

SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4022

SOL--90-15

DE91 001359

**A Discounted-Cost Continuous-Time Flexible
Manufacturing and Operator Scheduling Model
Solved by Deconvexification Over Time**

by

B. Curtis Eaves and Uriel G. Rothblum

TECHNICAL REPORT SOL 90-15

August 1990

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Research and reproduction of this report were partially supported by the Department of Energy Grant DE-FG03-87ER25028 and the National Science Foundation Grant NSF DMS-8902662.

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do NOT necessarily reflect the views of the above sponsors.

Reproduction in whole or in part is permitted for any purposes of the United States Government. This document has been approved for public release and sale; its distribution is unlimited.

MASTER
DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

**A DISCOUNTED-COST CONTINUOUS-TIME FLEXIBLE
MANUFACTURING AND OPERATOR SCHEDULING MODEL
SOLVED BY DECONVEXIFICATION OVER TIME**

by

B. Curtis Eaves and Uriel G. Rothblum

ABSTRACT

A discounted-cost, continuous-time, infinite-horizon version of a flexible manufacturing and operator scheduling model is solved. The solution procedure is to convexify the discrete operator-assignment constraints to obtain a linear program, and then to regain the discreteness and obtain an approximate manufacturing schedule by deconvexification of the solution of the linear program over time. The strong features of the model are the accommodation of linear inequality relations among the manufacturing activities and the discrete manufacturing scheduling, whereas the weak features are intra-period relaxation of inventory availability constraints, and the absence of inventory costs, setup times, and setup charges.

Acknowledgement: This research is supported in part by NSF DMS-8902662 and DOE Grant DE-FG03-87ER25028.

1. Introduction

In Eaves and Rothblum [1988, 1989], we described a procedure, called deconvexification over time, and used it for solving an average-cost, infinite-horizon, flexible-manufacturing and operator-scheduling problem for both discrete- and continuous-time versions. Our purpose in the current paper is to apply a new variant of the deconvexification over time to solve a discounted-cost, rather than average-cost, version of the manufacturing and scheduling problem. We refer to the flexible-manufacturing and operator-scheduling problem as FMOS, as in the above references.

The relative strengths and weaknesses of FMOS with discounted-costs, as it can be solved with deconvexification over time, remain, in character, the same as for the average-costs. On one hand, we capture discrete manufacturing and operator scheduling requirements as well as linear inequalities among manufacturing activities. On the other hand, the continuous-time inventory availability constraints are relaxed to periodic inventory availability constraints and, further, we cannot include inventory costs, setup times or setup charges. Nevertheless, all in all, we believe that deconvexification, potentially, offers a significant planning tool for scheduling in a complex environment.

Deconvexification over time, as it has been applied in the past and as it is applied here, is a polynomial-time algorithm. Inclusion of inventory costs, setup times or setup charges appears to push the model into an intractable class, that is, one requiring exponential effort to solve. No attempt, to date, has been made to use deconvexification over time in an exponential-time algorithm. At present, FMOS, as solvable by deconvexification over time, seems to have its greatest value as a planning tool in an environment where operator costs are high relative to inventory or setup costs, and where setup times are relatively small.

In the paragraph below we define the discounted-cost manufacturing and operator scheduling problem to which we will refer as FMOS. This statement represents our problem orientation; it is an important problem we know something about, but it is not a problem we can

solve. Following the description of FMOS, we define a relaxation of FMOS referred to as RFMOS. It is RFMOS, not FMOS, that we will solve with deconvexification over time.

Definition of FMOS:

Divisible input materials, that is, raw materials, unfinished goods, etc., enter the production system through, for example, purchase activities. In the system various divisible in-process goods are transformed into other in-process goods through the execution of activities. The rate and character of these activities depend upon the assigned operator. The operator pool may be composed of, for example, skilled and unskilled labor, robots and machines. The function of an activity with an assigned operator is to transform in-process goods into in-process goods in fixed proportions and at a rate not exceeding a bound which is determined by both the activity and the operator. A class of operator constraints are available, for example, an activity can be manned by at most one operator at any given time, an operator can be assigned at most one activity at any given time, or certain collections of activities may be conducted at a single workstation which may limit the number of operators present. It is assumed that operators can be moved from activity to activity with negligible setup times or costs and that activities can be started up with negligible delays or setup costs. **Generation of in-process goods, for example, through purchases or production, must precede their usage, for example, through selling or use in production.** It is assumed that in-process goods can be inventoried on the spot with negligible time delay. Finally, output goods, that is, finished goods, processed materials, rejections, etc. exit the production system. The rates and relative proportions of input materials entering and of output goods exiting the system vary and depend upon operator-assignments and activity-rate settings. The objective is to minimize costs (or maximize profits), that is, operator costs, activity costs, and **inventory costs** minus revenues therefrom, all discounted over the infinite horizon. ⊗

As mentioned earlier, FMOS as stated above represents our problem orientation, but it is not a problem which we can solve. The example below illustrates that FMOS may have no optimal solution, principally due to the lack of setup times and charges.

Buy-Sell Example for FMOS:

A trader buys and sells a single commodity. At any instant in time he must buy or sell, but he cannot do both simultaneously. He can buy at rate of one item per unit-time at \$1 per item, and he can sell at a rate of one item per unit-time for \$2 per item. He cannot sell short, that is, he must have positive inventory which he has bought in order to sell. Assume that he has an interest rate of 20% per unit-time. The trader's task is to buy and sell so as to minimize his discounted-cost (or maximize his discounted-profit).

As we shall see in Section 7 that there is no optimal policy to the above problem, however, there are "stuttering" ϵ -optimal policies, namely, for each positive ϵ , there is a positive Δ where the policy of iteratively buying and selling at consecutive intervals of length $\Delta/2$ (starting with buying) will result in a discounted cost which cannot be improved by any policy by more than ϵ . \otimes

We next define a relaxation of FMOS which we call RFMOS. It is RFMOS that will be solved with deconvexification over time not FMOS. In FMOS, inventories must be nonnegative at all times, and as we will show, this implies that discounted inventory levels must be nonnegative at all times. RFMOS is obtained from FMOS by relaxing the inventory availability and discounted inventory constraints to periodic constraints instead of constraints that must hold at all times. In particular, periods are mentioned in RFMOS whereas FMOS is continuous time. The length of the periods in RFMOS can be selected arbitrarily. We will see that the optimal objective value does not vary with the selected length of the periods, however, other features of the optimal policy probably will vary.

Definition of RFMOS:

RFMOS is FMOS with the emboldened sentence "Generation of in-process goods, for example, through purchases or production, must precede their usage, for example, through selling or use in production" replaced by "At the end of each period inventory levels and discounted inventory levels must be nonnegative," and the emboldened phrase "inventory costs" deleted. ⊗

FMOS requires that inventory levels are nonnegative at all times. These FMOS constraints are relaxed in RFMOS to the requirements that inventory levels and discounted inventory levels are nonnegative at the end of each period. Indeed, in RFMOS, we allow for inventory levels and discounted inventory levels to become negative within a period. Because inventory levels in RFMOS can be both negative and positive, we can no longer capture inventory carrying costs via the linear objective.

Buy-Sell Example for RFMOS:

Consider the buy-sell example of FMOS with the following modifications. The trader is allowed to sell short, that is, sell quantities he does not have in stock. But, he must cover all outstanding orders at the end of each period. Further, for the interest rate of 20% per unit time, the discounted inventory levels must be nonnegative at the end of each period. ⊗

The principle weakness of FMOS, as the example shows, is the absence of setup times and charges, otherwise, FMOS is a bonified model. However, we cannot solve it, nevertheless, we take it as our problem orientation. The relaxed version of FMOS, namely RFMOS, suffers from a weak treatment of inventories as well as the absence of setups. It is RFMOS that we solve with deconvexification over time.

In the next section we introduce a mathematical description of two dynamic programs which we call DP and RDP, and we point out how they model FMOS and RFMOS, respectively. Indeed, DP and RDP can be regarded as more precise formulations of FMOS and RFMOS. In Section 3, we state an auxiliary dynamic program RRDP and a linear program LP which assist the solution procedure. In Section 4, optimal policies are constructed for RRDP, RDP and hence

RFMOS. In Section 5 dependence of optimal policies and objective on interest rates and period lengths is investigated. In Section 6, an interpretation is given to the discounted inventory constraints. Finally, in Section 7, three matters are discussed further, but briefly.

2. Dynamic Programming Formulations

Two dynamic programs DP and RDP are defined in this section. DP and RDP are designed to model FMOS and RFMOS, respectively. As RFMOS is a relaxation of FMOS, RDP is a relaxation of DP.

Let Z and R denote the set of integer and real scalars, respectively, and let Z_+ and R_+ denote the set of nonnegative elements in Z and R , respectively. Let $R^{m \times n}$ be the set of $m \times n$ real matrices where $m = 1, 2, \dots$ and $n = 1, 2, \dots$. In particular, elements of $R^m \equiv R^{m \times 1}$ and $R^{1 \times n}$ are called *column* and *row vectors*, respectively. Superscript T is used to indicate the transpose of a matrix or a vector.

Throughout this paper dynamic programs, a linear program and an integer program are employed. They require certain data. We define such data here once and for all.

Data

Let k , m and n be three positive integers. Let ρ and Δ be two positive scalars. Let $a \in R^{1 \times k}$ and $b \in R^{1 \times n}$ be two row vectors. Let $C \in R^{n \times k}$ and $D \in R^{m \times n}$ be two matrices. Let E be a finite subset of R^k where $Dx \geq 0$ for every $x \in E$. \otimes

Let π be a function from R_+ into some topological space. We define π to be *piecewise continuous* if there is an increasing sequence of scalars $0 = t(0) < t(1) < t(2) < \dots$ in R_+ tending to ∞ , where π is continuous on $(t(s), t(s+1)]$ for each $s \in Z_+$. Note that the above definition of piecewise continuity implies left continuity.

Throughout the paper we use the term policy to indicate time-dependent decision variables for several dynamic programs. We define a policy here, once and for all.

Policy

A policy $\pi = (x(\cdot), y(\cdot))$ is defined to be a piecewise continuous function mapping R_+ into $R^k \times R^n$. \otimes

As a casual remark to assist the reader, we note that each and every policy $\pi = (x(.), y(.))$ in this paper will always have the additional properties $x(t) \in E$ and $0 \leq y(t) \leq Cx(t)$ for all $t \in R_+$. Also, the *objective value* of a policy π , as denoted by V_π , is always given by $V_\pi \equiv \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$.

We are now ready to introduce our first dynamic program.

Definition of DP:

DP is the following dynamic program where the search takes place over policies $\pi = (x(.), y(.))$:

$$\text{DP: } V_{\text{DP}} \equiv \inf_{\pi = (x(.), y(.))} \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$$

$$\text{subject to: } x(t) \in E \text{ for all } t \in R_+ \quad (\text{DP1})$$

$$0 \leq y(t) \leq Cx(t) \text{ for all } t \in R_+, \text{ and} \quad (\text{DP2})$$

$$\int_0^t Dy(\tau) d\tau \geq 0 \text{ for all } t \in R_+. \quad (\text{DP3})$$

Define $\pi = (x(.), y(.))$ to be a *policy for DP* if it is a policy satisfying the constraints of DP and let $V_\pi \equiv \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$ denote the *objective value* of such a policy. The problem in DP is to select an optimal policy for DP, i.e., select a policy π for DP which infimizes the objective value, namely has, $V_\pi = V_{\text{DP}}$. \otimes

We next show how DP can be used to formulate FMOS. This is accomplished by interpreting the data and the variables of DP in the context of FMOS. The discussion here is brief and a more detailed discussion can be found in Eaves and Rothblum [1988, 1989].

The variable $t \in R_+$ represents time in FMOS and the finite set E corresponds to the collection of possible operator-assignments in FMOS. The variable $x(t)$ is the vector of operator-assignments at time t , and the constraint (DP1) expresses the requirement that $x(t)$ at each time

instance be a possible operator-assignment. The coordinates of the vector $y(t)$ represent the rates at which the activities are executed at time t , and constraint (DP2) bounds the rate vector at each time instance by a linear function of the operator-assignment vector, i.e., given the operator-assignment vector $x(t)$ the activity-levels, represented by the coordinates of the vector $y(t)$, cannot exceed those of $Cx(t)$. The matrix D is the technology or netput matrix. Each column of D corresponds to some activity as buying, selling or transforming in-process goods into in-process goods. The rows of D correspond to in-process goods. For example, if column i of the matrix D is the vector $(3, -1, 1, -2, 0, \dots, 0)^T$ and the i -th component of $y(t)$ is 5 over the period from time 2 till time 3, i.e., $y(t)_i = 5$ for $2 \leq t \leq 3$, then the net change in the quantities of the first four in-process goods during that period due to this activity would be 15, -5, 5 and -10, respectively. In general, given activity-levels $y(t)$, the rate of transformation of in-process goods is $Dy(t)$. Constraint (DP3) corresponds to the requirement that inventory levels are never negative, or indeed, that transformation cannot take place without the constituent inputs.

The operator-assignment set E is typically defined by a unimodular linear inequality system wherein extreme points correspond to elements of E . That such a system can be constructed to accommodate the types of constraints that we have suggested in FMOS is fully discussed in Eaves and Rothbum [1988, 1989] and will not repeated here.

Examining the objective of DP, we see that discounted costs, which are linear in the operator-assignment vectors and the activity-level vectors, can be incorporated into the objective. We next show that linear inventory costs can actually be included in DP. Let b' be the vector whose coordinates are the inventory unit costs incurred per unit time for the different in-process goods. Inventory levels at time t are determined by the activity-levels applied up to time t , and are given by $z(t) = \int_0^t Dy(\tau)d\tau$, implying that the discounted inventory cost would be

$$\begin{aligned} \int_0^\infty e^{-\rho\tau} b' z(\tau) d\tau &= b' \int_0^\infty e^{-\rho\tau} \left[\int_0^\tau Dy(\sigma) d\sigma \right] d\tau = b' D \int_0^\infty y(\sigma) \left[\int_\sigma^\infty e^{-\rho\tau} d\tau \right] d\sigma \\ &= \rho^{-1} b' D \int_0^\infty e^{-\rho\tau} y(\tau) d\tau. \end{aligned}$$

Examining the objective of DP we see that we can incorporate discounted inventory costs by adding the term $\rho^{-1}b'D$ to b .

The problem of FMOS is to select a policy, that is, operator-assignments and achievable activity-rates over time which maintain nonnegative inventory levels, with the purpose of minimizing the discounted operator-assignment and activity-rates costs. This corresponds to finding a policy π for DP with $V_\pi = V_{DP}$. An interpretation of the units of the data and variables is given in the table:

| <u>Object</u> | <u>Units</u> |
|---------------|---|
| ρ | percent per unit time |
| a | \$ per assignment per unit time |
| b | \$ per activity rate per unit time |
| C | activity rate per assignment |
| D | units of in-process goods per unit time per activity rate |
| x | assignment |
| y | activity rate |

This completes our argument that DP is a dynamic programming formulation for FMOS.

We next return to the example presented in the Introduction and show how it is formulated by DP.

Buy-Sell Example for DP:

Let $k = 2$, $m = 1$, $n = 2$, $\rho = 0.2$, $a = (0, 0)$, $b = (1, -2)$,

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D = (1, -1), \quad E = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

$\pi = (x(\cdot), y(\cdot))$, $x(\cdot) = (x_1(\cdot), x_2(\cdot))$, $y(\cdot) = (y_1(\cdot), y_2(\cdot))$, $x_i(\cdot):R_+ \rightarrow R^1$ for $i = 1, 2$, and $y_i(\cdot):R_+ \rightarrow R^1$ for $i = 1, 2$. The dynamic program DP is then given by:

$$DP: \quad V_{DP} \equiv \inf_{\pi=(x,y)} \int_0^\infty e^{-\tau/5} [(0, 0)x(\tau) + (1, -2)y(\tau)] d\tau$$

$$\text{subject to: } x(t) \in E = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \text{ for all } t \in R_+ \quad (\text{DP1})$$

$$0 \leq y(t) \leq x(t) \text{ for all } t \in R_+ \quad (\text{DP2})$$

$$\int_0^t (1, -1)y(\tau)d\tau \geq 0 \text{ for all } t \in R_+. \quad \otimes \quad (\text{DP3})$$

The following lemma is used to derive implied constraints from DP.

Lemma 2.1.

Let $v: R_+ \rightarrow R$ be a bounded, piecewise continuous function satisfying

$$\int_0^t v(\tau)d\tau \geq 0 \text{ for all } t \in R_+. \quad (2.1)$$

Then for any $\rho > 0$,

$$\int_0^t e^{-\rho\tau}v(\tau)d\tau \geq 0 \text{ for all } t \in R_+. \quad (2.2)$$

Proof.

Let ∇_t denote the left derivative operator with respect to the variable t . Then, $\nabla_t[\int_0^t v(\tau)d\tau] = v(t)$, and the left continuity of $v(\cdot)$ and (2.1) imply that for all $t \in R_+$

$$\begin{aligned} \int_0^t e^{-\rho\tau}v(\tau)d\tau &= \int_0^t e^{-\rho\tau} \{ \nabla_\tau [\int_0^\tau v(\sigma)d\sigma] \} d\tau \\ &= e^{-\rho t} \int_0^t v(\tau)d\tau - \int_0^t [\nabla_\tau (e^{-\rho\tau})] [\int_0^\tau v(\sigma)d\sigma] d\tau \\ &= e^{-\rho t} \int_0^t v(\tau)d\tau + \rho \int_0^t e^{-\rho\tau} [\int_0^\tau v(\sigma)d\sigma] d\tau \geq 0. \quad \otimes \end{aligned}$$

Lemma 2.2.

The constraints of DP imply that

$$\int_0^t e^{-\rho\tau}Dy(\tau)d\tau \geq 0 \text{ for all } t \in R_+. \quad (2.3)$$

Proof.

Apply Lemma 2.1 to each component of the integral. \otimes

We refer to a constraint of the type found in (2.3) as a *discounted inventory constraint*, and we use such constraints in our relaxation RDP of DP.

Definition of RDP:

RDP is the following dynamic program where the search takes place over policies $\pi = (x(.), y(.))$:

$$\text{RDP: } V_{\text{RDP}} \equiv \inf_{\pi=(x(.), y(.))} \int_0^{\infty} e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$$

$$\text{subject to: } x(t) \in E \text{ for all } t \in R_+, \quad (\text{DP1})$$

$$0 \leq y(t) \leq Cx(t) \text{ for all } t \in R_+, \quad (\text{DP2})$$

$$\int_0^{\Delta s} Dy(\tau) d\tau \geq 0 \text{ for all } s \in Z_+, \text{ and} \quad (\text{DP3A})$$

$$\int_0^{\Delta s} e^{-\rho\tau} Dy(\tau) d\tau \geq 0 \text{ for all } s \in Z_+. \quad (\text{DP3B})$$

Define $\pi = (x(.), y(.))$ to be a *policy for RDP* if it is a policy satisfying the constraints of RDP and let $V_{\pi} \equiv \int_0^{\infty} e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$ denote the *objective value* of such a policy. The problem in RDP is to select an optimal policy for RDP, i.e., select a policy π for RDP which infimizes the objective value, namely, has $V_{\pi} = V_{\text{RDP}}$. \otimes

We observe that Lemma 2.2 implies that DP can be augmented by the discounted inventory constraints without altering the set of policies. As (DP3A) and (DP3B) are, respectively, the restriction of (DP3) and (2.3) to times $t = s\Delta$ for $s \in Z_+$, RDP is a relaxation of DP. In particular, we have proved the following lemma:

Lemma 2.3.

Each policy π of DP is a policy of RDP and $V_\pi \geq V_{\text{RDP}}$. Thus, if DP is feasible, then RDP is feasible, and further, the optimal objective value of DP is not less than that RDP, i.e., $V_{\text{DP}} \geq V_{\text{RDP}}$. \otimes

Though the discounted inventory constraints (2.3) are implied by the constraints of DP, they are not necessarily implied by the constraints of RDP. We do not expect to be able to use the next corollary, but we include it to maintain uniform development.

Corollary 2.4.

If π is a policy of DP with $V_\pi = V_{\text{RDP}}$, then π is an optimal policy for DP. \parallel

We do not know how to solve DP, but we can and do solve RDP. A policy π is defined to be Δ -stationary if $\pi(t) = \pi(t + \Delta)$ for all $t \in \mathbb{R}_+$. We shall compute an optimal policy for RDP which is Δ -stationary.

Now, let us indicate the correspondence between RDP and RFMOS. The period length of RFMOS is Δ . That inventory levels and discounted inventory levels be nonnegative at the end of each periods is captured by the constraints (DP3A) and (DP3B), respectively. Otherwise, the arguments that showed that DP models FMOS can be used to show that RDP models RFMOS. An interpretation of the discounted inventory constraints is given in Section 6.

Buy-Sell Example for RDP:

We continue to use the data, augment by $\Delta = 1$, and the notation for policies from the Buy-Sell Example for DP. The dynamic program RDP is then given by:

$$\text{RDP: } V_{\text{RDP}} \equiv \inf_{\pi=(x(\cdot), y(\cdot))} \int_0^\infty e^{-\tau/5} [(0, 0)x(\tau) + (1, -2)y(\tau)] d\tau$$

$$\text{subject to: } x(t) \in E = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \text{ for all } t \in \mathbb{R}_+ \quad (\text{DP1})$$

$$0 \leq y(t) \leq x(t) \text{ for all } t \in \mathbb{R}_+ \quad (\text{DP2})$$

$$\int_0^s (1, -1)y(\tau)d\tau \geq 0 \text{ for all } s \in Z_+, \text{ and} \quad (\text{DP3A})$$

$$\int_0^s e^{-\tau/5}(1, -1)y(\tau)d\tau \geq 0 \text{ for all } s \in Z_+ . \otimes \quad (\text{DP3B})$$

3. Auxiliary Programs

Towards our the solution of RDP, and hence of RFMOS, we define a dynamic program RRDP which is a relaxation of RDP and a linear program LP. The dynamic program RRDP is obtained by replacing (DP3A) and (DP3B) in RDP by a single constraint which, as the next lemma shows, is implied by the constraints of RDP. Although RRDP is a relaxation of RDP, as we shall see, certain optimal solutions for RRDP are optimal for RDP as well.

Lemma 3.1.

The constraints of RDP imply that

$$\int_0^{\infty} e^{-\rho\tau} Dy(\tau) d\tau \geq 0. \quad (3.1)$$

Proof.

From (DP1) and (DP2), we see that $Dy(\cdot)$ is bounded, that is, for some K we have that $\|Dy(t)\|_{\infty} \leq K$ for all $t \in \mathbb{R}_+$. Hence,

$$\left\| \int_0^r e^{-\rho\tau} Dy(\tau) d\tau \right\|_{\infty} \leq K \int_0^r e^{-\rho\tau} d\tau = K\rho^{-1} e^{-\rho t} [1 - e^{-\rho(r-t)}] \text{ for all } t \text{ and } r \text{ in } \mathbb{R}_+,$$

implying that the limit of $\int_0^t e^{-\rho\tau} Dy(\tau) d\tau$ exists as $t \rightarrow \infty$, i.e., the integral $\int_0^{\infty} e^{-\rho\tau} Dy(\tau) d\tau$ is well defined. In particular,

$$\lim_{s \rightarrow \infty} \int_0^{s\Delta} e^{-\rho\tau} Dy(\tau) d\tau = \int_0^{\infty} e^{-\rho\tau} Dy(\tau) d\tau,$$

and (3.1) follows from (DP3B) and the boundedness of $Dy(\cdot)$. \otimes

As the constraints of (3.1) are instances of those in (2.3) with $t = \infty$, we will also refer to the constraints of (3.1) as *discounted inventory constraints*. RRDP is next defined as the relaxation of RDP obtained by replacing the family of constraints (DP3A) and (DP3B) by the single constraint (3.1). We will construct optimal solutions of RRDP which will satisfy the

constraints of RDP, hence, as RRDP is a relaxation of RDP, such solutions will also be optimal for RDP.

Definition of RRDP:

RRDP is the following dynamic program where the search takes place over policies $\pi = (x(.), y(.))$:

$$\text{RRDP: } V_{\text{RRDP}} \equiv \inf_{\pi = (x(.), y(.))} \int_0^{\infty} e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$$

$$\text{subject to: } x(t) \in E \text{ for all } t \in \mathbb{R}_+ \quad (\text{DP1})$$

$$0 \leq y(t) \leq Cx(t) \text{ for all } t \in \mathbb{R}_+, \text{ and} \quad (\text{DP2})$$

$$\int_0^{\infty} e^{-\rho\tau} Dy(\tau) d\tau \geq 0. \quad (\text{DP3C})$$

Define $\pi = (x(.), y(.))$ to be a *policy* for RRDP, if it is a policy satisfying the constraints of RRDP and let $V_{\pi} \equiv \int_0^{\infty} e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$ denote the *objective value* of such a policy. The problem in RRDP is to select an optimal policy for RRDP, i.e., select a policy π for RRDP which infimizes the objective value, namely, has $V_{\pi} = V_{\text{RRDP}}$. \otimes

Lemma 3.2.

Each policy π of RDP is a policy of RRDP with $V_{\pi} \geq V_{\text{RRDP}}$. Thus, if RDP is feasible, then RRDP is feasible, and further the optimal objective value of RDP is not less than that of RRDP, i.e., $V_{\text{RDP}} \geq V_{\text{RRDP}}$. \otimes

Lemma 3.2 provides a simple sufficient test for optimality of policies of RDP.

Corollary 3.3.

If a policy π of RDP satisfies $V_{\pi} = V_{\text{RRDP}}$, then π is an optimal policy for RDP. \otimes

Buy-Sell Example for RRDP

We continue to use the data and notation for policies from the Buy-Sell Example for RDP.

The dynamic program RRDP is then given by:

$$\text{RRDP: } V_{\text{RRDP}} \equiv \inf_{\pi=((x(.), y(.)))} \int_0^{\infty} e^{-\tau/5} [(0, 0)x(\tau) + (1, -2)y(\tau)] d\tau$$

$$\text{subject to: } x(t) \in E = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \text{ for all } t \in \mathbb{R}_+ \quad (\text{DP1})$$

$$0 \leq y(t) \leq x(t) \text{ for all } t \in \mathbb{R}_+ \quad (\text{DP2})$$

$$\int_0^{\infty} e^{-\tau/5} (1, -1)y(\tau) d\tau \geq 0. \quad \otimes \quad (\text{DP3C})$$

In solving RDP and RRDP we employ the following linear program LP.

Definition of LP:

The decision variables of LP are vectors x in \mathbb{R}^k and y in \mathbb{R}^n . LP is the following linear program:

$$\text{LP: } V_{\text{LP}} \equiv \text{minimum}_{(x,y)} \rho^{-1}(ax + by)$$

$$\text{subject to: } x \in \text{conv}(E) \quad (\text{LP1})$$

$$0 \leq y \leq Cx, \text{ and} \quad (\text{LP2})$$

$$\rho^{-1}Dy \geq 0, \quad (\text{LP3})$$

where the notation $\text{conv}(E)$ is used to denote the convex hull of E . We call the pair of vectors (x, y) a *feasible solution* for LP, if it satisfies the constraints of LP and we let $V_{xy} \equiv \rho^{-1}(ax + by)$ denote the *objective value* of such a feasible solution. The problem in LP is to select an optimal policy for LP, i.e., select a feasible solution (x, y) for LP which minimizes the objective value, namely, has $V_{xy} = V_{\text{LP}}$. \otimes

The ρ^{-1} 's which appear in LP could, of course, be deleted without altering the sets of feasible and optimal solutions. We include the ρ^{-1} 's for subsequent notational convenience.

Lemma 3.4.

Let $\pi = (x(\cdot), y(\cdot))$ be a policy for RRDP. Then

$$x \equiv \rho \int_0^\infty e^{-\rho\tau} x(\tau) d\tau \text{ and } y \equiv \rho \int_0^\infty e^{-\rho\tau} y(\tau) d\tau$$

is a feasible solution to LP. Further, $V_\pi = V_{xy} \geq V_{LP}$.

Proof.

From (DP1) and (DP2) the policy $(x(\cdot), y(\cdot))$ are bounded, hence, x and y are well defined and finite. Next observe that $\rho \int_0^\infty e^{-\rho\tau} d\tau = 1$, hence, (DP1), (DP2) and (DP3C) imply that x and y satisfy (LP1) - (LP3). Further,

$$V_\pi \equiv \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau = \rho^{-1}(ax + by) = V_{xy} \geq V_{LP},$$

The last inequality following from the feasibility of (x, y) for LP. \otimes

Lemma 3.5

If RRDP is feasible, then LP is feasible. Further the optimal objective value of RRDP is not less than that of LP, i.e., $V_{RRDP} \geq V_{LP}$. \otimes

Proof.

It follows directly from Lemma 3.4 that if RRDP is feasible, then LP is feasible. Also, as the inequality $V_\pi \geq V_{LP}$ was established in Lemma 3.4 for each policy π of RRDP, we conclude that $V_{RRDP} \geq V_{LP}$. \otimes

Corollary 3.6.

If a policy π for RRDP satisfies $V_\pi = V_{LP}$, then π is an optimal policy for RRDP. \otimes

Buy-Sell Example for LP

We continue to use the data from the Buy-Sell Example for RDP. Let $x = (x_1, x_2)^T$ and $y = (y_1, y_2)^T$ be the decision variables. The linear program LP is then given by:

$$\text{LP: } V_{LP} \equiv \text{minimum}_{(x,y)} 5[(0, 0)x + (1, -2)y]$$

$$\text{subject to: } x \geq 0, (1, 1)x = 1 \quad (\text{LP1})$$

$$0 \leq y \leq x, \text{ and} \quad (\text{LP2})$$

$$5(1, -1)y \geq 0. \quad (\text{LP3})$$

The optimal solution of this linear program is given by $x_1 = x_2 = y_1 = y_2 = 2^{-1}$ and

$$V_{LP} = -2.5. \otimes$$

Call a policy $\pi = (x(\cdot), y(\cdot))$ *stationary* if $x(t)$ and $y(t)$ are invariant with time t .

Generally we do not expect RRDP to have a stationary optimal policy, but to clarify the point of our main effort in the next section, let us spend a moment on this issue. Consider the following discrete program IP.

Definition of IP:

In IP the search is over vectors x in R^k and y in R^n . IP is the following discrete program:

$$\text{IP: } V_{IP} \equiv \text{minimum}_{(x,y)} \rho^{-1}(ax + by)$$

$$\text{subject to: } x \in E \quad (\text{IP1})$$

$$0 \leq y \leq Cx, \text{ and} \quad (\text{LP2})$$

$$5\rho^{-1}Dy \geq 0, \quad (\text{LP3})$$

We call the pair of vectors (x, y) a *feasible solution* for IP if it satisfies the constraints of IP and we let $V_{xy} \equiv p^{-1}(ax + by)$ denote the *objective value* of such a feasible solution. The problem in IP is to select an optimal solution for LP, i.e., select a feasible solution (x, y) for IP which minimizes the objective value, namely, has $V_{xy} = V_{IP}$. \otimes

It can easily be verified that $\pi = (x(\cdot), y(\cdot))$ is a stationary policy for RRDP, if and only if $\pi(t) = (x, y)$ for all $t \in R_+$ and (x, y) is feasible for IP, in which case $V_\pi = V_{xy}$, see Lemma 3.4. Indeed, the best stationary policy is obtained by solving IP for an optimal solution (x, y) with objective value V_{IP} and forming a policy $\pi = (x(\cdot), y(\cdot))$ with $(x(t), y(t)) = (x, y)$ for all $t \in R_+$. Such a policy will have $V_\pi = V_{IP}$. We observe that IP need not be feasible when LP is, and even when both IP and LP are feasible it is typical that V_{IP} exceeds V_{LP} . We will construct a single policy π for both RRDP and RDP with objective value $V_\pi = V_{LP}$, and this policy will typically be better than the best stationary policy, that is, $V_\pi < V_{IP}$.

Buy-Sell Example for IP:

We continue to use the data and notation from the Buy-Sell Example for RDP and LP. The discrete program IP is then given by:

$$\text{IP: } V_{IP} \equiv \text{minimum}_{(x,y)} 5[(0, 0)x + (1, -2)y]$$

$$\text{subject to: } x \in E = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad (\text{IP1})$$

$$0 \leq y \leq x, \text{ and} \quad (\text{LP2})$$

$$5(1, -1)y \geq 0. \quad (\text{LP3})$$

The optimal solution of this discrete program is given by $x = (1, 0)^T$ or $x = (0, 1)^T$ and $y = (0, 0)$. Further, $0 = V_{IP} > V_{LP} = -2.5$. \otimes

4. Solving the Relaxed Dynamic Programs

We have seen in Lemma 3.4 that policies for RRDP result in feasible solution of LP. In the current section we use feasible solutions of LP to construct policies for RRDP which are also feasible for RDP. In particular, if we apply the construction to an optimal solution of LP we will conclude from Corollaries 3.3 and 3.6 that the resulting policy is optimal for RRDP and RDP.

The following lemma is used in our construction of policies from feasible solutions of LP.

Lemma 4.1.

Let a, b^1, \dots, b^q be nonnegative vectors in R^m with

$$a \leq \sum_{r=1}^q b^r.$$

Then, there exist vectors a^1, \dots, a^q in R^m with $0 \leq a^r \leq b^r$ for $r = 1, \dots, q$ and

$$\sum_{r=1}^q a^r = a.$$

Proof.

The required vectors a^1, \dots, a^q in R^m are selected recursively by

$$a_i^r = \min \{ b_i^r, a_i - \sum_{j=1}^{r-1} a_j^i \}, \quad r = 1, \dots, q \text{ and } i = 1, \dots, m,$$

see the proof and the discussion in Appendix A of Eaves and Rothblum [1988]. \otimes

Given the data and a feasible solution (x, y) for LP, *policy construction* defined below, generates a Δ -stationary policy π for RRDP and RDP with objective value V_{xy} .

Policy Construction.

Let (x, y) be a feasible solution of LP. As x is in the convex hull of E , there is a positive integer q , vectors x^1, \dots, x^q in E and positive scalars $\lambda_1, \dots, \lambda_q$ with $\sum_{i=1}^q \lambda_i = 1$ and

$$x = \sum_{i=1}^q \lambda_i x^i . \quad (4.1)$$

As $0 \leq y \leq Dx = \sum_{i=1}^q \lambda_i Dx^i$ and $Dx_i \geq 0$, Lemma 4.1 shows the existence of vectors y^1, \dots, y^q in R^n with

$$0 \leq y^i \leq Dx^i \text{ for } i = 1, \dots, q, \text{ and } y = \sum_{i=1}^q \lambda_i y^i. \quad (4.2)$$

Next, using the function T defined in the Appendix, apply Lemma A.1 to obtain the nested intervals

$$[0, \Delta] = I_q \supset I_{q-1} \supset \dots \supset I_1 \supset I_0 = \emptyset ,$$

where $I_i \equiv (t_i, t_i + \Lambda_i \Delta]$, $\Lambda_i = \sum_{j=1}^i \lambda_j$ and $t_i = \Delta T(\rho \Delta, \Lambda_i)$ for $i = 0, \dots, q$.

For $t \in R_+$ define the " Δ -remainder" function $r: R_+ \rightarrow (0, \Delta]$ by

$$\begin{aligned} r(t) &\equiv \text{minimize } t + s\Delta \\ &\text{subject to: } t + s\Delta > 0 . \\ &s \in Z . \end{aligned}$$

Define the function $\pi = (x(.), y(.))$ from R_+ into $R^k \times R^n$ by

$$\pi(t) = (x^i, y^i) \text{ if } r(t) \in I_i \setminus I_{i-1} , \quad (4.3)$$

where $i \in \{1, \dots, q\}$. See Figure 1. \otimes

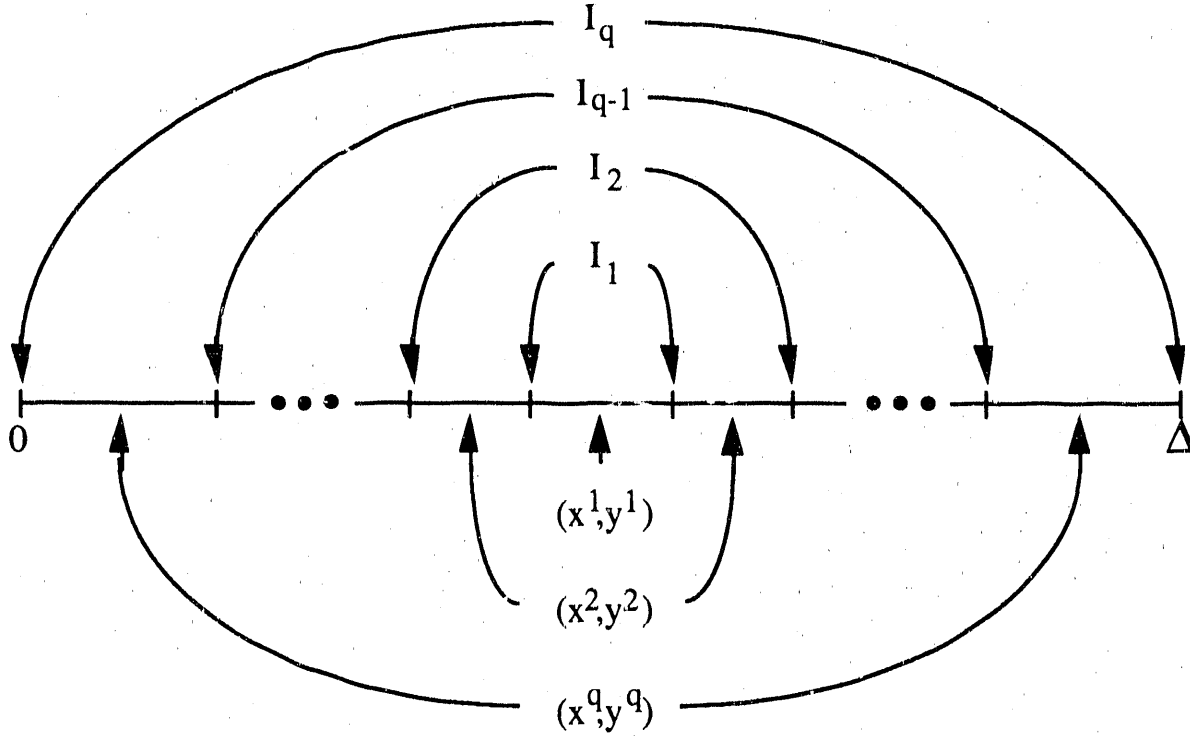


Figure 1

We proceed to show that the object π generated by policy construction from a feasible solution (x, y) of LP is a policy.

Lemma 4.2.

The object π generated by policy construction from a feasible solution (x, y) of LP is a Δ -stationary policy satisfying (DP1) and (DP2).

Proof.

As the intervals I_1, \dots, I_q are nested and $I_q = (0, \Delta]$, for each $t \in R_+$ we have that there exists a unique $i \in \{1, \dots, q\}$ for which $r(t) \in I_i \setminus I_{i-1}$. Hence, π is, in fact, a well defined function from R_+ into $R^k \times R^n$. As the intervals I_1, \dots, I_q are open on the left and closed on the right, π is left continuous. As $r(t + \Delta) = r(t)$ for all $t \in R_+$, we conclude that

π is Δ -stationary. As each $x^i \in E$, we have that $x(t) \in E$ for all $t \in \mathbb{R}_+$. Finally, as $0 \leq y^i \leq Dx^i$, we have that $0 \leq y(t) \leq Dx(t)$ for all $t \in \mathbb{R}_+$. \otimes

Lemma 4.3.

The policy $\pi = (x(\cdot), y(\cdot))$ generated by policy construction from the feasible solution (x, y) of LP has the following properties:

$$\int_t^{t+\Delta} x(\tau) d\tau = \Delta x \text{ for all } t \in \mathbb{R}_+, \quad (4.4)$$

$$\int_t^{t+\Delta} y(\tau) d\tau = \Delta y \text{ for all } t \in \mathbb{R}_+, \quad (4.5)$$

$$\int_{\Delta s}^{\Delta(s+1)} e^{-\rho\tau} x(\tau) d\tau = \rho^{-1} e^{-\rho\Delta s} (1 - e^{-\rho\Delta}) x \text{ for all } s \in \mathbb{Z}_+, \quad (4.6)$$

$$\int_{\Delta s}^{\Delta(s+1)} e^{-\rho\tau} y(\tau) d\tau = \rho^{-1} e^{-\rho\Delta s} (1 - e^{-\rho\Delta}) y \text{ for all } s \in \mathbb{Z}_+, \quad (4.7)$$

$$\int_0^\infty e^{-\rho\tau} x(\tau) d\tau = \rho^{-1} x, \quad (4.8)$$

$$\int_0^\infty e^{-\rho\tau} y(\tau) d\tau = \rho^{-1} y, \quad (4.9)$$

$$\int_{\Delta s}^{\Delta(s+1)} e^{-\rho\tau} Dy(\tau) d\tau \geq 0 \text{ for all } s \in \mathbb{Z}_+, \quad (4.10)$$

$$\int_0^\infty e^{-\rho\tau} Dy(\tau) d\tau \geq 0 \text{ and} \quad (4.11)$$

$$\int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau = V_\pi = V_{xy} \geq V_{LP}. \quad (4.12)$$

Proof.

The definition of the intervals I_0, \dots, I_q and Theorem A.2 imply that

$$\int_{I_j \setminus I_{j-1}} 1 d\tau = \Lambda_j \Delta - \Lambda_{j-1} \Delta = \lambda_j \Delta \quad (4.13)$$

and

$$\rho \int_{I_j N_{j-1}} e^{-\rho \tau} d\tau = (\Lambda_j - \Lambda_{j-1}) \rho \int_0^\Delta e^{-\rho \tau} d\tau = \lambda_j (1 - e^{-\rho \Delta}). \quad (4.14)$$

Using the change of variable $\tau \rightarrow \tau - t$, it follows from (4.13), the Δ -stationarity of π and (4.1) - (4.3) that for all $t \in \mathbb{R}_+$,

$$\int_t^{t+\Delta} x(\tau) d\tau = \int_0^\Delta x(\tau + t) d\tau = \int_0^\Delta x(\tau) d\tau = \sum_{j=1}^q x_j \left(\int_{I_j N_{j-1}} 1 d\tau \right) = \sum_{j=1}^q \lambda_j \Delta x_j = \Delta x$$

and

$$\int_t^{t+\Delta} y(\tau) d\tau = \int_0^\Delta y(\tau + t) d\tau = \int_0^\Delta y(\tau) d\tau = \sum_{j=1}^q y_j \left(\int_{I_j N_{j-1}} 1 d\tau \right) = \sum_{j=1}^q \lambda_j \Delta y_j = \Delta y,$$

establishing (4.4) and (4.5). Similarly, using the change of variable $\tau \rightarrow \tau - s\Delta$, it follows from (4.14) and (4.1) - (4.2) that for all $s \in \mathbb{Z}_+$,

$$\begin{aligned} \int_{\Delta s}^{\Delta(s+1)} e^{-\rho \tau} x(\tau) d\tau &= \int_0^\Delta e^{-\rho(\tau + \Delta s)} x(\tau + \Delta s) d\tau = e^{-\rho \Delta s} \int_0^\Delta e^{-\rho \tau} x(\tau) d\tau \\ &= e^{-\rho \Delta s} \left[\sum_{j=1}^q x_j \left(\int_{I_j N_{j-1}} e^{-\rho \tau} d\tau \right) \right] = \rho^{-1} e^{-\rho \Delta s} \left[\sum_{j=1}^q x_j \lambda_j (1 - e^{-\rho \Delta}) \right] \\ &= \rho^{-1} e^{-\rho \Delta s} (1 - e^{-\rho \Delta}) x, \end{aligned}$$

establishing (4.6). Using a corresponding string of equalities with (4.2) replacing (4.1) we similarly get that

$$\int_{\Delta s}^{\Delta(s+1)} e^{-\rho \tau} y(\tau) d\tau = \rho^{-1} e^{-\rho \Delta s} (1 - e^{-\rho \Delta}) y,$$

establishing (4.7). We next conclude from (4.6) that

$$\begin{aligned} \int_0^\infty e^{-\rho \tau} x(\tau) d\tau &= \sum_{s=0}^\infty \int_{\Delta s}^{\Delta(s+1)} e^{-\rho \tau} x(\tau) d\tau = \sum_{s=0}^\infty \rho^{-1} e^{-\rho \Delta s} (1 - e^{-\rho \Delta}) x = \\ &= \rho^{-1} (1 - e^{-\rho \Delta})^{-1} (1 - e^{-\rho \Delta}) x = \rho^{-1} x, \end{aligned}$$

establishing (4.8) , and similar arguments show that (4.9) follows from (4.7) . Next note that (LP3) implies that $Dy \geq 0$, hence, (4.10) and (4.11) follow directly from (4.7) and (4.9) , respectively. Finally, (4.12) follows from (4.8) and (4.9) , as

$$V_{\pi} = \int_0^{\infty} e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau = \rho^{-1}(ax + by) = V_{xy} \geq V_{LP} . \quad \otimes$$

The major results of this paper are now captured in the following theorem. Indeed, it shows one how to compute a policy which is simultaneously optimal for RRDP and RDP, further, the objective value of the policy is computed.

Theorem 4.4.

Given a feasible solution (x, y) of LP, policy construction yields a Δ -stationary policy π for RRDP and RDP with objective value $V_{\pi} = V_{xy}$. Furthermore, if (x, y) is an optimal solution of LP, then π is an optimal policy for RRDP and RDP .

Proof.

By Lemma 4.2, we see that π is a Δ -stationary policy satisfying (DP1) and (DP2) and (4.11) of Lemma 4.3 shows that π satisfies (DP3C) . Thus, π is a policy for RRDP. Also, (4.5) of Lemma 4.3 and the fact that $Dy \geq 0$ show that for all $r \in \mathbb{Z}_+$,

$$\int_0^{\Delta r} Dy(\tau) d\tau = D \sum_{s=0}^{r-1} \left[\int_{\Delta s}^{\Delta(s+1)} y(\tau) d\tau \right] = r D \Delta y \geq 0 ,$$

hence, (DP3A) is satisfied; further, (4.7) of Lemma 4.3 shows that for all $r \in \mathbb{Z}_+$

$$\begin{aligned} \int_0^{\Delta r} e^{-\rho\tau} Dy(\tau) d\tau &= D \sum_{s=0}^{r-1} \left[\int_{\Delta s}^{\Delta(s+1)} e^{-\rho\tau} y(\tau) d\tau \right] = D \sum_{s=0}^{r-1} [\rho^{-1} e^{-\rho\Delta s} (1 - e^{-\rho\Delta}) y] \\ &= \rho^{-1} (1 - e^{-\rho\Delta r}) Dy \geq 0 , \end{aligned}$$

hence, π satisfies (DP3B) . Thus, we see that π is a policy for RDP as well. The fact that $V_{\pi} = V_{xy}$ follows directly from part (4.12) of Lemma 4.3.

If (x, y) is an optimal solution of LP, then $V_\pi = V_{xy} = V_{LP}$, and Corollary 3.6 immediately implies that π is optimal for RRDP. Similarly, Corollary 3.3 further implies that π is optimal for RDP. \otimes

Theorem 4.4 shows that an optimal policy to RRDP and RDP can be computed by applying policy construction to any optimal solution of LP. In particular, it follows that the optimal objective value of RDP, RRDP and LP coincide. As V_{LP} and V_{RRDP} are independent of Δ , we conclude that so is V_{RDP} . These results are summarized in the next three corollaries.

Corollary 4.5.

One obtains an optimal policy π for RRDP and RDP by computing an optimal solution (x, y) for LP and then applying policy construction to (x, y) . \otimes

Corollary 4.6.

$$V_{RDP} = V_{RRDP} = V_{LP}. \quad \otimes$$

We note that as LP is a linear program, its objective value can be easily computed. Thus, Corollary 4.6 establishes a method for computing V_{RDP} and V_{RRDP} .

Corollary 4.7.

The optimal objective function of RDP is invariant with Δ . \otimes

The last corollary is counter-intuitive (upon first sight), but, it is explained by the absence of setups and inventory costs in RDP. Although the optimal objective of RDP is independent of Δ , the optimal solution of RDP computed by policy construction from a given optimal solution of LP is dependent on Δ .

On another matter, if we replace Δ and ρ by Δ' and ρ' , respectively, where $\rho\Delta = \rho'\Delta'$, notice that the intervals I'_0, I'_1, \dots, I'_q , are the intervals I_0, I_1, \dots, I_q scaled by Δ/Δ' , see the discussion in Section 5.

As RDP is a relaxation of DP we see that V_{LP} is a lower bound for the optimal objective value of DP, that is, $V_{DP} \geq V_{LP}$. We do not know how to construct a feasible policy for DP, however, for any positive ϵ , one can generate a policy π satisfying (DP1) and (DP2) and almost satisfying (DP3), that is,

$$\int_0^t Dy(\tau)d\tau \geq -\epsilon u \text{ for all } t \in R_+,$$

where $u = (1, \dots, 1)^T \in R^n$. Furthermore, $V_\pi \leq V_{DP}$. This policy almost satisfies the constraints of DP and has objective value which cannot be improved upon by a policy for DP. To obtain such a policy first select Δ so that $\Delta Dy \geq -\epsilon u$ for all y satisfying $0 \leq y \leq Cx$ for all $x \in E$. Then apply policy construction to an optimal solution (x, y) of LP to get the policy π . The trouble is that as ϵ tends to zero, so does Δ , and the desired policy stutters, that is jumps around from one pair (x, y) to another faster and faster as Δ tends to zero. All this leads one to suspect that DP, in interesting cases, will not have an optimal policy. Once again the culprit is the absence of setups and inventory costs.

In policy construction the pairs $(x^1, y^1), \dots, (x^q, y^q)$ can be ordered arbitrarily, and each order generates a different optimal policy. In fact, the order could be of considerable consequence, for example, in reducing stockouts. The multitude of optimal policies is largely explained by the absence of inventory costs in RDP. See other discussion of this matter in Eaves and Rothblum [1988].

Buy-Sell Example for RRDP and RDP, the Solution:

We continue to use the data and notation of the Buy-Sell Example for RDP and LP. The optimal solution of LP has $x = (x_1, x_2) = (2^{-1}, 2^{-1})$, $y = (y_1, y_2) = (2^{-1}, 2^{-1})$ and $V_{LP} = -2.5$. Using policy construction we represent x by $x = 2^{-1}(1, 0)^T + 2^{-1}(0, 1)^T$, thus, $q = 2$, $\lambda_1 = 2^{-1}$, $\lambda_2 = 2^{-1}$, $x^1 = (1, 0)^T$ and $x^2 = (0, 1)^T$, see Figure 2. Further, we compute $y^1 = (1, 0)^T$, $y^2 = (0, 1)^T$ and $\Delta T(p\Delta, \lambda_1) = T(5^{-1}, 2^{-1}) = .243\dots$. Thus, $I_1 = [.243\dots, .743\dots]$

and $I_2 = [0, 1]$. We can now define an optimal policy $\pi = (x(\cdot), y(\cdot))$ in the following way: for $s \in \mathbb{Z}_+$, set

$$(x(t), y(t)) = ((1, 0)^T, (1, 0)^T) \text{ for } s + .243.. < t \leq s + .743..$$

and

$$(x(t), y(t)) = ((0, 1)^T, (0, 1)^T) \text{ for } s < t \leq s + .243.. \text{ and } s + .743.. < t \leq s + 1.$$

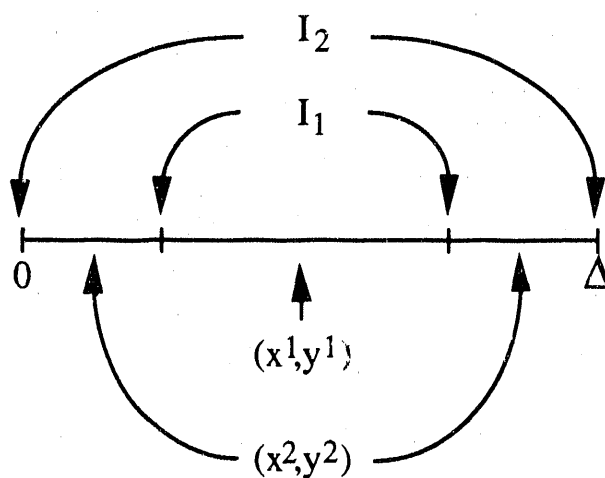


Figure 2

We observe that we could just as well have interchanged the selection of x^1 and x^2 , that is ordered the extreme points of E differently, in which case we would arrive at the alternative optimal solution $\pi' = (x'(\cdot), y'(\cdot))$ defined in the following way: for $s \in \mathbb{Z}_+$, set

$$(x'(t), y'(t)) = ((0, 1)^T, (0, 1)^T) \text{ for } s + .243.. < t \leq s + .743, \text{ and}$$

and

$$(x'(t), y'(t)) = ((1, 0)^T, (1, 0)^T) \text{ for } s < t \leq s + .243.. \text{ and } s + .743.. < t \leq s + 1. \otimes$$

Buy-Sell Example for RFMOS, the Solution:

The above solution of RRDP and RDP for the Buy-Sell Example generates two optimal policies π and π' . These solutions represent the sequence of activities "sell, buy, sell" and "buy, sell, buy", respectively, over shifts of the unit interval with breakpoints 0, .243..., .743..., 1. The maximum short positions of inventories are different for the two policies and equal .243.. for π and .256 for π' . The common objective value of these policies is -2.5 . \otimes

5. Sensitivity with the Interest Rate

In this section we investigate sensitivity of DP, RDP and RRDP on the interest rate and the period length. In particular, we obtain expressions for the dependence of policies and their objective values on the interest rate. We shall use suffixes " (ρ) " and " (ρ, Δ) " to indicate dependence on the interest rate ρ and the period length Δ , e.g., we refer to $DP(\rho)$, $RDP(\rho, \Delta)$, $RRDP(\rho)$ and $V(\rho)_\pi$.

For a given policy $\pi = (x(\cdot), y(\cdot))$ and $\alpha > 0$, define the time scaled policy $\pi\alpha \equiv (x\alpha(\cdot), y\alpha(\cdot))$ by $x\alpha(t) = x(\alpha t)$ and $y\alpha(t) = y(\alpha t)$ for all $t \in \mathbb{R}_+$. The next lemma follows by simple change of variables in the corresponding integrals.

Lemma 5.1..

Let ρ , Δ and α be positive numbers. Then:

- a) a policy π is feasible for $DP(\rho)$, if and only if $\pi\alpha$ is feasible for $DP(\alpha\rho)$,
- b) a policy π is feasible for $RDP(\rho, \Delta)$, if and only if $\pi\alpha$ is feasible for $RDP(\alpha\rho, \alpha^{-1}\Delta)$,
- c) a policy π is feasible for $RRDP(\rho)$, if and only if $\pi\alpha$ is feasible for $RRDP(\alpha\rho)$,
and
- d) for every policy π , $V(\rho)_\pi = \alpha V(\alpha\rho)_{\pi\alpha}$. \otimes

Corollary 5.2.

Let ρ , Δ and α be positive numbers. Then a policy π is optimal for $DP(\rho)$, $RDP(\rho, \Delta)$ and $RRDP(\rho)$ if and only if $\pi\alpha$ is optimal for $DP(\alpha\rho)$, $RDP(\alpha\rho, \alpha^{-1}\Delta)$ and $RRDP(\alpha\rho)$, respectively. \otimes

Select (ρ, Δ) and (ρ', Δ') with $\rho\Delta = \rho'\Delta'$. If policy construction generates the policy π for (ρ, Δ) , then policy construction can be used to generate $\pi(\Delta' / \Delta)$ for (ρ', Δ') , see Section 4. The above corollary shows that such time scaling is useful whether or not the policy is generated by policy construction.

Suppose we wanted to change the time scale and convert from one time unit to another, e.g., from months to years, where θ time units in the old system becomes 1 unit in the new system. Then the numerical value of the data is modified as follows: $\rho \leftarrow \theta\rho$, $a \leftarrow \theta a$, $b \leftarrow b$, $C \leftarrow \theta C$ and $D \leftarrow D$. We note that if $\pi = (x(\cdot), y(\cdot))$ is an optimal policy for the old data, then $\pi = (x(\theta\cdot), \theta y(\theta\cdot))$ is optimal for the new data.

6. The Discounted Inventory Constraints

Our next task is to give the discounted inventory constraints an interpretation. On the one hand, as the discounted inventory constraints are implied by the constraints of DP, they require no justification. On the other hand, to help appreciate the severity, or lack thereof, of the relaxation in moving from DP to RDP or RRDP we do want to understand the discounted inventory constraints. However, let us not forget that the most damaging aspect of the relaxation is that we lose inventory costs in the objective.

To pursue this matter we focus on the single discounted inventory constraint at infinity as in RRDP, rather than the periodic discounted inventory constraints as in RDP. Let us create a parametric version RRDP(p) of RRDP by moving the discounted inventory constraints into the objective with Lagrangian multiplier $p \in \mathbb{R}^{1 \times m}$. We shall interpret the multipliers as prices.

Definition of RRDP(p):

RRDP(p) is the following parametric dynamic program where the search takes place over policies $\pi = (x(\cdot), y(\cdot))$:

$$\text{RRDP}(p) : V_{\text{RRDP}(p)} \equiv \inf_{\pi = (x(\cdot), y(\cdot))} \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau - p \int_0^\infty e^{-\rho\tau} Dy(\tau) d\tau$$

$$\text{subject to: } x(t) \in E \text{ for all } t \in \mathbb{R}_+ \quad (\text{DP1})$$

$$0 \leq y(t) \leq Cx(t) \text{ for all } t \in \mathbb{R}_+ \quad (\text{DP2})$$

Define $\pi = (x(\cdot), y(\cdot))$ to be a *policy* for RRDP(p), if it is a policy satisfying the constraints of RRDP(p) and let $V_\pi(p) \equiv \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau - p \int_0^\infty e^{-\rho\tau} Dy(\tau) d\tau$ denote the *objective value* of such a policy. The problem in RRDP(p) is to select an optimal policy for RRDP(p), i.e., select a policy π for RRDP(p) which infimizes the objective value, namely, has $V_\pi(p) = V_{\text{RRDP}(p)}$. \otimes

Define $Y(t) \equiv \int_0^t y(\tau) d\tau$ and observe that $DY(t)$ is the inventory on hand at time t .

Using integration by parts, we may rewrite the objective function of RRDP(p) as

$$V_{\pi}(p) = \int_0^{\infty} e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau - \int_0^{\infty} e^{-\rho\tau} [\rho p DY(\tau)] d\tau .$$

At time t , the inventory levels are $DY(t)$ and at prices given by the coordinates of p the dollar value of the inventory is $pDY(t)$. Now, imagine a market for in-process goods which pays an interest rate ρ for the dollar value of in-process goods on deposit and charges an interest rate of ρ for each dollar value of in-process goods on loan. The payment rate at time t to the market is $\rho pDY(t)$. The task of $RRDP(p)$ is then to minimize discounted assignment costs plus discounted activity costs plus discounted interest payments to the market of in-process goods (minus corresponding revenues therefrom). We will show that there are prices p such that if a policy is optimal for $RRDP$ then it is optimal for $RRDP(p)$.

As the linear program LP assisted in our analysis of $RRDP$ we now introduce the parametric linear program $LP(p)$ to assist us in our analysis of $RRDP(p)$. As we created $RRDP(p)$ from $RRDP$, we create $LP(p)$ from LP . Specifically, $LP(p)$ is formed from LP by moving the constraints $\rho^{-1}Dy \geq 0$ into the objective via the use of prices given by the coordinates of p .

Definition of $LP(p)$:

The decision variables of $LP(p)$ are vectors x in R^k and y in R^n . $LP(p)$ is the following linear program:

$$LP(p): V_{LP}(p) \equiv \text{minimum}_{(x,y)} \rho^{-1}(ax + by) - \rho^{-1}pDy$$

$$\text{subject to:} \quad x \in \text{conv}(E) \quad (LP1)$$

$$0 \leq y \leq Cx . \quad (LP2)$$

We call the pair of vectors (x,y) a *feasible solution* for $LP(p)$, if it satisfies the constraints of $LP(p)$ and we let $V_{xy}(p) \equiv \rho^{-1}(ax + by) - \rho^{-1}pDy$ denote the *objective value* of such a feasible solution. The problem in $LP(p)$ is to select an optimal policy for $LP(p)$, i.e., select a feasible solution (x, y) for $LP(p)$ which minimizes the objective value, namely, has $V_{xy}(p) = V_{LP}(p)$. \otimes

The relationship between RRDP(p) and LP(p) is essentially as that of RRDP and LP, as the following three lemmas show.

Lemma 6.1.

Let $\pi = (x(\cdot), y(\cdot))$ be a policy for RRDP(p). Then

$$x \equiv \rho \int_0^\infty e^{-\rho\tau} x(\tau) d\tau \text{ and } y \equiv \rho \int_0^\infty e^{-\rho\tau} y(\tau) d\tau$$

is a feasible solution to LP(p). Further, $V_\pi(p) = V_{xy}(p) \geq V_{LP}(p)$ and $p \int_0^\infty e^{-\rho\tau} Dy(\tau) = \rho^{-1} p Dy$.

Proof.

Modify the proof of Lemma 3.4 in the obvious manner. \otimes

Lemma 6.2.

Given a feasible solution (x, y) for LP(p), policy construction yields a policy of RRDP(p) with $V_\pi(p) = V_{xy}(p)$ and $\int_0^\infty e^{-\rho\tau} Dy(\tau) = \rho^{-1} p Dy$.

Proof.

Modify the proof of Lemma 4.2 in the obvious manner. \otimes

Lemma 6.3.

$$V_{RRDP}(p) = V_{LP}(p).$$

Proof.

Lemma 6.1 shows that $V_{RRDP}(p) \geq V_{LP}(p)$ and Lemma 6.2 assures that $V_{RRDP}(p) \leq V_{LP}(p)$. \otimes

The relationship between LP and LP(p) is given in the next lemma.

Lemma 6.4.

The following three statements are equivalent:

- a) (x, y) is an optimal solution for LP and p is part of an optimal solution for the dual of LP corresponding to the constraints $\rho^{-1} p Dy \geq 0$.

b) (x, y) is a feasible solution for LP, (x, y) is an optimal solution for $LP(p)$, $p \geq 0$ and $pDy = 0$.

c) (x, y) is a feasible solution for $LP(p)$, $p \geq 0$ and

$$V_{xy}(p') \leq V_{xy}(p) \leq V_{x'y'}(p) \quad (6.1)$$

for all feasible solutions (x', y') for $LP(p)$ and for all $p' \geq 0$.

Further, if (a), (b) or (c) hold, then $V_{LP} = V_{LP}(p)$.

Proof.

The lemma is a special case of Theorem B.1 of Appendix B. \otimes

Buy Sell Example for $LP(p)$.

We continue to use the data from the Buy-Sell Example for RDP. Let $x = (x_1, x_2)^T$ and $y = (y_1, y_2)^T$ be the decision variables. The parametric linear program $LP(p)$ is then given by:

$$LP(p): V_{LP}(p) \equiv \text{minimum}_{(x,y)} 5[(0, 0)x + (1, -2)y] - p5(1, -1)y$$

$$\text{subject to:} \quad x \geq 0, (1, 1)x = 1 \quad (LP1)$$

$$0 \leq y \leq x, \text{ and} \quad (LP2)$$

The optimal dual variables p of the constraints $5(1, -1)y \geq 0$ of LP are $p = 1.5$ and, indeed, the optimal solution of LP, namely, $x_1 = x_2 = y_1 = y_2 = 2^{-1}$ is also optimal for $LP(1.5)$. The optimal objectives of LP and $LP(1.5)$ are equal, indeed, $V_{LP} = V_{LP}(1.5) = -2.5$. \otimes

We have seen in Sections 3 and 4 that RRDP and LP are closely related, namely, given a policy for RRDP we can obtain a feasible solution of LP, and vice versa, with coinciding objective values; $V_{RRDP} = V_{LP}$. Similarly, we have seen in the current section that $RRDP(p)$ and $LP(p)$ are closely related, namely, given a policy for $RRDP(p)$ we can obtain a feasible solution of $LP(p)$, and vice versa, with coinciding objective values; $V_{RRDP}(p) = V_{LP}(p)$. We next relate $RRDP$ and $RRDP(p)$ through LP and $LP(p)$ using the previous lemma.

Theorem 6.5.

The following four statements are equivalent:

a) π is an optimal policy for RRDP and p is part of an optimal solution of the dual of LP corresponding to the constraints $\rho^{-1}Dy \geq 0$.

b) $\pi = (x(\cdot), y(\cdot))$ is a policy for RRDP, π is an optimal policy for RRDP(p), $p \geq 0$ and $p \int_0^\infty e^{-\rho\tau} Dy(\tau) = 0$.

c) π is a policy for RRDP(p), $p \geq 0$ and

$$V_{\pi}(p') \leq V_{\pi}(p) \leq V_{\pi'}(p) \quad (6.2)$$

for all policies π' for RRDP(p) and for all $p' \geq 0$.

d) π is a policy for RRDP(p), $p \geq 0$ and

$$\sup_{\pi'} \inf_{\pi'} V_{\pi'}(p') = V_{\pi}(p) = \inf_{\pi'} \sup_{\pi'} V_{\pi'}(p')$$

where the suprema are taken over the set of policies of RRDP(p) and the infima are taken over the set of $p' \geq 0$.

Further, if (a), (b), (c) or (d) hold, then $V_{RRDP} = V_{RRDP}(p)$.

Proof.

Let $\pi = (x(\cdot), y(\cdot))$ be a policy and define $x \equiv \rho \int_0^\infty e^{-\rho\tau} x(\tau) d\tau$ and $y \equiv \rho \int_0^\infty e^{-\rho\tau} y(\tau) d\tau$.

The proof of the equivalence of (a) - (c) proceeds via a reduction to Lemma 6.4.

(a) \Rightarrow (b): Assume that (a) holds. As π is an optimal policy for RRDP, Corollary 4.6 and Lemma 3.4 show that $V_{LP} = V_{RRDP} = V_{\pi} = V_{xy}$; hence, (x, y) is optimal for LP. Thus, (a) here implies (a) of Lemma 6.5, and we conclude from that lemma that (x, y) is a feasible solution for LP, (x, y) is an optimal solution for LP(p), $p \geq 0$ and $pDy = 0$. In particular, $\int_0^\infty e^{-\rho\tau} Dy(\tau) d\tau = \rho^{-1}Dy \geq 0$ assuring that π is a policy for RRDP, $p \geq 0$ and $p \int_0^\infty e^{-\rho\tau} Dy(\tau) d\tau = pDy = 0$. Thus, it remains to show that π is an optimal policy for RRDP(p). Observing that $V_{\pi} = V_{xy}$ we get from $pDy = 0$, the optimality of (x, y) for LP(p) and Lemma 6.3 that, indeed,

$$V_{\pi} = V_{xy} = V_{xy} - p\rho^{-1}Dy = V_{xy}(p) = V_{LP}(p) = V_{RRDP}(p) .$$

(b) \Rightarrow (c): Assume that (b) holds. Then the right inequalities of (6.2) follows from the optimality of π for $RRDP(p)$. Further, as π is a policy for $RRDP$, the term $p' \int_0^{\infty} e^{-\rho\tau} Dy(\tau) d\tau$ is nonnegative for all $p' \geq 0$. As our assumption asserts that this term is zero for $p' = p$, we get the left inequalities of (6.2) .

(c) \Rightarrow (a): Assume that (c) holds. Lemmas 6.2 shows that each feasible solution (x', y') of $LP(p)$ can be used to construct a policy π' for $RRDP(p)$ with $V_{x'y'}(p) = V_{\pi'}(p)$. As $V_{xy}(p) = V_{\pi}(p)$, the right inequalities of (6.2) imply the right inequalities of (6.1) . Further, the right inequalities of (6.2) show that for every $p' \geq 0$

$$\rho^{-1}p'Dy = p' \int_0^{\infty} e^{-\rho\tau} Dy(\tau) d\tau \geq p \int_0^{\infty} e^{-\rho\tau} Dy(\tau) d\tau = \rho^{-1}pDy ,$$

implying the left inequalities of (6.1) . Hence, part (c) of Lemma 6.4 holds and that lemma implies that (x, y) is an optimal solution for LP and p is part of an optimal solution for the dual of LP corresponding to the constraints $\rho^{-1}Dy \geq 0$. In particular, $V_{\pi} = V_{xy} = V_{LP}$ and Corollary 3.6 implies that π is optimal for $RRDP$.

(c) \Leftrightarrow (d): This equivalence is a standard characterization of saddle points of arbitrary bivariate functions.

Finally, assume that (a) holds. Then $V_{\pi} = V_{RRDP}$. As the proof of the implication (a) \Rightarrow (b) shows that $V_{\pi} = V_{RRDP}(p)$ we have that $V_{RRDP} = V_{RRDP}(p)$. \otimes

Buy Sell Example for $RRDP(p)$.

We continue to use the data from the Buy-Sell Example for RDP . We have seen that $p = 1.5$ is the optimal dual variables of LP corresponding to the constraints $\rho^{-1}Dy \geq 0$. Thus, optimal policies for the Buy-Sell Example for $RRDP$ and RDP , e.g., those computed in Section 4, are optimal for $RRDP(1.5)$ with the same optimal objective value, namely, -2.5 . \otimes

From (d) of Theorem 6.5 we see that the prices for which $RRDP(p)$ and $RRDP$ have a common optimal solution are as unfavorable the manufacturer could face. Thus one could think of

the discounted inventory constraint as one which discourages the manufacturer from arbitraging in the in-process goods market.

7. Remarks and Extensions

An understanding of the dynamic program DP is outside our present reach. For the very simple case of the Buy-Sell Example of DP, we show that there is no optimal policy, but that stuttering policies can be ϵ -optimal. Although this is a very special case of DP, it is not pathological, and we hasten to speculate that it is quite representative or indicative of the typical nature of DP.

Buy-Sell Example of DP, an Analysis:

We continue to use the data and notation for polices of the Buy-Sell Example for DP. Lemma 2.3, Corollary 4.6 and the solution of RDP and RRDP for the Buy-Sell Example show that $V_{DP} \geq V_{RDP} = V_{LP} = 2.5$. We next demonstrate that this lower bound on V_{DP} is not attainable by policies of DP.

Assume that for a policy $\pi = (x(\cdot), y(\cdot))$, $V_\pi = -2.5$. Then,

$$\begin{aligned} -2.5 &= V_\pi = \int_0^\infty e^{-\tau/5} (1, -2) y(\tau) d\tau \\ &= (2, -2) \int_0^\infty e^{-\tau/5} y(\tau) d\tau - \int_0^\infty e^{-\tau/5} y_1(\tau) d\tau \\ &= (1, -1) \int_0^\infty e^{-\tau/5} y(\tau) d\tau - \int_0^\infty e^{-\tau/5} y_2(\tau) d\tau. \end{aligned}$$

Averaging, we get that

$$\begin{aligned} -2.5 &= 1.5(1, -1) \int_0^\infty e^{-\tau/5} y(\tau) d\tau - \int_0^\infty e^{-\tau/5} [0.5y_1(\tau) + 0.5y_2(\tau)] d\tau \\ &= 1.5(1, -1) \int_0^\infty e^{-\tau/5} y(\tau) d\tau - 0.5 \int_0^\infty e^{-\tau/5} (1, 1) y(\tau) d\tau \end{aligned} \quad (6.1)$$

Integrating by parts, we see that $\int_0^\infty e^{-\tau/5} y(\tau) d\tau = 5 \int_0^\infty e^{-\tau/5} [\int_0^\tau y(\sigma) d\sigma] d\tau$, and using (DP1) and (DP2) to get $(1, 1)y(\tau) \leq (1, 1)y(\tau) \leq 1$ we also see that $\int_0^\infty e^{-\tau/5} (1, 1)y(\tau) d\tau = \int_0^\infty e^{-\tau/5} d\tau = 5$. Hence, we obtain from (6.1) that

$$-2.5 \geq 0.3 \int_0^\infty e^{-\tau/5} [(1, -1) \int_0^\tau y(\sigma) d\sigma] d\tau - 2.5. \quad (6.1)$$

From (DP3) $[(1, -1) \int_0^\tau y(\sigma) d\sigma] \geq 0$ for all $\tau \in \mathbb{R}_+$. Hence, (6.1) implies that $[(1, -1) \int_0^\tau y(\sigma) d\sigma] = 0$ for all $\tau \in \mathbb{R}_+$, implying that $y_1(\sigma) - y_2(\sigma) = (1, -1)y(\sigma) = 0$ for all $\sigma \in \mathbb{R}_+$. Further, from (DP1) and (DP2), $0 \leq y_1(t)y_2(t) \leq x_1(t)x_2(t) = 0$ for all $t \in \mathbb{R}_+$. But $(1, -1)y(t) = 0$ and $y_1(t)y_2(t) = 0$, imply that $y_1(t) = y_2(t) = 0$. Hence, $y_1(t) = y_2(t) = 0$ for all $t \in \mathbb{R}_+$, implying that $V_\pi = 0$, a contradiction to the assertion that $V_\pi = -2.5$. Hence, we may conclude that $V_\pi > -2.5$ for all policies π for DP.

We next show that there are policies for DP whose objective value gets arbitrarily close to the unattainable bound -2.5. Let Δ be a positive number. Then, the "buy it then sell it" policy $\pi(\Delta) = ((x(.), y(.)))$, defined by

$$x_1(t) - 1 = x_2(t) = y_1(t) - 1 = y_2(t) = 0 \text{ if } s\Delta \leq t < (s + 2^{-1})\Delta \text{ for some } s \in \mathbb{Z}_+, \text{ and}$$

$$x_2(t) - 1 = x_1(t) = y_2(t) - 1 = y_1(t) = 0 \text{ if } (s + 2^{-1})\Delta \leq t < (s + 1)\Delta \text{ for some } s \in \mathbb{Z}_+$$

satisfies (DP1) - (DP3) and its objective value is easily computed and is given by

$$V_{\pi(\Delta)} = -5(1 - 2e^{-\Delta/10})(1 - e^{-\Delta/10})(1 - e^{-\Delta/5})^{-1}.$$

From L'Hopital's rule it next follows that

$$\lim_{\Delta \rightarrow 0} V_{\pi(\Delta)} = -2.5.$$

Thus, for every positive ϵ there exists a positive Δ for which $V_{\pi(\Delta)} \leq -2.5 + \epsilon$. Combining this fact with the earlier observation that -2.5 is an unattainable lower bound on V_{DP} , we conclude that $V_{DP} = -2.5$ and that each policy of DP is dominated by a policy of the form $\pi(\Delta)$. \otimes

Buy-Sell Example of RDP without (DP3B):

If we drop the constraints (DP3B) from RDP for the Buy-Sell Example for RDP, the "sell it then buy it" policy $\pi = ((x(.), y(.)))$, defined by

$$x_2(t) - 1 = x_1(t) = y_2(t) - 1 = y_1(t) = 0 \text{ if } s \leq t < s + 2^{-1} \text{ for some } s \in Z_+, \text{ and}$$

$$x_1(t) - 1 = x_2(t) = y_1(t) - 1 = y_2(t) = 0 \text{ if } s + 2^{-1} \leq t < s + 1 \text{ for some } s \in Z_+.$$

is optimal. This policy shorts inventory to the extent that is possible while covering the inventory constraint at times $s \in Z_+$. The corresponding optimal objective is then given by $V_\pi = 5(-2 + e^{-1/10})(1 - e^{-1/10})(1 - e^{-1/5})^{-1} = -2.874\dots$. This is an improvement of about 10% over the optimal objective value of RDP, however, the maximal shortage of inventories under this policy, namely, -0.5 , is about double that of the optimal policies of RDP that we constructed in Section 4. \otimes

In our definition of RDP we chose the period lengths to be equal. However, this special choice was motivated by notational convenience and was entirely unnecessary as we can solve the problem for any arbitrary sequence of period lengths. Further, the optimal objective value for the more general problem remains unchanged. To clarify these remarks, let $0 = t(0) < t(1) < \dots$ be any increasing sequence of scalars tending to ∞ and let us define a modification RDPU of the dynamic program RDP with period lengths $t(s+1) - t(s)$ for $s \in Z_+$; the "U" stands for "unequal".

Definition of RDPU:

RDPU is the following dynamic program where the search takes place over policies $\pi = (x(\cdot), y(\cdot))$:

$$\text{RDPU: } V_{\text{RDPU}} \equiv \inf_{\pi=(x(\cdot), y(\cdot))} \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$$

$$\text{subject to: } x(t) \in E \text{ for all } t \in R_+, \quad (\text{DP1})$$

$$0 \leq y(t) \leq Cx(t) \text{ for all } t \in R_+, \quad (\text{DP2})$$

$$\int_0^{t(s)} Dy(\tau) d\tau \geq 0 \text{ for all } s \in Z_+, \text{ and} \quad (\text{DP3AU})$$

$$\int_0^{t(s)} e^{-\rho\tau} Dy(\tau) d\tau \geq 0 \text{ for all } s \in Z_+. \quad (\text{DP3BU})$$

Define $\pi = (x(\cdot), y(\cdot))$ to be a *policy for RDPU* if it is a policy satisfying the constraints of RDPU and let $V_\pi \equiv \int_0^\infty e^{-\rho\tau} [ax(\tau) + by(\tau)] d\tau$ denote the *objective value* of such a policy. The problem in RDPU is to select an optimal policy for RDPU, i.e., select a policy π for RDPU which infimizes the objective value, namely, has $V_\pi = V_{\text{RDPU}}$. \otimes

To solve RDPU we first solve LP and compute an optimal solution (x, y) , and then obtain the $\lambda_1, \dots, \lambda_q, x_1, \dots, x_q, y_1, \dots, y_q$ as in policy construction. Now, we define the function $\pi = (x(\cdot), y(\cdot))$ from R_+ into $R^k \times R^n$. For $s = 0, 1, 2, \dots$ execute the following: Let $\Delta = t(s+1) - t(s)$ and construct the intervals

$$[0, \Delta] = I_q \supset I_{q-1} \supset \dots \supset I_1 \supset I_0 = \emptyset,$$

as in policy construction, and for $t \in (t(s), t(s+1)]$ let

$$\pi(t) = (x^i, y^i) \text{ if } t - t(s) \in I_i \setminus I_{i-1},$$

where $i \in \{1, \dots, q\}$. A modification of the analysis of Section 4 then shows that the policy π thus defined is optimal for RDPU and $V_{\text{RDP}} = V_{\text{RRDP}} = V_{\text{RDPU}}$. To see this we observe that, for example,

$$\int_{t(s)}^{t(s+1)} x(\tau) d\tau = [t(s+1) - t(s)] \text{ for all } s \in Z_+,$$

$$\int_{t(s)}^{t(s+1)} e^{-\rho\tau} x(\tau) d\tau = \rho^{-1} \{e^{-\rho t(s)} - e^{-\rho t(s+1)}\} x \text{ for all } s \in Z_+,$$

$$\int_0^\infty e^{-\rho\tau} x(\tau) d\tau = \rho^{-1} x.$$

The research of this paper can be continued in a number of directions. An especially interesting prospect is the study of DP itself in an attempt to identify ϵ -optimal policies. One approach is to solve the modification of DP where (DP1) is relaxed by the "convexified version

$$x(t) \in \text{conv}(E) \text{ for all } t \in \mathbb{R}_+.$$

This modified version of DP is a "continuous time linear program", see Perold [1978] and Anstreicher [1983]. Now, suppose a solution to this "continuous time linear program" is found. One will then attempt to gain discreteness by some deconvexification scheme that will preserve the inventory availability constraints (DP3) while giving up on ϵ objective value, see the analysis of the Buy Sell Example earlier in this section.

We finally note that due to the close relationship between the dynamic programs and the linear programs sensitivity analysis becomes available. Also, useful information can be obtained from the dual of the linear program LP, see Section 6.

Appendix A

The purpose of this appendix is to derive properties of the real valued function $T(., .)$ defined on $R_{++} \times R_+$ by

$$T(\theta, \Lambda) \equiv \begin{cases} \theta^{-1} \ln[\theta(1 - e^{-\theta})^{-1}] & \text{for } \Lambda = 0 \text{ and } \theta > 0, \text{ and} \\ \theta^{-1} \ln[(1 - e^{-\theta\Lambda})\Lambda^{-1}(1 - e^{-\theta})^{-1}] & \text{for } 0 < \Lambda \leq 1 \text{ and } \theta > 0, \end{cases} \quad (\text{A.1})$$

where $R_{++} \equiv \{\theta \in R : \theta > 0\}$. The following lemma will be useful.

Lemma A.1.

For every $z > 0$ we have $ze^{-z} + e^{-z} < 1$ and $z + e^{-z} > 1$.

Proof.

For the first statement we have equality at $z = 0$, and the derivative of the term on the left hand side of the inequality is negative. For the second statement we have equality at $z = 0$, and the derivative of the terms on the left hand-side of the inequality is positive. \otimes

Theorem A.2.

The function $T : R_{++} \times R_+ \rightarrow R_+$ is continuous.

For positive scalars ρ , Δ and Λ , the value $t = \Delta T(\rho\Delta, \Lambda)$ is the unique solution of

$$\int_t^{t+\Delta\Lambda} \rho e^{-\rho\tau} d\tau = \Lambda \int_0^\Delta \rho e^{-\rho\tau} d\tau. \quad (\text{A.2})$$

For every positive scalar θ , the function $T(\theta, .) : R_+ \rightarrow R_+$ has the following properties:

- a) $T(\theta, \Lambda)$ is strictly decreasing in $\Lambda \in [0, 1]$,
- b) $T(\theta, \Lambda) + \Lambda$ is strictly increasing in $\Lambda \in [0, 1]$,
- c) $T(\theta, 1) = 0$ and $T(\theta, 1) + 1 = 1$, and
- d) $T(\theta, \Lambda) > 0$ and $T(\theta, \Lambda) + \Lambda < 1$ for $0 \leq \Lambda < 1$.

Proof.

The function T is a composition of continuous functions on the set $R_{++} \times R_{++}$ which assures that it is continuous. Next, the continuity of $T(\theta, \cdot)$ at $\Lambda = 0$ follows from L'Hopital's Rule.

Let $\theta = \rho\Delta$. By integrating the left and right hand-side of (A.2) we have that this equation is equivalent to $e^{-\rho t}(1 - e^{-\theta\Lambda}) = \Lambda(1 - e^{-\theta})$. In particular, by taking natural logarithms, we see that for positive Λ , (A.2) is equivalent to the assertion $t = \Delta T(\theta, \Lambda) = \Delta T(\rho\Delta, \Lambda)$.

For positive Λ , the derivative of $T(\theta, \Lambda)$ with respect to Λ is given by

$$\theta^{-1}[(1 - e^{-\theta\Lambda})^{-1}\theta e^{-\theta\Lambda} - \Lambda^{-1}] = (\theta\Lambda)^{-1}(1 - e^{-\theta\Lambda})^{-1}[\Lambda\theta e^{-\theta\Lambda} - (1 - e^{-\theta\Lambda})],$$

and the first inequality of Lemma A.1 with $z = \Lambda\theta$ shows that this derivative is negative, hence (a).

For positive Λ , the derivative of $T(\theta, \Lambda) + \Lambda$ with respect to Λ is given by

$$\begin{aligned} & (\Lambda\theta)^{-1}(1 - e^{-\theta\Lambda})^{-1}[\Lambda\theta e^{-\theta\Lambda} - (1 - e^{-\theta\Lambda})] + 1 \\ &= (\Lambda\theta)^{-1}(1 - e^{-\theta\Lambda})^{-1}[\Lambda\theta e^{-\theta\Lambda} - (1 - e^{-\theta\Lambda}) + \Lambda\theta(1 - e^{-\theta\Lambda})] \\ &= (\Lambda\theta)^{-1}(1 - e^{-\theta\Lambda})^{-1}[-(1 - e^{-\theta\Lambda}) + \Lambda\theta] \end{aligned}$$

and the second inequality of Lemma A.1 with $z = \Lambda\theta$ shows that this derivative is positive; hence (b).

Direct substitution of $\Lambda = 1$ into (A.1) shows that $T(\theta, 1) = 0$, implying that $T(\theta, 1) + 1 = 1$; hence (c).

The conclusions of (d) follows immediately from (b) and (c). \otimes

Remark.

To understand the function T , note that the area under the curve $\rho e^{-\rho t}$ from $\Delta T(\rho\Delta, \Lambda)$ to $\Delta T(\rho\Delta, \Lambda) + \Delta\Lambda$ is Λ times the area under the curve from 0 to Δ , see Figure 3. That is,

for a fixed positive ρ and Δ , the length of the interval of integration and the computed area are both proportional to Λ .

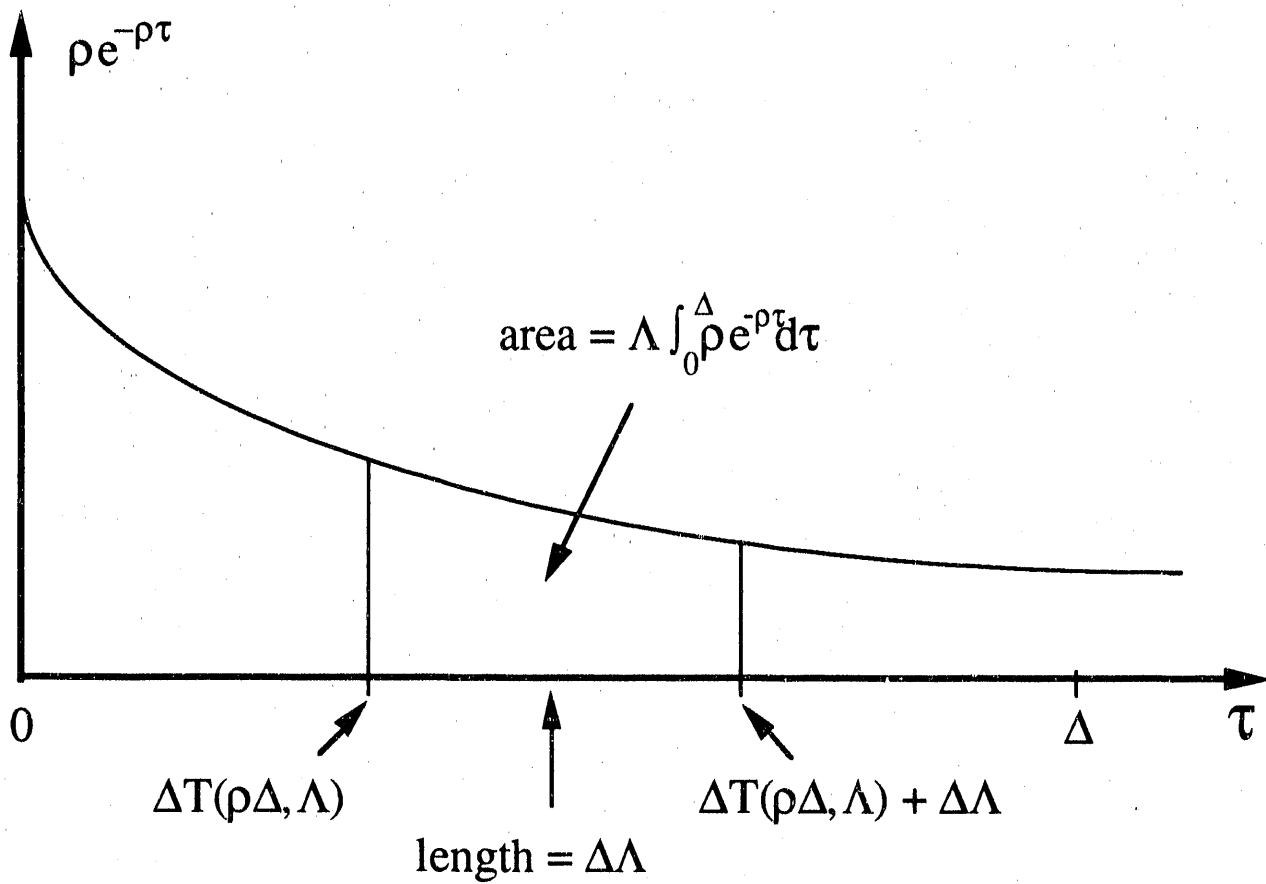


Figure 3

Appendix B

The results needed from linear programming for Lemma 6.4 are recorded here. Throughout this appendix let the data be $a \in R^m$, $b \in R^k$, $c \in R^{1 \times n}$, $A \in R^{m \times n}$ and $B \in R^{k \times n}$, and let the variables be $x \in R^n$, $p \in R^{1 \times m}$, and $q \in R^{1 \times k}$. We define a linear program LP and three other closely related linear programs DLP, LP(p) and DLP(p). DLP is the dual of LP. LP(p) is a parametric version of LP attained by moving the constraints $Ax \geq a$ into the objective. Finally, DLP(p) is the dual of LP(p).

$$\text{LP: } V_{\text{LP}} \equiv \text{minimum}_x cx$$

$$\text{subject to: } Ax \geq a$$

$$Bx \geq b,$$

$$\text{DLP: } V_{\text{DLP}} \equiv \text{maximum}_{p,q} pa + qb$$

$$\text{subject to: } pA + qB = c$$

$$p \geq 0, q \geq 0,$$

$$\text{LP(p): } V_{\text{LP}}(p) \equiv \text{minimum}_x cx - p(Ax - a)$$

$$\text{subject to: } Bx \geq b$$

and

$$\text{DLP(p): } V_{\text{DLP}}(p) \equiv \text{maximum}_q pa + qb$$

$$\text{subject to: } qB = c - pA$$

$$q \geq 0.$$

Note that we have included the constant pa in the objective of both $LP(p)$ and $DLP(p)$. For $x \in \mathbb{R}^n$ and $p \in \mathbb{R}^m$ define $V_x(p) = cx - p(Ax - a)$, see the objective for $LP(p)$.

Theorem B.1.

The following three statements are equivalent:

- a) x is an optimal solution for LP and p is part of an optimal solution for DLP.
- b) x is a feasible solution for LP, x is an optimal solution for $LP(p)$, $p \geq 0$ and $p(Ax - a) = 0$.
- c) x is a feasible solution for $LP(p)$, $p \geq 0$ and

$$V_x(p') \leq V_x(p) \leq V_{x'}(p) \quad (B.1)$$

for all feasible solutions x' for $LP(p)$ and all $p' \geq 0$.

Further, if (a), (b) or (c) hold, then $V_{LP} = V_{LP}(p)$.

Proof.

(a) \Rightarrow (b): Let x and (p, q) be optimal solutions for LP and DLP, respectively. In particular $pA + qB = c$ and, by strong duality, $cx = pa + qb$. It now follows from these facts and the nonnegativity of q , that for every x' which is feasible for $LP(p)$,

$$V_{LP} \leq cx = pa + qb \leq pa + qBx' = pa + (c - pA)x' = cx' - p(Ax' - a) = V_{x'}(p).$$

Hence, $V_{LP} \leq V_{LP}(p)$. By complementary slackness, we have that $p(Ax - a) = 0$ and, hence,

$$V_{LP} = cx = cx - p(Ax - a) = V_x(p) \geq V_{LP}(p).$$

Thus, $V_{LP} = V_x(p) = V_{LP}(p)$, that is, x is an optimal solution to $LP(p)$.

(b) \Rightarrow (c): Assume that (b) holds. As $V_x(p)$ is the objective value of x for $LP(p)$, the right inequalities of (B.1) follow from the fact that x is optimal for $LP(p)$. Also, as $Ax \geq a$, we have that $p'(Ax - a) \geq 0 = p(Ax - a)$ for all $p' \geq 0$, and the left inequalities of (B.1) follow as well.

(c) \Rightarrow (a): The left inequalities of (B.1) imply that $p'(Ax - a) \geq p(Ax - a)$ for all $p' \geq 0$, immediately implying that $Ax \geq a$. Thus, x is a feasible solution for LP. Further, by selecting $p' = 0$ we have that $0 \geq p(Ax - a)$. As $p \geq 0$ and $Ax \geq a$, $p(Ax - a) \geq 0$ and we conclude that $p(Ax - a) = 0$. Now, for any feasible solution x' for LP we have $p(Ax - a) = 0 \leq p(Ax' - a)$. The right inequalities of (B.1) imply that $cx - p(Ax - a) \leq cx' - p(Ax' - a)$. Combining these inequalities, we get that $cx \leq cx'$, establishing the optimality of x for LP. \otimes

References.

- Anstreicher, K.M.[1983], "Generation of feasible descent directions in continuous time linear programming," Technical Report SOL 83-18, Department of Operations Research, Stanford University.
- Eaves, B.C. and U.G. Rothblum [1988], "A discrete time average cost flexible manufacturing and operator scheduling problem solved by deconvexification over time," *Operations Research*, Vol. 36, pp. 242-257.
- Eaves, B.C. and U.G. Rothblum [1989], "A continuous-time average-cost flexible manufacturing and operator scheduling model solved by deconvexification over time", *Linear Algebra and Its Applications*, Vol. 114/115, pp. 417-428.
- Perold, F. P.[1978] "Fundamentals of a continuous time simplex method," Technical Report SOL 78-26, Department of Operations Research, Stanford University.

END

DATE FILMED

11 / 07 / 90

