Pioneer Journal of Advances in Applied Mathematics
Volume 19, Numbers 1-2, 2017, Pages 15-41
This paper is available online at http://www.pspchv.com/content_PJAAM.html

AN ALGORITHM FOR A DISCRETE DISCOUNTED STOCHASTIC PROGRAM WITH INCREASING LINKAGES BETWEEN STATE VARIABLES

MILAN HORNIAČEK

Comenius University in Bratislava
Faculty of Social and Economic Sciences
Mlynské luhy 4, SK-82105 Bratislava
Slovakia

Abstract

We develop an algorithm for solving an infinite horizon discrete time stochastic program with discounting of future single period payoffs and strictly increasing linkages between state variables. Expected value of each state variable over successors of the current state is strictly increasing in the current value of that variable, as well as in action affecting it taken at the current state. Uncertainty is captured by a Markov chain with countable, possibly infinite, state space. Sets of feasible actions are finite. Expected single period payoff at each state is strictly increasing in each state variable. A strategy assigns to each state a feasible action at it. An

Received June 6, 2017

2010 Mathematics Subject Classification: 90C15, 90C39, 90C40.

Keywords and phrases: stochastic program, Markov decision process, discounting, open-ended solution, strictly increasing linkages between state variables.

Research reported in this paper is financially supported by grant No. APVV-14-0020 of the Agency for Support of Research and Development of the Ministry of Education, Science, Research, and Sport of the Slovak Republic. Agency for Support of Research and Development did not play any role in design of the research or decision to submit its result for publication.

optimal strategy maximizes the sum of discounted expected single period payoffs. We develop a basic algorithm (and then its simplified version for special cases) that gives for each state the prescription of an optimal strategy for it, without specifying prescriptions for other states.

1. Introduction

Stochastic programs are widely used in decision-making. If they include development of random factors affecting outcomes of decisions, they usually have the form of a Markov decision process. (See, for example, [1], subchapter 1.5 for their characterization and Chapter 5 for examples of applications.) If there is not a final date of existence of an object of decision-making, they have infinite horizon, usually with discounting of future single period payoffs. Even in this case the set of states can be finite. (See [2] for an example.) In such a case a solution to the stochastic program can be obtained by solving a linear program (see, for example, [7]). Nevertheless, if decisions in future periods modify still prevailing effects of previous decisions, then states express cumulative effects of several decisions and the set of states is infinite. This includes, for example, optimization of any infinite horizon system with innovation, programs with decisions on advertising or investment by a monopolist, or models with use of durable goods. In this case solution to a stochastic program is open-ended. The whole solution can never be computed and recorded. For each state we have to compute the prescription of an optimal strategy for it, without computing prescriptions for following states. Such an approach can be useful and efficient also if the set of states is finite but large, especially if the decision-maker expects that in the future he will learn a more precise information about some parameters of the optimization program.

Model predictive control ([5], [6]) takes this approach. It replaces an infinite horizon stochastic program (or a stochastic program with a long finite time horizon) with a sequence of finite horizon programs (or programs with a shorter time horizon). Nevertheless, solutions obtained by it need not be ex ante optimal (i.e., optimal with respect to available information about random factors affecting the system in the future). Moreover, there can be some technical problems with its application. In its basic form, it assumes that finite horizon problems are time invariant ([5], p. 790). This assumption is violated, for example, if a firm minimizes some sort of costs or maximizes profit and prices or some parameters of inverse demand curve (that it cannot effect) change over time. Schildbach, and Morari [6] do not make such assumption but they assume that the resulting optimization problem would become convex if all uncertain variables were known and fixed ([6], Assumption 1, p. 543). This assumption is not satisfied if sets of feasible decisions are finite.

We take a different approach. For each state we compute the prescription of an optimal strategy (for the whole infinite horizon program) for it, without computing prescriptions for following states. This is possible because (besides discounting of future single period payoffs) we assume finite sets of feasible actions and there are strictly increasing linkages between state variables. Expected value of each state variable over successors of the current state is strictly increasing in the current value of that variable, as well as in action affecting it taken at the current state. Amortization of state variables is a special and, perhaps, the most important case of this. Amortization rate can be interpreted also as a probability of failure (or replacement for another reason) of an asset whose quantity is represented by a state variable.

The assumption that sets of feasible actions are finite is plausible. There is usually some smallest measurement unit that makes sense in a modelled environment (e.g., it does not make sense to consider smaller monetary units than cents, or in production of beverages there is no need to consider smaller volume units than milliliters).

Each state in the current period can be followed with a positive probability by a finite subset of the state space, which depends on the current state, an action taken at it, and random factors. Expected single period payoff at each state is strictly increasing in each state variable. Single period payoffs are uniformly bounded across all states. At each state, cost of taking each type of action (i.e., each component of a vector action) is strictly increasing in the value of this action and non-increasing in the taken value of any other type of action. The latter assumption is justified, for example, when cost of taking an action is additively separable in its components or all components of an action are obtained from one supplier who provides quantity discounts. We also assume that an increase in the value of one type of action at the current state does not reduce the expected value of maximal feasible action of any type in the following period (see part (a) of Assumption 2 in Section 2). The same holds for an increase in the value of a state variable (see part (b) of Assumption 2). This assumption is plausible if types of actions are expressed directly in terms of expenditures on them.

We compute, in a finite number of iterations, for each state the prescription of an optimal strategy for it, without computing prescriptions for other states. In each iteration, we solve a linear program with a finite number of variables and constraints. In iteration n we take into account only states that can occur with a positive probability in the first n periods. For each state that can occur with a positive probability in period $n+1$ but not in the first n periods, we replace the value of a strategy at it by zero.

Besides the basic algorithm (developed in Section 3) we give also (in Section 4) its simplified version that can be used under additional assumptions.

2. Stochastic Program

Throughout the paper, $\mathbb{N}$ denotes the set of positive integers and $\mathbb{R}$ is the set of real numbers. For $n \in \mathbb{N}$, we set $\mathbb{R}_+^n = [0, \infty)^n$. We endow each finite dimensional real vector space with the Euclidean topology and each infinite dimensional product of finite dimensional spaces with the product topology. For $n \in \mathbb{N}$, Δ_n is a simplex in $\mathbb{R}^n$. For a finite set A, $\#(A)$ is its cardinality. We use the following relations for vectors. For $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{a} \geq \mathbf{b}$ means that $a_j \geq b_j$ for each $j \in \{1, ..., n\}$, $\mathbf{a} > \mathbf{b}$ means that $\mathbf{a} \geq \mathbf{b}$ but $\mathbf{a} \neq \mathbf{b}$, and $\mathbf{a} \gg \mathbf{b}$ means that $a_j > b_j$ for each $j \in \{1, ..., n\}$. When a vector with specific numeric components is an argument of a function, we use only one pair of round brackets (e.g., we write $\eta(1, 1)$ instead of $\eta((1, 1))$).

The time horizon of the analyzed stochastic program is $\mathbb{N}$. The countable (possibly infinite) set of states is denoted by Ω. Each $\omega \in \Omega$ is an element of $\mathbb{R}^{m^*}$, $m^* \in \mathbb{N}$.

For state $\boldsymbol{\omega} \in \Omega$ a finite set of feasible actions at it is $X(\boldsymbol{\omega}) \in \mathbb{R}_+^{m^*}$. It has the form $X(\boldsymbol{\omega}) = \prod_{i=1}^{m^*} X_i(\boldsymbol{\omega})$. It has at least two elements and it includes the zero vector, interpreted as inaction. We set $X = \bigcup_{\boldsymbol{\omega} \in \Omega} X(\boldsymbol{\omega})$.

For $\boldsymbol{\omega} \in \Omega$, $\Omega(\boldsymbol{\omega})$ is the set of states that can occur with a positive probability in the following period when the state in the current period is $\boldsymbol{\omega}$. We assume that $\Omega(\boldsymbol{\omega})$ is finite and $\#(\Omega(\boldsymbol{\omega})) \geq 2$. For $\boldsymbol{\omega} \in \Omega$, function $\rho_{\boldsymbol{\omega}} : X(\boldsymbol{\omega}) \to \Delta_{\#(\Omega(\boldsymbol{\omega}))}$ assigns to $x \in X(\boldsymbol{\omega})$ a probability distribution on $\Omega(\boldsymbol{\omega})$. The symbol $\rho_{\boldsymbol{\omega}}(\mathbf{x})(\boldsymbol{\omega}')$ stands for the probability of occurrence of state $\boldsymbol{\omega}' \in \Omega(\boldsymbol{\omega})$ in the immediately following period when state in the current period is $\boldsymbol{\omega}$ and action $\mathbf{x}$ is taken at it.

Assumption 1. For each $x \in X$ taken in the current period and each $i \in \{1, \ldots, m^*\}$, the expected value of ω_i in the immediately following period is

(a) strictly increasing in x_i and non-decreasing in x_j, $j \in \{1, \ldots, m^*\} \backslash \{i\}$ and

(b) strictly increasing in the expected value of ω_i' in the current period and non-decreasing in the expected value of ω_j', $j \in \{1, \ldots, m^*\} \backslash \{i\}$ in the current period.

Assumption 2. (a) For each $\boldsymbol{\omega} \in \Omega$, each $(\mathbf{x}\ \mathbf{x}') \in (X(\omega))^2$ with $\mathbf{x} > \mathbf{x}'$, for each $\boldsymbol{\omega}' \in \Omega$ with $\rho_{\boldsymbol{\omega}}(\mathbf{x}')(\boldsymbol{\omega}') > 0$ there exists $\boldsymbol{\omega}'' > \boldsymbol{\omega}'$ such that $\rho_{\boldsymbol{\omega}}(\mathbf{x})(\boldsymbol{\omega}'') = \rho_{\boldsymbol{\omega}}(\mathbf{x}')(\boldsymbol{\omega}')$ and $X(\boldsymbol{\omega}') \subseteq X(\boldsymbol{\omega}'')$.

(b) For each $(\boldsymbol{\omega}, \boldsymbol{\omega}') \in \Omega^2$ with $\boldsymbol{\omega}' > \boldsymbol{\omega}$, each $\mathbf{x} \in X(\omega) \cap X(\boldsymbol{\omega}')$, for each $\tilde{\boldsymbol{\omega}} \in \Omega$ with $\rho_{\boldsymbol{\omega}}(\mathbf{x})(\tilde{\boldsymbol{\omega}}) > 0$ there exists $\tilde{\boldsymbol{\omega}}' > \tilde{\boldsymbol{\omega}}$ such that $\rho_{\boldsymbol{\omega}'}(\mathbf{x})(\tilde{\boldsymbol{\omega}}') = \rho_{\boldsymbol{\omega}}(\mathbf{x})(\tilde{\boldsymbol{\omega}})$ and $X(\tilde{\boldsymbol{\omega}}) \subseteq X(\tilde{\boldsymbol{\omega}}')$.

Assumption 2 is satisfied when (i) random factors affecting transition probabilities between states do not outweigh the effect of actions and the current state, and (ii) higher values of state variables do not prevent taking of actions that could be taken at lower values (although they can make taking of an action more costly).

Function

$$\pi : \{(\boldsymbol{\omega}, \mathbf{x}) \mid \boldsymbol{\omega} \in \Omega \text{ and } \mathbf{x} \in X(\boldsymbol{\omega})\} \to L,$$

where $L \subset \mathbb{R}$ is a nonempty compact set, assigns to $\boldsymbol{\omega} \in \Omega$ and $\mathbf{x} \in X(\boldsymbol{\omega})$ a single period payoff at state $\boldsymbol{\omega}$ when action $\mathbf{x}$ is taken at it.

Assumption 3. For each $(\boldsymbol{\omega}, \mathbf{x})$ from the domain of π

(a) $\pi(\boldsymbol{\omega}, \mathbf{x}) = \eta(\boldsymbol{\omega}) - \gamma(\mathbf{x})$, where $\eta(\boldsymbol{\omega})$ is a part of payoff at state $\boldsymbol{\omega}$ that is independent of x and $\gamma(\mathbf{x})$ is the cost of taking an action $\mathbf{x}$ independent of $\boldsymbol{\omega}$,

(b) function $\eta : \Omega \to L_1$ (where L_1 is a compact strict subset of L with $\max L_1 = \max L$ and $\min L_1 > \min L)$ is non-decreasing in ω_i for each $i \in \{1, ..., m^*\}$ at each $\boldsymbol{\omega} \in \Omega$ and strictly increasing in ω_i for each $i \in \{1, ..., m^*\}$ at each $\boldsymbol{\omega} \in \Omega$ with $\eta(\boldsymbol{\omega}) > 0$, and

(c) function $\gamma : X \to [0, \min L_1 - \min L]$ is strictly increasing in x_i for each $i \in \{1, ..., m^*\}$,

(d) for each $i \in \{1, ..., m^*\}$ and for each $(\mathbf{x}, \mathbf{x}') \in X$ with $x_i > x_i'$ and $x_j = x_j'$ for each $j \in \{1, ..., m^*\}\setminus\{i\}$, $\gamma(\mathbf{x}) - \gamma(\mathbf{x}')$ is non-increasing in $(x_j)_{j \in \{1, ..., m^*\}\setminus\{i\}}$, and

(e) $\gamma(\mathbf{0}) = 0$.

Part (d) of Assumption 3 can be justified by quantity discounts if materials for each action are supplied by the same firm. Also if actions are expressed in terms of expenditures, it can be justified by borrowing the whole amount from the same bank, or by covering expenditures from own financial sources.

Single period payoffs are discounted by discount factor $\delta \in (0, 1)$, without discounting the current period payoff.

We restrict attention to Markov strategies. In the analyzed stochastic program only current state is payoff relevant. (The set of feasible actions and their payoff consequences depend only on the current state.) Therefore, a Markov strategy (henceforth, “strategy”) is a function that assigns to each $\boldsymbol{\omega} \in \Omega$ an element of $X(\boldsymbol{\omega})$. We denote the set of strategies by Q and its generic element by q.

Function $\mu : \Omega^2 \times \mathbb{N} \times Q \to [0, 1]$ assigns to $(\boldsymbol{\omega}', \boldsymbol{\omega}) \in \Omega^2$, $t \in \mathbb{N}$, and $q \in Q$ the probability that state $\boldsymbol{\omega}$ occurs in period t under strategy q when the initial state (occurring in period one) is $\boldsymbol{\omega}'$. It is defined recursively by

$$\mu(\boldsymbol{\omega}', \boldsymbol{\omega}', 1, q) = 1, \tag{1}$$

$$\mu(\boldsymbol{\omega}', \boldsymbol{\omega}, 1, q) = 0, \quad \forall \boldsymbol{\omega} \in \Omega \setminus \{\boldsymbol{\omega}'\}, \tag{2}$$

$$\mu(\boldsymbol{\omega}', \boldsymbol{\omega}, t, q) = \sum_{\boldsymbol{\omega}'' \in \Omega : \mu(\boldsymbol{\omega}', \boldsymbol{\omega}'', t-1, q) > 0} \mu(\boldsymbol{\omega}', \boldsymbol{\omega}'', t-1, q) \rho_{\boldsymbol{\omega}''}(\mathbf{q}(\boldsymbol{\omega}''))(\boldsymbol{\omega}),$$

$$\forall \boldsymbol{\omega} \in \Omega, \ \forall t \in \mathbb{N} \setminus \{1\}. \tag{3}$$

For $k \in \mathbb{N}$ and $n \in \mathbb{N}$ we let

$$\Omega^{(k)}(\boldsymbol{\omega}') = \{\boldsymbol{\omega} \in \Omega \mid \exists\, q \in Q \text{ such that } \mu(\boldsymbol{\omega}', \boldsymbol{\omega}, k, q) > 0\}$$

and

$$\overline{Q}^{(n)}(\boldsymbol{\omega}') = \bigcup_{k=1}^{n} \Omega^{(k)}(\boldsymbol{\omega}').$$

Thus, $\Omega^{(k)}(\boldsymbol{\omega}')$ is the set of states that can occur under some strategy in period k when the initial state is $\boldsymbol{\omega}'$ and $\overline{Q}^{(n)}(\boldsymbol{\omega}')$ is the set of states that can occur under some strategy in the first n periods when the initial state is $\boldsymbol{\omega}'$. Recall that for each $\boldsymbol{\omega}' \in \Omega$, $\Omega(\boldsymbol{\omega}') = \Omega^{(1)}(\boldsymbol{\omega}')$ is finite. Therefore, sets $\Omega^{(k)}(\boldsymbol{\omega}')$ and $\overline{Q}^{(n)}(\boldsymbol{\omega}')$ are finite.

Since Ω is countable, we number its elements by positive integers in such a way that $\boldsymbol{\omega}(1)$ is the state in period 1 and for each $k > 1$

$$\Omega^{(k)}(\boldsymbol{\omega}(1)) = \left\{ \begin{array}{c} \sum_{j=1}^{k-1} \#\,(\Omega^{(j)}(\boldsymbol{\omega}(1))) + 1, \ldots, \\ \sum_{j=1}^{k-1} \#\,(\Omega^{(j)}(\boldsymbol{\omega}(1))) + \#\,(\Omega^{(k)}(\boldsymbol{\omega}(1))) \end{array} \right\}.$$

The analyzed stochastic program maximizes the sum of discounted expected single period payoffs subject to constraints stemming from sets $X(\boldsymbol{\omega})$, $\boldsymbol{\omega} \in \Omega$. For initial state $\boldsymbol{\omega}'$ it has the form

$$\max \sum_{t \in \mathbb{N}} \left(\delta^{t-1} \sum_{\boldsymbol{\omega} \in \Omega^{(t)}(\boldsymbol{\omega}')} \mu(\boldsymbol{\omega}', \boldsymbol{\omega}, t, q) \pi(\boldsymbol{\omega}, q(\boldsymbol{\omega})) \right) \tag{4}$$

subject to

$$q(\boldsymbol{\omega}) \in X(\boldsymbol{\omega}), \quad \forall \boldsymbol{\omega} \in \Omega. \tag{5}$$

Since future single period payoffs are discounted and the range of function π is bounded, the objective function (4) is well-defined.

Since Q is a countable product of finite (and, hence, compact) subsets of the Euclidean space, it is compact by Tychonoff's theorem. Since $X(\boldsymbol{\omega})$ is finite for each $\boldsymbol{\omega} \in \Omega$, function π is (trivially) continuous in $\mathbf{x}$ and functions $\rho_{\boldsymbol{\omega}}$, $\boldsymbol{\omega} \in \Omega$, are (trivially) continuous for each $\boldsymbol{\omega} \in \Omega$. Continuity of functions $\rho_{\boldsymbol{\omega}}$, $\boldsymbol{\omega} \in \Omega$, implies continuity of function μ in q. From this, continuity of and bounded range of function π, and discounting of future single period payoffs it follows that the objective function (4) continuous. Thus, program (4)-(5) is a maximization of a continuous function on a non-empty compact set, so it has an optimal solution.

For $i \in \{1, ..., m^*\}$, denote by $\overline{\omega}_i$ an arbitrarily chosen, but fixed, positive upper bound on ω_i and let $\overline{\boldsymbol{\omega}} = (\overline{\omega}_i)_{i \in \{1, ..., m^*\}}$.

We illustrate properties of the analyzed stochastic program described in this section by the following example.

Example 1. Consider a monopolist producing a single good, carrying out advertising that shifts upwards his inverse demand curve, and building a stock of capital that affects his marginal costs - an increase in the stock of capital decreases his marginal costs. We have $m^* = 2$, $\omega_1 \geq 0$ is the stock of goodwill created by advertising expenditures and subject to amortization rate $\alpha_1 = 0.1$, $\omega_2 \geq 0$ is the stock of capital subject to amortization rate $\alpha_2 = 0.7$. (The modelling of stock of goodwill created by advertising expenditures was used by [4] and the effect of stock of capital on marginal costs was analyzed by [3]. Nevertheless, both models are deterministic.) $X_1(\boldsymbol{\omega}) = \{0, 0.5\}$ is the set of feasible advertising expenditures and $X_2(\boldsymbol{\omega}) = \{0, 0.5, 1.5\}$ is the set of feasible investments into capital at each state $\boldsymbol{\omega}$. (We make sets of feasible actions "small" and "sparse" and use rather extreme values of depreciation rates in order to shorten computational Example 2.) For each $i \in \{1, 2\}$, the set of values of state i that can occur in the immediately following period when the current value of state i is ω_i' and taken type i of action at it is x_i, is

$$\Omega_i(\omega_i', x_i) = \begin{Bmatrix} \max\{(1-\alpha_i)\omega_i' + x_i - 0.1, 0\} \\ \\ (1-\alpha_i)\omega_i' + x_i \\ \\ (1-\alpha_i)\omega_i' + x_i + 0.1 \end{Bmatrix}$$

and probability distribution on it is $(0.25, 0.5, 0.25)$. Probability distributions on $\Omega_1(\omega_1', x_1)$ and $\Omega_2(\omega_2', x_2)$ are independent. Values of function $\rho_{\boldsymbol{\omega}'}$ are derived from them. Clearly, Assumptions 1 and 2 are satisfied.

Inverse demand function at state $\boldsymbol{\omega} \in \Omega$, $P_{\boldsymbol{\omega}} : [0, \omega_1] \to [0, \omega_1]$ has the form $P_{\boldsymbol{\omega}}(y) = \omega_1 - y$, where y is monopolist's output. ($P_{\boldsymbol{\omega}}(y)$ is the unit price of monopolist's good at which demand is equal to y. Of course, $y \in [0, \omega_1]$.) Costs function at state $\boldsymbol{\omega} \in \Omega$, $c_{\boldsymbol{\omega}} : [0, \omega_2] \to \mathbb{R}_+$, has the form $\gamma_{\boldsymbol{\omega}}(y) = \dfrac{y}{0.001 + \omega_2}$. (It expresses variable cost of production at state $\boldsymbol{\omega}$. Investments into capital are taken into account in definition of function π.)

We compute (non-attainable) upper bounds on state variables by assuming that the maximal actions affecting them are taken in each period since minus infinity up to the period immediately preceding the current one and the maximal random factor occurs in each period since minus infinity up to the current period. The stock of goodwill is bounded from above by

$$\overline{\omega}_1 = (0.5 + 0.1)\sum_{t=0}^{\infty}(1-\alpha_1)^t = \frac{0.6}{0.1} = 6$$

and stock of capital is bounded from above by

$$\overline{\omega}_2 = (0.5 + 0.1)\sum_{t=0}^{\infty}(1-\alpha_2)^t = \frac{1.6}{0.7} = 2.2857.$$

For each $\boldsymbol{\omega} \in \Omega$ function π has the form

$$\pi(\boldsymbol{\omega}, \mathbf{x}) = \left(\max\left\{\frac{\omega_1}{2} - \frac{1}{2(0.001 + \omega_2)}, 0\right\}\right)^2 - x_1 - x_2.$$

(If $\frac{\omega_1}{2} \geq \frac{1}{2(0.001+\omega_2)}$, the output maximizing profit from production at state $\boldsymbol{\omega} \in \Omega$ is $\frac{\omega_1}{2} - \frac{1}{2(0.001+\omega_2)}$ and maximized profit equals $\left(\frac{\omega_1}{2} - \frac{1}{2(0.001+\omega_2)}\right)^2$. If $\frac{\omega_1}{2} < \frac{1}{2(0.001+\omega_2)}$, the output maximizing profit from production at state $\boldsymbol{\omega} \in \Omega$ is 0, giving zero profit.) That is,

$$\eta(\boldsymbol{\omega}) = \left(\max\left\{\frac{\omega_1}{2} - \frac{1}{2(0.001+\omega_2)},\, 0\right\}\right)^2, \text{ and } \gamma(x_1, x_2) = x_1 + x_2.$$

Clearly, π satisfies Assumption 3.

We have $L = [-2, 7.7359]$, $L_1 = [0, 7.7359]$.

Discount factor is $\delta = 0.9$.

The initial state is (1, 2).

The following claim plays an important role in development of the algorithm.

Claim 1 (a) *If for some* $\boldsymbol{\omega}' \in \Omega$ *program* (4)-(5) *has* (*at least*) *two optimal strategies, prescribing actions* $\mathbf{x} \in X(\boldsymbol{\omega}')$ *and* $\mathbf{x}' \in X(\boldsymbol{\omega}') \setminus \{\mathbf{x}\}$ *at* $\boldsymbol{\omega}'$, *then either* $\mathbf{x} > \mathbf{x}'$ or $\mathbf{x}' > \mathbf{x}$. (b) *If at each state cost of taking actions at it is generic, then program* (4)-(5) *has the unique optimal strategy.*

Proof. (a) Suppose that

$$I_1 = \{i \in \{1, \ldots, m^*\} \mid x_i > x_i'\} \neq \emptyset \text{ and } I_2 = \{i \in \{1, \ldots, m^*\} \mid x_i < x_i'\} \neq \emptyset.$$

Take $\mathbf{x}''$ satisfying $x_i'' = x_i$ for each $i \in I_1$ and $x_i'' = x_i'$ for each $i \in I_2$. Since $X(\boldsymbol{\omega}')$ has the product structure, $\mathbf{x}'' \in X(\boldsymbol{\omega}')$. Using Assumption 1, Assumption 2, and parts (b) and (d) of Assumption 3, $\mathbf{x}''$ gives a higher sum of discounted expected single period payoffs than $\mathbf{x}$ and $\mathbf{x}'$. (With respect to Assumption 2, the same actions can be used since period two at states occurring with the same probability, but with higher components, when $\mathbf{x}''$ is used at $\boldsymbol{\omega}'$ as when $\mathbf{x}'$ is used at $\boldsymbol{\omega}'$. The same applies for comparison of $\mathbf{x}''$ and $\mathbf{x}$.) This contradiction with the assumption that both $\mathbf{x}$ and $\mathbf{x}'$ are prescriptions of optimal strategies of program (4)-(5) at $\boldsymbol{\omega}'$ shows that either $I_1 = \emptyset$ or $I_2 = \emptyset$, i.e., $\mathbf{x}' > \mathbf{x}$ or $\mathbf{x} > \mathbf{x}'$.

(b) Take (arbitrary) $\omega' \in \Omega$. Let

$$\xi(\omega') = (\gamma(\mathbf{x}))_{\mathbf{x} \in X(\omega')} \in [0, \min L_1 - \min L]^{\#(X(\omega'))}$$

be a vector of cost of actions that are feasible at state ω'. (We hold all other parameters of program (4)-(5) fixed.) If for some $\xi(\omega')$ each optimal strategy of program (4)-(5) prescribes the same action at ω', then there exists a neighborhood of $\xi(\omega')$ in $[0, \min L_1 - \min L]^{\#(X(\omega'))}$ such that for each $\xi'(\omega')$ from this neighborhood each optimal strategy of program (4)-(5) prescribes the same action at ω'. Thus, the set of vectors $\xi(\omega')$, for which each optimal strategy of program (4)-(5) prescribes the same action at ω', is an open subset of $[0, \min L_1 - \min L]^{\#(X(\omega'))}$. If for some $\xi(\omega')$ program (4)-(5) has two optimal strategies prescribing $\mathbf{x}$ and $\mathbf{x}' \neq \mathbf{x}$ at ω', then by part (a) of Claim 1 either $\mathbf{x}' > \mathbf{x}$ or $\mathbf{x} > \mathbf{x}'$. (The argument is analogous if there are more than two optimal strategies with different prescriptions at ω'.) Assume (without loss of generality) that $\mathbf{x} > \mathbf{x}'$. Then (using Assumption 1 and parts (b) and (c) of Assumption 3) higher cost of taking action $\mathbf{x}$ is exactly outweighed by higher state variables and higher single period payoffs generated by them in the following periods. Therefore, after an arbitrarily small decrease in $\gamma(\mathbf{x})$ only strategy prescribing $\mathbf{x}$ at ω' will be optimal. Thus, an arbitrarily small neighborhood of $\xi(\omega')$ in $[0, \min L_1 - \min L]^{\#(X(\omega'))}$ contains $\xi'(\omega')$ for which there is the unique prescription of an optimal strategy of program (4)-(5) at ω'. That is, the set of vectors $\xi(\omega')$, for which each optimal strategy of program (4)-(5) prescribes the same action at ω', is a dense subset of $[0, \min L_1 - \min L]^{\#(X(\omega'))}$. Thus, if cost of taking actions at ω' is generic, each optimal strategy of program (4)-(5) prescribes the same action at ω'. If this holds for each $\omega' \in \Omega$, program (4)-(5) has the unique optimal strategy. □

Let ε^* be the value of the smallest monetary unit. That is, if single period payoffs are expressed in basic monetary units and the basic monetary unit contains 100 smallest monetary units, then $\varepsilon^* = 0.01$.

Remark 1. In a finite number of computations it is not possible to verify whether arbitrarily small change in cost of some action at $\omega' \in \Omega$ will eliminate an alternative optimal prescription at $\boldsymbol{\omega}'$. Moreover, the existence of the smallest positive real number that can be represented by a digital computer prevents such verification. Therefore, in the algorithm given in the following section we will verify only whether the decrease in cost of some action at $\boldsymbol{\omega}' \in \Omega$ equal to ε^* will eliminate an alternative optimal prescription at $\boldsymbol{\omega}'$.

3. Basic Algorithm

For $\boldsymbol{\omega} \in \Omega$ let $v_{\boldsymbol{\omega}}(q)$ be the value of a strategy q at state $\boldsymbol{\omega}$ (i.e., the sum of discounted expected single period payoffs from a strategy q when the initial state is $\boldsymbol{\omega}$). That is,

$$v_{\boldsymbol{\omega}}(q) = \pi(\boldsymbol{\omega}, \mathbf{q}(\boldsymbol{\omega})) + \delta \sum_{\boldsymbol{\omega}' \in \Omega^{(2)}(\boldsymbol{\omega})} \rho_{\boldsymbol{\omega}}(\mathbf{q}(\boldsymbol{\omega}))(\boldsymbol{\omega}') v_{\boldsymbol{\omega}'}(q). \tag{6}$$

Applying Bellman's principle of optimality, a strategy $q^* \in Q$ is an optimal solution of program (4)-(5) if and only if

$$v_{\boldsymbol{\omega}'}(q^*) = \max_{\mathbf{x} \in X(\boldsymbol{\omega}')} \pi(\boldsymbol{\omega}', \mathbf{x}) + \delta \sum_{\boldsymbol{\omega} \in \Omega^{(2)}(\boldsymbol{\omega}')} \rho_{\boldsymbol{\omega}'}(\mathbf{x})(\boldsymbol{\omega}) v_{\boldsymbol{\omega}}(q^*), \quad \forall \boldsymbol{\omega}' \in \Omega. \tag{7}$$

Take (arbitrary) state $\boldsymbol{\omega}' \in \Omega$. In order to compute prescription of each strategy, which gives the sum of discounted expected single period payoffs within distance less than half of the smallest monetary unit from its maximum, for state $\boldsymbol{\omega}'$ (henceforth, only "prescription at $\boldsymbol{\omega}'$"), we proceed in a finite number of iterations. In iteration $n \in \mathbb{N} \setminus \{1\}$ we consider only actions in $X^{(n)}(\boldsymbol{\omega}') \subseteq X(\boldsymbol{\omega}')$. We set $X^{(2)}(\boldsymbol{\omega}') = X(\boldsymbol{\omega}')$; sets $X^{(n)}(\boldsymbol{\omega}')$ for $n > 2$ are defined below. Also, in iteration n we consider only states in $\overline{Q}^{(n)}(\boldsymbol{\omega}')$ and set $v_{\boldsymbol{\omega}} = 0$ for each $\boldsymbol{\omega} \in \Omega \setminus \overline{Q}^{(n)}(\boldsymbol{\omega}')$. Thus, (omitting q^*) in iteration n we replace (7) by

$$v_{\boldsymbol{\omega}'} = \max_{\mathbf{x} \in X^{(n)}(\boldsymbol{\omega}')} \pi(\boldsymbol{\omega}', \mathbf{x}) + \delta \sum_{\boldsymbol{\omega} \in \Omega^{(2)}(\boldsymbol{\omega}')} \rho_{\boldsymbol{\omega}'}(\mathbf{x})(\boldsymbol{\omega}) v_{\boldsymbol{\omega}}. \tag{8}$$

$$v_{\boldsymbol{\omega}''} = \max_{\mathbf{x} \in X(\boldsymbol{\omega}'')} \pi(\boldsymbol{\omega}'', \mathbf{x}) + \delta \sum_{\boldsymbol{\omega} \in \Omega^{(2)}(\boldsymbol{\omega}'')} \rho_{\boldsymbol{\omega}''}(\mathbf{x})(\boldsymbol{\omega}) v_{\boldsymbol{\omega}}, \quad \forall \boldsymbol{\omega}'' \in \overline{Q}^{(n)}(\boldsymbol{\omega}') \setminus \{\boldsymbol{\omega}'\}. \tag{9}$$

$$v_{\boldsymbol{\omega}} = 0, \text{ if } \boldsymbol{\omega} \notin \overline{Q}^{(n)}(\boldsymbol{\omega}'). \tag{10}$$

Solution of the system of equations (8)-(10) (with unknowns $v_{\boldsymbol{\omega}},\ \boldsymbol{\omega} \in \overline{Q}^{(n)}(\boldsymbol{\omega}')$) can be computed by solving the linear program

$$\min \sum_{\boldsymbol{\omega} \in \overline{Q}^{(n)}(\boldsymbol{\omega}')} v_{\boldsymbol{\omega}} \tag{11}$$

subject to

$$v_{\boldsymbol{\omega}'} \geq \pi(\boldsymbol{\omega}', \mathbf{x}) + \delta \sum_{\boldsymbol{\omega} \in \Omega^{(2)}(\boldsymbol{\omega}')} \rho_{\boldsymbol{\omega}'}(\mathbf{x})(\boldsymbol{\omega}) v_{\boldsymbol{\omega}}, \quad \forall \mathbf{x} \in X^{(n)}, \tag{12}$$

$$v_{\boldsymbol{\omega}''} \geq \pi(\boldsymbol{\omega}'', \mathbf{x}) + \delta \sum_{\boldsymbol{\omega} \in \Omega^{(2)}(\boldsymbol{\omega}'')} \rho_{\boldsymbol{\omega}''}(\mathbf{x})(\boldsymbol{\omega}) v_{\boldsymbol{\omega}},$$

$$\forall \boldsymbol{\omega}'' \in \overline{Q}^{(n)}(\boldsymbol{\omega}') \setminus \{\boldsymbol{\omega}'\}, \quad \forall \mathbf{x} \in X(\boldsymbol{\omega}''), \tag{13}$$

and (10).

For $\mathbf{x} \in X(\boldsymbol{\omega}')$ let $Q^{(k)}(\boldsymbol{\omega}', \mathbf{x})$ be the set of states that can occur under some strategy prescribing $\mathbf{x}$ at initial state $\boldsymbol{\omega}'$ in period k and $\overline{Q}^{(n)}(\boldsymbol{\omega}', \mathbf{x})$ be the set of states that can occur under some strategy prescribing $\mathbf{x}$ at initial state $\boldsymbol{\omega}'$ in the first n periods. That is,

$$Q^{(k)}(\boldsymbol{\omega}', \mathbf{x})$$

$$= \{\boldsymbol{\omega} \in \Omega \mid \exists\ q \in Q \text{ such that } q(\boldsymbol{\omega}') = \mathbf{x} \text{ and } \mu(\boldsymbol{\omega}', \boldsymbol{\omega}, k, q) > 0\} \tag{14}$$

and

$$\overline{Q}^{(n)}(\boldsymbol{\omega}', \mathbf{x}) = \bigcup_{k=1}^{n} \Omega^{(k)}(\boldsymbol{\omega}', \mathbf{x}). \tag{15}$$

In order to determine prescription at $\boldsymbol{\omega}'$, we need a solution of (10)-(13) for each fixed $\mathbf{x} \in X^{(n)}$ (i.e., replacing (12) by (17) below). Therefore, for each $\mathbf{x} \in X^{(n)}(\boldsymbol{\omega}')$ we solve the linear program

$$\min \sum_{\boldsymbol{\omega} \in \overline{Q}^{(n)}(\boldsymbol{\omega}', \mathbf{x})} v_{\boldsymbol{\omega}} \tag{16}$$

subject to

$$v_{\boldsymbol{\omega}'} = v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}) = \pi(\boldsymbol{\omega}', \mathbf{x}) + \delta \sum_{\boldsymbol{\omega} \in \Omega^{(2)}(\boldsymbol{\omega}', \mathbf{x})} \rho_{\boldsymbol{\omega}'}(\mathbf{x})(\boldsymbol{\omega}) v_{\boldsymbol{\omega}}, \tag{17}$$

$$v_{\boldsymbol{\omega}''} \geq \pi(\boldsymbol{\omega}'', \mathbf{x}) + \delta \sum_{\boldsymbol{\omega} \in \Omega^{(2)}(\boldsymbol{\omega}'')} \rho_{\boldsymbol{\omega}''}(\mathbf{x})(\boldsymbol{\omega}) v_{\boldsymbol{\omega}},$$

$$\forall \boldsymbol{\omega}'' \in \overline{Q}^{(n)}(\boldsymbol{\omega}', \mathbf{x}) \setminus \{\boldsymbol{\omega}'\}, \quad \forall \mathbf{x} \in X(\boldsymbol{\omega}''), \tag{18}$$

$$v_{\boldsymbol{\omega}} = 0, \text{ if } \boldsymbol{\omega} \notin \overline{Q}^{(n)}(\boldsymbol{\omega}', \mathbf{x}), \tag{19}$$

and set

$$v_{\boldsymbol{\omega}'}^{(n)} = \max_{\mathbf{x} \in X^{(n)}(\boldsymbol{\omega}')} v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}). \tag{20}$$

The following claim says how to identify the unique prescription of each optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$.

Claim 2. *Let* $\mathbf{x}^* \in X^{(n)}(\boldsymbol{\omega}')$ *satisfy* $v_{\boldsymbol{\omega}'}^{(n)} = v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}^*) > v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x})$ *for each* $\mathbf{x} \in X^{(n)}(\boldsymbol{\omega}') \setminus \{\mathbf{x}^*\}$ *and set*

$$\tilde{v}_{\boldsymbol{\omega}'}^{(n)} = \max\{v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}) \mid \mathbf{x} \in X^{(n)}(\boldsymbol{\omega}') \setminus \{\mathbf{x}^*\}\}. \tag{21}$$

If

$$v_{\boldsymbol{\omega}'}^{(n)} - \tilde{v}_{\boldsymbol{\omega}'}^{(n)} > \frac{2\delta^n (\max L_1 - \min L_1)}{1 - \delta}, \tag{22}$$

then $\mathbf{x}^*$ *is the unique prescription of each optimal strategy for program* (4)-(5) *for initial state* $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$.

Proof. In (22) we increased the sum of discounted expected single period payoffs since period $n + 1$ from a strategy giving $\tilde{v}_{\boldsymbol{\omega}'}^{(n)}$ by its unattainable upper bound $\frac{\delta^n (\max L_1 - \min L_1)}{1 - \delta}$ and decreased by the same amount the sum of discounted expected single period payoffs since period $n + 1$ from a strategy prescribing $\mathbf{x}^*$ at $\boldsymbol{\omega}'$. (This upper bound is unattainable for two reasons. First, the highest single period payoff is obtained when state variables are at their maximal values and zero actions are taken. This corresponds to the sum of discounted single period payoffs since period $n+1$ equal to $\frac{\delta^n \max L_1}{1-\delta}$. Second, the sum of discounted expected single period payoffs since period $n+1$ cannot be lower than $\frac{\delta^n \min L_1}{1-\delta}$ because the latter sum can be ensured by taking zero action in each period since $n+1$). If after such modifications the latter strategy still gives a higher sum of discounted expected single period payoffs, then it is the unique prescription of each optimal strategy for program (4)-(5) with initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. It is true that, moving form linear program (16)-(19) to program (4)-(5), the sum of discounted expected single period payoffs in the first n periods from the former strategy can increase. Nevertheless, if it increases because of using less costly actions (and, hence, actions with lower values), it decreases the sum of discounted expected single period payoffs since period $n+1$. Increase in the sum of discounted expected single period payoffs in the first n periods can continue up to the point when it equals the decrease in the sum of discounted expected single period payoffs since period $n+1$. Since we neglected such decrease in the sum of discounted expected single period payoffs since period $n+1$, we can neglect also increase in the sum of discounted expected single period payoffs in the first n periods. The sum of discounted expected single period payoffs in the first n periods cannot increase due to increases in values of state variables in the first n periods, because such increase would be possible also in linear program (16)-(19). □

If assumptions of Claim 2 are satisfied and $\boldsymbol{\omega}' = \boldsymbol{\omega}(r)$, we proceed with computations for state $\boldsymbol{\omega}(r+1)$. If time available for making a decision allows it, we wait until the actual state in the following period becomes known.

The relation between $X^{(n)}(\boldsymbol{\omega}')$ and $X^{(n+1)}(\boldsymbol{\omega}')$ is based on the following two claims.

Claim 3. *Let* $\mathbf{x}^* \in X(\boldsymbol{\omega}')$ *satisfy* $v_{\boldsymbol{\omega}'}^{(n)} = v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}^*)$. *If* $\mathbf{x} \in X(\boldsymbol{\omega}')$ *satisfies* $\mathbf{x} < \mathbf{x}^*$, *then* $v_{\boldsymbol{\omega}'}^{(k)} \neq v_{\boldsymbol{\omega}'}^{(k)}(\mathbf{x})$ *for each* $k \in \mathbb{N}$ *with* $k > n$ *and* $\boldsymbol{x}$ *is not the prescription of any optimal strategy for program* (4)-(5) *at state* $\boldsymbol{\omega}'$.

Proof. Proof follows from Assumption 1, Assumption 2, and part (b) of Assumption 3. With additional periods (in which we consider for all states occurring in them with a positive probability actual values of strategies at them) considered, the effect of $\mathbf{x}^*$ on increase of state variables (of each ω_i for which $x_i^* > x_i$, according to Assumption 1) and single period payoffs (according to part (b) of Assumption 3) in future periods occurs in additional periods. If for $k > n$ a strategy prescribing **x** at $\boldsymbol{\omega}'$ used actions that it does not use when only *n* periods are considered, these actions would be feasible (according to Assumption 2) also for a strategy prescribing $\mathbf{x}^*$ at $\boldsymbol{\omega}'$. Since this holds for each $k > n$, it holds also for infinite horizon program (4)-(5). □

Claim 4. *If* $\mathbf{x} \in X(\boldsymbol{\omega}')$ *satisfies*

$$v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}) - v_{\boldsymbol{\omega}'}^{(n)} < -\frac{2\delta^n(\max L_1 - \min L_1)}{1-\delta}, \tag{23}$$

then it is not the prescription of any optimal strategy for program (4)-(5) *for initial state* $\boldsymbol{\omega}'$ *at* $\boldsymbol{\omega}'$.

Proof. In (23) we increased the sum of discounted expected single period payoffs since period $n+1$ from a strategy prescribing **x** at $\boldsymbol{\omega}'$ by its unattainable upper bound $\dfrac{\delta^n(\max L_1 - \min L_1)}{1-\delta}$ and decreased by the same amount the sum of discounted expected single period payoffs since period $n+1$ from a strategy prescribing $\mathbf{x}^* \in X(\boldsymbol{\omega}')$ satisfying $v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}^*) = v_{\boldsymbol{\omega}'}^{(n)}$ at $\boldsymbol{\omega}'$. If with such modifications a strategy prescribing **x** at $\boldsymbol{\omega}'$ gives still lower sum of discounted expected single period payoffs than a strategy prescribing $\mathbf{x}^*$ at $\boldsymbol{\omega}'$, then the former strategy cannot solve program (4)-(5) with initial state $\boldsymbol{\omega}'$. A comment on an increase in the sum of discounted expected single period payoffs in the first *n* periods from the former strategy analogous to the one in the proof of Claim 2 applies also here. □

Claims 3 and 4 imply the following relation between $X^{(n)}(\boldsymbol{\omega}')$ and $X^{(n+1)}(\boldsymbol{\omega}')$:

$$X^{(n+1)}(\boldsymbol{\omega}') = X^{(n)}(\boldsymbol{\omega}') \backslash (\{\mathbf{x} \in X^{(n)}(\boldsymbol{\omega}') \mid \mathbf{x} < \mathbf{x}^* \text{ for some}$$

$$\mathbf{x}^* \in X^{(n)}(\boldsymbol{\omega}') \text{ with } v_{\boldsymbol{\omega}'}^{(n)} = v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}^*)\}$$

$$\cup \{\mathbf{x} \in X^{(n)}(\boldsymbol{\omega}') \mid x_i > x_i^* \text{ for some}$$

$$i \in \{i, \ldots, m^*\} \text{ and (23) holds}\}). \tag{24}$$

Thus, $X^{(n+1)}(\boldsymbol{\omega}')$ includes elements of $X^{(n)}(\boldsymbol{\omega}')$ that we have not yet ruled out as prescriptions of optimal strategies for program (4)-(5) for initial state $\boldsymbol{\omega}'$.

If $X^{(n+1)}(\boldsymbol{\omega}') = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}\}$,

$$\mathbf{x}^{(1)} > \mathbf{x}^{(2)} > \cdots > \mathbf{x}^{(k)}, \tag{25}$$

and

$$v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}^{(j)}) - v_{\boldsymbol{\omega}'}^{(n)} > \frac{2\delta^n(\max L_1 - \min L_1)}{1-\delta} - \varepsilon^*, \quad \forall j \in \{1, \ldots, k\}, \tag{26}$$

then - using Claim 1 and recalling Remark 1 - each element of $X^{(n+1)}(\boldsymbol{\omega}')$ is the prescription of some optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. In this case, for each $\mathbf{x} \in X^{(n+1)}(\boldsymbol{\omega}')$, reducing its cost by the smallest monetary unit turns it into the unique prescription of each optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$.

Remark 2. In (22), (23), and (26) we used the unattainable upper bound on increase and decrease in the sum of discounted expected single period payoffs since period $n+1$. This may increase the number of iterations needed for computation of the prescription of an optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$. Nevertheless, we use it for its simple computation. Computation of some lower upper bound (e.g., taking into account impact of states occurring with a positive probability in period n on future expected values of state variables) would be cumbersome and this would outweigh decrease in the number of iterations.

The last three claims in this section show that a finite number of iterations suffices for computation of the prescription of an optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$.

Claim 5. *If* $x^* \in X(\boldsymbol{\omega}')$ *is the prescription of each optimal strategy for program* (4)-(5) *for initial state* $\boldsymbol{\omega}'$ *at* $\boldsymbol{\omega}'$, *then there is* $n \in \mathbb{N} \setminus \{1\}$ *for which* (22) *holds.*

Proof. Let $\tilde{v}_{\boldsymbol{\omega}'} = \max\{v_{\boldsymbol{\omega}'}(x) \mid \mathbf{x} \in X(\boldsymbol{\omega}') \setminus \{\mathbf{x}^*\}\}$. Assumption of Claim 5 implies that there is $\psi > 0$ such that $v^*_{\boldsymbol{\omega}'} - \tilde{v}_{\boldsymbol{\omega}'} > \psi$. This further implies that there is $n_1 \in \mathbb{N} \setminus \{1\}$ such that $v^{(n)}_{\boldsymbol{\omega}'} - \tilde{v}^{(n)}_{\boldsymbol{\omega}'} > 0.5\psi$ for each $n \geq n_1$. There is also $n_2 \in \mathbb{N} \setminus \{1\}$ such that $\dfrac{2\delta^n(\max L_1 - \min L_1)}{1-\delta} \leq 0.5\psi$ for each $n \geq n_2$. Setting $n_0 = \max\{n_1, n_2\}$, (22) holds for each $n \geq n_0$. □

Claim 6. *If* $\mathbf{x} \in X(\boldsymbol{\omega}')$ *is not the prescription of any optimal strategy for program* (4)-(5) *for initial state* $\boldsymbol{\omega}'$ *at* $\boldsymbol{\omega}'$, *then there is* $n \in \mathbb{N} \setminus \{1\}$ *for which either* (23) *holds, or* $\mathbf{x} < \mathbf{x}^*$ *for some* $\mathbf{x}^*$ *satisfying* $v^{(n)}_{\boldsymbol{\omega}'} = v^{(n)}_{\boldsymbol{\omega}'}(\mathbf{x}^*)$.

Proof. If (for some $n \in \mathbb{N} \setminus \{1\}$) $\mathbf{x} < \mathbf{x}^*$ for some $\mathbf{x}^*$ satisfying $v^{(n)}_{\boldsymbol{\omega}'} = v^{(n)}_{\boldsymbol{\omega}'}(\mathbf{x}^*)$, then the claim holds. Thus, suppose that there are not such $n \in \mathbb{N} \setminus \{1\}$ and $\mathbf{x}^*$. Let $v^*_{\boldsymbol{\omega}'}$ be the value of an optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$ and $v_{\boldsymbol{\omega}'}(x)$ be the value at $\boldsymbol{\omega}'$ of a strategy solving (4)-(5) with the additional constraint $q(\boldsymbol{\omega}') = \mathbf{x}$. Since $\mathbf{x}$ is not the prescription of any optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$, there is $\psi > 0$ such that $v_{\boldsymbol{\omega}'}(\mathbf{x}) - v^*_{\boldsymbol{\omega}'} < -\psi$. Therefore, there is $n_1 \in \mathbb{N} \setminus \{1\}$ such that $v^{(n)}_{\boldsymbol{\omega}'}(\mathbf{x}) - v^{(n)}_{\boldsymbol{\omega}'} < -0.5\psi$ for each $n \geq n_1$. There is also $n_2 \in \mathbb{N} \setminus \{1\}$ such that $\dfrac{2\delta^n(\max L_1 - \min L_1)}{1-\delta} \leq 0.5\psi$ for each $n \geq n_2$. Setting $n_0 = \max\{n_1, n_2\}$, (23) holds for each $n \geq n_0$. □

Claim 7. *If for some* $k \in \mathbb{N} \setminus \{1\}$ *and each* $j \in \{1, ..., k\}\, x^{(j)}$ *is the prescription of some optimal strategy for program* (4)-(5) *for initial state* $\boldsymbol{\omega}'$ *at* $\boldsymbol{\omega}'$ *and* (25) *holds, then there is* $n \in \mathbb{N} \setminus \{1\}$ *such that* (26) *holds.*

Proof. By assumption of Claim 7, we have $v_{\boldsymbol{\omega}'}(x^{(j)}) = v^*_{\boldsymbol{\omega}'}$ for each $j \in \{1, ..., k\}$. This implies that for each $j \in \{1, ..., k\}$ there is $n_j \in \mathbb{N} \setminus \{1\}$ such that $v^{(n)}_{\boldsymbol{\omega}'}(x^{(j)}) - v^{(n)}_{\boldsymbol{\omega}'} > -0.5\varepsilon^*$ for each $n \geq n_j$. There also exists n_{k+1} such that $2\delta^n(\max L_1 - \min L_1)/(1-\delta) \leq 0.5\varepsilon^*$ for each $n \geq n_{k+1}$. Setting $n_0 = \max\{n_1, ..., n_k, n_{k+1}\}$, (26) holds for each $n \geq n_0$. □

The preceding results give rise to the following algorithm. In its description we use the assignment sign $(\rightarrow)$ whenever the equality sign $(=)$ would be mathematically incorrect.

Algorithm 1. Basic algorithm

Step 1. Set $r = 1$ and go to step 2.

Step 2. Set $\boldsymbol{\omega}' = \omega(r)$, $n = 2$, $X^{(2)}(\boldsymbol{\omega}') = X(\boldsymbol{\omega}')$, and go to step 3.

Step 3. Compute $v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x})$ for each $x \in X^{(n)}(\boldsymbol{\omega}')$, $v_{\boldsymbol{\omega}'}^{(n)}$ according to (20), and go to step 4.

Step 4. If there is $\mathbf{x}^* \in X^{(n)}(\boldsymbol{\omega}')$ satisfying $v_{\boldsymbol{\omega}'}^{(n)} = v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x}^*) > v_{\boldsymbol{\omega}'}^{(n)}(\mathbf{x})$ for each $\mathbf{x} \in X^{(n)}(\boldsymbol{\omega}') \setminus \{\mathbf{x}^*\}$ and (22), then $\mathbf{x}^*$ is the unique prescription of each optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. Go to step 8. Otherwise, go to step 5.

Step 5. Compute $X^{(n+1)}(\boldsymbol{\omega}')$ according to (24) and go to step 6.

Step 6. If $X^{(n+1)}(\boldsymbol{\omega}') = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}\}$, and (25) and (26) are satisfied, then each element of $X^{(n+1)}(\boldsymbol{\omega}')$ is the prescription of some optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. Go to step 8. Otherwise, go to step 7.

Step 7. Let $n + 1 \rightarrow n$ and go to step 3.

Step 8. Set $r + 1 \rightarrow r$ and go to step 2.

Steps 3 - 7 form one iteration. A return to step 3 starts a new iteration.

4. Simplified Algorithm

In this section, we develop a simplified algorithm. It is based on claims 8 and 9 below. It can be used under the following three additional assumptions and assumptions contained in Claims 8 and 9.

Assumption 4. $\Omega \subset \mathbb{R}_+^{m^*}$.

We use here an extension of function η from its original domain Ω to

$$W = \{\boldsymbol{\omega} \in \mathbb{R}^{m^*} \mid 0 \ll \omega \ll \overline{\omega} + (0.01, \ldots, 0.01), \eta(\omega) > 0\}.$$

We can use it if the formula specifying η on Ω is well defined also for each $\boldsymbol{\omega} \in W$. In order to avoid additional notation, we use symbol η also for extension of original function η to W.

Assumption 5. Function η is twice continuously differentiable on W.

Assumption 6. For each $\boldsymbol{\omega}' \in \Omega$ and each $\mathbf{x} \in X(\boldsymbol{\omega}')$, each $\boldsymbol{\omega} \in \Omega^{(1)}(\boldsymbol{\omega}', \mathbf{x})$ satisfies

$$\omega_i = (1 - \alpha_i)\omega_i' + x_i + \varepsilon_i(\boldsymbol{\omega}'), \quad i = 1, \ldots, m^*,$$

where $\alpha_i \in (0, 1)$ is a depreciation rate for component i of the state vector and $\varepsilon_i(\boldsymbol{\omega}')$ is a realization of a random variable with finite support and zero mean.

Claim 8. *Assume that for some* $i \in \{1, \ldots, m^*\}$

(i) *function* η *is convex in* ω_i *at each* $\boldsymbol{\omega} \in W$,

(ii) *for some* $\boldsymbol{\omega}' \in \Omega \cap W$, $\mathbf{x} \in X(\boldsymbol{\omega}')$ *and* $\mathbf{x}' \in X(\boldsymbol{\omega}')$ *are such that* $x_i > x_i'$ *and* $x_j = x_j'$ *for each* $j \in \{1, \ldots, m^*\} \setminus \{i\}$,

(iii) *for each* $j \in \{1, \ldots, m^*\}$, $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_j}$ *has the same sign at each* $\boldsymbol{\omega} \in W$,

(iv) $\boldsymbol{\omega}''$ *satisfies* $\omega_i'' = \omega_i'$, *if* $j \neq i$ *and* $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_j} \leq 0$, *then* $\omega_j'' = \overline{\omega}_j$, *and if* $j \neq i$ *and* $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_j} > 0$, *then* $\omega_j'' = \omega_j'(1 - \alpha_j)^q$, *where q is the highest non-negative integer k satisfying* $\eta(\tilde{\omega}) > 0$, *when* $\tilde{\omega}_i = \omega_i'$, $\tilde{\omega}_r = \omega_r'$ *for each* $r \in \{1, \ldots, m^*\} \setminus \{i\}$ *with* $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_r} \leq 0$, *and*

$$\tilde{\omega}_r = (1 - \alpha_r)^k \omega_r' \textit{ for each } r \in \{1, \ldots, m^*\} \setminus \{i\} \textit{ with } \frac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_r} > 0,$$

(v)

$$\frac{\partial \eta(\boldsymbol{\omega}'')}{\partial \omega_i}(x_i - x_i')\frac{\delta}{1-\delta(1-\alpha_i)} > \gamma(x) - \gamma(x'). \tag{27}$$

Then $\mathbf{x}'$ *is not a prescription of any optimal strategy for program* (4)-(5) *for initial state* $\boldsymbol{\omega}'$ *at* $\boldsymbol{\omega}'$.

Proof. Since future payoffs are discounted, if it is optimal to use action **x** and not to use action $\mathbf{x}'$ in the first period, then it cannot be optimal to let ω_i fall below ω_i' in any future period. Also, it cannot be optimal to let $\eta(\boldsymbol{\omega})$ fall to zero at any state $\boldsymbol{\omega}$ occurring with a positive probability in the future.

With respect to Assumption 6 and part (a) of Assumption 2, after taking action **x** in the first period and after taking action $\mathbf{x}'$ in the first period, in any period $n \geq 2$ and for any succession of realizations of random factors in periods $2, \ldots, n,$ the same action can be taken. Then the left hand side of (27) expresses a lower bound on increase in the sum of average discounted single period payoffs since period two, because of determination of ω'' in (iv) and the fact that convexity of η in ω_i implies that

increase in $\eta(\boldsymbol{\omega})$ evaluated along a tangent line is lower or equal than its true increase,

expected value of increase in $\eta(\boldsymbol{\omega})$ in any period is higher or equal than its increase computed from expected value of ω_i,

$\frac{\partial \eta(\boldsymbol{\omega})}{\partial \omega_i}$ increases or remains unchanged with increasing ω_i.

If this lower bound exceeds the extra cost of taking action **x** rather than action $\mathbf{x}'$ (as stated in (27)), then $\mathbf{x}'$ cannot be a prescription of any optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. □

Claim 9. *Assume that for some* $i \in \{1, \ldots, m^*\}$

(i) *for some* $\boldsymbol{\omega}' \in \Omega \cap W$, $\mathbf{x} \in X(\boldsymbol{\omega}')$ *and* $\mathbf{x}' \in X(\boldsymbol{\omega}')$ *are such that*

$$x_i > x_i' \textit{ and } x_j = x_j' \textit{ for each } j \in \{1, \ldots, m^*\} \setminus \{i\},$$

(ii) *function* η *is concave in* ω_i *either* (iia) *at each* $\boldsymbol{\omega} \in W$, *or* (iib) *at each* $\boldsymbol{\omega} \in W$ *with* $\boldsymbol{\omega} \geq \boldsymbol{\omega}'$ *and for each* $j \in \{1, ..., m^*\} \backslash \{i\}$ *all actions with* $x_j = 0$ *were eliminated on the basis of Claim* 8,

(iii) *for each* $j \in \{1, ..., m^*\}$, $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_j}$ *has the same sign at each* $\boldsymbol{\omega} \in W$ (*in case* (iia)) *or at each* $\boldsymbol{\omega} \in W$ *with* $\boldsymbol{\omega} \geq \boldsymbol{\omega}'$ (*in case* (iib)),

(iv) $\boldsymbol{\omega}''$ *satisfies* $\omega_i'' = \omega_i'$, *if* $j \neq i$ *and* $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_j} > 0$, *then* $\omega_j'' = \overline{\omega}_j$, *and if* $j \neq i$ *and* $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_j} \leq 0$, *then* $\omega_j'' = \omega_j'(1 - \alpha_j)^q$, *where* q *is the highest non-negative integer* k *satisfying* $\eta(\tilde{\boldsymbol{\omega}}) > 0$, *when* $\tilde{\omega}_i = \omega_i'$, $\tilde{\omega}_r = \omega_r'$ *for each* $r \in \{1, ..., m^*\} \backslash \{i\}$ *with* $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_r} > 0$, *and* $\tilde{\omega}_r = (1 - \alpha_r)^k \omega_r'$ *for each* $r \in \{1, ..., m^*\} \backslash \{i\}$ *with* $\dfrac{\partial^2 \eta(\omega)}{\partial \omega_i \partial \omega_r} \leq 0$,

(v) *the set of feasible actions is the same for each state*,

(vi)

$$\frac{\partial \eta(\boldsymbol{\omega}'')}{\partial \omega_i}(x_i - x_i')\frac{\delta}{1 - \delta(1 - \alpha_i)} < \gamma(x) - \gamma(x'). \tag{28}$$

`*Then* $\mathbf{x}$ *is not a prescription of any optimal strategy for program* (4)-(5) *for initial state* $\boldsymbol{\omega}'$ *at* $\boldsymbol{\omega}'$.

Proof. Since future payoffs are discounted, if it is optimal to use action $\mathbf{x}$ and not to use action $\mathbf{x}'$ in the first period, then it cannot be optimal to let ω_i fall below ω_i' in any future period. For an analogous reason, in case (iib) it cannot be optimal to let ω_j fall below ω_j' for some $j \in \{1, ..., m^*\} \backslash \{i\}$. Also, it cannot be optimal to let $\eta(\boldsymbol{\omega})$ fall to zero at any state $\boldsymbol{\omega}$ occurring with a positive probability in the future. With respect to Assumption 6 and part (v) of assumptions of this claim, after taking action $\mathbf{x}$ in the first period and after taking action $\mathbf{x}'$ in the first period, in any period $n \geq 2$ and for any succession of realizations of random factors in periods $2, ..., n$, the same action can be taken. Then the left hand side of (28) expresses an upper bound on increase in the sum of average discounted single period payoffs since period two, because of determination of ω'' in (iv) and the fact that concavity of η in ω_i implies that

increase in $\eta(\boldsymbol{\omega})$ evaluated along a tangent line is higher or equal than its true increase,

expected value of increase in $\eta(\boldsymbol{\omega})$ in any period is lower or equal than its increase computed from expected value of ω_i,

$\dfrac{\partial\eta(\boldsymbol{\omega})}{\partial\omega_i}$ decreases or remains unchanged with increasing ω_i.

If this upper bound is lower than the extra cost of taking action $\mathbf{x}$ rather than action $\mathbf{x}'$ (as stated in (28)), then $\mathbf{x}$ cannot be a prescription of any optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. □

Algorithm 2. Simplified algorithm

Between steps 2 and 3 of the basic algorithm we insert steps 2a, 2b, and 2c. From step 2 we go to step 2a.

Step 2a. If there exists pair $(\mathbf{x}, \mathbf{x}') \in X^{(2)}(\boldsymbol{\omega}')$ satisfying assumptions of Claim 8, let $X^{(2)}(\boldsymbol{\omega}')\backslash\{\mathbf{x}'\} \to X^{(2)}(\boldsymbol{\omega}')$ and return to step 2a. Otherwise, go to step 2b.

Step 2b. If there exists pair $(\mathbf{x}, \mathbf{x}') \in X^{(2)}(\boldsymbol{\omega}')$ satisfying assumptions of Claim 9, let $X^{(2)}(\boldsymbol{\omega}')\backslash\{\mathbf{x}'\} \to X^{(2)}(\boldsymbol{\omega}')$ and return to step 2b. Otherwise, go to step 2c.

Step 2c. If $X^{(2)}(\boldsymbol{\omega}') = \{\mathbf{x}^*\}$, then $\mathbf{x}^*$ is the unique prescription of each optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. Go to step 8. Otherwise, go to step 3.

We end this section by an example illustrating the use of the simplified algorithm for setting defined in Example 1. In it, we compute a prescription of an optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}' = (\mathbf{1}, \mathbf{2})$ at $\boldsymbol{\omega}'$.

Example 2. The stochastic program in Example 1 satisfies Assumptions 4- 6.

Step 1. We set $r = 1$ and go to step 2.

Step 2. We set $\boldsymbol{\omega}' = \boldsymbol{\omega}(1) = (1, 2)$, $n = 2$,

$$X^{(2)}(\boldsymbol{\omega}') = \{(0, 0), (0, 0.5), (0, 1.5), (0.5, 0), (0.5, 0.5), (0.5, 1.5)\},$$

and go to step 2a.

Step 2a. Since $\dfrac{\partial^2\eta(\boldsymbol{\omega})}{\partial\omega_1^2} = \dfrac{\partial^2\left(\left(\dfrac{\omega_1}{2} - \dfrac{1}{2(0.001+\omega_2)}\right)^2\right)}{\partial\omega_1^2} = \dfrac{1}{2}$ and

$$\frac{\partial^2\eta(\boldsymbol{\omega})}{\partial\omega_1\partial\omega_2} = \frac{\partial^2\left(\left(\dfrac{\omega_1}{2} - \dfrac{1}{2(0.001+\omega_2)}\right)^2\right)}{\partial\omega_1\partial\omega_2}$$

$$= \frac{1}{2}\frac{1}{(0.001+\omega_2)^2} > 0 \quad \forall\boldsymbol{\omega} \in W,$$

pair $((0.5, 0), (0, 0))$ satisfies assumptions (i)-(iii) of Claim 8. Since inequality $\dfrac{1}{2} - \dfrac{1}{2(0.001 + 2\times 0.3^k)} > 0$ has solution $(-\infty, 0.57655)$ we set $q = 0$ and $\boldsymbol{\omega}'' = \boldsymbol{\omega}' = (1, 2)$. We have $\dfrac{\partial\eta(\boldsymbol{\omega}'')}{\partial\omega_1} = 0.25012$ and

$$0.25012 \times 0.5 \times \frac{0.9}{1 - 0.9\times 0.9} = 0.59239 > 0.5 = \gamma(0.5, 0) - \gamma(0, 0),$$

$$X^{(2)}(\boldsymbol{\omega}')\setminus\{(0, 0)\} = \{(0, 0.5), (0, 1.5), (0.5, 0), (0.5, 0.5), (0.5, 1.5)\} \to X^{(2)}(\boldsymbol{\omega}')$$

and we return to step 2a.

Step 2a again. By the same reasoning as in the first application of step 2a, pair $((0.5, 0.5), (0, 0.5))$ satisfies all assumptions of Claim 8. Therefore,

$$X^{(2)}(\boldsymbol{\omega}')\setminus\{(0, 0.5)\} = \{(0, 1.5), (0.5, 0), (0.5, 0.5), (0.5, 1.5)\} \to X^{(2)}(\boldsymbol{\omega}')$$

and we return to step 2a.

Step 2a again. By the same reasoning as in the first application of step 2a, pair $((0.5, 1.5), (0, 1.5))$ satisfies all assumptions of Claim 8. Therefore,

$$X^{(2)}(\boldsymbol{\omega}')\backslash\{(0,\ 1.5)\} = \{(0.5,\ 0),\ (0.5,\ 0.5),\ (0.5,\ 1.5)\} \to X^{(2)}(\boldsymbol{\omega}')$$

and we return to step 2a.

Step 2a again. There is no pair of elements of $X^{(2)}(\boldsymbol{\omega}')$ satisfying assumptions of Claim 8. Therefore, we go to step 2b.

Step 2b. Since

$$\frac{\partial^2 \eta(\boldsymbol{\omega})}{\partial \omega_2^2} = \frac{\partial^2 \left(\left(\frac{\omega_1}{2} - \frac{1}{2(0.001 + \omega_2)} \right)^2 \right)}{\partial \omega_2^2}$$

$$= \frac{1}{(0.001 + \omega_2)^3} \left(\frac{3}{2} \frac{1}{0.001 + \omega_2} - \omega_1 \right) < 0$$

$$\forall\, \boldsymbol{\omega} \in W \ \text{ with } \ \boldsymbol{\omega} \geq (1,\ 2),$$

all actions with $x_1 = 0$ were eliminated in repeated applications of step 2a,

$$\frac{\partial^2 \eta(\boldsymbol{\omega})}{\partial \omega_1 \partial \omega_2} = \frac{\partial^2 \left(\left(\frac{\omega_1}{2} - \frac{1}{2(0.001 + \omega_2)} \right)^2 \right)}{\partial \omega_1 \partial \omega_2}$$

$$= \frac{1}{2} \frac{1}{(0.001 + \omega_2)^2} > 0, \quad \forall \boldsymbol{\omega} \in W,$$

the set of feasible actions is the same for each state, for $\boldsymbol{\omega}'' = (6,\ 2)$ pair of actions $((0.5,\ 0),\ (0.5,\ 0.5))$ satisfies assumptions (i), (iib), (iii), (iv), and (v) of Claim 9. We have $\frac{\partial \eta(\boldsymbol{\omega}'')}{\partial \omega_2} = 0.686\,84$ and

$$0.686\,84 \times 0.5 \times \frac{0.9}{1 - 0.9 \times 0.3} = 0.4239 < 0.5 = \gamma(0.5,\ 0.5) - \gamma(0.5,\ 0),$$

$$X^{(2)}(\boldsymbol{\omega}')\backslash\{(0.5,\ 0.5)\} = \{(0.5,\ 0),\ (0.5,\ 1.5)\} \to X^{(2)}(\boldsymbol{\omega}'),$$

and we return to step 2b.

Step 2b again. By the same reasoning as in the preceding application of step 2b, for $\boldsymbol{\omega}'' = (6, 2)$ pair of actions $((0.5, 0), (0.5, 1.5))$ satisfies assumptions (i), (iib), (iii), (iv), and (v) of Claim 9. We have

$$0.68684 \times 1.5 \times \frac{0.9}{1 - 0.9 \times 0.3} = 1.2702 < 1.5 = \gamma(0.5, 1.5) - \gamma(0.5, 0),$$

$$X^{(2)}(\boldsymbol{\omega}') \setminus \{(0.5, 1.5)\} = \{(0.5, 0)\} \to X^{(2)}(\boldsymbol{\omega}'),$$

and we return to step 2b.

Step 2b again. Since there does not exist pair $(\mathbf{x}, \mathbf{x}') \in X^{(2)}(\boldsymbol{\omega}')$ satisfying assumptions of Claim 9, we go to step 2c.

Step 2c. Since $X^{(2)}(\boldsymbol{\omega}') = \{(0.5, 0)\}$, $(0.5, 0)$ is the unique prescription of each optimal strategy for program (4)-(5) for initial state $\boldsymbol{\omega}'$ at $\boldsymbol{\omega}'$. We go to step 8.

Step 8. We set $r = 2$ and go to step 2. This starts computation for another initial state.

5. Conclusions

We have developed two algorithms - the basic one and a simplified one applicable in special cases - for gradual computation of an open-ended solution of a stochastic program with strictly increasing linkages between state variables. A stochastic program with amortization of state variables is a typical case of it. Such problems emerge whenever investments (with physical depreciation) or innovations (with moral depreciation) are involved in decision making. Thus, our results are applicable not only in decision making in business and economics but also (for example) in optimal design of air defence systems, coastal guard systems, and emergency help systems.

References

[1] W. K. Ching, X. Huang, M. K. Ng and T. K. Siu, Markov Chains, Springer, Heidelberg, 2013.

[2] W. K. Ching, M. K. Ng, K. Wong and E. Atlman, Customer lifetime value: a stochastic programming approach, J. Oper. Res. Soc. 55 (2004), 860-868.

[3] M. T. Flaherty, Industry structure and cost-reducing investment, Econometrica 48(5) (1980), 1187-1210.

[4] J. Friedman, Oligopoly Theory, Cambridge University Press, Cambridge, 1980.

[5] D. Q. Mayne, J. B. Rawlings, C. V. Rao and P. O. M. Scokaert, Constrained model predictive control: stability and optimality, Automatica J. IFAC 36(6) (2000), 789-814.

[6] G. Schildbach and M. Morari, Scenario-based model predictive control for multi-echelon supply chain management, European J. Oper. Res. 252(2) (2016), 540-549.

[7] D. White, Markov Decision Processes, Wiley, Chichester, 1993.