In particular, we show by construction, the existence of a forecast horizon (and hence, a solution horizon) for all such well-posed problems. We illustrate the theory and algorithms developed by solving every well-posed instance of the time-varying version of the classic asset selling problem.

---

1991 *Mathematics Subject Classification*. Primary 90C40 Secondary 90B15, 90C39.

*Key words and phrases*. Nonhomogeneous Markov decision process, infinite horizon optimization, monotone policy.

This research was partially supported by the National Science Foundation under Grants DMI-0322114 and DMI-9713723. The second author was partially supported by an Oakland University Research Fellowship.

Typeset by  $\mathcal{A}\mathcal{M}\mathcal{S}$ - $\mathcal{T}\mathcal{E}\mathcal{X}$

## 1. Introduction and Preliminaries

From a mathematical modeling perspective, the class of infinite horizon nonhomogeneous MDP's is general enough to encompass a wide variety of applications. We are interested primarily in problems which are truly nonhomogeneous, as opposed to problems in which there are predictable trends in the data limited to a finite set of parameters. Much research effort has been devoted to determining when (and how) it is possible to solve an infinite horizon nonhomogeneous MDP by making use of only a finite amount of data. In practice, a class of algorithms, known as "rolling (or receding) horizon" algorithms, is typically the tool of choice when solving infinite horizon nonhomogeneous MDP's. Intuitively, these algorithms rely on the premise that "early" optimal decisions are effectively decoupled from the infinite "tail" of problem data. Much of the published research on this issue has indeed confirmed that this intuitive premise underlying rolling horizon algorithms holds, if there is a *unique* infinite horizon optimal solution. (See Cruz-Suarez et al [2004] for conditions for uniqueness in the context of MDP problems.) Formally, uniqueness ensures the existence of a "solution" horizon, i.e. a finite horizon such that the first period optimal decision of the finite horizon problem is arbitrarily close to the infinite horizon first period optimal decision. Even though the existence of a solution horizon means only finite computation is called for when solving for the first period optimal decision of the underlying infinite nonhomogeneous MDP, an infinite forecast of data may be necessary to discover such a horizon. This is the motivation for the stronger notion of a "forecast horizon," which is a horizon sufficiently distant that data beyond it cannot affect the optimality of a first period policy. Uniqueness of a infinite horizon optimal first policy also insures existence of a forecast horizon.

In this paper, monotonicity of optimal policies is exploited to establish the existence and discovery of solution and forecast horizons in a class of nonhomogeneous Markov decision processes *without* requiring uniqueness of an infinite horizon optimal first policy. We instead require a much weaker condition we call well-posedness. In fact, this condition simply requires that the problem be solvable. An infinite horizon nonhomogeneous MDP will be called *well-posed* if, roughly speaking, a first-period optimal decision can be determined by the data associated with a finite number of periods of problem data (See section 6 for a formal definition.) We use this notion to characterize the sub-class of problems that can be solved with recourse to finite amounts of data. We show that if a problem is well-posed, monotonicity of optimal policies ensures the existence of an algorithm to compute a forecast horizon. We provide a simple detection procedure to identify a forecast, and hence solution, horizon in finite time, thereby constructing an optimal initial policy for the infinite horizon problem.

A substantial literature has developed over the years which focuses on attempts to establish existence of algorithms for determining both solution and forecast horizons (see Chand et al (2002) for a complete and up-to-date review). In cases where solution horizons fail to exist, Bean et al (1987), Alden and Smith (1984), and Lee and Denardo (1986) provide bounds on the error (as a function of horizon) caused by solving finite horizon versions of the problem.

Most of the results on the existence and discovery of solution and forecast horizons have been established in the case of uniqueness of infinite horizon optima. (See, for example, Bean and Smith (1984,1993), Bean et al (1990), Bhaskaran and Sethi (1985), Hopp (1989), Hopp et al (1987), and Ryan et al (1992).) In the context of finite state non-homogeneous MDP problems, Bes and Lasserre [1986] proposed a finite procedure for finding an optimal first policy under the assumption of uniqueness of an optimal first policy. Hernandez-Lerma and Lasserre [1988] extended the stopping rule employed to the case of stationary MDP problems with Borel state spaces and finite action sets, again under the assumption of a unique optimal policy (the stationary model is relevant to us since any non-stationary MDP problem may be transformed into a stationary MDP with infinite state space by augmenting the state variable with time). However, uniqueness of an infinite horizon optimal solution is usually difficult, if not impossible, to establish (see for example, Ryan and Bean (1989)).

One can, in the absence of a uniqueness assumption on optimal policies, construct a series of approximating finite MDP problems which approximate the underlying infinite MDP problems optimal *value* (see for example, Th. 4.8 of Hernandez-Lerma [1989]), but our goal is more ambitious, namely to approximate the MDP's optimal first period *policy*. Without assuming uniqueness of an optimal first decision, Smith and Zhang (1998) and Cheevaprawatdomrong and Smith (2004) established the existence of solution and forecast horizons, and provided closed-form formulas for a forecast horizon in the case of infinite horizon

production planning problems under deterministic and stochastic demand, respectively. The key property exploited was monotonicity of the first decision with respect to problem horizons. (See also Altman and Koole [1995] where monotonicity is used for queueing control problems.) Garcia and Smith (2000) used a similar approach to establish the existence and discovery of forecast horizons for nonstationary infinite horizon dynamic optimization problems. In their paper, it was assumed that there exists a “forecast index” such that the first-period optimal decision is monotonically non-decreasing in that index. However, the algorithm for detecting forecast horizons reported in their paper may not terminate finitely in the presence of multiple infinite horizon optimal first decisions.

This paper is organized as follows. In section 2, our infinite horizon nonhomogeneous Markov decision process problem is formulated, and the assumptions required for establishing results in subsequent sections are stated. In section 3, finite horizon versions of the infinite horizon problem are formulated and their monotonicity properties are established. In section 4, we prove value and policy convergence for the infinite horizon problem relative to its finite horizon versions. In section 5, we discuss sufficient conditions for policy optimality. In section 6, sufficient conditions for existence of forecast and solution horizons are established. An algorithm for detecting forecast horizons is presented with conditions under which it is guaranteed to terminate within a finite number of steps. Finally, in section 7, we present an application of our main results to asset selling under time-varying distributions for offers.

### An Asset Selling Problem which is not Well-Posed

Before we provide a formal mathematical model for the class of problems of interest, we discuss an example from asset selling that illustrates the challenges posed in solving time-varying infinite horizon problems. In particular, we construct an instance of this problem that is not well-posed and hence, a problem for which no algorithm exists to provide its solution. We will return to the general class of asset selling problems (in the last section), as an application of the theory and algorithms developed. An algorithm is presented which can solve any well-posed problem in this class. More generally, we provide an algorithm that will solve every solvable problem.

Suppose we have an asset which is available for sale. At the beginning of each period, an offer is presented for our asset. This offer may be zero (corresponding to no offer presented that period) or any positive integer up to some largest possible value. We can either accept or decline the offer. If we decline, an asset holding cost (eg., maintenance) must be paid before we continue to the next period to await the next offer. If we accept, we receive the offer, the asset is sold, remains sold forever, and no further offers are received. We assume that offers in all periods are independent random variables with known but period dependent probability distributions (for example the offers may stochastically decline as time wears on). Our objective is to make a decision (accept or reject) in each period, so as to maximize the expected total discounted value of the asset over the infinite horizon. In each period  $n \in \mathbb{N}$ , the state  $i$  is an element of the set  $\{0, 1, \dots, g\}$ , where  $g \geq 0$ , and  $i$  represents the random amount offered in period  $n$ , so that  $g$  is the greatest possible offer in any period; or the state  $i = g + 1$ , a dummy state, which represents the state of the system after sale of the asset. Consequently, the state space is  $S = \{0, \dots, g, g + 1\}$ . We assume that we are given the probability mass function  $q_{n-1}$  of the random offer at the beginning of each period  $n \geq 1$ . Thus, in the event that an offer  $i$  is made at the start of period  $n + 1$ ,  $q_n(i)$  represents the probability that the offer is  $i$ , where  $i \in \{0, \dots, g\}$ . In particular,  $q_0$  is the distribution of the initial state. In general, for  $n \geq 0$ ,  $\sum_{i=0}^g q_n(i) = 1$  and  $q_n(g + 1) = 0$

Corresponding to each state  $i \in S$ , the decision set  $D_i$  is given by

$$D_i = \begin{cases} \{0, 1\}, & \text{if } 0 \leq i \leq g, \\ \{1\}, & \text{if } i = g + 1, \end{cases}$$

where  $k = 0$  (respectively,  $k = 1$ ) represents *rejection* (respectively, *acceptance*), if the current offer is  $0 \leq i \leq g$ . If the asset is sold, that is  $i = g + 1$ , then we see that  $k = 1$  is the only decision choice, i.e.,  $k = 0$  implies that  $0 \leq i \leq g$ . For each period  $n$ , the reward function is defined by

$$\rho_n(i, k) = \begin{cases} -h_n, & \text{if } 0 \leq i \leq g, \quad k = 0, \\ i, & \text{if } 0 \leq i \leq g, \quad k = 1, \\ 0, & \text{if } i = g + 1, \quad k = 1, \end{cases}$$

where  $h_n \geq 0$  is the  $n$ -th period holding cost, which is assumed to be bounded, i.e.,  $\sup_n h_n < \infty$ .

Let  $\{c_n : n \geq 1\}$  denote the collection of optimal thresholds, i.e., at time period  $n$ , the optimal decision is to reject the current offer  $i$  if  $i < c_n$ , to accept if  $i > c_n$  and to either accept or reject if  $i = c_n$ . In principle, the computation of  $c_n$ , for a particular value of  $n$ , may require precise knowledge of an infinite amount of data, namely, knowledge of an infinite subset of the time-indexed collection of probability distributions of offers and holding costs, i.e.  $\{(q_n, h_n) : n \geq 1\}$ . If this is indeed the case, the asset selling problem is not well-posed.

To illustrate, consider the case with  $g = 2$ ,  $h > 0$ ,  $0 < a < 1$ ,

$$q(i) = \begin{cases} 0, & \text{if } i = 0, 3 \\ 1 - a, & \text{if } i = 1, \\ a, & \text{if } i = 2, \end{cases}$$

with stationary strategies  $\delta^1$ , accept any offer above or at 1; and  $\delta^2$ , accept any offer above or at 2. For appropriate choices of the discount factor  $\alpha$  and  $h_n = h$ ,  $\forall n$  (see Cheevaprawatdomrong et al (2005) for the details), the strategies  $\delta^1$  and  $\delta^2$  are both optimal (non-uniqueness). Now, consider a forecast in which a slight perturbation of the holding cost from a given point in time in the future is introduced. This perturbation will break the tie in favor of either  $\delta^1$  or  $\delta^2$ . Hence, an optimal decision for the first period is inextricably tied to which particular infinite forecast is identified.

In the interest of brevity, we will omit many details and proofs throughout the manuscript. In particular, proofs of all lemmas and corollaries have been omitted. Long proofs of theorems have been presented in outline form. We refer the reader to the technical report by Cheevaprawatdomrong et al (2005) for all of the missing details and proofs. In what follows, it will also be convenient to refer to this manuscript as CSSG [2005].

## 2. Problem Formulation and Assumptions

In each period  $n = 1, 2, \dots$ , a dynamic system may be in one of the stochastic states belonging to the finite (discrete) set  $S = \{l, \dots, u\}$ , where  $l \leq u$  are integers which are assumed to be period-independent. Hence,  $s \equiv u - l + 1$  denotes the number of stochastic states in any period. For each state  $i \in S$ , there is a finite (non-empty) set of feasible decisions (actions)  $D_i$ , which is assumed to be period-independent and (without loss of generality) in the form of consecutive integers, i.e.,  $D_i = \{a_i, \dots, b_i\}$ , for some integers  $a_i \leq b_i$ , with  $a \equiv \min_{i \in S} a_i$  and  $b \equiv \max_{i \in S} b_i$ . The *policy space*  $D$  in period  $n$  is then given by the finite (discrete) space  $D \equiv D_l \times \dots \times D_u$ , so that it is also period-independent. Thus, the data sets  $S$  and  $D_i$ ,  $i \in S$ , are constant over discrete time.

Now let  $K$  denote the set of all feasible decisions, i.e.,  $K = \cup_{i \in S} D_i$ , so that  $K \subseteq \{a, \dots, b\}$ . Also let  $L$  denote the set of all feasible pairs of states and decisions, i.e.,

$$L = \{(i, k) \in S \times K : k \in D_i\},$$

which is period-independent. In general,  $L$  is partially ordered by the restricted product order on the lattice  $S \times K$ . In general,  $L$  need not be a sublattice, even though  $S$  and  $K$  are chains. (See section 7 for an example of this.)

**Assumption 1.** The partially ordered set  $L$  is a sublattice of  $S \times K$ .

This assumption, although restrictive, can often be met by an appropriate reformulation of the problem.

We have seen that  $S$  and  $D$  do not vary with  $n$ . However, this will not be the case for the remaining data, i.e., the state transition and expected reward structures. This non-homogenous problem could be transformed into a stationary problem by augmenting the state variable to include the period we are in, but this transformation would result in an infinite state space problem (see for example, Hernandez-Lerma [1989].) We define a *state transition probability function* to be a non-negative, real-valued function  $p$  on  $L \times S$ , denoted by  $((i, k), j) \rightarrow p(i, j; k)$ ,  $\forall (i, k) \in L$ ,  $\forall j \in S$ , satisfying  $\sum_{j=l}^u p(i, j; k) = 1$ , for each  $(i, k)$

in  $L$ . The quantity  $p(i, j; k)$  is interpreted as the probability that the system transitions from state  $i$  to state  $j$ , given that decision  $k$  in  $D_i$  is chosen.

Given such  $p$ , for each  $d \in D$ , let  $P(d)$  denote the corresponding  $s \times s$  stochastic state transition matrix given by

$$P(d) = \begin{bmatrix} p(l, l; d_l) & \cdots & p(l, u; d_l) \\ \vdots & \vdots & \vdots \\ p(u, l; d_u) & \cdots & p(u, u; d_u) \end{bmatrix}.$$

The resulting *state transition probability matrix mapping*  $P$  corresponding to  $p$  is a mapping of  $D$  into the space  $\mathcal{P}_s$  of  $s \times s$  stochastic matrices with matrix norm defined by

$$\|A\| = \max_{1 \leq i \leq s} \sum_{j=1}^s |a_{ij}| = 1,$$

for all such  $A$ , where  $\|Ax\| \leq \|A\| \|x\|$ , relative to the supremum norm  $\|x\| = \|x\|_\infty$ , for  $x \in \mathbb{R}^s$ . The correspondence  $p \rightarrow P$  is one-to-one.

For each  $n$ , if the dynamic system is in state  $i \in S$  at the start of period  $n$ , and decision  $k \in D_i$  is chosen, then the system transitions to state  $j$  at the start of period  $n + 1$  with probability  $p_n(i, j; k)$ . If  $p_n \rightarrow P_n$  as above, then for policy  $d \in D$ ,  $P_n(d)$  denotes the resulting  $s \times s$  stochastic state transition matrix in period  $n$  corresponding to decision  $d$ . Thus, as above,  $P_n : D \rightarrow \mathcal{P}_s$ ,  $\forall n$ . Consequently, the state transition structure consists of a sequence  $(p_n)_{n=1}^\infty$  of state transition probabilities, one for each period, with corresponding sequence  $(P_n)_{n=1}^\infty$  of state transition mappings. Note that it is automatically the case that  $\|P_n\| = 1$ ,  $\forall n$ , since the  $P_n$  are stochastic matrices.

Analogously, we define an *expected reward function* to be a function  $\rho : L \rightarrow \mathbb{R}$ . For each  $(i, k) \in L$ , the quantity  $\rho(i, k)$  is interpreted as the expected reward associated with choosing decision  $k$  in  $D_i$ , when the system is in state  $i$ . Given such  $\rho$ , for each  $d \in D$ , let  $r(d)$  denote the corresponding  $s$ -dimensional vector given by

$$r(d) = [\rho(l, d_l) \cdots \rho(u, d_u)]^t, \quad \forall d = (d_l, \dots, d_u) \in D,$$

where  $x^t$  denotes the transpose of the vector  $x \in \mathbb{R}^s$ . The resulting *expected reward vector-valued mapping*  $r$  is a mapping of  $D$  into  $\mathbb{R}^s$ , and the correspondence  $\rho \rightarrow r$  is one-to-one. Define

$$\|r\| = \max_{d \in D} \|r(d)\| = \max_{i \in S} \max_{d_i \in D_i} |\rho(i, d_i)| = \max_{(i, k) \in L} |\rho(i, k)| = \|\rho\|.$$

Now let  $\rho_n$  denote the expected reward function at the start of period  $n$ , with  $\rho_n \rightarrow r_n$ ,  $\forall n$ , as above. Thus,  $r_n : D \rightarrow \mathbb{R}^s$ ,  $\forall n$ , and the expected reward structure consists of a sequence  $(\rho_n)_{n=1}^\infty$  of real-valued functions defined on  $L$ , one for each period, with corresponding sequence  $(r_n)_{n=1}^\infty$  of expected reward vector functions. The following assumption requires marginal returns in actions be increasing in states.

**Assumption 2.** For each  $n$ , the function  $\rho_n$  has *increasing differences on  $L$*  (Topkis (1998)), i.e., if  $(i, k), (i', k'), (i', k), (i, k') \in L$  with  $k < k'$  and  $i < i'$ , then  $\rho_n(i', k') - \rho_n(i', k) \geq \rho_n(i, k') - \rho_n(i, k)$ .

**2.1 Lemma.** *Suppose Assumptions 1 and 2 hold. Then, for each  $n$ , the function  $\rho_n$  is supermodular on the lattice  $L$ , i.e.,*

$$\rho_n((i, k) \vee (i', k')) + \rho_n((i, k) \wedge (i', k')) \geq \rho_n(i, k) + \rho_n(i', k'),$$

for all  $(i, k), (i', k') \in L$ .

Next, let  $J_n(i, k)$  denote the random state of the system in period  $n + 1$ , given that it is in state  $i$  at the beginning of period  $n$ , and decision  $k \in D_i$  is chosen. Note that each  $J_n$  is a function from  $L$  into the space of random variables with values in  $S$ . Also, the probability that  $J_n(i, k)$  equals  $j$  in  $S$  is  $p_n(i, j; k)$ , i.e.,  $j \rightarrow p_n(i, j; k)$  is the probability mass function for  $J_n(i, k)$ , so that  $\text{Prob}[J_n(i, k) = j] = p_n(i, j; k)$ ,  $\forall n \in \mathbb{N}$ , and  $\forall (i, k) \in L$ . Let  $i_k$  denote the smallest  $i \in S$  for which  $k \in D_i$ ,  $\forall k \in K$ .

**Assumption 3.** For each  $n$ ,  $p_n$  is constant with respect to  $i$ , i.e.,  $p_n(i, j; k) = p_n(i_k, j; k)$ , so that  $\text{Prob}[J_n(i, k) = j] = \text{Prob}[J_n(i_k, k) = j]$ ,  $\forall (i, k) \in L$ ,  $\forall j \in S$ .

Assumption 3 requires the probability distribution over the next state be dependent only on the action  $k$  taken out of the preceding state  $i$  and in particular be independent of  $i$ . This assumption can often be met by an artful formulation of the problem. For example, it is met in production planning if we choose action  $k$  to be the produce-up-to quantity as opposed to the quantity ordered  $k - i$ .

Let  $G_n(k)$  denote the cumulative distribution function for the random variable  $J_n(i_k, k)$ , that is,

$$G_n(k; x) = \text{Prob}[J_n(i_k, k) \leq x] = \begin{cases} 0, & \text{for all } x < l, \\ \sum_{j=l}^{\lfloor x \rfloor} p_n(i_k, j; k), & \text{for all } l \leq x < u, \\ 1, & \text{for all } u \leq x, \end{cases}$$

where  $\lfloor x \rfloor$  is the largest integer less than or equal to  $x$ . Note that  $J_n(i, k)$  is distributed as  $G_n(k)$ ,  $\forall (i, k) \in L$  under Assumption 3.

The probability transition functions  $p_n$  are required also to satisfy the following assumption which requires the state transitioned into be stochastically increasing in action  $k$ . For example, inventory beginning the next period is stochastically increasing in after production inventory from the previous period.

**Assumption 4.** For each  $n$ ,  $J_n(i_k, k)$  is *stochastically increasing* in  $k$ , that is, the distribution function  $G_n(k)$  is stochastically increasing in  $k$ , i.e., if  $k < k' \in K$ , then  $1 - G_n(k') \geq 1 - G_n(k)$ . Equivalently,  $G_n(k') \leq G_n(k)$ , as real-valued functions, i.e.,

$$\sum_{j=l}^{\lfloor x \rfloor} p_n(i_{k'}, j, k') \leq \sum_{j=l}^{\lfloor x \rfloor} p_n(i_k, j, k), \quad \forall l \leq x < u,$$

or

$$\sum_{j=l}^u g(j) p_n(i_{k'}, j, k') \geq \sum_{j=l}^u g(j) p_n(i_k, j, k), \quad \forall k, k' \in K,$$

for each increasing real-valued function  $g$  on  $S$  (Topkis (1998), Corollary 3.9.1(a)).

Now fix a discount factor  $0 < \alpha < 1$ .

**Assumption 5.** The expected reward structure  $(\rho_n)_{n=1}^\infty$  is *dominated* by the discount factor  $\alpha$  in the sense that  $\sum_{n=1}^\infty \alpha^{n-1} \|\rho_n\| < \infty$ . Note, in particular, that if the expected reward structure is uniformly bounded, i.e.,  $\sup_n \|\rho_n\| < \infty$ , then the expected reward structure is dominated by any discount factor.

In summary, the *problem data* for our nonhomogeneous MDP consist of:

- (i)  $\alpha$ ,  $S$  and the  $D_i$ ,  $\forall i \in S$ , with lattice  $L$ , all of which we assume to be constant through time, as well as
- (ii) the function sequence  $(\rho_n)_{n=1}^\infty$  dominated by  $\alpha$ , each  $\rho_n$  with increasing differences, and
- (iii) the function sequence  $(p_n)_{n=1}^\infty$ , each  $p_n$  constant with respect to  $i$  and stochastically increasing with respect to  $k$ .

Note that the pair  $(p_n, \rho_n)$  constitutes the variable part of the problem data for period  $n$ .

**Definition (forecast).** We define a forecast for our problem to be a sequence  $(p_n, \rho_n)_{n=1}^\infty$  as above, with reward structure dominated by the discount factor  $\alpha$ , and with corresponding sequence  $(P_n, r_n)_{n=1}^\infty$ . A forecast will be called *homogeneous* if it is a constant sequence, i.e., both  $(p_n)_{n=1}^\infty$  and  $(\rho_n)_{n=1}^\infty$  are constant sequences.

For convenience, define  $\phi_n = (p_n, \rho_n)$ ,  $\forall n \in \mathbb{N}$ , so that a forecast  $\phi$  may also be viewed as a sequence  $(\phi_n)_{n=1}^\infty$ , where  $\phi_n$  represents the state transition and expected reward data, i.e., the period forecast, for

period  $n$ . Much of what follows will be forecast-dependent. However, since  $\phi$  is our only forecast at this point, it will be convenient for the time being to suppress notational reference to it.

A (feasible) *strategy*  $\pi = (\pi_n)_{n=1}^\infty$  is a sequence of policies, one for each period, so that  $\pi_n \in D$ ,  $\forall n$ . A policy  $\pi_n$  in period  $n$  is a vector of decisions, i.e.,  $\pi_n = [\pi_n(l) \cdots \pi_n(u)]^t$ , one for each state, where, for  $l \leq i \leq u$ , the “ $i^{\text{th}}$  component”  $\pi_n(i)$  of  $\pi_n$  denotes the decision chosen in period  $n$ , if strategy  $\pi$  is implemented, given that the system is in state  $i$ . We denote by  $\Pi$  the space of all strategies, so that  $\Pi = D^\infty$ ; then  $\Pi$  is a compact topological space in the product topology. Note that, under our assumptions, *all* strategies are feasible and the strategy space  $\Pi$  is independent of the forecast  $\phi$ . Given a strategy  $\pi$ , the stochastic matrix  $P_n(\pi_n)$  is then (relative to  $\phi$ ) the corresponding state transition matrix for period  $n$ . Let  $Q_n(\pi)$  denote the total transition matrix beginning in period 1, and terminating at the beginning of period  $n \geq 2$ , when strategy  $\pi$  is implemented, i.e.,  $Q_n(\pi) = P_1(\pi_1) \cdot P_2(\pi_2) \cdots P_{n-1}(\pi_{n-1})$ ,  $\forall \pi \in \Pi$ . For convenience, define  $Q_1(\pi)$  to be the  $s \times s$  identity matrix  $I$ . Therefore, the element  $Q_n(\pi)_{ij}$  in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column of  $Q_n(\pi)$  is the probability that the system is in state  $j$  at the start of period  $n$ , given that it was in state  $i$  at the start of period 1, and strategy  $\pi$  is implemented, for  $l \leq i, j \leq u$ . It is easy to see that each  $Q_n(\pi)$  is also a stochastic matrix. For convenience, we will denote the  $i^{\text{th}}$  row of the block matrix  $Q_n(\pi)$  by  $Q_n(\pi)_i$ ,  $\forall i = l, \dots, u$ , so that  $Q_n(\pi)_i = [Q_n(\pi)_{il} \cdots Q_n(\pi)_{iu}]$  and

$$Q_n(\pi) = \begin{bmatrix} Q_n(\pi)_l \\ \vdots \\ Q_n(\pi)_u \end{bmatrix}.$$

It is not difficult to see that, for each  $n$ , the resulting mapping  $Q_n : \Pi \rightarrow \mathcal{P}_s$  is continuous relative to the norm  $\|\cdot\|$  on  $\mathcal{P}_s$  and the product topology on  $\Pi$ , with  $\|Q_n(\pi_n)\| = 1$ ,  $\forall \pi \in \Pi$ .

Clearly,  $r_n(\pi_n)$  is the vector of expected rewards in period  $n$  under strategy  $\pi$ . Rewards received in period  $n$  are to be discounted by  $\alpha^{n-1}$ . Let  $R_N(\pi; i)$  denote the total expected *discounted* reward incurred in periods 1 through  $N$  under strategy  $\pi$ , given that the initial state is  $i \in S$ , that is

$$R_N(\pi; i) = \sum_{n=1}^N \alpha^{n-1} Q_n(\pi)_i \cdot r_n(\pi_n), \quad \forall i \in S,$$

with  $R_N(\pi) = [R_N(\pi; l) \cdots R_N(\pi; u)]^t$ , the total expected discounted reward vector for periods 1 through  $N$ . Thus, we obtain a sequence  $(R_N)_{N=1}^\infty$  of vector-valued functions of the form  $R_N : \Pi \rightarrow \mathbb{R}^s$ , each given by

$$R_N(\pi) = \sum_{n=1}^N \alpha^{n-1} Q_n(\pi) \cdot r_n(\pi_n), \quad \forall \pi \in \Pi.$$

In addition, let

$$R(\pi; i) = \lim_{N \rightarrow \infty} R_N(\pi; i) = \sum_{n=1}^{\infty} \alpha^{n-1} Q_n(\pi)_i \cdot r_n(\pi_n),$$

for all  $i \in S$ , which converges in  $\mathbb{R}$  by Assumption 5. Then we obtain a mapping  $R : \Pi \rightarrow \mathbb{R}^s$  given by

$$R(\pi) = \lim_{N \rightarrow \infty} R_N(\pi) = \sum_{n=1}^{\infty} \alpha^{n-1} Q_n(\pi) \cdot r_n(\pi_n), \quad \forall \pi \in \Pi,$$

which converges in  $\mathbb{R}^s$ , with  $\|R(\pi)\| \leq \sum_{n=1}^{\infty} \alpha^{n-1} \|\rho_n\|$ . The vector  $R(\pi) = [R(\pi; 1) \cdots R(\pi; s)]^t$  is the infinite horizon total expected discounted reward vector indexed by the different starting states in  $S = \{l, \dots, u\}$ . If we let  $q_0$  denote the distribution of the initial states, a probability mass function on  $S$ , then the infinite horizon total expected reward function  $\hat{R} : \Pi \rightarrow \mathbb{R}$  is given by  $\hat{R}(\pi) = \sum_{i=l}^u R(\pi; i) q_0(i)$ .

Given that  $Q_n$  and  $r_n$  are continuous and bounded, we have:

**2.2 Lemma.** *Suppose Assumption 5 holds. Then, for each  $N \in \mathbb{N}$ , the function  $R_N : \Pi \rightarrow \mathbb{R}^s$  is continuous, as are the functions  $R : \Pi \rightarrow \mathbb{R}^s$  and  $\tilde{R} : \Pi \rightarrow \mathbb{R}$ .*

For the MDP (with given forecast  $\phi$ ), we wish to find a (feasible) strategy that maximizes the expected total discounted reward over the infinite horizon, i.e., we wish to solve the following (vector) infinite horizon optimization problem  $\langle \phi \rangle$  with forecast  $\phi$ :  $v \equiv \max_{\pi \in \Pi} R(\pi)$ , defined component-wise. Note that  $v$  is an element of  $\mathbb{R}^s$  whose  $i^{\text{th}}$  component  $v(i) = \max_{\pi \in \Pi} R(\pi; i)$ ,  $\forall i \in S$ , denotes the maximum expected net present value of rewards incurred from initial state  $i$ , under forecast  $\phi$ . Moreover,  $|v(i)| \leq \sum_{n=1}^{\infty} \alpha^{n-1} \|\rho_n\|$ . Analogously, consider the (scalar) infinite horizon optimization problem given by  $\tilde{v} = \max_{\pi \in \Pi} \tilde{R}(\pi)$ . For convenience, define the vector  $\Delta v$  of *marginal* optimal values to  $\langle \phi \rangle$  to be the element of  $\mathbb{R}^{s-1}$  whose  $i^{\text{th}}$  component is given by  $\Delta v(i) \equiv v(i) - v(i-1)$ ,  $\forall i = l+1, \dots, u$ . Since  $\Pi$  is compact and  $R$  (respectively  $\tilde{R}$ ) is continuous on  $\Pi$ , the above maximum is attained component-wise, i.e., the non-empty space of  $\phi$ -optimal strategies  $\Omega$  (respectively,  $\tilde{\Omega}$ ) is given by  $\Omega = \{\pi \in \Pi : R(\pi) = v\}$  (respectively,  $\tilde{\Omega} = \{\pi \in \Pi : \tilde{R}(\pi) = \tilde{v}\}$ ). Clearly,  $\Omega \subseteq \tilde{\Omega}$ . Observe that the previous discussion, results and notation are applicable to *any* generic forecast  $\theta = (\theta_n)_{n=1}^{\infty}$  which satisfies Assumption 5.

For  $n \in \mathbb{N}$  and  $i \in S$ , define  $v_n(i)$  to be the maximum expected net present value, beginning period  $n$ , of rewards incurred forever, given that the system is in state  $i$  at the beginning of period  $n$ . Also let  $v_n \in \mathbb{R}^s$  be defined by  $v_n = [v_n(l) \cdots v_n(u)]^t$ ,  $\forall n \in \mathbb{N}$ . Note that, in particular,  $v(i) = v_1(i)$ , for all  $i$  in  $S$ .

Denote by  $f_n(i, k)$  the maximum expected net present value of rewards incurred forever, starting in period  $n$ , given that the system is in state  $i$  at the beginning of period  $n$ , and decision  $k \in D_i$  is chosen. Specifically, for  $(i, k) \in L$ , we have (where  $E[\cdot]$  denotes expectation)

$$f_n(i, k) = \rho_n(i, k) + \alpha E[v_{n+1}(J_n(i, k))],$$

i.e., each function  $f_n : L \rightarrow \mathbb{R}$  is given by  $f_n = \rho_n + \alpha E[v_{n+1} \circ J_n]$ . Analogously, define  $f_n : D \rightarrow \mathbb{R}^s$  by  $f_n(d) = [f_n(l, d_l) \cdots f_n(u, d_u)]^t$ ,  $\forall d \in D$ , and  $f_1 : D \rightarrow \mathbb{R}$  by  $f_1(d) = \sum_{i=l}^u f_1(i, d_i) q_0(i)$ ,  $\forall d \in D$ .

Consider the composition mapping  $v_{n+1} \circ J_n$  from  $K$  into the space of real random variables. Then, for  $(i, k) \in L$ ,  $(v_{n+1} \circ J_n)(i, k) = v_{n+1}(J_n(i, k))$  is the maximum expected net present value, beginning period  $n+1$ , of rewards incurred forever, given that the system is in state  $i$  beginning period  $n$ , in which decision  $k$  is chosen. For fixed  $n$  and  $k$ , the probability mass function of  $J_n(i, k)$  is given by  $j \rightarrow p_n(i, j; k)$ , for  $j \in S$ . Consequently,  $v_{n+1}(J_n(i, k))$  is given by

$$E[v_{n+1}(J_n(i, k))] = \sum_{j=l}^u v_{n+1}(j) p_n(i, j; k) = \sum_{j=l}^u v_{n+1}(j) p_n(i_k, j; k) = E[v_{n+1}(J_n(i_k, k))]$$

for all  $(i, k) \in L$ .

**2.3 Lemma.** *Suppose Assumptions 1 and 2 hold. Then, for each  $n \in \mathbb{N}$ , the function  $f_n : L \rightarrow \mathbb{R}$  is supermodular, i.e.,*

$$f_n((i, k) \vee (i', k')) + f_n((i, k) \wedge (i', k')) \geq f_n(i, k) + f_n(i', k'), \quad \forall (i, k), (i', k') \in L.$$

The previous discussion, results and notation hold for *any* forecast  $\theta$  which satisfies Assumptions 1 through 5. We now find it necessary to consider a *collection* of period forecasts  $(p, \rho)$ . We assume we have a non-empty collection  $F$  of *admissible* period forecast data  $(p, \rho)$  where  $\rho$  has increasing differences on  $L$ ,  $p$  is constant with respect to  $i$  and stochastically increasing in  $k$ . Consider the subset  $\Phi$  of  $F^\infty$  consisting of those infinite horizon forecasts  $\theta$  which satisfy Assumption 5, i.e., which are dominated by  $\alpha$ . Hence,  $\phi \in \Phi$ . Note that  $\Phi$  depends on  $F$  and  $\alpha$  which are fixed and hence, suppressed. We think of  $\Phi$  as the set of forecasts relative to  $S$  and  $D$  (as well as  $F$  and  $\alpha$ ) which define a particular nonhomogeneous MDP problem class  $\langle \Phi \rangle = \{\langle \theta \rangle : \theta \in \Phi\}$ . Of course, all the previous results hold for each forecast  $\theta$  in  $\Phi$ . For the remainder of this paper, we assume that we have such a problem class with fixed forecast  $\phi \in \Phi$  and generic forecast  $\theta \in \Phi$ .



In what follows, it will be necessary to consider different forecasts simultaneously. Consequently, if  $\theta$  is a generic forecast satisfying Assumptions 2,3,4,5, then we will incorporate  $\theta$  in the notation which we have introduced in order to distinguish between constructs depending on different forecasts. For example, we write  $v(\theta; i) = v_1(\theta; i)$  for the optimal value of optimization problem  $\langle \theta \rangle$  with initial state  $i \in S$ . In particular,  $v(\phi; i) = v(i) = v_1(i) = v_1(\phi; i)$ . Wherever practical, we will suppress the reference to  $\phi$  for convenience. Thus, if a forecast is not specified, then we will assume it is the underlying fixed forecast  $\phi$ .

Our final assumption requires that there exist two special *homogeneous* forecasts (from  $F$ ) with *maximum* (respectively, *minimum*) marginal optimal values with respect to changes of problem data in the *first* period only.

**Assumption 6.** There exist  $(\underline{p}, \underline{\rho}), (\bar{p}, \bar{\rho}) \in F$ , with corresponding *homogeneous* forecasts  $\bar{\psi} = ((\bar{p}, \bar{\rho}))_{n=1}^{\infty}$  and  $\underline{\psi} = ((\underline{p}, \underline{\rho}))_{n=1}^{\infty}$  (in  $\Phi$ ), and corresponding optimal vector values  $\bar{w} = v(\bar{\psi})$  and  $\underline{w} = v(\underline{\psi})$  having the following properties. Suppose  $(p, \rho) \in F$ , with associated forecasts

$$\bar{\psi}(p, \rho) \equiv ( (p, \rho), (\bar{p}, \bar{\rho}), (\bar{p}, \bar{\rho}), \dots ) \quad \text{and} \quad \underline{\psi}(p, \rho) \equiv ( (p, \rho), (\underline{p}, \underline{\rho}), (\underline{p}, \underline{\rho}), \dots )$$

(also in  $\Phi$ ). These satisfy Assumptions 1,2. Consider the corresponding optimization problems  $\langle \bar{\psi}(p, \rho) \rangle$  and  $\langle \underline{\psi}(p, \rho) \rangle$ , with corresponding optimal vector values  $\bar{w}(p, \rho) = v(\bar{\psi}(p, \rho))$  and  $\underline{w}(p, \rho) = v(\underline{\psi}(p, \rho))$ , and corresponding marginal optimal vector values  $\Delta \bar{w}(p, \rho)$  and  $\Delta \underline{w}(p, \rho)$ . Then we assume that

$$\bar{w}(p, \rho; i') - \bar{w}(p, \rho; i) \leq \bar{w}(i') - \bar{w}(i) \quad \text{and} \quad \underline{w}(p, \rho; i') - \underline{w}(p, \rho; i) \geq \underline{w}(i') - \underline{w}(i),$$

for all  $l \leq i < i' \leq u$ , and for all  $(p, \rho) \in F$ . It is easy to see that this is equivalent to assuming that the marginal optimal values of problems  $\langle \bar{\psi} \rangle$  and  $\langle \underline{\psi} \rangle$  are related to those of problems  $\langle \bar{\psi}(p, \rho) \rangle$  and  $\langle \underline{\psi}(p, \rho) \rangle$  as follows:

$$\Delta \bar{w}(p, \rho; i) \leq \Delta \bar{w}(i) \quad \text{and} \quad \Delta \underline{w}(p, \rho; i) \geq \Delta \underline{w}(i),$$

for all  $i = l + 1, \dots, u$ , and for all  $(p, \rho) \in F$ .

To summarize, we assume that we have a forecast class  $\Phi$  defined by a period forecast set  $F$  as above, with typical forecast  $\theta \in \Phi$ , a fixed arbitrary forecast  $\phi \in \Phi$  and two homogeneous forecasts  $\bar{\psi}, \underline{\psi} \in \Phi$  for the given state space  $S$  and decision sets  $D_i, \forall i \in S$ , all of which satisfy Assumptions 1 through 6. Furthermore, if  $n \in \mathbb{N}$ , and forecast  $\theta = (\theta_1, \dots, \theta_n, \dots) \in \Phi$ , let  $\theta^{(n)}$  denote the forecast given by  $\theta^{(n)} = (\theta_n, \theta_{n+1}, \dots)$ , so that  $\theta^{(1)} = \theta$ , in particular, and  $\theta^{(n)} \in \Phi, \forall n = 2, 3, \dots$ . Then, the optimal value  $v(\theta^{(n)})$  of problem  $\langle \theta^{(n)} \rangle$  satisfies  $v(\theta^{(n)}) = v_n(\theta)$ . In particular, for  $\theta = \phi$ , we have  $v(\phi^{(n)}) = v_n(\phi) = v_n$ .

By the Principle of Optimality, we have that the  $i^{\text{th}}$  component of the  $s$ -tuple  $v_n = v_n(\phi)$  is given by

$$v_n(i) = \max_{k \in D_i} f_n(i, k), \quad \forall i \in S, \quad \forall n \in \mathbb{N}.$$

Since each  $D_i$  is finite, the set

$$K_n(i) \equiv \{k \in D_i : f_n(i, k) = v_n(i)\} = \arg \max_{k \in D_i} f_n(i, k)$$

is a *non-empty* finite subset of  $D_i \subseteq K$ , which depends *only* on the forecast  $\phi^{(n)} = (\phi_n, \phi_{n+1}, \dots)$ , since

$$f_n(i, k) = \rho_n(i, k) + \alpha \mathbb{E}[v_{n+1}(J_n(i, k))] = \rho_n(i, k) + \alpha \sum_{j=l}^u v_{n+1}(j) p_n(i, j; k),$$

and  $v_{n+1}(j) = v_{n+1}(\phi; j) = v(\phi^{(n+1)}; j)$ . Hence, we may define  $\gamma_n(i)$  to be the *greatest* such optimal solution, i.e.,  $\gamma_n(i) = \max\{k : k \in K_n(i)\}$ , and  $\lambda_n(i)$  to be the *least* such optimal solution, i.e.,  $\lambda_n(i) = \min\{k : k \in K_n(i)\}$ , so that  $\lambda_n(i) \leq \gamma_n(i)$ , and  $f_n(i, \lambda_n(i)) = f_n(i, \gamma_n(i)) = v_n(i), \forall i \in S$ . Observe that  $\lambda_n$  and  $\gamma_n$  are policies in  $D$  with  $\lambda_n \leq \gamma_n, \forall n$ . Also,  $\lambda \equiv (\lambda_n)_{n=1}^{\infty}$  and  $\gamma \equiv (\gamma_n)_{n=1}^{\infty}$  are strategies in  $\Pi$  with  $\lambda \leq \gamma$ .

Let  $d \in D$  and  $n \in \mathbb{N}$ . We will say that the *policy*  $d$  satisfies the optimality equations for the forecast  $\phi$  in period  $n$  (alternately,  $d$  is optimal for problem  $\langle \phi \rangle$  in period  $n$ ) if  $d_i \in K_n(i)$ , that is  $v_n(i) = f_n(i, d_i)$ , for all  $i \in S$ . Let  $\Omega_n(\phi)$  denote the non-empty set of  $\phi$ -optimal policies in period  $n$ , so that by definition,

$$\Omega_n(\phi) = K_n(l) \times \cdots \times K_n(u), \quad \forall n \in \mathbb{N}.$$

Observe that for each  $n$ ,  $\lambda_n$  and  $\gamma_n$  are  $\phi$ -optimal for period  $n$ , i.e.,  $\lambda_n, \gamma_n \in \Omega_n(\phi)$ .

We will also say that the *strategy*  $\pi \in \Pi$  is *strongly  $\phi$ -optimal* if  $\pi^{(n)} \equiv (\pi_n, \pi_{n+1}, \dots)$  is  $\phi^{(n)}$ -optimal, for each  $n \in \mathbb{N}$ . Of course, strongly optimal strategies are also optimal, since  $\pi^{(1)} = \pi$  and  $\phi^{(1)} = \phi$ . Also let

$$\Omega^{(n)}(\phi) \equiv \left\{ \pi \in \prod_{m=n}^{\infty} D : \pi^{(m)} \in \Omega(\phi^{(m)}) \right\}, \quad \forall m \geq n, \quad \forall n \in \mathbb{N},$$

so that,  $\Omega^{(1)}(\phi)$  is the set of all strongly  $\phi$ -optimal strategies, and  $\pi \in \Omega^{(1)}(\phi)$  if and only if  $\pi^{(m)} \in \Omega(\phi^{(m)})$ ,  $\forall m \geq 1$ . Moreover,  $\Omega^{(1)}(\phi)$  is a proper (non-empty) subset of  $\Omega(\phi) = \Omega(\phi^{(1)})$  in general, i.e.,  $\phi$ -optimal strategies need not be strongly optimal. See CSSG [2005] for an example.

The following theorem is essentially the Principle of Optimality for nonhomogeneous infinite horizon MDPs.

**2.4 Theorem.** *Suppose Assumption 5 holds. Let  $\pi \in \Pi$ . Then  $\pi$  is strongly  $\phi$ -optimal if and only if  $\pi_n$  is  $\phi$ -optimal in each period, i.e.,  $\pi \in \Omega^{(1)}(\phi)$  if and only if  $\pi_n \in \Omega_n(\phi)$ ,  $\forall n \in \mathbb{N}$ .*

**Outline of Proof.** Suppose  $\pi_n$  is  $\phi$ -optimal in each period. By induction, we first show that, for each  $m \geq 1$ ,

$$v(i) = \sum_{n=1}^m \alpha^{n-1} Q_n(\pi)_i \cdot r_n(\pi_n) + \alpha^m Q_{m+1}(\pi)_i \cdot v_{m+1}, \quad \forall i \in S.$$

For  $m = 1$ , by hypothesis, we have that  $v(i) = Q_1(\pi)_i \cdot r_1(\pi_1) + \alpha Q_2(\pi)_i \cdot v_2$ , i.e., the claim is valid for  $m = 1$ . Similarly, for  $m > 1$ , we have that

$$\alpha^m Q_{m+1}(\pi)_i \cdot v_{m+1} = \alpha^m Q_{m+1}(\pi)_i \cdot r_{m+1}(\pi_{m+1}) + \alpha^{m+1} Q_{m+2}(\pi)_i \cdot v_{m+2}.$$

Now suppose the above claim is true for  $m$ . Then

$$v(i) = \sum_{n=1}^{m+1} \alpha^{n-1} Q_n(\pi)_i \cdot r_n(\pi_n) + \alpha^{m+1} Q_{m+2}(\pi)_i \cdot v_{m+2},$$

i.e., the claim is true for  $m + 1$  as well. But  $\|v_m\| \leq \sum_{n=1}^{\infty} \alpha^{n-1} \|\rho_n\|$ ,  $\forall m$ , i.e., the  $\|v_m\|$  are bounded (Assumption 5). Consequently,

$$v(i) = \sum_{n=1}^{\infty} \alpha^{n-1} Q_n(\pi)_i \cdot r_n(\pi_n) = R(\pi; i), \quad \forall i \in S,$$

so that  $\pi$  is  $\phi$ -optimal.

Now fix  $m \in \mathbb{N}$ . Then the infinite horizon forecast  $\phi^{(m)}$  does not depend on the period forecasts  $\phi_1, \dots, \phi_{m-1}$ . Thus, for each  $n \geq m$ ,  $\pi_n$  is also optimal for  $\phi^{(m)}$  in period  $n$ . Therefore, applying the previous result to  $\phi^{(m)}$ , we have that  $\pi^{(m)}$  is  $\phi^{(m)}$ -optimal. Since,  $m$  is arbitrary,  $\pi$  is strongly  $\phi$ -optimal.

Conversely, suppose that  $\pi$  is strongly  $\phi$ -optimal. Fix  $n \geq 1$ . Then  $\pi^{(n)}$  is  $\phi^{(n)}$ -optimal and  $\pi^{(n+1)}$  is  $\phi^{(n+1)}$ -optimal. Moreover,  $\pi_n$  is  $\phi$ -optimal in period  $n$ , since  $f_n(i, \pi_n(i)) = v_n(i)$ ,  $\forall i \in S$ . (See CSSG [2005] for the details.) Since  $n$  is arbitrary, the proof is complete.  $\square$

**Remark.** By the previous theorem, if  $d^n \in \Omega_n(\phi)$ ,  $\forall n$ , then the strategy  $(d^n)_{n=1}^{\infty}$  is strongly  $\phi$ -optimal. Thus, (strongly)  $\phi$ -optimal strategies exist, and every  $\phi$ -optimal policy in any period can be completed to such a strategy.

Next we establish Theorem 2.4 for finite horizon problems with salvage value. Fix  $N \in \mathbb{N}$  and  $\nu \in \mathbb{R}^s$ . Consider the following  $N$ -horizon optimization problem with forecast  $\phi(N, \nu) = (\phi_1, \dots, \phi_N; \nu)$ , where each  $\phi_n$ ,  $1 \leq n \leq N$ , belongs to  $F$ . The vector  $\nu \in \mathbb{R}^s$  is viewed as a terminal reward (salvage value) after period  $N$ , whose contribution will be computed as though it were awarded in period  $N + 1$ . As above, we have the correspondences  $p_n \rightarrow P_n$ ,  $\rho_n \rightarrow r_n$ ,  $\forall n = 1, \dots, N$ , as well as the shifted forecasts

$$\phi(N, \nu)^{(n)} = (\phi_n, \dots, \phi_N; \nu), \quad \forall 1 \leq n \leq N.$$

The  $N$ -horizon optimization problem  $\langle \phi(N, \nu) \rangle$  is then given as follows:

$$\text{maximize} \quad \sum_{n=1}^N \alpha^{n-1} Q_n(\pi^N) \cdot r_n(\pi_n) + \alpha^N Q_{N+1}(\pi^N) \cdot \nu$$

subject to

$$\pi^N = (\pi_1, \dots, \pi_N) \in D^N,$$

where  $Q_n(\pi_1, \dots, \pi_N) \equiv P_1(\pi_1) \cdots P_{n-1}(\pi_{n-1})$ ,  $\forall 1 \leq n \leq N + 1$ , and  $Q_1(\pi_1, \dots, \pi_N) = I$ , in particular. Let  $\Omega(\phi(N, \nu))$  denote the corresponding set of  $N$ -horizon optimal strategies in  $D^N$ , and  $\Omega^{(1)}(\phi(N, \nu))$  the corresponding set of strongly optimal  $N$ -horizon optimal strategies, where  $\pi^N$  is strongly  $(\phi(N, \nu))$ -optimal if  $(\pi_n, \dots, \pi_N)$  is  $\phi(N, \nu)^{(n)}$ -optimal for each  $n = 1, \dots, N$ . Moreover, the policy  $d \in D$  is optimal in period  $1 \leq n \leq N$  for  $\phi(N, \nu)$  if  $d_i$  belongs to

$$K_n(\phi(N, \nu); i) = \{k \in D_i : f_n(\phi(N, \nu); i, k) = v_n(\phi(N, \nu); i)\}, \quad \forall i \in S,$$

where  $f_n(\phi(N, \nu); i, k)$  and  $v_n(\phi(N, \nu); i)$  are defined as above, but only up to horizon  $N$ . Let  $\Omega_n(\phi(N, \nu))$  denote the corresponding set of optimal policies in period  $n$ ,  $\forall 1 \leq n \leq N$ . Then we have the following finite dimensional version of Theorem 2.4.

**2.5 Theorem.** *Suppose Assumptions 3, 4 and 5 hold. Let  $\pi^N \in D^N$ . Then  $\pi^N$  is strongly  $\phi(N, \nu)$ -optimal if and only if  $\pi_n$  is  $\phi(N, \nu)$ -optimal for each period  $n = 1, \dots, N$ , i.e.,  $\pi^N \in \Omega^{(1)}(\phi(N, \nu))$  if and only if  $\pi_n \in \Omega_n(\phi(N, \nu))$ ,  $\forall 1 \leq n \leq N$ .*

**Outline of Proof.** For  $\nu \in \mathbb{R}^s$  and  $\rho_\nu : L \rightarrow \mathbb{R}$  given by  $\rho_\nu(i, k) = \nu(i)$ ,  $\forall (i, k) \in L$ , consider the infinite horizon forecasts

$$\phi^N(\nu)^{(n)} = (\phi_n, \dots, \phi_N, (p_{N+1}, \rho_\nu), (p_{N+2}, \rho_0), (p_{N+3}, \rho_0), \dots), \quad \forall 1 \leq n \leq N,$$

where  $p_{N+k}$  is any transition probability satisfying Assumptions 3, 4,  $\forall k \geq 1$ . All such forecasts belong to  $\Phi$ . (For convenience, let  $\phi^N(\nu) = \phi^N(\nu)^{(1)}$ .) Then  $(\pi_n, \dots, \pi_N)$  is  $\phi(N, \nu)^{(n)}$ -optimal if and only if any completion  $(\pi_n, \dots, \pi_N, \pi_{N+1}, \dots)$  of  $(\pi_n, \dots, \pi_N)$  to an element of  $\Pi_{m=n}^\infty D$  is strongly  $\phi^N(\nu)^{(n)}$ -optimal, i.e.,

$$\Omega(\phi^N(\nu)^{(n)}) = \Omega(\phi(N; \nu)^{(n)}) \times \Pi_{m=N+1}^\infty D, \quad \forall 1 \leq n \leq N.$$

Consequently,  $\pi^N = (\pi_n, \dots, \pi_N)$  is strongly  $\phi(N, \nu)$ -optimal if and only if any completion

$$\pi = (\pi_1, \dots, \pi_N, \pi_{N+1}, \dots)$$

of  $\pi^N$  to an element of  $\Pi$  is strongly  $\phi^N(\nu)$ -optimal, i.e.,

$$\Omega^{(1)}(\phi^N(\nu)) = \Omega^{(1)}(\phi(N, \nu)) \times \Pi_{m=N+1}^\infty D.$$

By Theorem 2.4, this is true if and only if  $\pi_n$  is optimal in period  $n$  for  $\phi^N(\nu)$ ,  $\forall n \geq 1$ . But this is equivalent to  $\pi_n$  is optimal in period  $n$  for  $\phi(N, \nu)$ ,  $\forall 1 \leq n \leq N$ , since any policy is optimal in period  $n \geq N + 1$  for the forecast  $((p_{N+1}, \rho_\nu), (p_{N+2}, \rho_0), (p_{N+3}, \rho_0), \dots)$ ,  $\forall \nu \in \mathbb{R}^s$ .  $\square$

**2.6 Theorem.** *The  $N$ -horizon forecast  $\phi(N, \nu)$  and the infinite horizon forecast*

$$\phi^N(\nu) = (\phi_1, \dots, \phi_N, (p_{N+1}, \rho_\nu), (p_{N+2}, \rho_0), (p_{N+3}, \rho_0))$$

*have the same optimal policies in each period  $1 \leq n \leq N$ , i.e.,  $\Omega_n(\phi(N, \nu)) = \Omega_n(\phi^N(\nu))$ ,  $\forall 1 \leq n \leq N$ .*

**Proof.** Observe that  $f_N(\phi(N, \nu); i, k) = f_N(\phi^N(\nu); i, k)$ ,  $\forall (i, k) \in L$ . Consequently,

$$v_N(\phi(N, \nu)) = v_N(\phi^N(\nu)), \quad K_N(\phi(N, \nu)) = K_N(\phi^N(\nu)) \quad \text{and} \quad \Omega_N(\phi(N, \nu)) = \Omega_N(\phi^N(\nu)).$$

Hence, the theorem is true for  $n = N$ . In the same way, we may show that the theorem holds for  $n = N - 1$ . Continue in this way down to  $n = 1$ .  $\square$

### 3. Policy Monotonicity for Finite Horizon Problems

In this section, for the given forecast  $\phi$ , and for each positive integer  $N$ , we formulate certain  $(N + 1)$ -horizon “upper and lower bound” forecasts  $\bar{\phi}^N$  and  $\underline{\phi}^N$ , respectively. Each is a finite horizon truncation, after period  $N$ , of forecast  $\phi$ , with salvage value equal to  $\underline{w} = v(\underline{\psi})$  or  $\bar{w} = v(\bar{\psi})$ , treated as the reward vector at the start of period  $N + 1$ . We will show that, for any horizon  $N$ , and any period  $1 \leq n \leq N$ , there exists an optimal policy  $\bar{d}_n^N$  for problem  $\langle \bar{\phi}^N \rangle$  such that, for each  $n \in \mathbb{N}$ , the sequence  $(\bar{d}_n^N)_{N=n}^\infty$  is *decreasing* in  $N$ , i.e.,  $n \leq N < M$  implies  $\bar{d}_n^N \geq \bar{d}_n^M$ . Similarly, for any horizon  $N$ , and any period  $1 \leq n \leq N$ , there exists an optimal policy  $\underline{d}_n^N$  for problem  $\langle \underline{\phi}^N \rangle$  such that, for each  $n \in \mathbb{N}$ , the sequence  $(\underline{d}_n^N)_{N=n}^\infty$  is *increasing* in  $N$ , i.e.,  $n \leq N < M$  implies  $\underline{d}_n^N \leq \underline{d}_n^M$ .

Recall that our fixed infinite horizon forecast  $\phi = (\phi_n)_{n=1}^\infty$  is such that  $\phi_n = (p_n, \rho_n)$ , with  $p_n \rightarrow P_n$  and  $\rho_n \rightarrow r_n$  as in section 2. Also fix a positive integer  $N$  and recall the period forecasts  $(\underline{p}, \underline{\rho})$  and  $(\bar{p}, \bar{\rho})$  in  $F$  with

$$\rho_{\underline{w}} = \underline{\rho} \rightarrow \underline{r} = r_{\underline{w}}, \quad \rho_{\bar{w}} = \bar{\rho} \rightarrow \bar{r} = r_{\bar{w}}, \quad \underline{p} \rightarrow \underline{P}, \quad \bar{p} \rightarrow \bar{P},$$

also defined as in section 2. Define the hybrid infinite horizon forecasts

$$\bar{\phi}^N = (\phi_1, \dots, \phi_N, (\bar{p}, \rho_{\bar{w}}), (\bar{p}, \rho_0), (\bar{p}, \rho_0), \dots),$$

and

$$\underline{\phi}^N \equiv \phi^N(\underline{w}) = (\phi_1, \dots, \phi_N, (\underline{p}, \rho_{\underline{w}}), (\underline{p}, \rho_0), (\underline{p}, \rho_0), \dots),$$

all of which belong to  $\Phi$ . Note that

$$\Omega_n(\phi(N, \bar{w})) = \Omega_n(\phi^N(\bar{w})) = \Omega_n(\bar{\phi}^N),$$

$$\Omega_n(\phi(N, \underline{w})) = \Omega_n(\phi^N(\underline{w})) = \Omega_n(\underline{\phi}^N),$$

for all  $1 \leq n \leq N$ , by Theorem 2.6. Thus, for periods  $n = 1, \dots, N$ , finding the optimal policies of the infinite horizon problems  $\langle \bar{\phi}^N \rangle$  and  $\langle \underline{\phi}^N \rangle$  is equivalent to finding the optimal policies of the finite horizon problems  $\langle \phi(N, \bar{w}) \rangle$  and  $\langle \phi(N, \underline{w}) \rangle$ , respectively. More generally, we may repeat the above procedure for any  $\theta \in \Phi$  to get  $\underline{\theta}^N$  and  $\bar{\theta}^N$  in  $\Phi$ .

Consequently, we may apply the results of section 2 to the infinite horizon forecasts  $\underline{\phi}^N$  and  $\bar{\phi}^N$ . If we replace  $\phi$  by  $\bar{\phi}^N$  in section 2, we obtain analogues of problem  $\langle \phi \rangle$ , and of the quantities

$$R = R(\phi), \quad \Omega = \Omega(\phi), \quad J_n = J_n(\phi), \quad f_n = f_n(\phi),$$

$$v_n = v_n(\phi), \quad K_n = K_n(\phi), \quad \gamma_n = \gamma_n(\phi), \quad \lambda_n = \lambda_n(\phi),$$

for  $1 \leq n \leq N$ , which we denote respectively by problem  $\langle \bar{\phi}^N \rangle$  and

$$\bar{R}^N = R(\bar{\phi}^N), \quad \bar{\Omega}^N = \Omega(\bar{\phi}^N), \quad \bar{J}_n^N = J_n(\bar{\phi}^N), \quad \bar{f}_n^N = f_n(\bar{\phi}^N),$$

$$\bar{v}_n^N = v_n(\bar{\phi}^N), \quad \bar{K}_n^N = K_n(\bar{\phi}^N), \quad \bar{\gamma}_n^N = \gamma_n(\bar{\phi}^N), \quad \bar{\lambda}_n^N = \lambda_n(\bar{\phi}^N),$$

where

$$\bar{R}^N(\pi) = \sum_{n=1}^N \alpha^{n-1} Q_n(\pi) \cdot r_n(\pi_n) + \alpha^N Q_{N+1}(\pi) \cdot \bar{w}; \quad \langle \bar{\phi}^N \rangle$$

$\bar{v}^N = \bar{v}_1^N = \max_{\pi \in \Pi} \bar{R}^N(\pi)$ ;  $\bar{\Omega}^N = \{\pi \in \Pi : \bar{R}^N(\pi) = \bar{v}^N\}$ ;  $\bar{J}_n^N(i, k)$  is the state ending period  $n$ , under forecast  $\bar{\phi}^N$ , starting in state  $i$  and choosing decision  $k$ , so that  $\bar{J}_n^N(i, k) = J_n(i, k)$ ,  $\forall 1 \leq n \leq N + 1$ ;  $\bar{J}_n^N(i, k)$  is the maximum expected present net value of rewards incurred in problem  $\langle \bar{\phi}^N \rangle$  through period

$N + 1$ , starting in period  $1 \leq n \leq N$ , given that the system is in state  $i$  at the beginning of period  $n$ , and decision  $k \in D_i$  is chosen;  $\bar{v}_n^N(i)$  is the maximum expected net present value, beginning period  $n$  of problem  $\langle \bar{\phi}^N \rangle$ , of rewards incurred through period  $N + 1$ , given that the system is in state  $i$  at the beginning of period  $n$ ;

$$\bar{f}_n^N(i, k) = \bar{\rho}_n^N(i, k) + \alpha E[\bar{v}_{n+1}^N(\bar{J}_n^N(i, k))] = \rho_n(i, k) + \alpha E[\bar{v}_{n+1}^N(J_n(i, k))],$$

for all  $1 \leq n \leq N$ , since forecasts  $\bar{\phi}^N$  and  $\phi$  agree in periods  $1 \leq n \leq N$ ;

$$\bar{v}_n^N(i) = \begin{cases} \max\{\bar{f}_n^N(i, k) : k \in D_i\}, & 1 \leq n \leq N, \\ \bar{w}(i), & n = N + 1, \end{cases}$$

by the Principle of Optimality; for all  $i \in S$ , and for all  $1 \leq n \leq N$ ,

$$\bar{K}_n^N(i) = \{k \in D_i : \bar{f}_n^N(i, k) = \bar{v}_n^N(i)\} = \arg \max_{k \in D_i} \bar{f}_n^N(i, k);$$

$$\bar{\gamma}_n^N(i) = \max\{k : k \in \bar{K}_n^N(i)\} \quad \text{and} \quad \bar{\lambda}_n^N(i) = \min\{k : k \in \bar{K}_n^N(i)\}.$$

Clearly, the previous quantities are the same, for  $1 \leq n \leq N + 1$ , irrespective of whether we view  $\bar{\phi}^N$  as an infinite horizon forecast or as an  $(N + 1)$ -horizon forecast. We have the following analogue of Theorem 2.5 for forecast  $\bar{\phi}^N$ .

**3.1 Lemma.** *Suppose Assumptions 5 and 6 hold. The strategy  $\pi \in \Pi$  is strongly  $\bar{\phi}^N$ -optimal (i.e.,  $\pi \in \Omega^{(1)}(\bar{\phi}^N)$ ) if and only if  $\pi_n$  is a  $\bar{\phi}^N$ -optimal policy in period  $n$  (i.e.,  $\pi_n \in \Omega_n(\bar{\phi}^N)$ ), for each  $n \leq N$ , in which case,  $\bar{\lambda}_n^N(i) \leq \pi_n(i) \leq \bar{\gamma}_n^N(i)$ , for all  $1 \leq n \leq N$ , and all  $i \in S$ . In particular, this is the case for arbitrary extensions of  $(\bar{\lambda}_1^N, \dots, \bar{\lambda}_N^N)$  and  $(\bar{\gamma}_1^N, \dots, \bar{\gamma}_N^N)$  to elements of  $\Pi$ .*

Similarly, if we replace  $\phi$  by  $\underline{\phi}^N$  in section 2, then we obtain the  $(N + 1)$ -horizon problem  $\langle \underline{\phi}^N \rangle$  and the entire previous discussion for problem  $\langle \bar{\phi}^N \rangle$  is valid for  $\langle \underline{\phi}^N \rangle$ , including Lemma 3.1.

**3.2 Lemma.** *Suppose Assumptions 5 and 6 hold. The strategy  $\pi \in \Pi$  is strongly  $\underline{\phi}^N$ -optimal, (i.e.,  $\pi \in \Omega^{(1)}(\underline{\phi}^N)$ ) if and only if  $\pi_n$  is a  $\underline{\phi}^N$ -optimal policy in period  $n$  (i.e.,  $\pi_n \in \Omega_n(\underline{\phi}^N)$ ), for all  $1 \leq n \leq N$ , in which case,  $\underline{\lambda}_n^N(i) \leq \pi_n(i) \leq \underline{\gamma}_n^N(i)$ , for all  $1 \leq n \leq N$ , and all  $i \in S$ . In particular, this is the case for arbitrary extensions of  $(\underline{\lambda}_1^N, \dots, \underline{\lambda}_N^N)$  and  $(\underline{\gamma}_1^N, \dots, \underline{\gamma}_N^N)$  to elements of  $\Pi$ .*

The following theorem is the main result of this section. .

**3.3 Theorem.** *Suppose all Assumptions 1 through 6 hold. Fix  $N \in \mathbb{N}$  and let  $H_N = \{N, N + 1\}$ . Then we have the following for problem  $\langle \underline{\phi}^N \rangle$ . For each  $1 \leq n \leq N$ :*

(i) *the real-valued function  $(M, (i, k)) \rightarrow \underline{f}_n^M(i, k)$  is supermodular on the lattice  $H_N \times L$ , i.e.,*

$$\underline{f}_n^{M \vee M'}(i \vee i', k \vee k') + \underline{f}_n^{M \wedge M'}(i \wedge i', k \wedge k') \geq \underline{f}_n^M(i, k) + \underline{f}_n^{M'}(i', k'),$$

for all  $M, M' \in H_N$  and for all  $(i, k), (i', k') \in L$ .

(ii) *the integer-valued functions  $(M, j) \rightarrow \underline{\gamma}_n^M(j)$  and  $(M, j) \rightarrow \underline{\lambda}_n^M(j)$  are increasing on  $H_N \times S$ . Equivalently, if  $l \leq j \leq j' \leq u$  and  $N \leq M \leq M' \leq N + 1$ , then  $\underline{\gamma}_n^M(j) \leq \underline{\gamma}_n^{M'}(j')$  and  $\underline{\lambda}_n^M(j) \leq \underline{\lambda}_n^{M'}(j')$ .*

(iii) *the real-valued function  $(M, j) \rightarrow \underline{v}_n^M(j)$  has increasing differences on  $H_N \times S$ . Equivalently, if  $l \leq j \leq j' \leq u$  and  $N \leq M \leq M' \leq N + 1$ , then  $\underline{v}_n^M(j') - \underline{v}_n^M(j) \leq \underline{v}_n^{M'}(j') - \underline{v}_n^{M'}(j)$ .*

**Outline of Proof.** The proof is by finite reverse induction from  $n = N$  down to  $n = 1$ . First we show that the real-valued function  $(M, j) \rightarrow \underline{v}_{N+1}^M(j)$  has increasing differences on  $H_N \times S$ , i.e., if  $j < j'$ , then

$\underline{v}_{N+1}^N(j') - \underline{v}_{N+1}^N(j) \leq \underline{v}_{N+1}^{N+1}(j') - \underline{v}_{N+1}^{N+1}(j)$ , so that part (iii) of the Theorem is true for  $n = N + 1$  also. Equivalently, we show that

$$\underline{v}_{N+1}^N(j) - \underline{v}_{N+1}^N(j-1) \leq \underline{v}_{N+1}^{N+1}(j) - \underline{v}_{N+1}^{N+1}(j-1), \quad \forall j = l+1, \dots, u.$$

Consider the (infinite horizon) forecasts  $\underline{\phi}^N$ ,  $\underline{\psi}$ ,  $\underline{\psi}(\phi_{N+1}) = (\phi_{N+1}, (\underline{p}, \underline{\rho}), (\underline{p}, \underline{\rho}), \dots)$  and

$$\underline{\phi}^{N+1} = (\phi_1, \dots, \phi_N, \phi_{N+1}, (\underline{p}, \underline{\rho}_{\underline{w}}), (\underline{p}, \underline{\rho}_0), (\underline{p}, \underline{\rho}_0), \dots).$$

Then, for  $i \in S$ , using forecast  $\underline{\phi}^{N+1}$ ,

$$v_{N+1}(\underline{\phi}^{N+1}; i) = \max_{k \in D_i} \{ \rho_{N+1}(i, k) + \alpha \sum_{j=l}^u \underline{w}(j) p_{N+1}(i, j; k) \}.$$

On the other hand, by definition, using forecast  $\underline{\psi}(\phi_{N+1})$ ,

$$\underline{w}(\phi_{N+1}; i) = \max_{k \in D_i} \{ \rho_{N+1}(i, k) + \alpha \sum_{j=l}^u \underline{w}(j) p_{N+1}(i, j; k) \}.$$

Consequently,  $\underline{v}_{N+1}^{N+1}(i) = \underline{w}(\phi_{N+1}; i)$ ,  $\forall i \in S$ . By Assumption 6 applied to the forecasts  $\underline{\psi}$  and  $\underline{\psi}(\phi_{N+1})$  (with  $\phi_{N+1} \in F$ ), which satisfy Assumptions 2,3,4,5,

$$\underline{v}_{N+1}^{N+1}(j) - \underline{v}_{N+1}^{N+1}(j-1) \geq \underline{v}_{N+1}^N(j) - \underline{v}_{N+1}^N(j-1),$$

(using forecast  $\underline{\phi}^N$ ), for all  $j = l+1, \dots, u$ , as required.

Next we show that part (i) of the Theorem is true for  $n = N$ . We do this in steps. Our first step is to show that the function  $y : H_N \times K \rightarrow \mathbb{R}$  defined (for convenience) by  $y(M, k) = \alpha E[\underline{v}_{N+1}^M(J_N(i_k, k))]$  (using forecast  $\underline{\phi}^M$ ) is supermodular, i.e.,

$$y(M \wedge M', k \wedge k') + y(M \vee M', k \vee k') \geq y(M, k) + y(M', k'), \quad \forall (M, k), (M', k') \in H_N \times K.$$

Without loss of generality, we may assume that  $k < k'$ , so that  $k \wedge k' = k$  and  $k \vee k' = k'$ . Then  $y$  is supermodular if

$$y(M \wedge M', k) + y(M \vee M', k') \geq y(M, k) + y(M', k'),$$

which is trivially true if  $M \leq M'$  also, in which case  $M \vee M' = M'$  and  $M \wedge M' = M$ . Thus, assume in addition that  $M > M'$ , i.e.,  $M = N + 1$  and  $M' = N$ , in which case  $M \wedge M' = N$  and  $M \vee M' = N + 1$ . We next show that

$$y(N+1, k) - y(N, k) \leq y(N+1, k') - y(N, k').$$

On the one hand,

$$y(N+1, k) - y(N, k) = \alpha \sum_{j=l}^u (\underline{v}_{N+1}^{N+1}(j) - \underline{v}_{N+1}^N(j)) p_N(i_k, j; k).$$

Analogously, by Assumption 4,

$$y(N+1, k') - y(N, k') \geq y(N+1, k) - y(N, k),$$

since the ‘‘integrand’’  $g(j) \equiv \underline{v}_{N+1}^{N+1}(j) - \underline{v}_{N+1}^N(j)$  is an increasing function of  $j \in S$  by the first claim of this proof. To conclude that the function  $y$  is supermodular, apply Corollary 2.6.1 of Topkis (1998), together with our hypothesis.

Now consider the real-valued function on  $H_N \times L$  defined by  $(M, (i, k)) \rightarrow \alpha \mathbb{E}[\underline{v}_{N+1}^M(\underline{J}_N^M(i, k))]$  (using forecast  $\underline{\phi}^M$ ). This function is simply the canonical “lifting” of the previous function from  $H_N \times K$  to  $H_N \times L$ . We leave it to the reader to verify that it inherits the supermodularity property. Next, consider the real-valued function defined by  $(M, (i, k)) \rightarrow \rho_N(i, k)$ , whose domain is also  $H_N \times L$ . We have seen that  $\rho_N$  is supermodular on  $L$  (Lemma 2.1). Consequently, as above, the lifting of  $\rho_N$  from  $L$  to  $H_N \times L$  is also supermodular. Finally, observe that the required function

$$(M, (i, k)) \rightarrow \underline{f}_N^M(i, k) = \rho_N(i, k) + \alpha \mathbb{E}[\underline{v}_{N+1}^M(\underline{J}_N^M(i, k))],$$

which is the lifting of  $\underline{f}_N^M$  from  $L$  to  $H_N \times L$ , is the sum of the two previous supermodular functions. Thus, it is also supermodular (Topkis (1998), Lemma 2.6.1). This completes the proof of part (i) of the Theorem for the case  $n = N$ .

We next prove part (ii) for  $n = N$ . Observe that, in the notation of Chapter 2 of Topkis (1998):

- (a)  $X \equiv K$ ,  $T \equiv H_N \times S$  and  $X \times T = K \times (H_N \times S) \cong H_N \times (S \times K)$  are lattices (relative to the product order), where  $x = k$  and  $t = (M, j)$ ;
- (b)  $Z \equiv \{(k, (M, j)) \in K \times (H_N \times S) : k \in D_j\}$  is a sublattice of  $X \times T$  which is order isomorphic to the sublattice  $H_N \times L$  of  $H_N \times (S \times K)$ ;
- (c)  $Z_t = \{x \in X : (x, t) \in Z\}$ , i.e.,  $Z_{(M, j)} = \{k \in K : (M, (j, k)) \in H_N \times L\} = D_j$  is the (finite) section of  $Z = H_N \times L$  at  $t = (M, j)$  in  $T = H_N \times S$ ;
- (d) the (continuous) function  $g : X \times T \rightarrow \mathbb{R}$ , i.e.,  $g : K \times (H_N \times S) \rightarrow \mathbb{R}$ , defined by

$$g(x, t) = g(k, (M, j)) = \underline{f}_N^M(j, k)$$

is supermodular in  $(x, t) = (k, (M, j))$  on  $Z = H_N \times L$ .

Then, by Theorem 2.8.3(a) of Topkis (1998),  $\operatorname{argmax}_{x \in Z_t} g(x, t)$ , i.e.,

$$\operatorname{argmax}_{k \in D_j} \underline{f}_N^M(j, k) = \{k \in D_j : \underline{f}_N^M(j, k) = \underline{v}_N^M(j)\} \equiv \underline{K}_N^M(j)$$

is a non-empty sublattice of  $X = K$ , and the functions  $(M, j) \rightarrow \underline{\gamma}_N^M(j)$  and  $(M, j) \rightarrow \underline{\Delta}_N^M(j)$  are increasing in  $t = (M, j)$  on  $T = H_N \times S$ , which gives part (ii) for  $n = N$ .

Next we show part (iii) for  $n = N$ . In addition to (a), (b), (c), (d) above, we also have:

- (e) the function  $g(x, t)$  above is supermodular (by part (i)) on the sublattice  $Z = H_N \times L$  of the lattice  $X \times T = K \times (H_N \times S) \cong H_N \times (S \times K)$ ;
- (f)  $\Pi_T Z \equiv \{t \in T : Z_t \neq \emptyset\}$ , i.e.,  $\Pi_{H_N \times S}(H_N \times L) = H_N \times S$  is the projection of  $H_N \times L$  onto  $H_N \times S$ ; and
- (g) the function  $t \rightarrow \max_{x \in Z_t} g(x, t)$ , i.e.,  $(M, j) \rightarrow \underline{v}_N^M(j) = \max_{k \in D_j} \underline{f}_N^M(j, k)$  is finite, and consequently supermodular, on  $\Pi_T Z = H_N \times S$  by Theorem 2.7.6 of Topkis (1998). Finally, this function has increasing differences on  $H_N \times S$  by Theorem 2.6.1 of Topkis (1998).

This completes the proof of all three parts of the Theorem for the case  $n = N$ . They were obtained as consequences of the fact that the function  $(M, j) \rightarrow \underline{v}_{N+1}^M(j)$  has increasing differences. For the case  $n = N - 1$ , repeat all the above steps, beginning with the fact that the function  $(M, j) \rightarrow \underline{v}_N^M(j)$  has increasing differences. We leave it to the interested reader to verify that these steps are valid for this case. In view of the resulting validity of part (iii) for  $n = N - 1$ , i.e., the function  $(M, j) \rightarrow \underline{v}_{N-1}^M(j)$  has increasing differences, we may continue in this way until we have verified parts (i), (ii) and (iii) down to the case  $n = 1$ . This completes the proof of Theorem 3.3.  $\square$

We next observe that, in each period, the extreme optimal policies  $\underline{\Delta}_n^N$  and  $\underline{\gamma}_n^N$  for the problems  $\langle \phi^N \rangle$  are monotonically non-decreasing in horizon length and state. As consequences of Theorem 3.3, we have:

**3.4. Corollary.** *Suppose Assumptions 1 through 6 hold. Fix arbitrary  $n \in \mathbb{N}$ . Then the infinite policy sequences  $(\underline{\gamma}_n^N)_{N=n}^\infty$  and  $(\underline{\lambda}_n^N)_{N=n}^\infty$  are*

(1) *component-wise increasing in the finite horizon  $N \geq n$ , i.e.,  $\forall j \in S$ , if  $1 \leq n \leq N < M$ , then  $\underline{\gamma}_n^N(j) \leq \underline{\gamma}_n^M(j)$  and  $\underline{\lambda}_n^N(j) \leq \underline{\lambda}_n^M(j)$ , and (2) increasing on  $S$ , i.e.,  $\forall N \geq n$ , if  $l \leq j < i \leq u$  then  $\underline{\gamma}_n^N(j) \leq \underline{\gamma}_n^N(i)$  and  $\underline{\lambda}_n^N(j) \leq \underline{\lambda}_n^N(i)$ .*

By replacing all the lower bars in Theorem 3.3 by upper bars, we obtain:

**3.5. Theorem.** *Suppose Assumptions 1 through 6 hold. Fix  $N \in \mathbb{N}$  and let  $H_N = \{N, N+1\}$ . Then we have the following for problem  $\langle \bar{\phi}^N \rangle$ . For each  $1 \leq n \leq N$ :*

(i) *the real-valued function  $(M, (i, k)) \rightarrow \bar{f}_n^M(i, k)$  satisfies*

$$\bar{f}_n^{M \wedge M'}(i \vee i', k \vee k') + \bar{f}_n^{M \vee M'}(i \wedge i', k \wedge k') \geq \bar{f}_n^M(i, k) + \bar{f}_n^{M'}(i', k'),$$

for all  $(M, (i, k))$  and  $(M', (i', k'))$  in  $H_N \times L$ ;

(ii) *the integer-valued functions  $(M, j) \rightarrow \bar{\gamma}_n^M(j)$  and  $(M, j) \rightarrow \bar{\lambda}_n^M(j)$  are decreasing on  $H_N \times L$ . Equivalently, if  $l \leq j \leq j' \leq u$  and  $N \leq M \leq M' \leq N+1$ , then  $\bar{\gamma}_n^{M'}(j') \leq \bar{\gamma}_n^M(j)$  and  $\bar{\lambda}_n^{M'}(j') \leq \bar{\lambda}_n^M(j)$ ;*

(iii) *the real-valued function  $(M, j) \rightarrow \bar{v}_n^M(j)$  has decreasing differences on  $H_N \times S$ . Equivalently, if  $l \leq j < j' \leq u$  and  $N \leq M \leq M' \leq N+1$ , then  $\bar{v}_n^M(j') - \bar{v}_n^M(j) \geq \bar{v}_n^{M'}(j') - \bar{v}_n^{M'}(j)$ .*

**Proof.** The proof for Theorem 3.3 can be applied verbatim with the order in  $H_N$  reversed. In short,  $M$  and  $M'$  can be interchanged in the right hand sides of the claims in Theorem 3.3. Note that this proof requires the other half of Assumption 6.  $\square$

As in Corollary 3.4, we have that the extreme optimal policies for the problems  $\langle \bar{\phi}^N \rangle$  are decreasing in horizon length, and decreasing in state.

**3.6. Corollary.** *Suppose Assumptions 1 through 6 hold. Fix arbitrary  $n \in \mathbb{N}$ . Then the infinite policy sequences  $(\bar{\gamma}_n^N)_{N=n}^\infty$  and  $(\bar{\lambda}_n^N)_{N=n}^\infty$  are (1) component-wise decreasing in the finite horizon  $N \geq n$ , i.e., for all  $j \in S$ ,  $1 \leq n \leq N < M$  implies  $\bar{\gamma}_n^N(j) \geq \bar{\gamma}_n^M(j)$  and  $\bar{\lambda}_n^N(j) \geq \bar{\lambda}_n^M(j)$ , and (2) decreasing on  $S$ , i.e.,  $\forall 1 \leq n \leq N$ ,  $l \leq j < i \leq u$  implies  $\bar{\gamma}_n^N(j) \geq \bar{\gamma}_n^N(i)$  and  $\bar{\lambda}_n^N(j) \geq \bar{\lambda}_n^N(i)$ .*

#### 4. Value and Policy Convergence

In this section, we show that the  $\phi$ -optimal value vector  $v$  is the limit of the sequence of  $\phi^N$ -optimal value vectors  $\underline{v}^N$ . The same is true for the forecasts  $\bar{\phi}^N$  and the  $\bar{v}^N$ , i.e.,  $\underline{v}^N \rightarrow v$  and  $\bar{v}^N \rightarrow v$  (*value convergence*). Moreover, there exists a sequence  $(\underline{\delta}^N)_{N=1}^\infty$  of strongly  $\phi^N$ -optimal strategies which converges to a strongly  $\phi$ -optimal strategy  $\underline{\delta}$ . Similarly for the forecasts  $\bar{\phi}^N$ , i.e., there exists a sequence  $(\bar{\delta}^N)_{N=1}^\infty$  of strongly  $\bar{\phi}^N$ -optimal strategies which converges to a strongly  $\phi$ -optimal strategy  $\bar{\delta}$ , i.e.,  $\underline{\delta}^N \rightarrow \underline{\delta}$  and  $\bar{\delta}^N \rightarrow \bar{\delta}$  in  $\Pi$  (*policy convergence*).

Part (1) of Corollary 3.4 states that, for each period  $n$ , both the largest and the smallest optimal policy functions  $N \rightarrow \underline{\gamma}_n^N$  and  $N \rightarrow \underline{\lambda}_n^N$  are pointwise increasing in  $N \geq n$ , i.e.,  $1 \leq n \leq N < M$  implies  $\underline{\gamma}_n^N \leq \underline{\gamma}_n^M$  and  $\underline{\lambda}_n^N \leq \underline{\lambda}_n^M$ , as functions on  $S$ . Therefore, we may let  $(\underline{\delta}_n^N)_{N=n}^\infty$  represent a generic component-wise increasing sequence of *optimal* policies to the problems  $\langle \phi^N \rangle_{N=n}^\infty$  (for example,  $\underline{\delta}_n^N = \underline{\gamma}_n^N$  or  $\underline{\lambda}_n^N$ ). Thus,  $\underline{\delta}_n^N(i) \in \underline{K}_n^N(i)$ ,  $\forall 1 \leq n \leq N$ ,  $\forall i \in S$ , so that  $\underline{\delta}_n^N \in \Omega_n(\phi(N, \underline{w}))$ . Analogously, we may let  $(\bar{\delta}_n^N)_{N=n}^\infty$  represent a generic component-wise decreasing sequence of optimal policies to the problems  $\langle \bar{\phi}^N \rangle_{N=n}^\infty$  (for example,  $\bar{\delta}_n^N = \bar{\gamma}_n^N$  or  $\bar{\lambda}_n^N$ ). Thus,  $\bar{\delta}_n^N(i) \in \bar{K}_n^N(i)$ ,  $\forall i \in S$ , so that  $\bar{\delta}_n^N \in \Omega_n(\phi^N(\bar{w}))$ ,  $\forall 1 \leq n \leq N$ . Consequently, for each  $N$ , we obtain the  $N$ -tuples

$$(\underline{\delta}_1^N, \dots, \underline{\delta}_N^N) \in \Omega^{(1)}(\phi(N, \underline{w})) \quad \text{and} \quad (\bar{\delta}_1^N, \dots, \bar{\delta}_N^N) \in \Omega^{(1)}(\phi(N, \bar{w}))$$



by Theorem 2.5. For convenience, we fix an arbitrary policy  $\zeta$  in  $D$ , and define

$$\underline{\delta}^N = (\underline{\delta}_1^N, \dots, \underline{\delta}_N^N, \zeta, \zeta, \dots) \quad \text{and} \quad \bar{\delta}^N = (\bar{\delta}_1^N, \dots, \bar{\delta}_N^N, \zeta, \zeta, \dots).$$

Note that *any* policy in a period after period  $N$  is  $\underline{\phi}^N$ -optimal and  $\bar{\phi}^N$ -optimal, since rewards in these periods are constant, and consequently do not distinguish between policies. Then  $\underline{\delta}^N$  and  $\bar{\delta}^N$  belong to  $\Pi$  and (by Theorems 2.4, 2.5) are strongly  $\underline{\phi}^N$ -optimal and strongly  $\bar{\phi}^N$ -optimal, respectively, i.e.,  $\underline{\delta}^N \in \Omega^{(1)}(\underline{\phi}^N)$  and  $\bar{\delta}^N \in \Omega^{(1)}(\bar{\phi}^N)$ , so that  $\underline{R}^N(\underline{\delta}^N) = \underline{v}^N$  and  $\bar{R}^N(\bar{\delta}^N) = \bar{v}^N$ , for all  $N \in \mathbb{N}$ .

**4.1 Lemma.** *Suppose Assumptions 1 through 6 hold. Then, for each  $n \in \mathbb{N}$ , the policy sequences  $(\underline{\delta}_n^N)_{N=n}^\infty$  and  $(\bar{\delta}_n^N)_{N=n}^\infty$  converge in  $D$ . In particular, the policy sequences  $(\underline{\lambda}_n^N)_{N=n}^\infty$ ,  $(\bar{\lambda}_n^N)_{N=n}^\infty$ ,  $(\underline{\gamma}_n^N)_{N=n}^\infty$ ,  $(\bar{\gamma}_n^N)_{N=n}^\infty$  converge in  $D$ .*

For each  $i \in S$  and  $n \in \mathbb{N}$ , let  $\underline{\delta}_n(i) \equiv \lim_{N \rightarrow \infty} \underline{\delta}_n^N(i)$ ,  $\bar{\delta}_n(i) \equiv \lim_{N \rightarrow \infty} \bar{\delta}_n^N(i)$ , in  $D_i$ , and observe that there exists  $M$  sufficiently large so that  $\underline{\delta}_n(i) = \underline{\delta}_n^N(i)$  and  $\bar{\delta}_n(i) = \bar{\delta}_n^N(i)$ ,  $\forall N \geq M$ . Also let

$$\underline{\delta}_n = (\underline{\delta}_n(l), \dots, \underline{\delta}_n(u)) \quad \text{and} \quad \bar{\delta}_n = (\bar{\delta}_n(l), \dots, \bar{\delta}_n(u)) \quad \text{in} \quad D,$$

and let  $\underline{\delta} = (\underline{\delta}_n)_{n=1}^\infty$  and  $\bar{\delta} = (\bar{\delta}_n)_{n=1}^\infty$ , in  $\Pi$ . Clearly,  $\lim_{N \rightarrow \infty} \underline{\delta}^N = \underline{\delta}$  and  $\lim_{N \rightarrow \infty} \bar{\delta}^N = \bar{\delta}$ , in  $\Pi$ . In particular, this is the case for  $\underline{\delta} = \underline{\gamma}$  or  $\underline{\lambda}$ , and  $\bar{\delta} = \bar{\gamma}$  or  $\bar{\lambda}$ .

Next, we will show that the strategies  $\underline{\delta}$  and  $\bar{\delta}$  are  $\phi$ -optimal, i.e.,  $\underline{\delta}, \bar{\delta} \in \Omega = \Omega(\phi)$ .

**4.2 Lemma.** *Suppose Assumptions 5 and 6 hold. Then the mapping sequences  $(\underline{R}^N)_{N=1}^\infty$  and  $(\bar{R}^N)_{N=1}^\infty$  converge pointwise to  $R$  on  $\Pi$ , i.e., for each  $\pi \in \Pi$ ,  $\lim_{N \rightarrow \infty} \underline{R}^N(\pi) = \lim_{N \rightarrow \infty} \bar{R}^N(\pi) = R(\pi)$  in  $\mathbb{R}^s$ .*

**4.3 Lemma.** *Suppose Assumptions 5 and 6 hold. Then the sequence of optimal values  $\{\underline{v}^N\}_{N=1}^\infty$  for problems  $\langle \underline{\phi}^N \rangle_{N=1}^\infty$  converges to the total discounted expected reward for problem  $\langle \phi \rangle$  of strategy  $\underline{\delta}$ , i.e.,  $\lim_{N \rightarrow \infty} \underline{v}^N = \lim_{N \rightarrow \infty} \underline{R}^N(\underline{\delta}^N) = R(\underline{\delta})$ . Similarly for the sequence  $(\bar{v}^N)_{N=1}^\infty$  relative to problems  $\{\langle \bar{\phi}^N \rangle\}_{N=1}^\infty$ , i.e.,*

$$\lim_{N \rightarrow \infty} \bar{v}^N = \lim_{N \rightarrow \infty} \bar{R}^N(\bar{\delta}^N) = R(\bar{\delta}).$$

In particular,

$$\lim_{N \rightarrow \infty} \underline{R}^N(\underline{\lambda}^N) = R(\underline{\lambda}), \quad \lim_{N \rightarrow \infty} \bar{R}^N(\bar{\lambda}^N) = R(\bar{\lambda}), \quad \lim_{N \rightarrow \infty} \underline{R}^N(\underline{\gamma}^N) = R(\underline{\gamma}), \quad \lim_{N \rightarrow \infty} \bar{R}^N(\bar{\gamma}^N) = R(\bar{\gamma}).$$

We are now ready to prove that optimal strategies and optimal values to the finite horizon lower and upper bound problems converge, as horizons go to infinity, to optimal strategies and optimal values respectively for the infinite horizon problem, that is,  $\underline{\delta}$  and  $\bar{\delta}$  are  $\phi$ -optimal (i.e.,  $\underline{\delta}, \bar{\delta} \in \Omega = \Omega(\phi)$ ). Recall that  $\underline{\delta}^N \in \Omega^{(1)}(\underline{\phi}^N)$  and  $\bar{\delta}^N \in \Omega^{(1)}(\bar{\phi}^N)$ , for all horizons  $N$ .

**4.4 Theorem.** *Suppose Assumptions 5 and 6 hold. Then optimal strategy (policy) convergence holds, i.e.,  $\underline{\delta}$  and  $\bar{\delta}$  are optimal strategies for problem  $\langle \phi \rangle$ ,  $\lim_{N \rightarrow \infty} \underline{\delta}^N = \underline{\delta}$ , and  $\lim_{N \rightarrow \infty} \bar{\delta}^N = \bar{\delta}$ , in  $\Pi$ . In particular,  $\underline{\lambda}$ ,  $\bar{\lambda}$ ,  $\underline{\gamma}$ ,  $\bar{\gamma}$  are optimal for problem  $\langle \phi \rangle$ ,  $\lim_{N \rightarrow \infty} \underline{\lambda}^N = \underline{\lambda}$ ,  $\lim_{N \rightarrow \infty} \bar{\lambda}^N = \bar{\lambda}$ ,  $\lim_{N \rightarrow \infty} \underline{\gamma}^N = \underline{\gamma}$ , and  $\lim_{N \rightarrow \infty} \bar{\gamma}^N = \bar{\gamma}$ . Also, optimal value convergence holds, i.e.,  $\lim_{N \rightarrow \infty} \bar{v}^N = \lim_{N \rightarrow \infty} \underline{v}^N = v$ , as functions on  $S$ . More generally, for each  $n \in \mathbb{N}$ ,  $\lim_{N \rightarrow \infty} \bar{v}_n^N = \lim_{N \rightarrow \infty} \underline{v}_n^N = v_n$ , as functions on  $S$ .*

**Outline of Proof.** For each  $N \in \mathbb{N}$ , we have that  $\underline{R}^N(\underline{\delta}^N) \geq \underline{R}^N(\pi)$ ,  $\forall \pi \in \Pi$ , since  $\underline{\delta}^N$  is optimal for  $\langle \underline{\phi}^N \rangle$ . By Lemma 4.3,  $\lim_{N \rightarrow \infty} \underline{R}^N(\underline{\delta}^N) = R(\underline{\delta})$ , and by Lemma 4.2,  $\lim_{N \rightarrow \infty} \underline{R}^N(\pi) = R(\pi)$ ,  $\forall \pi \in \Pi$ . Hence,  $R(\underline{\delta}) \geq R(\pi)$ ,  $\forall \pi \in \Pi$ , so that for problem  $\langle \phi \rangle$ ,  $\underline{\delta}$  is an optimal strategy and  $R(\underline{\delta})$  is the optimal value  $v$ .

Since  $\lim_{N \rightarrow \infty} \underline{\delta}^N = \underline{\delta}$ , and  $\lim_{N \rightarrow \infty} \bar{\delta}^N = \bar{\delta}$ , we have optimal policy convergence. Moreover, by Lemma 4.3, we also have  $\lim_{N \rightarrow \infty} \underline{v}^N = \lim_{N \rightarrow \infty} \underline{R}^N(\underline{\delta}^N) = R(\underline{\delta}) = v = v_1$ , so that optimal value convergence holds as well.

For the last part, observe that although forecast  $\phi$  is fixed, it is also arbitrary in  $\Phi$ . Moreover,  $\underline{v}^N = v(\underline{\phi}^N)$ ,  $\forall N \in \mathbb{N}$ . Thus, we have actually shown that  $\lim_{N \rightarrow \infty} v(\underline{\theta}^N) = v(\theta) = \lim_{N \rightarrow \infty} v(\bar{\theta}^N)$ , for each  $\theta$  in  $\Phi$ . Also, for  $1 \leq n \leq N$ , we have that  $\underline{v}_n^N = v_n(\underline{\phi}^N) = v((\underline{\phi}^N)^{(n)}) = v(\overline{(\phi^{(n)})}^{N-n+1})$ . Consequently, letting  $\theta = \phi^{(n)} \in \Phi$  and  $M = N - n + 1$  in the previous limits, so that  $\underline{\theta}^N = \overline{(\phi^{(n)})}^{N-n+1}$ , it follows that

$$v(\phi^{(n)}) = v_n(\phi) = v_n = \lim_{N \rightarrow \infty} v(\overline{(\phi^{(n)})}^{N-n+1}) = \lim_{N \rightarrow \infty} \underline{v}_n^N,$$

for each  $n \in \mathbb{N}$ . For the remaining case, this proof can be repeated verbatim with lower bars replaced by upper bars.  $\square$

**4.5 Lemma.** *Suppose Assumptions 5 and 6 hold. Then, for each  $n \in \mathbb{N}$ , the function sequences  $(\underline{f}_n^N)_{N=n}^\infty$  and  $(\bar{f}_n^N)_{N=n}^\infty$  converge pointwise on  $L$  to  $f_n$ , i.e.,  $\lim_{N \rightarrow \infty} \underline{f}_n^N(i, k) = \lim_{N \rightarrow \infty} \bar{f}_n^N(i, k) = f_n(i, k)$ ,  $\forall (i, k) \in L$ .*

**4.6 Theorem.** *Suppose Assumptions 5 and 6 hold. Then the strategies  $\underline{\delta}$  and  $\bar{\delta}$  are strongly  $\phi$ -optimal, i.e.,  $\underline{\delta}, \bar{\delta} \in \Omega^{(1)}(\phi)$ . In particular, the strategies  $\underline{\lambda}, \bar{\lambda}, \underline{\gamma}, \bar{\gamma}$  are strongly  $\phi$ -optimal, i.e.,  $\underline{\lambda}, \bar{\lambda}, \underline{\gamma}, \bar{\gamma} \in \Omega^{(1)}(\phi)$ .*

**Proof.** Since  $\underline{\delta}_n^N(i) \in \underline{K}_n^N(i)$ , we have that  $\underline{v}_n^N(i) = \underline{f}_n^N(i, \underline{\delta}_n^N(i))$ ,  $\forall n \in \mathbb{N}$ ,  $\forall N \geq n$ ,  $\forall i \in S$ . But  $\lim_{N \rightarrow \infty} \underline{v}_n^N(i) = v_n(i)$ , by Theorem 4.4,  $\underline{\delta}_n^N(i) = \underline{\delta}_n(i)$ , for sufficiently large  $N$ , and  $\lim_{N \rightarrow \infty} \underline{f}_n^N(i, \underline{\delta}_n(i)) = f_n(i, \underline{\delta}_n(i))$ , by Lemma 4.5. Hence,  $v_n(i) = f_n(i, \underline{\delta}_n(i))$ ,  $\forall i \in S$ , so that  $\underline{\delta}_n \in \Omega_n(\phi)$ ,  $\forall n$ . Consequently,  $\underline{\delta} \in \Omega^{(1)}(\phi)$  by Theorem 2.4. Similarly for  $\bar{\delta} \in \Omega^{(1)}(\phi)$ .  $\square$

**4.7 Lemma.** *Suppose Assumptions 1 through 6 hold. Then, for each  $N \in \mathbb{N}$ , if  $l \leq j < i \leq u$ , then*

$$\underline{v}_{N+1}^N(i) - \underline{v}_{N+1}^N(j) \leq v_{N+1}(i) - v_{N+1}(j) \leq \bar{v}_{N+1}^N(i) - \bar{v}_{N+1}^N(j).$$

For the following, recall (from section 2) the meaning of  $\lambda$  and  $\gamma$ , and the fact that  $\lambda_n$  and  $\gamma_n$  are the smallest and largest policies (respectively) in  $\Omega_n$ ,  $\forall n \in \mathbb{N}$ .

**4.8 Theorem.** *Suppose Assumptions 1 through 6 hold. Fix  $n \in \mathbb{N}$ . Then,  $\forall N \geq n$ , we have  $\underline{\lambda}_n^N \leq \lambda_n \leq \bar{\lambda}_n^N$  and  $\underline{\gamma}_n^N \leq \gamma_n \leq \bar{\gamma}_n^N$ . Consequently,  $\underline{\lambda} \leq \lambda \leq \bar{\lambda}$  and  $\underline{\gamma} \leq \gamma \leq \bar{\gamma}$ .*

**Proof.** In the proof of Theorem 3.3, replace  $\{N, N+1\}$ ,  $\underline{v}_n^M(j)$ ,  $\underline{\lambda}_n^M$  and  $\underline{\gamma}_n^M$  by  $\{N, \infty\}$ ,  $v_n(j)$ ,  $\lambda_n$  and  $\gamma_n$ , respectively. Beginning with the fact (Lemma 4.7) that  $\underline{v}_{N+1}^N(i) - \underline{v}_{N+1}^N(j) \leq v_{N+1}(i) - v_{N+1}(j)$ , we see that the entire proof of Theorem 3.3 is valid for  $N+1$  replaced by  $\infty$ . Then, it follows from part (ii) of the resulting version of Theorem 3.3, that  $\underline{\lambda}_n^N \leq \lambda_n$  and  $\underline{\gamma}_n^N \leq \gamma_n$ . Analogously, repeat this proof with lower bars replaced by upper bars.

**4.9 Corollary.** *Suppose Assumptions 1 through 6 hold. Then, for any  $n \in \mathbb{N}$  and any  $N \geq n$ :*

- (1) either  $\underline{\lambda}_n^N \leq \bar{\lambda}_n^N \leq \underline{\gamma}_n^N \leq \bar{\gamma}_n^N$  or  $\underline{\lambda}_n^N \leq \underline{\gamma}_n^N \leq \bar{\lambda}_n^N \leq \bar{\gamma}_n^N$ ;
- (2) if  $\underline{\lambda}_n^N = \bar{\lambda}_n^N$ , then  $\underline{\lambda}_n^M = \bar{\lambda}_n^M$ , for all  $M \geq N$ ;
- (3) if  $\underline{\gamma}_n^N = \bar{\gamma}_n^N$  then  $\underline{\gamma}_n^M = \bar{\gamma}_n^M$ , for all  $M \geq N$ .

## 5. Conditions for Optimal Policies

Let  $d \in D$  and  $n \in \mathbb{N}$ . Recall that *policy  $d$  is  $\phi$ -optimal in period  $n$*  if  $d \in \Omega_n(\phi)$ , i.e.,  $d_i \in K_n(\phi; i)$ ,  $\forall l \leq i \leq u$ . In this section, we investigate sufficient conditions for this to happen, particularly for the case  $n = 1$ , i.e., for  $d$  to be a first-period optimal policy. However, we begin with a necessary condition.

**5.1 Theorem.** *Suppose Assumptions 1 through 6 hold. If  $d \in D$  is  $\phi$ -optimal in period  $n$ , then  $\underline{\lambda}_n^N \leq d \leq \overline{\gamma}_n^N$  component-wise on  $S$ ,  $\forall N \geq n$ .*

**Proof.** By hypothesis,  $\lambda_n \leq d \leq \gamma_n$ . Hence, by Theorem 4.8,  $\underline{\lambda}_n^N \leq \lambda_n \leq d \leq \gamma_n \leq \overline{\gamma}_n^N$ ,  $\forall N \geq n$ , as required.  $\square$

Next we turn to finding a sufficient condition. First, we require some preliminary results. Fix  $i \in S$ . Then  $D_i = \{a_i, \dots, b_i\}$  with cardinality  $|D_i| = b_i - a_i + 1$ . If  $|D_i| \geq 3$  and  $g$  is any real-valued function on  $D_i$ , then we will say that  $g$  is *concave* if  $g(i, k'') - g(i, k) \leq g(i, k) - g(i, k')$ , whenever  $a_i \leq k' \leq k \leq k'' \leq b_i$ .

**Assumption 7.** Assume that for each  $N \in \mathbb{N}$ , and for each  $i \in S$  for which  $|D_i| \geq 3$ ,

(i) the real-valued functions  $k \rightarrow \underline{f}_1^N(i, k)$  are concave on  $D_i$ , and

(ii) the real-valued functions  $k \rightarrow \overline{f}_1^N(i, k)$  are concave on  $D_i$ .

Note that Assumption 7 is automatically met in the important special case where  $|D_i| = 2$ , all  $i$  as in the asset selling problem of section 7.

**5.2 Lemma.** *Suppose Assumptions 6 and 7 hold. Let  $1 \leq n \leq N$ .*

(i) *A policy  $d \in D$  is optimal for problem  $\langle \phi^N \rangle$  in period  $n$  if and only if  $\underline{\lambda}_n^N(i) \leq d_i \leq \underline{\gamma}_n^N(i)$ , i.e.,  $\underline{K}_n^N(i)$  is the integer interval  $\{\underline{\lambda}_n^N(i), \dots, \underline{\gamma}_n^N(i)\}$ ,  $\forall i \in S$ .*

(ii) *Furthermore, for each  $i \in S$  for which  $|D_i| \geq 2$ ,*

$$\underline{f}_n^N(i, k) - \underline{f}_n^N(i, k-1) \quad \text{is} \quad \begin{cases} > 0, & \text{if } a_i < k \leq \underline{\lambda}_n^N(i), \\ = 0, & \text{if } \underline{\lambda}_n^N(i) < k \leq \underline{\gamma}_n^N(i), \\ < 0, & \text{if } \underline{\gamma}_n^N(i) < k \leq b_i. \end{cases}$$

Analogously, we have the following:

**5.3 Lemma.** *Suppose Assumptions 6 and 7 hold. Let  $1 \leq n \leq N$ .*

(i) *Then a policy  $d \in D$  is optimal for problem  $\langle \overline{\phi}^N \rangle$  in period  $n$  if and only if  $\overline{\lambda}_n^N(i) \leq d_i \leq \overline{\gamma}_n^N(i)$ ,  $\forall i \in S$ , i.e.,  $\overline{K}_n^N(i)$  is the integer interval  $\{\overline{\lambda}_n^N(i), \dots, \overline{\gamma}_n^N(i)\}$ ,  $\forall i \in S$ .*

(ii) *Furthermore, for each  $i \in S$  for which  $|D_i| \geq 2$ ,*

$$\overline{f}_n^N(i, k) - \overline{f}_n^N(i, k-1) \quad \text{is} \quad \begin{cases} > 0, & \text{if } a_i < k \leq \overline{\lambda}_n^N(i), \\ = 0, & \text{if } \overline{\lambda}_n^N(i) < k \leq \overline{\gamma}_n^N(i), \\ < 0, & \text{if } \overline{\gamma}_n^N(i) < k \leq b_i. \end{cases}$$

**Remark.** Lemma 5.2 requires Assumption 7 part (i), while Lemma 5.3 requires Assumption 7 (ii).

The following is our sufficient condition for policy optimality in a period.

**5.4 Theorem.** *Suppose Assumptions 1 through 6 hold as well as either part (i) or (ii) of Assumption 7 (not necessarily both). Let  $n \in \mathbb{N}$  and  $d \in D$ . If there exists  $M \geq n$  such that  $\overline{\lambda}_n^M \leq d \leq \underline{\gamma}_n^M$  component-wise on  $S$ , then  $d$  is  $\theta$ -optimal in period  $n$ , for all  $\theta \in \Phi$  satisfying  $\theta_i = \phi_i$ ,  $\forall i = 1, \dots, M$ .*

**Proof.** By Corollary 3.6, the function  $N \rightarrow \overline{\lambda}_n^N$  is component-wise monotonically non-increasing. Therefore, by hypothesis and Theorem 4.8, we have that  $1 \leq n \leq M \leq N$  implies  $\underline{\lambda}_n^N \leq \overline{\lambda}_n^N \leq \overline{\lambda}_n^M \leq d$ , as functions on  $S$ . Analogously, by Corollary 3.4, the function  $N \rightarrow \underline{\gamma}_n^N$  is component-wise monotonically non-decreasing on  $S$ . Therefore, by hypothesis and Theorem 4.8, we have  $1 \leq n \leq M \leq N$  implies  $d \leq \underline{\gamma}_n^M \leq \underline{\gamma}_n^N \leq \overline{\gamma}_n^N$ , as functions on  $S$ . Thus,  $\underline{\lambda}_n^N \leq d \leq \underline{\gamma}_n^N$ ,  $\forall N \geq M$ , as functions on  $S$ . By Lemma 5.2(i),  $d$  is an optimal policy for period  $n$  of problem  $\langle \phi^N \rangle$ , for all  $N \geq M$ . Hence, by definition of optimal policy,

$\underline{f}_n^N(i, d_i) = \underline{v}_n^N(i)$ ,  $\forall i \in S$ ,  $\forall N \geq M$ . But, by Lemma 4.5,  $\lim_{N \rightarrow \infty} \underline{f}_n^N(i, d_i) = f_n(i, d_i)$ ,  $\forall i \in S$ , and, by Theorem 4.4,  $\lim_{N \rightarrow \infty} \underline{v}_n^N(i) = v_n(i)$ ,  $\forall i \in S$ . Consequently,  $f_n(i, d_i) = v_n(i)$ ,  $\forall i \in S$ , so that  $d$  is an optimal policy for problem  $\langle \phi \rangle$  in period  $n$ . But  $\bar{\lambda}_n^M$  and  $\underline{\gamma}_n^M$  depend only on the period forecasts  $\phi_1, \dots, \phi_M$  and the quantities  $\bar{w}$  and  $\underline{w}$ . Consequently,  $d$  is also an optimal policy for each such problem  $\langle \theta \rangle$  in period  $n$ . The proof for the other case is similar, using Lemma 5.3(i) in place of Lemma 5.2(i).  $\square$

## 6. Solution and Forecast Horizons

In this section we give sufficient conditions for the existence of solution and forecast horizons, as well as algorithms for their determination.

**Definition (solution horizon).** *We define a solution horizon for problem  $\langle \phi \rangle$  to be a finite horizon  $M$  for which there exists a vector  $\nu \in \mathbb{R}^s$  and a first-period optimal policy  $d$  in  $D$  for the finite horizon problem  $\langle \phi(N, \nu) \rangle$ , i.e.,  $d \in \Omega_1(\phi(N, \nu)) = \Omega_1(\phi^N(\nu))$ ,  $\forall N \geq M$ . Of course, each  $N \geq M$  is also a solution horizon.*

As a result, there exists a finite horizon version  $\langle \phi^M(\nu) \rangle$  of problem  $\langle \phi \rangle$  which yields a  $\phi$ -optimal first-period policy. Note that this definition applies to any forecast satisfying Assumption 2. Moreover, all homogeneous problems have solution horizons.

Suppose we are given the problem class  $\Phi$  as above with fixed forecast  $\phi$ .

**Definition (forecast horizon).** *We define a forecast horizon for problem  $\langle \phi \rangle$  in class  $\Phi$  to be a finite horizon  $M$  for which there exists a first-period  $\phi$ -optimal policy which is also first-period  $\theta$ -optimal for any forecast  $\theta \in \Phi$  whose forecasts in periods  $1 \leq n \leq M$  are the same as those of  $\phi$ , i.e., there exists  $d \in \Omega_1(\phi)$  such that  $d \in \Omega_1(\theta)$ , for all  $\theta \in \Phi$  such that  $\theta_n = \phi_n$ ,  $\forall 1 \leq n \leq M$ . Note that any horizon greater than  $M$  is also a forecast horizon.*

Note that our definition of forecast horizon is very weak in that it is problem dependent but algorithm independent. If we can identify solution horizons  $M_n$  with vectors  $\nu_n$ , one for each forecast  $\phi^{(n)}$ , we can apply a “rolling horizon” procedure to recursively solve problem  $\langle \phi \rangle$ . That is, we first solve problem  $\langle \phi(M_1, \nu_1) \rangle$  and implement an optimal first policy for  $\phi$ , which is guaranteed to exist. After that, roll forward to the next period, solve problem  $\langle \phi(M_2, \nu_2) \rangle$  and implement an optimal first policy for  $\phi^{(2)}$ , which is also guaranteed to exist. Repeat this procedure ad infinitum. In this way, we are able to recursively construct a sequence of  $\phi$ -optimal policies (i.e., a strategy) which is guaranteed to be strongly  $\phi$ -optimal (Theorem 2.5). Moreover, we shall see (Theorem 6.2) that solution horizons exist for all the  $\phi^{(n)}$ , for all  $n \geq 1$ .

**6.1 Lemma.** *Suppose Assumptions 5 and 6 hold. If  $M$  is a forecast horizon for problem  $\langle \phi \rangle$  and class  $\Phi$ , then it is also a solution horizon for  $\langle \phi \rangle$ .*

We next show the existence of solution horizons, and present an algorithm for determining such a horizon. The crux of our algorithm is to successively solve problems  $\langle \underline{\phi}^N \rangle$  and  $\langle \bar{\phi}^N \rangle$ , with increasing horizons  $N$ , until we obtain optimal strategies whose first-period policy decisions satisfy our sufficient condition. We conclude this part of the section by, in turn, providing a sufficient condition for this algorithm to terminate in a finite number of steps.

Recall the meanings of  $\underline{\delta}$  and  $\bar{\delta}$  from section 4, as well as Assumptions 1-6.

**6.2 Theorem.** *Suppose Assumptions 1 through 6 hold. Then there exists a solution horizon for problem  $\langle \phi \rangle$ . Moreover, there exists a horizon  $M$  sufficiently large such that, if  $N \geq M$ , then  $\underline{\delta}_1^N = \underline{\delta}_1 \in \Omega_1(\phi)$  and  $\bar{\delta}_1^N = \bar{\delta}_1 \in \Omega_1(\phi)$ , which are both  $\phi$ -optimal first-period policies. In particular, this is true for  $\underline{\delta} = \underline{\lambda}, \underline{\gamma}$  and  $\bar{\delta} = \bar{\lambda}, \bar{\gamma}$ . The previous is also true for each  $\phi^{(n)}$ ,  $n \geq 1$ .*

**Proof.** We have that the sequence  $(\underline{\delta}^N)_{N=1}^\infty$  converges to  $\underline{\delta}$  in the product topology of  $\Pi$  (Theorem 4.4). Consequently, in particular,  $(\underline{\delta}_1^N)_{N=1}^\infty$  converges to  $\underline{\delta}_1$  in  $D$ , i.e., the sequence of integers  $(\underline{\delta}_1^N(i))_{N=1}^\infty$  converges to  $\underline{\delta}_1(i)$  in  $\mathbb{N}$ , for each  $i \in S$ . Thus, eventually  $\underline{\delta}_1^N(i)$  must equal  $\underline{\delta}_1(i)$ , for each  $i$  in the finite set  $S$ . Hence,

there exists  $M$  sufficiently large such that, if  $N \geq M$ , then  $\underline{\delta}_1^N(i) = \underline{\delta}_1(i)$ ,  $\forall i \in S$ , i.e.,  $\underline{\delta}_1^N = \underline{\delta}_1$ . We have the analogous result for  $\bar{\delta}$ . The proof is then completed by observing that  $\underline{\delta}_1$  and  $\bar{\delta}_1$  belong to  $\Omega_1(\phi)$  (Theorem 4.6). The last part follows from the fact that  $\phi^{(n)} \in \Phi$ ,  $\forall n$ .  $\square$

**6.3 Theorem.** *Suppose Assumptions 1 through 7 hold. If there exists  $M \geq 1$  such that  $\bar{\lambda}_1^M \leq \underline{\gamma}_1^M$  component-wise on  $S$ , then  $M$  is a solution horizon for problem  $\langle \phi \rangle$ .*

**Proof.** This follows from Theorem 5.4 for  $n = 1$  and the definition of solution horizon.  $\square$

### SOLUTION HORIZON ALGORITHM

Suppose we are given  $\phi \in \Phi$  as in section 2 satisfying Assumptions 1 through 7.

**INITIALIZE:** Solve the infinite horizon homogeneous problems  $\langle \underline{\psi} \rangle$  and  $\langle \bar{\psi} \rangle$  to get  $\underline{w}$  and  $\bar{w}$ . Set  $N=1$ .

**ITERATE:** Solve the resulting problems  $\langle \underline{\phi}^N \rangle$  and  $\langle \bar{\phi}^N \rangle$  for  $\underline{\gamma}_1^N$  and  $\bar{\lambda}_1^N$ .

**TERMINATE:** If  $\bar{\lambda}_1^N \leq \underline{\gamma}_1^N$ , stop. Then  $N$  is a solution horizon. Otherwise, replace  $N$  by  $N + 1$  and Iterate.

**Remark.** Note that the proof of Theorem 6.3 and the Solution Horizon Algorithm require only that either (i) or (ii) of Assumption 7 holds.

We next turn to forecast horizons.

**6.4 Lemma.** *Suppose Assumptions 1 through 6 hold. Suppose also that there exists a horizon  $N$  and a period  $1 \leq n \leq N$  for which  $\underline{\lambda}_n^N = \bar{\lambda}_n^N$  (respectively,  $\underline{\gamma}_n^N = \bar{\gamma}_n^N$ ). Then  $\underline{\lambda}_n^M = \bar{\lambda}_n^M = \lambda_n = \lambda_n(\phi)$  (respectively,  $\underline{\gamma}_n^M = \bar{\gamma}_n^M = \gamma_n = \gamma_n(\phi)$ ), for all  $M \geq N$ .*

**6.5 Theorem.** *Suppose Assumptions 1 through 6 hold. Suppose also that there exists a horizon  $N$  for which  $\underline{\lambda}_1^N = \bar{\lambda}_1^N$  (respectively,  $\underline{\gamma}_1^N = \bar{\gamma}_1^N$ ). Then such  $N$  is a forecast horizon for problem  $\langle \theta \rangle$  and class  $\Phi$ , for all  $\theta \in \Phi$  such that  $\theta_i = \phi_i$ ,  $\forall i = 1, \dots, N$ .*

**Proof.** Suppose  $\underline{\lambda}_1^N = \bar{\lambda}_1^N$ . Let  $\theta \in \Phi$  be such that  $\theta_n = \phi_n$ ,  $\forall 1 \leq n \leq N$ . For such  $\theta$ , we have  $\underline{\theta}^N = \underline{\phi}^N$  and  $\bar{\theta}^N = \bar{\phi}^N$ . Consequently, by hypothesis,  $\lambda_1(\underline{\theta}^N) = \lambda_1(\underline{\phi}^N) = \underline{\lambda}_1^N = \bar{\lambda}_1^N = \lambda_1(\bar{\phi}^N) = \lambda_1(\bar{\theta}^N)$ . Lemma 6.4 is valid for fixed, but arbitrary,  $\phi \in \Phi$ . Hence, applying Lemma 6.4 to  $\theta$  and  $\phi$ , with  $M = N$ , we obtain  $\lambda_1(\phi) = \lambda_1 = \bar{\lambda}_1^N = \underline{\lambda}_1^N = \lambda_1(\underline{\theta}^N) = \lambda_1(\theta)$ . Therefore,  $d \equiv \lambda_1$  is in  $\Omega_1(\phi)$  and in each of the required  $\Omega_1(\theta)$ . The second part is proved similarly.  $\square$

For notational convenience, given  $N \geq 1$  and  $\phi \in \Phi$ , we let

$$\Phi^N(\phi) = \{\theta \in \Phi : \theta_i = \phi_i, \forall i = 1, \dots, N\}.$$

**Definition (well-posedness).** *Define  $\phi \in \Phi$  to be well-posed if there exists  $N$  (depending on  $\phi$ ) such that  $\bigcap_{\theta \in \Phi^N(\phi)} \Omega_1(\theta) \neq \emptyset$ , i.e., there exists a policy which is first-period optimal for each forecast  $\theta \in \Phi^N(\phi)$ .*

Note that a problem is well-posed if and only if it admits a forecast horizon and that all stationary forecasts are well-posed relative to the set of all stationary forecasts. In section 2, we gave an example of a problem which is not well-posed. We next characterize forecast horizons for well-posed problems.

**6.6 Theorem.** *Suppose Assumptions 1 through 7 hold. Let  $\phi \in \Phi$  be well-posed. Then:*

- (i) *the set of positive integers  $\mathbb{N}(\phi) = \{N \in \mathbb{N} : \bar{\lambda}_1^N \leq \underline{\gamma}_1^N\}$  is non-empty;*
- (ii) *for each  $N \in \mathbb{N}(\phi)$ , we have  $\{d \in D : \bar{\lambda}_1^N \leq d \leq \underline{\gamma}_1^N\} \subseteq \Omega_1(\theta)$ ,  $\forall \theta \in \Phi^N(\phi)$ ;*
- (iii) *for each  $N \in \mathbb{N}(\phi)$ ,  $N$  is a forecast horizon.*

*Proof.* (i) Suppose that for each  $N$ ,  $\bar{\lambda}_1^N \not\leq \underline{\gamma}_1^N$ . Then, for each  $N$ , there exists  $i_N \in S$  for which  $\bar{\lambda}_1^N(i_N) > \underline{\gamma}_1^N(i_N)$ . However, since  $S$  is finite, there must exist a state  $i$  and a subsequence  $(N_n)_{n=1}^\infty$  such that  $\bar{\lambda}_1^{N_n}(i) > \underline{\gamma}_1^{N_n}(i)$ ,  $\forall n$ . But  $\lim_{n \rightarrow \infty} \bar{\lambda}_1^{N_n}(i) = \bar{\lambda}_1(i)$  and  $\lim_{n \rightarrow \infty} \underline{\gamma}_1^{N_n}(i) = \underline{\gamma}_1(i)$  (Theorem 4.4). Hence,  $\bar{\lambda}_1(i) \geq \underline{\gamma}_1(i)$ . If  $\bar{\lambda}_1(i) > \underline{\gamma}_1(i)$ , then for all  $n$ ,  $\Omega_1(\bar{\phi}^{N_n}) \cap \Omega_1(\underline{\phi}^{N_n}) = \emptyset$ , which is a contradiction by well-posedness and the fact that  $\underline{\phi}^{N_n}$  and  $\bar{\phi}^{N_n}$  both belong to  $\Phi^{N_n}(\phi)$ . If  $\bar{\lambda}_1(i) = \underline{\gamma}_1(i)$ , then necessarily  $\bar{\lambda}_1^{N_n}(i) = \underline{\gamma}_1^{N_n}(i)$ , for large  $n$ , since  $\bar{\lambda}_1(i)$  and  $\underline{\gamma}_1(i)$  are integers. This is also a contradiction.

(ii) For each  $N \in \mathbb{N}(\phi)$ , we have that  $\bar{\lambda}_1^N \leq \underline{\gamma}_1^N$ , so that by Theorem 5.4,

$$\{d \in D : \bar{\lambda}_1^N \leq d \leq \underline{\gamma}_1^N\} \subseteq \Omega_1(\theta), \quad \forall \theta \in \Phi^N(\phi).$$

Consequently,

$$\bigcap_{\theta \in \Phi^N(\phi)} \Omega_1(\theta) \supseteq \{d : \bar{\lambda}_1^N \leq d \leq \underline{\gamma}_1^N\} \neq \emptyset.$$

(iii) By definition, each  $N \in \mathbb{N}(\phi)$  is a forecast horizon. □

When the problem is well-posed (and only then), a forecast horizon exists. In this case, the Solution Horizon Algorithm terminates in the discovery of a forecast horizon.

### FORECAST HORIZON ALGORITHM

Suppose we are given a  $\phi \in \Phi$  which is well-posed.

**INITIALIZE:** Solve the infinite horizon homogeneous problems  $\langle \underline{\psi} \rangle$  and  $\langle \bar{\psi} \rangle$  to get  $\underline{w}$  and  $\bar{w}$ . Set  $N=1$ .

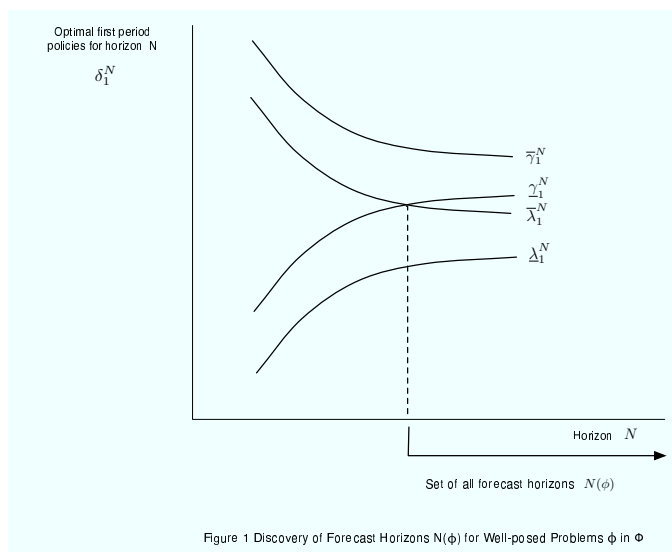
**ITERATE:** Solve the resulting problems  $\langle \underline{\phi}^N \rangle$  and  $\langle \bar{\phi}^N \rangle$  for  $\underline{\gamma}_1^N$  and  $\bar{\lambda}_1^N$

**TERMINATE:** Stop if  $\bar{\lambda}_1^N \leq \underline{\gamma}_1^N$ . Then  $N$  is a forecast horizon. Otherwise, replace  $N$  by  $N + 1$  and Iterate.

**6.7 Theorem.** *Suppose all seven assumptions hold. If problem  $\langle \phi \rangle$  is well-posed, then the Forecast Horizon Algorithm terminates in finite time.*

*Proof.* This follows from parts (i) and (ii) of Theorem 6.6. □

Theorem 6.7 assures us that the Forecast Horizon Algorithm will successfully determine with finite data an infinite horizon optimal first policy. See Figure 1 for an illustration of this discovery for a well-posed problem (the graphs are drawn along continuous time for visual clarity).



Recall (Theorem 6.2) that solution horizons exist in general, and  $D_i = \{a_i, \dots, b_i\}$ ,  $\forall i \in S$ . The following theorems provide conditions which are sufficient for a problem to be well-posed so that a forecast horizon exists and the Forecast Horizon Algorithm terminates in finitely many iterations. Uniqueness of an infinite horizon optimal first policy is sufficient but not necessary for this condition to hold.

**6.8 Theorem.** *Suppose Assumptions 1 through 7 hold. Suppose also that if  $i \in S$  is such that  $\underline{\gamma}_1(i) < b_i$ , then either  $f_1(i, \underline{\gamma}_1(i) + 1) \neq f_1(i, \underline{\gamma}_1(i))$ , or  $\bar{F}_1^N(i, \underline{\gamma}_1^N(i) + 1) \neq \bar{F}_1^N(i, \underline{\gamma}_1^N(i))$ , for sufficiently large  $N$ . Then problem  $\langle \phi \rangle$  is well-posed. In particular, if there exists a unique infinite horizon optimum for problem  $\langle \phi \rangle$ , then it is well-posed, since the condition is then satisfied.*

**Proof.** Recall that by setting  $\delta = \gamma$  in section 4, we have, in particular, that for each  $i \in S$ ,  $\underline{\gamma}_1(i) = \lim_{N \rightarrow \infty} \underline{\gamma}_1^N(i)$ . Thus, in view of the discreteness of  $K$ , there must exist  $M$  sufficiently large so that  $N \geq M$  implies  $\underline{\gamma}_1^N(i) = \underline{\gamma}_1(i)$ ,  $\forall i \in S$ . Hence, for all  $i \in S$  for which  $\underline{\gamma}_1(i) = b_i$ , we have  $N \geq M$  implies  $\underline{\gamma}_1^N(i) \leq \bar{\gamma}_1^N(i) \leq b_i = \underline{\gamma}_1(i) = \underline{\gamma}_1^N(i)$ , i.e.,  $\underline{\gamma}_1^N(i) = \bar{\gamma}_1^N(i)$ . On the other hand, for  $i \in S$  for which  $\underline{\gamma}_1(i) < b_i$ , we have  $\underline{\gamma}_1(i) + 1 \leq b_i$ , so that, by hypothesis,  $f_1(i, \underline{\gamma}_1(i) + 1) < f_1(i, \underline{\gamma}_1(i))$ , since  $\underline{\gamma}_1 \in \Omega_1(\phi)$ . Necessarily, by Lemma 4.5 and hypothesis (i) of Theorem 6.6, we may assume that  $M$  is sufficiently large such that  $N \geq M$  implies  $\bar{F}_1^N(i, \underline{\gamma}_1(i) + 1) < \bar{F}_1^N(i, \underline{\gamma}_1(i))$ , and that  $\bar{F}_1^N(i, \cdot)$  is concave (if  $|D_i| \geq 3$ ), for all  $N \geq M$ . Consequently, for any such  $N$ , by Lemma 5.3(ii), it follows that  $\bar{\gamma}_1^N(i) < \underline{\gamma}_1^N(i) + 1 = \underline{\gamma}_1(i) + 1$ . Therefore, for such  $N$ ,  $\bar{\gamma}_1^N(i) \leq \underline{\gamma}_1^N(i)$ . However, by Theorem 4.8, it is also the case that  $\bar{\gamma}_1^N(i) \geq \underline{\gamma}_1^N(i)$ , so that  $\bar{\gamma}_1^N(i) = \underline{\gamma}_1^N(i)$ . Since the stopping condition of Theorem 6.5 is satisfied, the algorithm will stop at  $N$ , if not earlier. The remaining part is proved similarly.  $\square$

Analogously, we have the following.

**6.9 Theorem.** *Suppose Assumptions 1 through 7 hold. Suppose also that if  $i \in S$  is such that  $a_i < \bar{\lambda}_1(i)$ , then either  $f_1(i, \bar{\lambda}_1(i) - 1) \neq f_1(i, \bar{\lambda}_1(i))$ , or  $\bar{F}_1^N(i, \bar{\lambda}_1^N(i) - 1) \neq \bar{F}_1^N(i, \bar{\lambda}_1^N(i))$ , for sufficiently large  $N$ . Then problem  $\langle \phi \rangle$  is well-posed. In particular, if there exists a unique infinite horizon optimum for problem  $\langle \phi \rangle$ , then it is well-posed.*

## 7. Application to Asset Selling

We return to the asset selling problem introduced in section 1. We will show how to solve every well-posed problem in this class. In this setting, we have that  $K = \{0, 1\}$ ,

$$i_k = \begin{cases} 0, & \text{if } k = 0, \\ 1, & \text{if } k = 1, \end{cases}$$

and

$$L = \{(i, k) : 0 \leq i \leq k, \quad k = 0, 1\} \cup \{(g+1, 1)\}$$

is a lattice (Assumption 1). (Note that it would not be a lattice had we chosen  $D_{g+1} = \{0\}$  instead.)

For each period  $n$ , the reward function is defined by

$$\rho_n(i, k) = \begin{cases} -h_n, & \text{if } 0 \leq i \leq g, \quad k = 0, \\ i, & \text{if } 0 \leq i \leq g, \quad k = 1, \\ 0, & \text{if } i = g+1, \quad k = 1, \end{cases}$$

where  $h_n \geq 0$  is the  $n$ -th period holding cost, which is assumed to be bounded; thus,  $\sup_n h_n < \infty$ . Hence, the sequence  $(\rho_n)_{n=1}^\infty$  is indexed by the bounded sequence  $(h_n)_{n=1}^\infty$  of non-negative real numbers with  $\rho_n \leftrightarrow h_n$ , and clearly, each function  $\rho_n(\cdot, 0)$  is independent of the rejected offer  $0 \leq i \leq g$ . Of course, once the asset is sold, the reward in each later period is zero. The reward functions are easily seen to have increasing differences on  $L$  (Assumption 2).

In order to define the transition structure, let

$$p_n(i, j; k) = \begin{cases} q_n(j), & \text{if } 0 \leq j \leq g, \quad k = 0, \\ 0, & \text{if } j = g+1, \quad k = 0; \text{ or } 0 \leq j \leq g, \quad k = 1, \\ 1, & \text{if } j = g+1, \quad k = 1. \end{cases}$$

Note that these probabilities are independent of  $i \in S$  (Assumption 3), and the sequence  $(p_n)_{n=1}^\infty$  is indexed by the sequence  $(q_n)_{n=1}^\infty$  of probability mass functions on  $\{0, 1, \dots, g\}$  with  $p_n \leftrightarrow q_n$ ,  $\forall n$ .

For each  $n$ , let  $p_n, \rho_n$  be as above, and let  $\phi = (\phi_n)_{n=1}^\infty = (p_n, \rho_n)_{n=1}^\infty$  be the fixed forecast. Clearly,  $\phi$  is parametrized by (1) the bounded sequence  $(h_n)_{n=1}^\infty$  of non-negative real numbers, and (2) the sequence  $(q_n)_{n=1}^\infty$  of probability mass functions on the set  $\{0, 1, \dots, g\}$ .

From the definition of the transition structure, we see that for each  $k \in K$ , the cumulative distribution functions  $G_n(0)$  and  $G_n(1)$  are given by

$$G_n(0; x) = \begin{cases} 0, & \text{if } x < 0, \\ \sum_{j=0}^{\lfloor x \rfloor} q_n(j), & \text{if } 0 \leq x < g+1, \\ 1, & \text{if } x \geq g+1, \end{cases}$$

and

$$G_n(1; x) = \begin{cases} 0, & \text{if } x < g+1, \\ 1, & \text{if } x \geq g+1, \end{cases}$$

from which it follows that the  $G_n$  are increasing in  $k \in \{0, 1\}$  (Assumption 4). Note that every forecast of the form of  $\phi$  is dominated by  $0 < \alpha < 1$  (Assumption 5), since the  $\rho_n$  are bounded.

We next define  $F$  and  $\Phi$ . Let  $F$  be the set of all  $(p, \rho)$  such that  $p$  is of the form of  $p_n$  above for some probability mass function  $q$  ( $p \leftrightarrow q$ ),  $\rho$  is of the form of  $\rho_n$  above for some  $h \geq 0$  ( $\rho \leftrightarrow h$ ), and  $\Phi$  is defined as in section 2. We think of  $\Phi$  as the ‘‘asset selling’’ class of forecasts.

Next, in order to define the *homogeneous* forecasts  $\bar{\psi}$  and  $\underline{\psi}$  in  $\Phi$ , define the transition probabilities  $\bar{p}_n = \bar{p}$  and  $\underline{p}_n = \underline{p}$  to be of the same form as the  $p_n$ ,  $\forall n$ , with the corresponding probability mass functions  $\bar{q}_n = \bar{q}$  and  $\underline{q}_n = \underline{q}$  given by

$$\bar{q}(j) = \begin{cases} 1, & \text{if } j = 0, \\ 0, & \text{if } 1 \leq j \leq g, \end{cases}$$

and

$$\underline{q}(j) = \begin{cases} 0, & \text{if } 0 \leq j \leq g-1, \\ 1, & \text{if } j = g, \end{cases}$$

for each  $n$  in  $\mathbb{N}$ . Then  $\bar{p} \leftrightarrow \bar{q}$ , i.e.,

$$\bar{p}_n(i, j; k) = \bar{p}(i, j; k) = \begin{cases} 1, & \text{if } j = 0, \quad k = 0; \text{ or } j = g+1, \quad k = 1, \\ 0, & \text{if } 1 \leq j \leq g+1, \quad k = 0; \text{ or } 0 \leq j \leq g, \quad k = 1, \end{cases}$$



and  $\underline{p} \leftrightarrow \underline{q}$ , i.e.,

$$\underline{p}_n(i, j; k) = \underline{p}(i, j; k) = \begin{cases} 0, & \text{if } j \neq g, k = 0; \text{ or } 0 \leq j \leq g, k = 1, \\ 1, & \text{if } j = g, k = 0; \text{ or } j = g + 1, k = 1. \end{cases}$$

Also define the reward functions  $\bar{\rho}_n = \bar{\rho}$  and  $\underline{\rho}_n = \underline{\rho}$  to be of the same form as the  $\rho_n$ , with  $\bar{h}_n = 1$  and  $\underline{h}_n = 0$ , for all  $n$ . Then  $(\underline{p}, \underline{\rho}), (\bar{p}, \bar{\rho}) \in F$  and the resulting homogeneous forecasts  $\bar{\psi} = (\bar{p}, \bar{\rho})_{n=1}^\infty$  and  $\underline{\psi} = (\underline{p}, \underline{\rho})_{n=1}^\infty$  clearly belong to  $\Phi$ .

Next, consider the homogeneous infinite horizon problem  $\langle \bar{\psi} \rangle$ , as in section 2, with associated  $\bar{w}$  and  $\Delta \bar{w}$ . In each period  $n$ , if decision  $k = 0$  is chosen, no matter what the offer  $0 \leq i \leq g$  is, the process will transition to state  $j = 0$  in the next period. In this case, future offers and rewards are zero. Thus, it is clearly optimal to accept *any* offer in the first period, since it is worse to pay holding costs and suffer the effects of discounting, only to be faced with zero offers in the future. Thus,

$$\bar{w}(i) = v(\bar{\psi}; i) = \begin{cases} i, & \text{if } 0 \leq i \leq g, \\ 0, & \text{if } i = g + 1, \end{cases}$$

which implies that

$$\Delta \bar{w}(i) = \Delta v(\bar{\psi}; i) = \begin{cases} 1, & \text{if } 1 \leq i \leq g, \\ -g, & \text{if } i = g + 1. \end{cases}$$

For  $(p, \rho) \in F$ , consider the infinite horizon problem corresponding to the forecast

$$\bar{\psi}(p, \rho) = ((p, \rho), (\bar{p}, \bar{\rho}), (\bar{p}, \bar{\rho}), \dots).$$

Let  $v(\bar{\psi}(p, \rho); \cdot)$  and  $\Delta v(\bar{\psi}(p, \rho); \cdot)$  be as in section 2 for the forecast  $\bar{\psi}(p, \rho)$ , where  $p \leftrightarrow q$  and  $\rho \leftrightarrow h$ . For convenience, let  $\hat{q} = \sum_{i=0}^{g+1} iq(i)$ , so that  $\hat{q} < g$  and

$$\alpha \hat{q} - h = f_1(\bar{\psi}(p, \rho); 0, 0) = f_1(\bar{\psi}(p, \rho); i, 0), \quad \forall 0 \leq i \leq g.$$

Then we may show that

$$\bar{w}(p, \rho; i) = v(\bar{\psi}(p, \rho); i) = \begin{cases} \alpha \hat{q} - h, & \text{if } 1 \leq i \leq \alpha \hat{q} - h, \\ i, & \text{if } \alpha \hat{q} - h < i \leq g, \\ 0, & \text{if } i = g + 1, \end{cases}$$

from which it follows that

$$\Delta \bar{w}(p, \rho; i) = \Delta v(\bar{\psi}(p, \rho); i) \leq \begin{cases} 1, & \text{if } 1 \leq i \leq g, \\ -g, & \text{if } i = g + 1. \end{cases}$$

Consequently,  $\Delta \bar{w}(p, \rho; i) \leq \Delta \bar{w}(i)$ ,  $\forall i = 1, \dots, g + 1$ , as required.

Analogously, starting with problem  $\langle \underline{\psi} \rangle$ , we can show that

$$\underline{w}(i) = v(\underline{\psi}; i) = \begin{cases} \alpha g, & \text{if } 0 \leq i \leq \alpha g, \\ i, & \text{if } \alpha g < i \leq g, \\ 0, & \text{if } i = g + 1, \end{cases}$$

which implies that  $\Delta \underline{w}(p, \rho; i) \geq \Delta \underline{w}(i)$ ,  $\forall i = 1, \dots, g + 1$  (Assumption 6). Assumption 7 is automatically satisfied since each  $D_i$  contains only two elements. Thus, since all of our assumptions hold for the asset selling problem, we may conclude all the results established in the previous sections.

Now fix  $n \in \mathbb{N}$ . Then  $f_n(g + 1, 1) = 0$ ,  $f_n(i, 1) = i$ ,  $\forall 0 \leq i \leq g$ , and, if  $0 \leq i \leq i' \leq g$ , then  $f_n(i', k) \geq f_n(i, k)$ , for all  $k$ .

We may define  $c_n = f_n(i, 0) = f_n(0, 0)$ , since it is independent of  $0 \leq i \leq g$ . Necessarily,

$$v_n(i) = \max_{k \in D_i} f_n(i, k) = \max\{c_n, i\} \geq i,$$

so that, in particular,

$$v(i) = v_1(i) = \begin{cases} c_1, & \text{if } 0 \leq i \leq c_1, \\ i, & \text{if } c_1 < i \leq g, \\ 0, & \text{if } i = g + 1, \end{cases}$$

Also define  $S_n^*$  to be the set of acceptable offers in period  $n$ , i.e.,

$$S_n^* = \{i \in S : f_n(i, 1) \geq f_n(i, 0)\} = \{i \in S : i \geq c_n\}.$$

Since  $f_n(g, 1) = g \geq \alpha g \geq f_n(g, 0)$ , we see that  $g \in S_n^*$ . Define  $i_n^*$  to be the smallest element (*threshold*) of  $S_n^*$ , so that  $0 \leq i_n^* \leq g$ . (The threshold is the smallest acceptable offer in the period.) In fact,  $S_n^* = \{i_n^*, \dots, g\}$ , since  $f_n(\cdot, k)$  is increasing in  $i$ . Consequently,

$$K_n(i) = \begin{cases} \{0\}, & \text{if } 0 \leq i < i_n^*, \\ \{0, 1\}, & \text{if } i = i_n^* = c_n, \\ \{1\}, & \text{if } i_n^* < i \leq g + 1. \end{cases}$$

Moreover, optimal policies are *threshold policies*, i.e.,

$$\Omega_n = \{(0, 0, \dots, 0, 1, 1, \dots, 1, 1)\},$$

for  $n$  such that  $i_n^* \neq c_n$ , with the first  $k = 1$  in the  $i_n^*$ -th position, and

$$\Omega_n = \{(0, \dots, 0, 1, 1, \dots, 1, 1), (0, \dots, 0, 0, 1, \dots, 1, 1)\},$$

for  $n$  such that  $i_n^* = c_n$ , with the first  $k = 1$  in the  $i_n^*$ -th or the  $(i_n^* + 1)$ -th position. Thus,  $\Omega_n$  is a singleton if  $i_n^* \neq c_n$ ; otherwise, it has two elements.

Since Assumptions 1-7 hold, we have

**7.1 Theorem.** *Consider the asset selling problem class  $\Phi$  with well-posed instance  $\phi \in \Phi$ . There exists a forecast horizon for the associated problem  $\langle \phi \rangle$  which may be determined in finite time by the Forecast Horizon Algorithm.*

**Proof.** This follows from Theorem 6.7. □

We have from Theorem 7.1 that an optimal initial policy for every well-posed instance of the time-varying asset selling problem can be computed with a finite number of periods of forecasted data.

## REFERENCES

1. Alden, J. M. and Smith, R. L., *Rolling Horizon Procedures in Nonhomogeneous Markov Decision Processes*, Operations Research **40** (1984), S183-S194.
2. Altman, E. and Koole, G., *On submodular value functions of dynamic programming*, Technical Report Number 2658, INRIA, France (1995), 1-23.
3. Bean, J. C., Birge, J. R., and Smith, R. L., *Aggregation in Dynamic Programming*, Operations Research **35** (1987), 215-220.
4. Bean, J. C. and Smith, R. L., *Conditions for the Existence of Planning Horizons*, Mathematics of Operations Research **9** (1984), 391-401.
5. Bean, J. C. and Smith, R. L., *Conditions for the Discovery of Solution Horizons*, Mathematical Programming **59** (1993), 215-229.
6. Bean, J. C., Smith, R. L., and Lasserre, J. B., *Denumerable State Nonhomogeneous Markov Decision Processes*, Journal of Mathematical Analysis and Applications **153** (1990), 64-77.
7. Bes, C. and Sethi, S., *Concepts of Forecast and Decision Horizons: Applications to Dynamic Stochastic Optimization Problems*, Mathematics of Operations Research **13** (1988), 295-310.

8. Bes, C. and Lasserre, J.B., *An On-Line Procedure in Discounted Infinite-Horizon Stochastic Optimal Control*, Journal of Optimization Theory and Applications **50** (1986), 61-67.
9. Bhaskaran, S. and Sethi, S., *Conditions for the existence of Decision Horizons for Discounted Problems in Stochastic Environment: A Note*, Operations Research Letters **4** (1985), 61-65.
10. Billingsley, Patrick., *Probability and Measure, 3rd edition.*, John Wiley & Sons, New York, 1995.
11. Chand, S., Hsu, V., and Sethi, S., *Forecast, Solution, and Rolling Horizons in Operations Management Problems: A Classified Bibliography*, Manufacturing & Service Operations Management- **4** (2002), 25-43.
12. Cheevaprawatdomrong, T. and Smith, R. L., *A Paradox in Equipment Replacement under Technological Improvement*, Operations Research Letters **31** (2003), 77-82.
13. Cheevaprawatdomrong, T. and Smith, R. L., *Infinite Horizon Production Scheduling in Time-varying Systems under Stochastic Demand*, Operations Research **52** (2004), 105-115.
14. Cheevaprawatdomrong, T., Schochetman, I. E., Smith, R. L. and Garcia, A., *Solution and Forecast Horizons for Infinite Horizon Nonhomogeneous Markov Decision Processes*, Technical Report 05-10, Department of Industrial and Operations Engineering, University of Michigan, (February 4, 2005), 1-32.
15. Cruz-Suarez, D., Montes-de-Oca, R. and Salem-Silva, F., *Conditions for the uniqueness of optimal policies of discounted Markov decision processes*, Mathematical Methods of Operations Research **60** (2004), 415-436.
16. Garcia, S. and Smith, R. L., *Solving Nonstationary Infinite Horizon Dynamic Optimization Problems*, Journal of Mathematical Analysis and Applications . **244** (2000), 304-317.
17. Hernandez-Lerma, O., *Adaptive Markov Control Processes*, Springer, New York, 1989.
18. Hernandez-Lerma, O. and Lasserre, J. B., *A Forecast Horizon and a Stopping Rule for General Markov Decision Processes*, Journal of Mathematical Analysis and Applications **132** (1988), 388-400.
19. Heyman, D. P. and Sobel, M. J., *Stochastic Models in Operations Research*, McGraw-Hill, New York, 1982.
20. Hopp W. J., *Identifying Forecast Horizons in Nonhomogeneous Markov Decision Processes*, Operations Research **37** (1989), 339-343.
21. Hopp W. J., Bean, J. C., and Smith, R. L., *A New Optimality Criterion for Nonhomogeneous Markov Decision Processes*, Operations Research **35** (1987), 875-883.
22. Horn, R. A. and Johnson, C. R., *Matrix Analysis*, Cambridge University Press, Cambridge, 1988.
23. Lee, C. and Denardo, E., *Rolling Planning Horizons: Error Bounds for the Dynamic Lot Size Model*, Mathematics of Operations Research **11** (1986), 423-432.
24. Munkres, J. R., *Topology: A First Course*, Prentice Hall, New Jersey, 1975.
25. Ross, S. M., *Introduction to Stochastic Dynamic Programming*, Academic Press, New York, 1983.
26. Ryan S. M. and Bean, J. C., *Degeneracy in Infinite Horizon Optimization*, Mathematical Programming **43** (1989), 305-316.
27. Ryan S. M., Bean, J. C. and Smith, R. L., *A Tie-breaking Algorithm for Discrete Infinite Horizon Optimization*, Operations Research **40** (1992), S117-S126.
28. Schochetman, E. and Smith, R., *Finite Dimensional Approximation in Infinite Dimensional Mathematical Programming*, Mathematical Programming **54** (1992), 307-333.
29. Smith, R. L. and Zhang, R. Q., *Infinite Horizon Production Planning in Time Varying Systems with Convex Production and Inventory Costs*, Management Science **44** (1998), 1313-1320.
30. Sobel, M. J., *Myopic Solution of Markov Decision Processes and Stochastic Games*, Operations Research **29** (1981), 995-1009.
31. Topkis, D. M., *Supermodularity and Complementarity*, Princeton University Press, New Jersey, 1998.