DYNAMIC AND STOCHASTIC MODELS WITH FREIGHT DISTRIBUTION APPLICATIONS

A Thesis
Submitted to the Faculty

of

Purdue University

by

Anton J. Kleywegt

In Partial Fulfillment of the
Requirements for the Degree

of

Doctor of Philosophy

August 1996

Acknowledgments

Thanks to my advisor, Jason Papastavrou, for his support and guidance, and especially for the freedom he gave me to pursue research in my own field of interest. He created a flexible and pleasant work atmosphere, without which this study would not have been such a rewarding experience. Thanks to the other members of my Ph.D. advisory committee, Bruce Schmeiser, Colm O'Cinneide, Tom Morin and Samer Madanat, for their interest and willingness to answer my numerous questions. Thanks to all the other professors who have played such a great role during this academic adventure. Kumares Sinha, my master's advisor, Jon Fricker, and Mike Cassidy taught me transportation, and guided me through my master's studies. Ron Rardin, Reha Uzsoy, Nagabhushana Prabhu, and Dong Shaw, together with my committee members, taught me operations research, and made my graduate studies such an interesting and worthwhile experience. David Drasin, Rodrigues Bañuelos, Steve Lalley, Tom Sellke, Tom Kuczek, and James Berger taught me mathematics and statistics, without which this thesis would not have been possible.

Thanks to all my friends at Purdue, who made the past four years so enjoyable. Mike Swihart, my academic brother, shared many hours with me in semi-philosophical discussion, and taught me the law of the infinitely divisible lacunae and the rules of the glass bead game. Thanks to my office mates, Elham Ghashghai, Yee-yin Choong, Te-min Chang, Mel Leeb, Daniel Hiquebran, Jihong Jin, and Peichun Lin, my room mates, Sanjay Mehta, Fethi Çalisir, and Keyvan Esfarjani, and other friends, Mark and Gladys Scott, Graeme and Yolanda Warren, Darel and Susan Benysh, Srikanth Rajagopalan, Farhad Zaerpoor, Beril Toktay, John Shewchuk, Huifen Chen, Jin Wang, Sarah Morris, Pam Niedzwiecki, Sara McComb, Joy Chang, Su-Hsia Yang, and Farqad Alkhal for their friendship and the fun times we had together.

Thanks to my family; my father Robert and my mother Johanna, and my sisters Maryke and Annaleen, for their encouragement, correspondence and visits, and for having me in their thoughts so often. It is their love and support through the years that made this achievement possible.

DISCARD THIS PAGE

Table of Contents

Page

List of Tables . . . vii

List of Figures . . . viii

Nomenclature . . . x

Abstract . . . xiv

1. **Introduction** . . . 1

I THE DYNAMIC AND STOCHASTIC KNAPSACK PROBLEM **5**

2. **Introduction** . . . 6

3. **Literature Review** . . . 9

4. **Problem Definition** . . . 14

4.1 Problem Description . . . 14
4.2 Dynamic and Stochastic Knapsack Problem Formulation . . . 15
4.3 A Related Markov Decision Process . . . 19
4.3.1 Markov Decision Process Formulation . . . 19
4.3.2 Optimality Results for Markov Decision Processes . . . 24

5. **The Infinite Horizon DSKP** . . . 26

5.1 Optimality of Stationary Policies . . . 26
5.2 The Value of a Policy . . . 27
5.3 The Optimal Value Function . . . 29
5.4 An Optimal Policy for the MDP and DSKP . . . 30
5.4.1 Preliminary Results . . . 30
5.4.2 An Optimal Threshold Policy for the MDP and DSKP . . . 31
5.5 A Recursion for the Optimal Expected Value . . . 34
5.5.1 Preliminary Results . . . 34
5.5.2 Definition of the Recursive Policies . . . 36
5.5.3 Optimality Results for the Recursive Policies . . . 37

5.6 An Efficient Algorithm for the Infinite Horizon DSKP 39

6. **The Finite Horizon DSKP** 41

6.1 The Value of a Policy 41
6.1.1 Derivation of the Bellman-Hamilton-Jacobi Equation 42
6.1.2 Solution of the Bellman-Hamilton-Jacobi Equation 43
6.1.3 Continuity Results 48
6.2 The Optimal Value and an Optimal Policy for the MDP 49
6.2.1 The Optimality Equation 49
6.2.2 Solution of the Optimality Equation 51
6.2.3 Continuity Results 52
6.3 The Optimal Value and an Optimal Policy for the DSKP 55
6.4 Convergence of Discrete-Time Approximations 58
6.5 Examples 62
6.6 Structural Characteristics 67
6.6.1 A Fundamental Relationship 69
6.6.2 A Recursion for the Optimal Expected Value 74
6.6.3 Monotonicity and Convexity Properties 85

II THE DYNAMIC AND STOCHASTIC DISTRIBUTION PROBLEM 106

7. **Introduction** 107

8. **Literature Review** 109

8.1 Vehicle Routing Problems 109
8.2 Vehicle and Crew Scheduling Problems 110
8.3 Dynamic Vehicle Allocation Problems 110
8.4 Service Network Design Problems 110
8.5 Control of Bulk Service Queues 112

9. **Problem Definition** 115

9.1 Problem Description 115
9.2 Dynamic and Stochastic Distribution Problem Formulation 116

10. **A Markov Decision Process Model** 120

10.1 Assumptions 120
10.2 Markov Decision Process Formulation 120
10.3 The Value of a Policy 123
10.4 Structural Properties 125

Page

10.4.1 The Spectral Radius of the MDP 126
10.4.2 Dispatching at Most One Vehicle in Each State 129
10.5 The Optimal Value and an Optimal Policy for the DSDP 131
10.5.1 An Optimal Threshold Policy for the DSDP 131
10.5.2 Monotonicity Result 132

11. **Solution Methods** 133

11.1 Value Iteration 134
11.1.1 Jacobi Value Iteration 134
11.1.2 Gauss-Seidel Value Iteration 140
11.2 Modified Policy Iteration 144
11.2.1 Jacobi Modified Policy Iteration 144
11.2.2 Gauss-Seidel Modified Policy Iteration 148
11.3 Two-Decisions Algorithm 151
11.3.1 Pre-Jacobi Two-Decisions Algorithm 151
11.3.2 Pre-Gauss-Seidel Two-Decisions Algorithm 154
11.4 Numerical Results 159
11.4.1 Comparison of Algorithms 159
11.4.2 Example 160

12. **Research Directions** 165

12.1 Dynamic and Stochastic Resource Allocation Problems 165
12.2 Dynamic and Stochastic Freight Distribution Problems 166

13. **Conclusion** 170

List of References 172

Appendices

Appendix A: Proof of Lemma 1 184
Appendix B: Proof of Lemma 5 186
Appendix C: Proof of Lemma 7 190
Appendix D: Proof of Lemma 14 192
Appendix E: Proof of Lemma 15 194
Appendix F: Proof of Lemma 17 196
Appendix G: Proof of Lemma 20 198
Appendix H: Proof of Proposition 2 201
Appendix I: Proof of Proposition 3 204
Appendix J: Proof of Lemma 25 205
Appendix K: Proof of Lemma 30 211
Appendix L: Proof of Proposition 6 212

Page

Vita 223

DISCARD THIS PAGE

List of Tables

Table Page

0.1 Summary of notation for the DSKP x

0.2 Summary of notation for the DSDP xii

11.1 Input data of example problem 163

DISCARD THIS PAGE

List of Figures

Figure | Page

6.1 Example with Exponential Rewards and Variable Holding Cost Per Unit Time 64

6.2 Example with Exponential Rewards and Constant Holding Cost Per Unit Time 65

6.3 Example with Uniform Rewards and Constant Holding Cost Per Unit Time . . 68

6.4 Example to show that Proposition 4 does not hold if λ is an arbitrary decreasing function of n 71

6.5 Example to show that Proposition 4 does not hold if λ is an arbitrary increasing function of n 72

6.6 Example to show that Proposition 4 does not hold if p is an arbitrary increasing function of n 72

6.7 Example to show that Proposition 4 does not hold if c is an arbitrary increasing function of n 73

6.8 Example to show that Proposition 4 does not hold if v is an arbitrary decreasing function of n 74

6.9 Example to show that Corollary 8 does not hold if v is an arbitrary increasing function of n 78

6.10 Example to show that Theorem 8 does not hold if λ is an arbitrary decreasing function of n 81

6.11 Example to show that Theorem 8 does not hold if λ is an arbitrary increasing function of n 82

6.12 Example to show that Theorem 8 does not hold if p is an arbitrary increasing function of n 83

6.13 Example to show that Theorem 8 does not hold if c is an arbitrary increasing function of n 83

6.14 Example to show that Theorem 8 does not hold if v is an arbitrary decreasing function of n 84

6.15 Example to show that Theorem 8 does not hold if v is an arbitrary increasing function of n 85

6.16 Example to show that Theorem 9 does not hold if λ is an arbitrary decreasing function of n 92

6.17 Example to show that Theorem 9 does not hold if λ is an arbitrary increasing function of n 93

6.18 Example to show that Theorem 9 does not hold if p is an arbitrary decreasing function of n 94

6.19 Example to show that Theorem 9 does not hold if p is an arbitrary increasing function of n 95

6.20 Example to show that Theorem 9 does not hold if c is an arbitrary decreasing function of n 97

6.21 Example to show that Theorem 9 does not hold if c is an arbitrary increasing function of n 98

11.1 Computation times vs. state space size for pre-Gauss-Seidel value iteration, pre-Jacobi modified policy iteration, pre-Gauss-Seidel modified policy iteration, and pre-Gauss-Seidel two-decisions algorithm 160

11.2 Optimal expected value vs. number of loads at terminal 3 destined to terminal 2 for different numbers of vehicles m available at terminal 3 161

11.3 Optimal acceptance threshold vs. number of loads at terminal 3 destined to terminal 2 for different numbers of vehicles m available at terminal 3 162

12.1 Optimal acceptance threshold vs. number of loads at terminal 3 destined to terminal 2 for different numbers of vehicles m available at terminal 3 168

Nomenclature

1. Dynamic and Stochastic Knapsack Problem Notation

Table 0.1 Summary of notation for the DSKP

$DSKP$	Dynamic and Stochastic Knapsack Problem
MDP	Markov Decision Process
N_0	initial capacity
T	deadline
$\mathcal{T}$	stopping time
n	remaining capacity
t	time
$N(t)$	remaining capacity at time t
$\lambda(n,t)$	item arrival rate if remaining capacity is n at time t, $\lambda \in (0, \overline{\lambda}]$
A_i	arrival time of item i
R_i	reward of item i
S_i	size of item i
$F_{R,S}$	joint probability distribution of item rewards R and sizes S
F_R	probability distribution of item rewards R
F_S	probability distribution of item sizes S
$F_{R\|S}$	conditional probability distribution of item rewards R given item sizes S
$F_{S\|R}$	conditional probability distribution of item sizes S given item rewards R

Table 0.1 Summary of notation for the DSKP (continued)

$c(n,t)$	holding cost per unit time if remaining capacity is n at time t
$p(n,t)$	penalty for rejecting an item if remaining capacity is n at time t
$v(n)$	terminal value with $N(\mathcal{T}^+) = n$
α	discount rate
π	policy
Π	set of policies
Π^{HD}	set of history-dependent deterministic policies
Π^{MD}	set of memoryless deterministic policies
Π^{SD}	set of stationary deterministic policies
$D^\pi(n,t,r,s)$	acceptance decision rule of policy π
$\mathcal{R}_1^\pi(n,t)$	set of rewards and sizes accepted under policy π
$\mathcal{R}_0^\pi(n,t)$	set of rewards and sizes rejected under policy π
$I^\pi(n,t)$	stopping decision rule of policy π
$\mathcal{I}_s$	set of unit step functions $f_s : (0,T] \mapsto \{0,1\}$
$\mathcal{I}_I$	set of Borel measurable functions $f_I : (0,T] \mapsto \{0,1\}$
$\mathcal{I}_{RS}$	set of Borel measurable functions $f_{RS} : \Re \times [0,N_0] \mapsto \{0,1\}$
$V^\pi(n,t)$	expected value of policy π
$x^\pi(n,t,s)$	reward acceptance threshold used by policy π
$z^\pi(n,t,r)$	size acceptance threshold used by policy π
V^*	optimal expected value
π^*	optimal policy

2. Dynamic and Stochastic Distribution Problem Notation

Table 0.2 Summary of notation for the DSDP

K	number of terminals
$\mathcal{K}$	set of terminals
$\overline{N}(k)$	capacity of terminal k
$\overline{M}$	number of vehicles
$\mathcal{M}$	set of vehicles
$\overline{N}_v$	vehicle capacity
$s(k,l)$	size of items of class (k,l)
λ	request arrival rate, $\lambda \in (0,\infty)$
A_i	arrival time of request i
B_i	batch composition of request i
R_i	reward of request i
F_B	probability distribution of batch compositions B
U_i	arrival time of vehicle arrival i
W_i	origin-destination pair of vehicle arrival i
$F_{R\|B}$	conditional probability distribution of item rewards R given batch B
$m(k)$	number of vehicles available at terminal k
$\tilde{m}(k,l)$	number of vehicles en route from terminal k to terminal l
$n(k,l)$	number of waiting loads of class (k,l)
$c(m,\tilde{m},n)$	holding cost per unit time if the state is $(m,\tilde{m},n)$
p	penalty for rejecting an item
$\varphi(k,l,q)$	dispatching/transportation cost if q loads are dispatched from k to l
α	discount rate

Table 0.2 Summary of notation for the DSDP (continued)

π	policy
Π	set of policies
Π^{HD}	set of history-dependent deterministic policies
Π^{SD}	set of stationary deterministic policies
$D^{\pi}(m, \tilde{m}, n, r, b)$	acceptance decision rule of policy π
$\mathcal{R}_1^{\pi}(m, \tilde{m}, n)$	set of rewards and sizes accepted under policy π
$\mathcal{R}_0^{\pi}(m, \tilde{m}, n)$	set of rewards and sizes rejected under policy π
$L^{\pi}(m, \tilde{m}, n)$	origin-destination pairs of vehicles dispatched under policy π
$Q^{\pi}(m, \tilde{m}, n)$	number of loads dispatched under policy π
$\mathcal{G}_D$	set of Borel-measurable functions $f : \Re \times \mathcal{Z}_+^{K \times K} \mapsto \{0, 1\}$
$\mathcal{G}_L$	set of lists $\{(k_j, l_j)\}_{j=1}^{\ell}, (k_j, l_j) \in \mathcal{K} \times \mathcal{K}, k_j \neq l_j, \ell \in \{0, 1, \ldots, \overline{M}\}$
$\mathcal{G}_Q$	set of lists $\{q_j\}_{j=1}^{\ell}, q_j \in \mathcal{Z}_+, \ell \in \{0, 1, \ldots, \overline{M}\}$
$\mathcal{M}$	pre-Jacobi value iteration operator
$\mathcal{M}_{GV}$	pre-Gauss-Seidel value iteration operator
$\mathcal{M}_{JM}$	pre-Jacobi modified policy iteration operator
$\mathcal{M}_{GM}$	pre-Gauss-Seidel modified policy iteration operator
$\mathcal{M}_{JT}$	operator of the pre-Jacobi two-decisions algorithm
$\mathcal{M}_{GT}$	operator of the pre-Gauss-Seidel two-decisions algorithm
$V^{\pi}(m, \tilde{m}, n)$	expected value of policy π
$x^{\pi}(m, \tilde{m}, n, b)$	reward acceptance threshold used by policy π
V^*	optimal expected value
π^*	optimal policy

Abstract

Kleywegt, Anton J., Ph.D., Purdue University, August 1996. Dynamic and Stochastic Models with Freight Distribution Applications. Major Professor: Jason D. Papastavrou.

This research is motivated by the need for decision support tools to manage complex freight distribution operations. The fact that freight distribution is such an important economic activity, makes this research both worthwhile and rewarding.

In the first part of this research, a generic resource allocation problem in a dynamic and stochastic environment is studied. This problem is motivated by the issues facing a manager of a transportation operation regarding the acceptance of loads and the dispatching of a vehicle. Because it has applications in many other areas, such as the scheduling of batch processors, the selling of assets, and the selection of investment projects, it is called the Dynamic and Stochastic Knapsack Problem (DSKP). In the DSKP, demands for a resource arrive in time according to a Poisson process. Each demand requires an amount of resource, and has an associated reward. The resource requirements and rewards are jointly distributed according to a known probability distribution, and become known at the time of the demand's arrival. Demands can be either accepted or rejected. If a demand is accepted, the associated reward is received, and if a demand is rejected, a penalty is incurred. The problem can be stopped at any time, at which time a terminal value is received, that may depend on the amount of resource remaining. A holding cost, that may depend on the amount of resource remaining, is incurred until the process is stopped. The objective is to determine an optimal policy that maximizes the expected value (rewards minus costs) accumulated. The DSKP is analyzed for both the infinite horizon and the finite horizon cases. It is shown that an optimal acceptance rule is given by an easily computed threshold rule. The optimal stopping time is also easily determined, and has the convenient property that under typical conditions, only those times right after a demand has been accepted need to be considered as potential stopping times. Efficient algorithms for computing optimal solutions are proposed. These algorithms have to compute the value corresponding to each state of the process only once, in contrast with the classical iterative algorithms that are used to solve dynamic programming problems. It is established that the optimal value and optimal threshold have a number of interesting monotonicity and convexity properties.

In the second part of this research, a distribution problem with a number of terminals and a fleet of vehicles is studied. This problem is called the Dynamic and Stochastic Distribution Problem (DSDP). Customers request the transportation of batches of loads between different origins and destinations. A request can be accepted or rejected; if the request is accepted, a reward is received. Holding costs for vehicles and loads at terminals as well as transportation costs are included in the model. The objective is to determine a policy for accepting transportation requests and for dispatching vehicles that maximizes the expected value of operating the distribution system. A Markov decision process model is developed, and optimal policies are characterized. A number of algorithms for computing optimal solutions are studied. It is shown that the classical algorithms for solving Markov decision processes converge if applied to the DSDP, in spite of the fact that the DSDP does not satisfy all the assumptions made in the traditional convergence proofs. An algorithm that exploits the structure of the DSDP is developed, and is compared with some of the classical algorithms in computational experiments. The proposed algorithm consistently outperforms the classical algorithms. A number of directions for future research are discussed, especially with the aim of obtaining more efficient solution methods for large systems.

1. Introduction

Freight distribution is an activity of great economic importance. Revenues attributed to freight transportation in the U.S.A. in 1993 were estimated at $400 billion, of which $313 billion, or 78%, was accounted for by trucking [1]. Improved operations, brought about by operations research techniques, can lead to substantial cost savings, even if the efficiency is improved by only a few percent.

Freight distribution operations are typically very complex, involving the interaction of many loads with different origins and destinations, a geographically distributed network of terminals with different facilities and capacitites, a number of vehicles with different characteristics, drivers with their work rules and personal preferences, and customer requirements such as pickup and delivery time windows and storage and handling requirements. All these factors change rapidly in a dynamic and stochastic environment. Managing these distribution operations is therefore a challenging task, and sophisticated decision support tools are needed to improve efficiency.

The objective of this research is to develop a number of models that can be used to improve freight distribution operations. The models to be developed aim to capture the dynamics and uncertainties involved in distribution operations, and specifically address the development of policies for load acceptance and vehicle dispatching. This emphasis is chosen for the following reasons.

Most of the existing research effort in this area has been focused on problems of a deterministic and static nature, although a substantial number of dynamic and stochastic models have been formulated and studied. The models developed in this study are of a dynamic and stochastic nature, because such models better capture the true problems faced in the industry.

A number of researchers mention load-acceptance as an important issue in freight distribution operations. However, none of the studies encountered in the literature incorporates the load-acceptance problem into the analysis; it is assumed that all arrivals are accepted, and therefore that the loads to be transported are determined exogenously. One of the contributions of this research is to develop a method to identify good load acceptance decisions.

In the first part of this research the *Dynamic and Stochastic Knapsack Problem* (DSKP) is formulated and studied. This research is motivated by a typical situation in the transportation industry, where ships, trains, aircraft and trucks often carry loads for different clients. Transportation requests arrive stochastically over time, and rewards are offered (or negotiated) for servicing demands (transporting loads). Loads may have different sizes, such as packages or different numbers of rail freight cars to be loaded, or the same size, as in container shipping or truckload trucking. Both the DSKP with constant sized items, which has a number of special properties, and the DSKP with random sizes are studied. Demands can be accepted or rejected, and cannot be recalled. Rejection of demands may bring about a penalty which takes potential losses of customer goodwill and future sales into account. Often there is a fixed schedule for moving vehicles, and hence a deadline after which loads cannot be accepted for a specific shipment. Even when there is not a fixed schedule, an incentive exists to consolidate and dispatch shipments with high frequency, to maintain short delivery times (and therefore high levels of customer service), and to maximize the rate at which revenue is earned with the given investment in capital and labor costs. This incentive is modeled with a holding cost per unit time that is incurred until the shipment is dispatched. The cost may depend on the loads accepted, but not yet dispatched. The carrier may therefore want to dispatch the shipment some time before the deadline. Costs are also incurred when loads are dispatched and transported, and these costs may depend on the number of loads in the shipment. The carrier strives to maximize the expected profit.

Many other problems are also characterized by the allocation of limited resources to competing demands. The demands usually arrive randomly over time, and are associated with rewards as well as resource requirements. The problem is to decide which demands to accept and which to reject in order to optimally allocate the limited resources. For example, consider a salesperson trying to sell a number of assets, such as a real-estate agent selling new condominiums. The salesperson receives offers stochastically over time, which can be accepted or rejected. There is also an opportunity cost associated with the capital tied into the unsold assets, which causes a holding cost per unit time to exist. A similar situation is encountered by a scheduler of a batch processor, where jobs with random rewards and capacity requirements arrive over time. A fixed schedule or customer commitments lead to deadlines. Pressure to increase the utilization of equipment and labor, and to maintain a high level of customer service, leads to a holding cost per unit time. A portfolio manager who wishes to invest a certain amount of funds faces a similar problem. The manager is presented with investment opportunities with random arrival times, funding requirements, and returns.

The limited resource can be viewed as a knapsack of fixed capacity, and the demands for the resource can be viewed as the items to be loaded into the knapsack. The items arrive stochastically over time and have random sizes and rewards. Therefore, to capture the wide applicability of this problem, it is called the Dynamic and Stochastic Knapsack Problem (DSKP).

It is shown that an optimal acceptance rule for the DSKP is given by a simple threshold rule. It is also shown that under typical conditions the optimal stopping time (dispatching time) is right after an arrival has been accepted. This gives an optimal policy, consisting of an optimal acceptance rule and an optimal stopping time, with a simple, intuitive structure.

For the infinite horizon case (with no deadline), the optimal expected value is the unique fixed point of a certain function. An algorithm is proposed that computes the optimal expected value and optimal policy with very little effort. For the finite horizon case (with a deadline), the optimal expected value is given by a differential equation. This differential equation is shown to have a unique solution which can be computed arbitrarily accurately by a discrete-time approximation under very general conditions. In computational experiments the differential equation was solved, and the optimal expected value and optimal policy were obtained with little computational effort. A number of intuitive monotonicity and convexity results are shown to hold for the optimal expected value and the optimal acceptance rule.

In the second part of this research the *Dynamic and Stochastic Distribution Problem* (DSDP) is formulated and analyzed. The DSDP models a distribution operation with a network of terminals and a fleet of vehicles. Customers call in at random times, requesting transportation of their loads from their origins to their destinations. Each customer's request may involve a batch of loads that have to go from a number of origins to a number of destinations. A price is negotiated with the customer for transporting the loads, and this price, minus the costs incurred for servicing the customer, determines the reward that the carrier receives if the request is accepted. If the reward is not attractive enough, the request may be rejected, in which case a penalty is incurred for the loss of customer goodwill and future sales. Rejected requests cannot be recalled. At any point in time there may be a number of loads waiting at each terminal to be transported to their destinations. There are also a number of vehicles available at each terminal, and a number of vehicles en route between the terminals. There is a holding cost per unit time that depends on the number of vehicles as well as the number of loads waiting at each terminal, as well as a dispatching and transportation cost that depends on the origin-destination pair and the number of loads transported. The objective of the carrier is to maximize the expected value earned from managing the distribution operation.

A Markov decision process formulation of the DSDP is studied in detail. As for the DSKP, an optimal acceptance rule is given by a simple threshold rule. Together with an optimal dispatching rule it gives an optimal policy for the DSDP. Optimality equations are derived for the optimal expected value. A number of classical algorithms for computing the optimal expected value and optimal policy, such as value iteration and modified policy iteration, are studied. It is shown that these algorithms converge when applied to the DSDP, although the DSDP does not satisfy the assumptions made in the traditional convergence proofs. An algorithm that exploits the structure of the DSDP is developed, and its convergence to the optimal solution is shown. The performance of this algorithm is compared with some of the classical algorithms in computational experiments.

There are many issues in freight distribution that remain to be resolved; some directions for future research are given at the conclusion of this study.

Part I

THE DYNAMIC AND STOCHASTIC KNAPSACK PROBLEM

2. Introduction

The *knapsack problem* is a generic resource allocation problem with many applications. Items to be loaded into a knapsack with fixed capacity are selected from a given set of items with known sizes and rewards. The objective is to maximize the total reward, subject to capacity constraints. This problem is static and deterministic, because all the items are considered at a point in time, and their sizes and rewards are known when the decisions are made. However, in many practical applications, knapsack-type resource allocation problems are encountered in an uncertain and dynamically changing environment. Furthermore, there are often costs associated with delays that are not captured in the static knapsack problem. Applications of the dynamic and stochastic counterpart of the knapsack problem include:

1. In the transportation industry, ships, trains, aircraft or trucks often carry loads for different clients. Transportation requests arrive stochastically over time, and prices are offered or negotiated for transporting loads. If a load is accepted, costs are incurred for picking up and handling the load, and for the administrative activities involved. These costs are specific to the load, and can be subtracted from the price to give the "reward" of the load. If a load is rejected, some customer goodwill (possible future sales) is lost, which can be taken into account with a penalty for rejecting loads. Loads may have different sizes, such as parcels, or the same size, such as containers. Often there is a fixed schedule for moving vehicles and a deadline after which loads cannot be accepted for a specific shipment. Even when there is not a fixed schedule, an incentive exists to consolidate and dispatch shipments with high frequency, to maintain short delivery times, and to maximize the rate at which revenue is earned with the given investment in capital and labor costs. This incentive can be modeled with a discount rate, and a holding cost per unit time that is incurred until the shipment is dispatched. The holding cost may be constant or may depend on the number of loads accepted, but not yet dispatched. The dispatcher can decide to dispatch a vehicle at any time before the deadline. There is also a dispatching and transportation cost that

is incurred for the shipment as a whole, that may depend on the number of loads in the shipment.

2. A scheduler of a batch processor has to schedule jobs with random capacity requirements and rewards as they arrive over time. Fixed schedules or customer commitments lead to deadlines. The pressure to increase the utilization of equipment and labor, and to maintain a high level of customer service, lead to a holding cost per unit time. The cost of running the batch processor may depend on the number of jobs in the batch.

3. A real estate agent selling new condominiums receives offers stochastically over time and may want to sell the assets before winter or before the new tax year. Hence, the agent faces a deadline, possibly with a salvage value for the unsold assets. There is also an opportunity cost associated with the capital tied into the unsold assets, and property taxes, which cause a holding cost per unit time to exist.

4. An investor who wishes to invest a certain amount of funds faces a similar problem. The investor is presented with investment projects with random arrival times, funding requirements, and returns. The opportunity cost of unutilized capital is represented by a holding cost per unit time, and the objective is to maximize the expected value earned from investing the funds.

5. A ticket office sells tickets for a performance or a sports event. Customers interested in purchasing tickets arrive randomly over time. Unsold tickets after the event are worthless, and thus there is a deadline after which tickets cannot be sold. Although it is not common practice, ticket prices might be dynamically adjusted according to the sales volume and the approaching deadline. The objective is to maximize the profit from the sale of the available tickets.

These problems are characterized by the allocation of limited resources to competing demands that arrive randomly over time. Demands are associated with resource requirements as well as rewards, which may include any demand-specific costs incurred. Usually the arrival times, resource requirements, and rewards are unknown before arrival, and become known upon arrival. Arriving demands can be either accepted or rejected. Incentives such as a deadline after which arriving demands cannot be accepted, discounting, and a holding cost per unit time, serve to encourage the timely acceptance of demands. The problem can be stopped at any time before or at the deadline. There may also be a cost associated with the group of accepted demands as a whole, or a salvage value for unused resources, which may depend on the amount of resources allocated to the accepted demands. A typical objective is to maximize the expected total value (rewards minus costs). We call problems of this general nature the *Dynamic and Stochastic Knapsack Problem* (DSKP).

Infinite horizon and finite horizon versions of the DSKP are studied. Optimality equations are derived, which are shown to have the optimal expected value functions as their unique solutions. It is shown that an optimal acceptance rule is given by a simple threshold rule. It is also shown how an optimal stopping time is determined. Conditions are established under which the optimal value function and the optimal acceptance threshold exhibit intuitive structural characteristics. Algorithms for computing solutions are proposed. Examples with closed-form solutions, as well as numerical examples, are discussed.

In Chapter 3, previous research on similar problems is reviewed. The DSKP as well as a related Markov decision process are defined in Chapter 4. The DSKP without a deadline is studied in Chapter 5, and the DSKP with a deadline follows in Chapter 6.

3. Literature Review

The static and deterministic knapsack problem has been studied extensively in operations research [109]. Items are selected from a set of items with known sizes and rewards, to be loaded into a knapsack with fixed capacity. The objective is to maximize the total reward, subject to capacity constraints. This problem is static and deterministic, because all the items are considered at a point in time, and their sizes and rewards are known a priori.

Stochastic versions of the knapsack problem have also been studied; these can be classified as static or dynamic. In static stochastic knapsack problems the set of items is given, but the rewards and/or sizes are unknown. Steinberg and Parks [172] proposed a preference order dynamic programming algorithm, with a preference ordering on the distributions of the item rewards, instead of a real valued return. Sniedovich [167, 168] further investigated the use of preference order dynamic programming, with weak and strong criteria. He pointed out that the weak relation, as used by Steinberg and Parks, is complete, but not necessarily monotone, and therefore may lead to suboptimal solutions. The problem with the weak criterion can also be stated as a regular deterministic optimization problem. The strong relation may not be complete, and this may lead to the failure of an optimal solution to exist, or to a trivial optimal solution. Henig [85] combined dynamic programming and a search procedure to solve stochastic knapsack problems where the items have known sizes and independent normally distributed rewards. The objective is to maximize the probability that a target value will be achieved, or some percentile of the total value distribution, or to optimize a linear combination of the mean and variance of the total value distribution. Caraway *et al.* [35] proposed a hybrid dynamic programming/branch-and-bound algorithm for a stochastic knapsack problem similar to that of Henig, with an objective that maximizes the probability of target achievement.

In dynamic stochastic knapsack problems the items arrive over time, and the rewards and/or sizes are unknown before arrival. There is a time deadline after which items cannot be accepted, and/or a holding cost per unit time. Items are accepted or rejected upon arrival, and recall of rejected items (at a cost) may be allowed.

Some stopping time problems and best choice (optimal selection) problems are similar to the DSKP. A well-known example is the *secretary problem*, where candidates arrive over time, and the objective is to maximize the probability of choosing the best candidate or k best candidates from a given or random number of candidates, or to maximize the expected value of the chosen candidates. These problems have been studied by Presman and Sonin [134], Stewart [173], Freeman [74], Yasuda [186], Bruss [26, 27], Bruss and Rogers [28], Nakai [112], Sakaguchi [156], and Tamaki [176, 177].

The problem of selling a single asset, where offers arrive periodically [147], or according to a renewal process [108], with a fixed holding cost, with or without a deadline, has been studied widely. Albright [4] studied a house selling problem where a given number n of offers are received, and $k \leq n$ houses are to be sold. Asymptotic properties of an optimal policy for the house selling problem with discrete time periods, as the deadline and number of houses become large, were derived by Saario [155].

A more general problem is the *Sequential Stochastic Assignment Problem* (SSAP). Derman, Lieberman and Ross [60] defined the problem as follows: n persons, with known values $p_i, i = 1, \ldots, n$, are to be assigned sequentially to n jobs, which arrive one at a time. The jobs have values $X_j, j = 1, \ldots, n$, which are unknown before arrival, but become known upon arrival, and which are i.i.d. with known probability distribution F. If a person with value p_i is assigned to a job with value x_j, the reward is $p_i x_j$. The objective is to maximize the expected total reward. They showed that an optimal policy is characterized by partitioning the real line into n intervals, independent of the p-values of the men, but depending on the distribution F and the number n. The man with the i^{th} highest p-value is then assigned to an arriving job if the x-value of the job falls in the i^{th} highest interval. The proof of the optimality of this policy is based on Hardy's inequality. Albright and Derman [5] studied the asymptotic behavior of the optimal intervals as $n \to \infty$. Albright [3] also investigated a sequential stochastic assignment problem where the reward is $r(t)p_i x_j$ if a job with value x_j arrives at time t and is assigned to a man with value p_i. An interval type policy, such as the one described above, is also optimal in this case, but the intervals now depend on time. Sakaguchi [157, 158] studied sequential stochastic assignment problems with an unknown number of jobs, but with a known upper bound for the number of jobs. He showed that if the jobs arrive according to a "fishing process", then an interval type policy, where the intervals depend on time, is optimal. A sequential stochastic assignment problem, where jobs arrive in discrete time, the option to reject a job exists, and the distribution of the job values changes according to a stationary Markov chain, was studied by Nakai [114]. An interval type policy, with the intervals depending on the state of the Markov chain, is optimal. Nakai [113] also showed that if the states of the Markov chain are not observed, but an a priori distribution of the states is given, then an interval type policy is not necessarily optimal. For the case where the job values are dependent, Kennedy [95] showed that an interval type policy is optimal, with the intervals depending on the job values.

Some resource allocation problems are similar to the dynamic and stochastic knapsack problem. Mendelson *et al.* [110] investigated the problem of allocating a given amount of resource to a set of activity classes, with demands following a known probability distribution, and arriving according to a renewal process. The objective is to maximize the expected time until the resource allocated to an activity is depleted. Righter [143] studied a resource allocation problem, which is a version of the sequential stochastic assignment problem. Resources arrive according to a Poisson process in time, and have random values, which become known upon arrival. Resources are assigned to activities, which have known values. The return from assigning a resource to an activity is the product of the resource value and the activity value, as in the SSAP described above. Activities have exponentially distributed deadlines. Characteristics of optimal policies are derived if some of the parameters of the problem, such as the resource arrival rate, the activity values, and the activity deadline rates, change according to independent continuous time Markov chains. The characteristics, such as the optimality of an interval type policy, and the monotonicity and convexity of optimal expected values, are similar to those derived for the SSAP.

Many investment problems can be regarded as DSKP's. For example, Prastacos [133] studied the problem of allocating a given amount of resource before a deadline to irreversible investment opportunities that arrive according to a geometric process in discrete time. The quality of an investment opportunity becomes known upon arrival, and the objective is to maximize the total expected value. We incorporate a holding cost, in addition to the issues taken into account by Prastacos, and arrivals either occur according to a geometric process in discrete time, or according to a Poisson process in continuous time. Also, Prastacos assumed that each investment opportunity was large enough to absorb all the available capital, but in our problem the sizes of the investment opportunities are random, and cannot be chosen.

Other versions of the DSKP have been studied for communications applications. Kaufman [92] studied policies for a system where K customer types, each with given resource requirements, arrive according to a Poisson process to a system with fixed capacity, and are blocked if sufficient remaining capacity is not available. Ross and Tsang [148, 149] investigated control policies to optimize a system similar to that of Kaufman. They proposed a dynamic programming algorithm to find the optimal partitioning of the resource between the classes. This is a static solution, because it does not adjust to the dynamically changing state of the system, and hence is not an optimal policy.

A class of problems similar to the DSKP have been termed *Perishable Asset Revenue Management* (PARM) problems by Weatherford and Bodily [183, 184]. These problems are often called *yield management* problems, and have been studied widely, with specific application to airline seat inventory control and hotel yield management, by Rothstein [150, 151, 152], Shlifer and Vardi [162], Silver [163], Alstrup *et al.* [6], Belobaba [11, 12], Dror *et al.* [69], Curry [45], Brumelle *et al.* [25, 24], Wollmer [185], Lee and Hersh [103], and Robinson [144]. In most of these problems there are a number of different fare classes, which are usually assumed to be given due to competition. The objective is to dynamically assign the available capacity to the different fare classes to maximize expected revenues. In the DSKP, the available capacity is dynamically assigned to arriving demands with random rewards and random resource requirements. Another type of PARM problem, in which an inventory has to be sold before a deadline, has been studied by Kincaid and Darling [97], Stadje [171], and Gallego and Van Ryzin [75]. In their problems customers arrive according to a Poisson process, with price-dependent probability of purchasing. The major differences with our model is that in our model offers arrive, and the offers can be accepted or rejected, as is typical with large contracts such as the selling of real estate, whereas in the models of Stadje and Gallego and Van Ryzin prices are set and all demands are accepted as long as supplies last, which is typical in retail; also, we incorporate a holding cost and an option to stop before the deadline. Kinberg and Rao [96] studied models of price promotion.

Papastavrou, Rajagopalan and Kleywegt [117] studied a discrete-time version of the DSKP, without holding costs.

4. Problem Definition

In this chapter, first a description of the DSKP is given in Section 4.1. This is followed by a mathematical formulation of the DSKP in Section 4.2. Then a Markov decision process model that is closely related to the DSKP is introduced in Section 4.3.

4.1 Problem Description

Items arrive according to a Poisson process in time. Each item has an associated size and reward. The size and reward of an item are unknown prior to arrival, and become known upon arrival. The joint distribution of the sizes and rewards is known, and is independent of the arrival times and of the sizes and rewards of other arrivals. The size and the reward of an item can be dependent. In typical applications we would expect a positive correlation between the size and the reward of an item. The items are to be included in a knapsack of known capacity. Each arriving item can be either accepted or rejected. If the item size is too large to fit into the remaining capacity of the knapsack, the item has to be rejected. If an item is accepted, the reward associated with the item is received, and if the item is rejected, a penalty is incurred. Once an item is rejected, it cannot be recalled.

There is a known deadline (possibly infinite) after which items can no longer be accepted. It is allowed to stop waiting for arrivals before the capacity is exhausted or the deadline is reached. In a freight consolidation application, this occurs when a vehicle is dispatched before it is filled to capacity and before the scheduled deadline is reached. There is a holding cost per unit time that depends on the total size of items already accepted, or equivalently, on the remaining capacity of the knapsack. A terminal value is earned that depends on the remaining capacity of the knapsack at the stopping time. Rewards and costs may be discounted. The objective is to determine a policy for accepting items and for stopping that maximizes the expected total (discounted) value (rewards minus costs) accumulated, among the class of nonanticipatory policies.

4.2 Dynamic and Stochastic Knapsack Problem Formulation

Let N_0 denote the initial capacity of the knapsack. Let $T \in (0, \infty]$ denote the deadline for accepting items, and let $\mathcal{T} \in [0, T]$ denote the chosen stopping time. Let $\lambda : [0, N_0] \times (0, T) \mapsto (0, \overline{\lambda}]$ be Borel-measurable and bounded. Let $\{A_i\}_{i=1}^{\infty}$ denote the arrival times of a (doubly stochastic) Poisson process on $(0, T)$ with rate $\lambda(n, t)$ at time t if the remaining capacity $N(t) = n$. Let S_i denote the size of arrival i, and let R_i denote the reward of arrival i. Assume that $\{(R_i, S_i)\}_{i=1}^{\infty}$ is an i.i.d. sequence, independent of $\{A_i\}_{i=1}^{\infty}$. Let $(\Omega, \mathcal{F}, P)$ be a probability space satisfying these assumptions. Let $F_{R,S}$ denote the joint probability distribution of R and S, F_S the probability distribution of S, F_R the probability distribution of R, $F_{R|S}$ a conditional probability distribution of R given S, and $F_{S|R}$ a conditional probability distribution of S given R. Assume that $\int_0^{N_0} \int_{-\infty}^{\infty} |r| \; F_{R|S}(dr|s) F_S(ds) < \infty$. Let D_i denote the decision whether to accept or reject arrival i, defined as follows:

$$D_i \equiv \begin{cases} 1 & \text{if arrival } i \text{ is accepted} \\ 0 & \text{if arrival } i \text{ is rejected} \end{cases}$$

Let $\mathcal{I}_s$ denote the set of all unit step functions $I_s : (0, T] \mapsto \{0, 1\}$ of the form

$$I_s(t) \equiv \begin{cases} 1 & \text{if } \; t \in (0, \tau] \\ 0 & \text{if } \; t \in (\tau, T] \end{cases}$$

for some $\tau \in [0, T]$. For $I_s \in \mathcal{I}_s$, let $I_s(t)$ denote whether t is less than the chosen stopping time $\mathcal{T}$ or not, defined as follows:

$$I_s(t) \equiv \begin{cases} 1 & \text{if } t \leq \mathcal{T} \\ 0 & \text{if } t > \mathcal{T} \end{cases}$$

The class Π_{DSKP}^{HD} of history-dependent deterministic policies for the DSKP is defined as follows. For any $t \in [0, \infty)$, let $\mathcal{H}_t$ be the history of the process $\{A_i, R_i, S_i\}$ up to time t, i.e., the σ-algebra generated by $\{(A_i, R_i, S_i) : A_i \leq t\}$, denoted by $\mathcal{H}_t \equiv \sigma\left(\{(A_i, R_i, S_i) : A_i \leq t\}\right)$. Let $\mathcal{H}_{t^-} \equiv \sigma\left(\{(A_i, R_i, S_i) : A_i < t\}\right)$, and let $\mathcal{H}_{A_i} \equiv \sigma\left(\{(A_j, R_j, S_j)\}_{j=1}^{i}\right)$. Let $\mathcal{A} \equiv \{\{A_i\}_{i=1}^{\infty} : 0 < A_1 < A_2 < \ldots < \infty\}$, let $\mathcal{R} \equiv \{\{R_i\}_{i=1}^{\infty} : R_i \in \Re\}$, let $\mathcal{S} \equiv \{\{S_i\}_{i=1}^{\infty} : S_i \in \Re_+\}$, where $\Re_+ \equiv [0, \infty)$, and let $\mathcal{D} \equiv \{\{D_i\}_{i=1}^{\infty} : D_i \in \{0, 1\}\}$. Define Π_{DSKP}^{HD} as the set of all Borel-measurable functions $\pi : \mathcal{A} \times \mathcal{R} \times \mathcal{S} \mapsto \mathcal{D} \times \mathcal{I}_s$ which satisfy the conditions

$$
\begin{gathered}
D_i^\pi \text{ is } \mathcal{H}_{A_i} \text{ measurable, i.e., } \{D_i^\pi = 1\} \in \mathcal{H}_{A_i} \\
\text{or } D_i^\pi : \{\{(A_j, R_j, S_j)\}_{j=1}^i\} \mapsto \{0,1\}, \text{ for all } i \in \{1,2,\ldots\} \\
I^\pi(t) \text{ is } \mathcal{H}_{t^-} \text{ measurable, i.e., } \{I^\pi(t) = 1\} \in \mathcal{H}_{t^-} \text{ for all } t \in (0,T] \\
\textstyle\sum_{\{i:A_i \le \mathcal{T}^\pi\}} D_i^\pi S_i \;\; \le \;\; N_0
\end{gathered}
$$

where $(\{D_i^\pi\}, I^\pi) \equiv \pi(\{A_i\},\{R_i\},\{S_i\})$, and the stopping time $\mathcal{T}^\pi$ is given by

$$
I^\pi(t) \;\; = \;\; \begin{cases} 1 & \text{if } \; t \in (0, \mathcal{T}^\pi] \\ 0 & \text{if } \; t \in (\mathcal{T}^\pi, T] \end{cases}
$$

Let $N^\pi(t)$ denote the remaining capacity under policy π at time t, where N^π is defined to be left-continuous, i.e.,

$$
N^\pi(t) \;\; \equiv \;\; N_0 - \sum_{\{i:A_i < t\}} D_i^\pi I^\pi(A_i) S_i \tag{4.1}
$$

and let the remaining capacity just after time t be denoted by

$$
N^\pi(t^+) \;\; \equiv \;\; N_0 - \sum_{\{i:A_i \le t\}} D_i^\pi I^\pi(A_i) S_i \tag{4.2}
$$

Let $c : [0, N_0] \times (0,T) \mapsto \Re$ be Borel-measurable and bounded, where $c(n,t)$ denotes the holding cost per unit time at time t if the remaining capacity $N^\pi(t) = n$. Let $p : [0, N_0] \times (0,T) \mapsto \Re$ be Borel-measurable and bounded, where $p(n,t)$ denotes the penalty that is incurred if an item is rejected at time t (i.e., the item's arrival time $A_i = t$), when the remaining capacity $N^\pi(t) = n$. Let $v : [0, N_0] \mapsto \Re$ be Borel-measurable and bounded, where $v(n)$ denotes the terminal value that is earned at time $\mathcal{T}^\pi$ if the remaining capacity $N^\pi(\mathcal{T}^{\pi+}) = n$. Let $\alpha \ge 0$ be the discount rate; if $T = \infty$, it is required that $\alpha > 0$.

Let V_{DSKP}^π denote the expected total discounted value under policy $\pi \in \Pi_{DSKP}^{HD}$, i.e.,

$$
\begin{aligned}
V_{DSKP}^\pi \;\; \equiv \;\; E\Bigg[& \sum_{\{i:A_i \le \mathcal{T}^\pi\}} e^{-\alpha A_i} \left[R_i D_i^\pi - p(N^\pi(A_i), A_i)\,(1 - D_i^\pi) \right] \\
& - \int_0^{\mathcal{T}^\pi} e^{-\alpha\tau} c(N^\pi(\tau), \tau)\, d\tau \; + \; e^{-\alpha \mathcal{T}^\pi} v(N^\pi(\mathcal{T}^{\pi+})) \Bigg| N^\pi(0^+) = N_0 \Bigg]
\end{aligned}
$$

Lemma 1

$$\begin{aligned} V^{\pi}_{DSKP} &= E\left[\sum_{\{i:A_i\le T\}} e^{-\alpha A_i}\left[R_i D^{\pi}_i - p(N^{\pi}(A_i),A_i)\left(1-D^{\pi}_i\right)\right] I^{\pi}(A_i)\right. \\ &\quad + \int_0^T e^{-\alpha\tau}\left[-c(N^{\pi}(\tau),\tau)I^{\pi}(\tau) + \alpha v(N^{\pi}(\tau))\left(1-I^{\pi}(\tau)\right)\right] d\tau \\ &\quad \left. + \, e^{-\alpha T} v(N^{\pi}(T^+)) \,\middle|\, N^{\pi}(0^+) = N_0 \right] \end{aligned} \tag{4.3}$$

The proof of Lemma 1 is given in Appendix A.

The objective is to find the optimal expected value V^*_{DSKP}, i.e.,

$$V^*_{DSKP} \equiv \sup_{\pi\in\Pi^{HD}_{DSKP}} V^{\pi}_{DSKP}$$

and to find an optimal policy $\pi^* \in \Pi^{HD}_{DSKP}$ that achieves this optimal value, if such a policy exists.

Let $\mathcal{I}_{RS}$ denote the set of all Borel-measurable functions $D_{RS} : \Re \times \Re_+ \mapsto \{0,1\}$. The class Π^{MD}_{DSKP} of memoryless deterministic policies for the DSKP is defined as the set of all Borel-measurable functions $\pi : [0, N_0] \times (0,T] \mapsto \mathcal{I}_{RS} \times \{0,1\}$, where $\pi \equiv (D^{\pi}, I^{\pi})$, and D^{π} and I^{π} are defined as follows. $D^{\pi}(n,t,r,s)$ denotes the decision under policy π whether to accept or reject an arrival i at time $A_i = t$ with size $S_i = s$ and reward $R_i = r$ if the remaining capacity $N^{\pi}(t) = n$, defined as follows:

$$D^{\pi}(n,t,r,s) \equiv \begin{cases} 1 & \text{if } s \le n \text{ and arrival } i \text{ is accepted} \\ 0 & \text{if } s > n \text{ or arrival } i \text{ is rejected} \end{cases}$$

Let the acceptance set for policy π be denoted by $\mathcal{R}^{\pi}_1(n,t) \equiv \{(r,s) \in \Re\times\Re_+ : D^{\pi}(n,t,r,s) = 1\}$, and let the rejection set be denoted by $\mathcal{R}^{\pi}_0(n,t) \equiv \{(r,s) \in \Re \times \Re_+ : D^{\pi}(n,t,r,s) = 0\}$. $I^{\pi}(n,t)$ denotes the decision under policy π whether the process has been stopped by time t if the remaining capacity $N^{\pi}(t) = n$, defined as follows:

$$I^{\pi}(n,t) \equiv \begin{cases} 1 & \text{if the process has not been stopped before time } t \\ 0 & \text{if the process has been stopped before time } t \end{cases}$$

It has to be ensured that $I^{\pi}(n,t)$ remains 0 once the decision has been made to stop, i.e., that $I^{\pi}(N^{\pi}(t)(\omega),t) = 0$ for all $t > \mathcal{T}^{\pi}(\omega)$ for each $\omega \in \Omega$. To ensure this property, it is sufficient to require that $I^\pi(n,\cdot) \in \mathcal{I}_s$, i.e.,

$$I^\pi(n,t) = \begin{cases} 1 & \text{if } t \in (0, \tau^\pi(n)] \\ 0 & \text{if } t \in (\tau^\pi(n), T] \end{cases}$$

for some $\tau^\pi(n) \in [0,T]$, for each $n \in [0, N_0]$ and each $\pi \in \Pi^{MD}_{DSKP}$. Then it follows that each policy $\pi \in \Pi^{MD}_{DSKP}$ is admissible for the DSKP, i.e., that $\Pi^{MD}_{DSKP} \subset \Pi^{HD}_{DSKP}$. The remaining capacity corresponding to policy π is given by

$$N^\pi(t) = N_0 - \sum_{\{i:A_i<t\}} D^\pi(N^\pi(A_i), A_i, R_i, S_i) I^\pi(N^\pi(A_i), A_i) S_i \tag{4.4}$$

$$N^\pi(t^+) = N_0 - \sum_{\{i:A_i\le t\}} D^\pi(N^\pi(A_i), A_i, R_i, S_i) I^\pi(N^\pi(A_i), A_i) S_i \tag{4.5}$$

Let $V^\pi_{DSKP}(n,t)$ denote the expected total discounted value under policy $\pi \in \Pi^{MD}_{DSKP}$ from time t until time T, if the remaining capacity $N^\pi(t^+) = n$, i.e.,

$$\begin{aligned} V^\pi_{DSKP}(n,t) \equiv E\Bigg[& \sum_{\{i:A_i\in(t,\mathcal{T}^\pi]\}} e^{-\alpha(A_i-t)} \left[R_i D^\pi(N^\pi(A_i), A_i, R_i, S_i)\right. \\ & \left. - p(N^\pi(A_i), A_i)\left(1 - D^\pi(N^\pi(A_i), A_i, R_i, S_i)\right)\right] \\ & - \int_t^{\mathcal{T}^\pi} e^{-\alpha(\tau-t)} c(N^\pi(\tau))\, d\tau \\ & + e^{-\alpha(\mathcal{T}^\pi-t)} v(N^\pi(\mathcal{T}^{\pi+})) \Bigg| N^\pi(t^+) = n \Bigg] \end{aligned} \tag{4.6}$$

Lemma 2

$$\begin{aligned} V^\pi_{DSKP}(n,t) &= E\Bigg[\sum_{\{i:A_i\in(t,T]\}} e^{-\alpha(A_i-t)} \left[R_i D^\pi(N^\pi(A_i), A_i, R_i, S_i)\right. \\ &\quad \left. - p(N^\pi(A_i), A_i)\left(1 - D^\pi(N^\pi(A_i), A_i, R_i, S_i)\right)\right] I^\pi(N^\pi(A_i), A_i) \\ &\quad + \int_t^T e^{-\alpha(\tau-t)} \left[-c(N^\pi(\tau),\tau) I^\pi(N^\pi(\tau),\tau)\right. \\ &\quad \left. + \alpha v(N^\pi(\tau))\left(1 - I^\pi(N^\pi(\tau),\tau)\right)\right] d\tau \\ &\quad + e^{-\alpha(T-t)} v(N^\pi(T^+)) \Bigg| N^\pi(t^+) = n \Bigg] \end{aligned} \tag{4.7}$$

$$\begin{aligned} &= E\Bigg[\int_t^T e^{-\alpha(\tau-t)} \Bigg\{ \lambda(N^\pi(\tau),\tau) \Bigg[\iint_{\mathcal{R}^\pi_1(N^\pi(\tau),\tau)} r\, F_{R|S}(dr|s) F_S(ds) \\ &\quad - p(N^\pi(\tau),\tau) \iint_{\mathcal{R}^\pi_0(N^\pi(\tau),\tau)} F_{R|S}(dr|s) F_S(ds) \Bigg] I^\pi(N^\pi(\tau),\tau) \end{aligned}$$

$$-c(N^\pi(\tau),\tau)I^\pi(N^\pi(\tau),\tau) + \alpha v(N^\pi(\tau))\left(1 - I^\pi(N^\pi(\tau),\tau)\right)\Big\} d\tau$$
$$+ e^{-\alpha(T-t)}v(N^\pi(T^+))\Big| N^\pi(t^+) = n\Big]$$

Proof: The proof of the first equality is similar to that of Lemma 1. The second equality follows from an integration theorem for point processes; see for example Brémaud [23] Theorem II.T8. □

Let $V^*_{DSKP}(n,t)$ be the corresponding optimal expected value, i.e.,

$$V^*_{DSKP}(n,t) \equiv \sup_{\pi\in\Pi^{MD}_{DSKP}} V^\pi_{DSKP}(n,t)$$

4.3 A Related Markov Decision Process

If the DSKP did not include the option to stop at any time before the deadline, then the problem could be formulated as a continuous time Markov decision process (MDP). However, the option of choosing a stopping time in continuous time for the DSKP introduces a complexity into the problem that is not modeled in a tractable way by an MDP. If a *stopped* state is introduced, then the process should make an infinite rate transition to this state as soon as the decision is made to stop. However, most of the results for MDP's assume that transition rates are bounded. Also see the comment by Boel and Varaiya [21] (Remark 3.1.(iv)) to this effect. Therefore an MDP is studied which is a relaxation of the DSKP, in that the MDP can switch *off* and *on* multiple times, instead of stopping only once. This MDP can be modeled with bounded transition rates. It is shown in Chapter 5 for the infinite horizon case (no deadline) and in Chapter 6 for the finite horizon case (finite deadline) that there exist optimal policies for the MDP which stop only once, and hence which are admissible and optimal for the DSKP.

4.3.1 Markov Decision Process Formulation

The MDP has state space $[0, N_0]$. The policy spaces Π^{HD}_{MDP}, Π^{MD}_{MDP}, and Π^{SD}_{MDP}, are defined hereafter, where superscript HD denotes history-dependent deterministic policies, MD denotes memoryless deterministic policies, and SD denotes stationary deterministic policies. Let $\mathcal{I}_I$ denote the set of all Borel-measurable functions $I_I : (0,T] \mapsto \{0,1\}$. Note that $\mathcal{I}_I \supset \mathcal{I}_s$. The class Π^{HD}_{MDP} is defined as the set of all Borel-measurable functions $\pi : \mathcal{A} \times \mathcal{R} \times \mathcal{S} \mapsto \mathcal{D} \times \mathcal{I}_I$ which satisfy the conditions

$$\begin{aligned} & D_i^\pi \text{ is } \mathcal{H}_{A_i} \text{ measurable for all } i \in \{1,2,\ldots\} \\ & I^\pi(t) \text{ is } \mathcal{H}_{t-} \text{ measurable for all } t \in (0,T] \\ & \textstyle\sum_{\{i:A_i \leq T\}} D_i^\pi I^\pi(A_i) S_i \quad \leq \quad N_0 \end{aligned}$$

where $(\{D_i^\pi\}, I^\pi) \equiv \pi(\{A_i\},\{R_i\},\{S_i\})$. The remaining capacity under policy π at time t is given by Equations (4.1) and (4.2). Note that the MDP is allowed to switch *on* $(I^\pi(t) = 1)$ and *off* $(I^\pi(t) = 0)$ multiple times, in contrast with the DSKP, which has to remain *off* once it stops. Hence, with the expected value V^π_{MDP} of the MDP properly defined, the MDP is a relaxation of the DSKP. The optimal expected value of the MDP is therefore at least as good as that of the DSKP. This is the result of Lemma 3, which follows after the definitions of V^π_{MDP} and V^*_{MDP}.

V^π_{MDP} denotes the expected total discounted value under policy $\pi \in \Pi^{HD}_{MDP}$, given by

$$\begin{aligned} V^\pi_{MDP} \quad \equiv \quad E \Bigg[& \sum_{\{i:A_i \leq T\}} e^{-\alpha A_i} \left[R_i D_i^\pi - p(N^\pi(A_i), A_i)\,(1 - D_i^\pi) \right] I^\pi(A_i) \\ & + \int_0^T e^{-\alpha\tau} \left[-c(N^\pi(\tau),\tau) I^\pi(\tau) + \alpha v(N^\pi(\tau))\,(1 - I^\pi(\tau)) \right] d\tau \\ & + e^{-\alpha T} v(N^\pi(T^+)) \Bigg| N^\pi(0^+) = N_0 \Bigg] \end{aligned} \tag{4.8}$$

The optimal expected value V^*_{MDP} is given by

$$V^*_{MDP} \quad \equiv \quad \sup_{\pi \in \Pi^{HD}_{MDP}} V^\pi_{MDP}$$

Lemma 3 $V^*_{MDP} \quad \geq \quad V^*_{DSKP}$.

Proof: $\mathcal{I}_s \subset \mathcal{I}_I$ implies that $\Pi^{HD}_{DSKP} \subset \Pi^{HD}_{MDP}$. Comparing Equations (4.3) and (4.8), it follows that for any $\pi \in \Pi^{HD}_{DSKP}$, $V^\pi_{MDP} = V^\pi_{DSKP}$. Hence,

$$V^*_{MDP} \quad \equiv \quad \sup_{\pi \in \Pi^{HD}_{MDP}} V^\pi_{MDP} \quad \geq \quad \sup_{\pi \in \Pi^{HD}_{DSKP}} V^\pi_{MDP} \quad = \quad \sup_{\pi \in \Pi^{HD}_{DSKP}} V^\pi_{DSKP} \quad \equiv \quad V^*_{DSKP} \qquad \square$$

The class Π^{MD}_{MDP} of memoryless deterministic policies for the MDP is defined as the set of all Borel-measurable functions $\pi : [0, N_0] \times (0, T] \mapsto \mathcal{I}_{RS} \times \{0,1\}$, where $\pi \equiv (D^\pi, I^\pi)$, D^π

is defined as for $\pi \in \Pi^{MD}_{DSKP}$, and I^π is defined as follows. $I^\pi(n,t)$ denotes the decision under policy π whether to be switched *on* or *off* at time t if the remaining capacity $N^\pi(t) = n$, defined as follows:

$$I^\pi(n,t) \equiv \begin{cases} 1 & \text{if switched } on \text{ at time } t \\ 0 & \text{if switched } off \text{ at time } t \end{cases}$$

As for policies $\pi \in \Pi^{HD}_{MDP}$, policies $\pi \in \Pi^{MD}_{MDP}$ are allowed to switch *on* and *off* multiple times. It is easy to show that $\Pi^{MD}_{MDP} \subset \Pi^{HD}_{MDP}$. The remaining capacity corresponding to policy π is given by Equations (4.4) and (4.5).

The MDP can be modeled with transition rates

$$\lambda(n,t,\pi(n,t)) \equiv \lambda(n,t)I^\pi(n,t)$$

and Markov kernels

$$\begin{aligned}
P[n \mid n, \pi(n,t)] &\equiv \iint_{\mathcal{R}_0^\pi(n,t)} F_{R|S}(dr|s)F_S(ds) \\
&\quad + \iint_{\mathcal{R}_1^\pi(n,t)\cap(\Re\times\{0\})} F_{R|S}(dr|s)F_S(ds) \\
P[n - ds \mid n, \pi(n,t)] &\equiv \iint_{\mathcal{R}_1^\pi(n,t)\cap(\Re\times ds)} F_{R|S}(dr|s)F_S(ds) \\
\text{i.e., } P[n - B \mid n, \pi(n,t)] &\equiv \iint_{\mathcal{R}_1^\pi(n,t)\cap(\Re\times B)} F_{R|S}(dr|s)F_S(ds)
\end{aligned}$$

for any Borel set $B \subseteq (0,\infty)$.

Let $V^\pi_{MDP}(n,t)$ be the expected total discounted value under policy $\pi \in \Pi^{MD}_{MDP}$ from time t until time T, if the remaining capacity $N^\pi(t^+) = n$, i.e.,

$$\begin{aligned}
V^\pi_{MDP}(n,t) &\equiv E\Bigg[\sum_{\{i:A_i\in(t,T]\}} e^{-\alpha(A_i-t)}\left[R_i D^\pi(N^\pi(A_i),A_i,R_i,S_i)\right. \\
&\qquad \left. -p(N^\pi(A_i),A_i)\left(1 - D^\pi(N^\pi(A_i),A_i,R_i,S_i)\right)\right] I^\pi(N^\pi(A_i),A_i) \\
&\qquad + \int_t^T e^{-\alpha(\tau-t)}\left[-c(N^\pi(\tau),\tau)I^\pi(N^\pi(\tau),\tau)\right. \\
&\qquad\qquad \left. +\alpha v(N^\pi(\tau))\left(1 - I^\pi(N^\pi(\tau),\tau)\right)\right] d\tau \\
&\qquad + e^{-\alpha(T-t)}v(N^\pi(T^+)) \Bigg| N^\pi(t^+) = n\Bigg] \qquad (4.9) \\
&= E\Bigg[\int_t^T e^{-\alpha(\tau-t)}\Bigg\{\lambda(N^\pi(\tau),\tau)\Bigg[\iint_{\mathcal{R}_1^\pi(N^\pi(\tau),\tau)} r\, F_{R|S}(dr|s)F_S(ds) \\
&\qquad -p(N^\pi(\tau),\tau)\iint_{\mathcal{R}_0^\pi(N^\pi(\tau),\tau)} F_{R|S}(dr|s)F_S(ds)\Bigg] I^\pi(N^\pi(\tau),\tau)
\end{aligned}$$

$$-c(N^\pi(\tau),\tau)I^\pi(N^\pi(\tau),\tau)+\alpha v(N^\pi(\tau))\,(1-I^\pi(N^\pi(\tau),\tau))\Big\}\,d\tau$$
$$\left. + \; e^{-\alpha(T-t)}v(N^\pi(T^+))\right| N^\pi(t^+)=n\Big]$$

The equality again follows from an integration theorem for point processes as given in Brémaud [23] Theorem II.T8. Thus, the MDP can be modeled with a reward rate per unit time

$$\begin{aligned}\xi(n,t,\pi(n,t)) \;\; &\equiv \;\; \lambda(n,t)\left[\int\int_{\mathcal{R}_1^\pi(n,t)} r\,F_{R|S}(dr|s)F_S(ds)\right.\\ &\qquad\qquad \left. -p(n,t)\int\int_{\mathcal{R}_0^\pi(n,t)} F_{R|S}(dr|s)F_S(ds)\right] I^\pi(n,t)\\ &\qquad -\; c(n,t)I^\pi(n,t) \;+\; \alpha v(n)\,(1-I^\pi(n,t))\end{aligned}$$

(Most continuous time MDP's in the literature are formulated with a reward rate per unit time $\xi(n,t,\pi(n,t))$ and not with rewards r that are received according to a point process.) The assumptions that $\int_0^{N_0}\int_0^\infty r\,F_{R|S}(dr|s)F_S(ds)<\infty$, and that λ, p, c and v are bounded, imply that $\sup_{\pi\in\Pi_{MDP}^{MD}}\xi(n,t,\pi(n,t))$ is bounded. Also note that if $I^\pi(n,t)=0$ for all $t\in(t_1,T)$, then $V_{MDP}^\pi(n,t)=v(n)$ for all $t\in[t_1,T]$, as for the DSKP.

Let $V_{MDP}^*(n,t)$ be the corresponding optimal expected value, i.e.,

$$V_{MDP}^*(n,t) \;\;\equiv\;\; \sup_{\pi\in\Pi_{MDP}^{MD}} V_{MDP}^\pi(n,t)$$

Lemma 4 $V_{MDP}^*(n,t) \;\;\geq\;\; V_{DSKP}^*(n,t)$ *for all* n *and* t.

Proof: For any policy $\pi\in\Pi_{DSKP}^{MD}$, it is required that $I^\pi(n,\cdot)\in\mathcal{I}_s$ for all $n\in[0,N_0]$, and this is not required of policies $\pi\in\Pi_{MDP}^{MD}$. Thus $\Pi_{DSKP}^{MD}\subset\Pi_{MDP}^{MD}$. Comparing Equations (4.7) and (4.9), it follows that for any $\pi\in\Pi_{DSKP}^{MD}$, and for any n and t, $V_{MDP}^\pi(n,t)=V_{DSKP}^\pi(n,t)$. Hence,

$$\begin{aligned}V_{MDP}^*(n,t) \;\;&\equiv\;\; \sup_{\pi\in\Pi_{MDP}^{MD}} V_{MDP}^\pi(n,t) \;\;\geq\;\; \sup_{\pi\in\Pi_{DSKP}^{MD}} V_{MDP}^\pi(n,t)\\ &=\;\; \sup_{\pi\in\Pi_{DSKP}^{MD}} V_{DSKP}^\pi(n,t) \;\;\equiv\;\; V_{DSKP}^*(n,t) \qquad\qquad \square\end{aligned}$$

Note that $V_{MDP}^*(n,t)\geq v(n)$ for all n and t, because the policy $\pi\in\Pi_{MDP}^{MD}$ with $I^\pi=0$ has $V_{MDP}^\pi(n,t)=v(n)$ for all n and t.

Intuitively it is expected that if λ, p and c are independent of t, then there is an optimal memoryless policy $\pi^* \in \Pi_{MDP}^{MD}$ for the MDP such that $\pi^* \in \Pi_{DSKP}^{MD}$, i.e., such that $I^*(n, \cdot) \in \mathcal{I}_s$ for each $n \in [0, N_0]$, for the following reason. For any $t_1 \in [0, T]$ such that $V_{MDP}^*(n, t_1) > v(n)$, it holds that $V_{MDP}^*(n, t) > v(n)$ for all $t \in [0, t_1]$, because $V_{MDP}^*(n, t)$ is nonincreasing in t as is shown in Proposition 2. Hence, if the remaining capacity is n, it is optimal to continue waiting (i.e., $I^*(n, t) = 1$) for all $t \in (0, t_1]$. Similarly, for any $t_2 \in [0, T]$ such that $V_{MDP}^*(n, t_2) = v(n)$, it holds that $V_{MDP}^*(n, t) = v(n)$ and it is optimal to stop (i.e., $I^*(n, t) = 0$) for all $t \in [t_2, T]$. It is shown in Chapters 5 and 6 that there are optimal memoryless policies $\pi^* \in \Pi_{MDP}^{MD}$ for the MDP such that $\pi^* \in \Pi_{DSKP}^{MD}$, and hence which are admissible and optimal for the DSKP among all policies $\pi \in \Pi_{DSKP}^{MD}$. Under certain conditions, it also follows that such a memoryless policy π^* is optimal for the MDP among all history-dependent policies $\pi \in \Pi_{MDP}^{HD}$. From this it follows that π^* is also an optimal policy for the DSKP among all policies $\pi \in \Pi_{DSKP}^{HD}$.

A policy $\pi \in \Pi_{DSKP}^{MD}$ or $\pi \in \Pi_{MDP}^{MD}$ is said to be a threshold policy if it has a threshold acceptance rule D^π with a reward threshold $x^\pi : [0, N_0] \times (0, T] \times \Re_+ \mapsto \Re$. If an item that arrives at time t when the remaining capacity $N^\pi(t) = n$, has size $s \leq n$ and reward $r \geq x^\pi(n, t, s)$, then the item is accepted; otherwise the item is rejected. That is,

$$D^\pi(n, t, r, s) \quad \equiv \quad \begin{cases} 1 & \text{if} \quad s \leq n \text{ and } r \geq x^\pi(n, t, s) \\ 0 & \text{if} \quad s > n \text{ or } r < x^\pi(n, t, s) \end{cases}$$

The following argument suggests that reward threshold $x^*(n, t, s) = V^*(n, t) - V^*(n-s, t) - p(n, t)$ gives an optimal acceptance rule. Suppose an item with size $s \leq n$ and reward r arrives at time t when the remaining capacity $N^*(t) = n$, and $I^*(n, t) = 1$. If the item is accepted, the optimal expected value from then on is $r + V^*(n-s, t)$. If the item is rejected, the optimal expected value from then on is $V^*(n, t) - p(n, t)$. Hence, the item is accepted if $r + V^*(n - s, t) \geq V^*(n, t) - p(n, t)$, i.e., if $r \geq V^*(n, t) - V^*(n - s, t) - p(n, t)$; otherwise the item is rejected. It is shown that there is a threshold policy π^* with reward threshold $x^*(n, t, s) = V^*(n, t) - V^*(n - s, t) - p(n, t)$ that is optimal among all policies $\pi \in \Pi_{MDP}^{MD}$.

Under typical conditions the reward threshold x^* is equivalent to a size threshold $z^* : [0, N_0] \times (0, T] \times \Re \mapsto [0, N_0]$, defined by

$$z^*(n, t, r) \quad \equiv \quad \sup\{s \in [0, n] : r \geq V^*(n, t) - V^*(n - s, t) - p(n, t)\}$$

(where $\sup \emptyset \equiv -\infty$). If an item that arrives at time t when the remaining capacity $N^*(t) = n$, has reward r and size $s \leq z^*(n,t,r) \leq n$, then the item is accepted; otherwise the item is rejected. That is,

$$D^*(n,t,r,s) \equiv \begin{cases} 1 & \text{if } s \leq z^*(n,t,r) \\ 0 & \text{if } s > z^*(n,t,r) \end{cases}$$

It is shown that under conditions that usually hold in practical applications, the size threshold z^* is equivalent to the reward threshold x^*.

The class Π^{SD}_{DSKP} (Π^{SD}_{MDP}) of stationary deterministic policies for the DSKP (MDP) is the subset of Π^{MD}_{DSKP} (Π^{MD}_{MDP}) of policies π which do not depend on t. Stationary policies $\pi \in \Pi^{SD}_{MDP}$ have unit step function stopping rules with $\tau^\pi(n) = T$ if $I^\pi(n) = 1$, and $\tau^\pi(n) = 0$ if $I^\pi(n) = 0$. Therefore, $\Pi^{SD}_{DSKP} = \Pi^{SD}_{MDP}$, and for any $\pi \in \Pi^{SD}_{MDP}$, $V^\pi_{DSKP} = V^\pi_{MDP}$.

4.3.2 Optimality Results for Markov Decision Processes

Dynamic programming, including Markov decision processes, has long been studied by researchers. A small number of the important results are highlighted in this section.

Blackwell [19] showed that there is a stationary policy that is optimal among all memoryless deterministic policies for a discrete time infinite horizon discounted MDP with finite state and action spaces and an additive objective function. He also considered the undiscounted case. It was shown by Derman and Strauch [61] (Theorem 2) for a discrete time MDP with finite state and action spaces, and an arbitrary objective function, that there is a memoryless deterministic policy that is optimal among all history-dependent randomized policies. It was also shown (Theorem 1) that if the time horizon is infinite, and the Markov chain is irreducible for each stationary deterministic policy, then for a very general class of objective functions, including the expected total discounted value, there is a stationary deterministic policy that is optimal among all history-dependent randomized policies.

It was shown by Miller [111] (Sections 2 and 3) for a continuous time MDP with finite state and action spaces, that for each memoryless deterministic policy, the resulting stochastic process is a (measurable) Markov process with a unique absolutely continuous Markov transition function. A similar result was shown by Pliska [120] (Theorem 6) with a state space that is a Polish space and an action space that is a compact subset of a Polish space.

Rykov [154] (Theorems 1, 4 and 5) showed optimality results for a continuous time MDP with finite state and action spaces, and expected total discounted value as well as long run average value as objectives, for the finite and infinite horizon cases, considering measurable history-dependent deterministic policies. Yushkevich and Feinberg [187] (Theorem 2) extended these results to the case with a countable state space, an arbitrary measurable action space, and expected total discounted value as objective, also considering measurable history-dependent deterministic policies. Pliska [120] (Theorem 20) obtained necessary and sufficient conditions for a memoryless deterministic policy to be optimal, and sufficient conditions for the existence of such an optimal policy, for a finite horizon continuous time MDP with a Polish state space, an action space that is a compact subset of a Polish space, and total value as objective. For the infinite horizon case, he showed that there is a stationary policy that is optimal among all memoryless deterministic policies, for the transient as well as discounted cases.

Boel and Varaiya [21] studied more general controlled jump processes, with partial as well as complete information, and total value as objective. Necessary and sufficient conditions for a history-dependent deterministic policy to be optimal were obtained. These results were specialized to the Markovian case, for which necessary and sufficient conditions were found for a memoryless deterministic policy to be optimal among all history-dependent deterministic policies.

5. The Infinite Horizon DSKP

In this chapter the Dynamic and Stochastic Knapsack Problem with no deadline is studied. In Section 5.1 it is shown that if λ, p and c are independent of t, then attention can be restricted to the class of stationary policies Π_{MDP}^{SD}. In Section 5.2 equations are derived for the expected value V^{π} of a policy $\pi \in \Pi_{MDP}^{SD}$. The optimality equation, giving the optimal expected value function V^*, is introduced in Section 5.3. In Section 5.4 a threshold policy π^* is obtained that is optimal for the MDP among all policies $\pi \in \Pi_{MDP}^{HD}$. This policy π^* is also optimal in Π_{DSKP}^{HD} for the DSKP. An intuitive recursive form that gives the optimal value function V^* is investigated in Section 5.5. This leads to a simple structure for π^* and an efficient algorithm for computing V^* and π^*, which is given in Section 5.6.

5.1 Optimality of Stationary Policies

Assume that λ, p and c are independent of t. For this case it was shown by Yushkevich and Feinberg [187] (Theorem 2) for an MDP with a countable state space, a measurable action space, and an infinite horizon, that if $\alpha > 0$ and $\sup_{\pi \in \Pi_{MDP}^{SD}} \sup_{n \in [0,N_0]} \xi(n, \pi(n)) < \infty$, then for any $\varepsilon > 0$ there exists a stationary deterministic policy $\pi \in \Pi_{MDP}^{SD}$ that is ε-optimal among all history-dependent deterministic policies $\pi \in \Pi_{MDP}^{HD}$. Therefore, attention is restricted to the class of policies Π_{MDP}^{SD}. Because $V_{DSKP}^{\pi} = V_{MDP}^{\pi}$ for any $\pi \in \Pi_{MDP}^{SD}$, the subscripts of V are dropped in this chapter. For $\pi \in \Pi_{MDP}^{SD}$, D^{π} is a function of n, r and s only, and I^{π} and V^{π} are functions of n only. This also means that stopping times are restricted to the starting time and the times when the state (remaining capacity) changes, i.e., those arrival times when items are accepted. Because decisions can be restricted to the starting time and arrival times, an intuitive result is that with an infinite horizon, a continuous time Markov decision process is equivalent to a discrete time Markov decision process, in the sense that for any stationary policy $\pi \in \Pi_{MDP}^{SD}$, the expected value $V^{\pi}(n)$ is the same for both processes for all n. It was shown by Serfozo [159] that this holds for countable state and action spaces (and it can readily be extended to more general spaces) if the transition rates are bounded. He also showed how to construct the transformation from a continuous time Markov decision process to a discrete time Markov decision process if there are lumped rewards as well as continuously earned rewards.

5.2 The Value of a Policy

If $I^\pi(n) = 0$, then $V^\pi(n) = v(n)$. If $I^\pi(n) = 1$, then by conditioning on the arrival time A_k, the size S_k, and the reward R_k of the first arrival k after time t, it follows that

$$
\begin{aligned}
V^\pi(n) &= V^\pi(n,t) \\
&= \int_0^\infty \int_{-\infty}^\infty \int_t^\infty \lambda(n) e^{-\lambda(n)(a_k - t)} E\Bigg[\sum_{\{i : A_i \in (t, \mathcal{T}^\pi]\}} e^{-\alpha(A_i - t)} \times \\
&\qquad [R_i D^\pi(N^\pi(A_i), R_i, S_i) - p(N^\pi(A_i)) (1 - D^\pi(N^\pi(A_i), R_i, S_i))] \\
&\qquad - \int_t^{\mathcal{T}^\pi} e^{-\alpha(\tau - t)} c(N^\pi(\tau))\, d\tau \;+\; e^{-\alpha(\mathcal{T}^\pi - t)} v\left(N^\pi\left(\mathcal{T}^{\pi+}\right)\right) \Bigg| \\
&\qquad N^\pi(t^+) = n, A_k = a_k, R_k = r_k, S_k = s_k \Bigg] \, da_k \, F_{R|S}(dr_k|s_k) F_S(ds_k) \\
&= \int_0^\infty \int_{-\infty}^\infty \int_t^\infty \lambda(n) e^{-\lambda(n)(a_k - t)} \Bigg[e^{-\alpha(a_k - t)} [r_k D^\pi(n, r_k, s_k) - p(n) (1 - D^\pi(n, r_k, s_k))] \\
&\qquad - \int_t^{a_k} e^{-\alpha(\tau - t)} c(n)\, d\tau \;+\; E\Bigg[\sum_{\{i : A_i \in (a_k, \mathcal{T}^\pi]\}} e^{-\alpha(A_i - t)} \times \\
&\qquad [R_i D^\pi(N^\pi(A_i), R_i, S_i) - p(N^\pi(A_i)) (1 - D^\pi(N^\pi(A_i), R_i, S_i))] \\
&\qquad - \int_{a_k}^{\mathcal{T}^\pi} e^{-\alpha(\tau - t)} c(N^\pi(\tau))\, d\tau \;+\; e^{-\alpha(\mathcal{T}^\pi - t)} v\left(N^\pi\left(\mathcal{T}^{\pi+}\right)\right) \Bigg| \\
&\qquad N^\pi(t^+) = n, A_k = a_k, R_k = r_k, S_k = s_k \Bigg]\Bigg] \, da_k \, F_{R|S}(dr_k|s_k) F_S(ds_k) \\
&= -\frac{1}{\alpha + \lambda(n)} c(n) \;+\; \frac{\lambda(n)}{\alpha + \lambda(n)} \Bigg[-p(n) \int\!\!\int_{\mathcal{R}_0^\pi(n)} F_{R|S}(dr_k|s_k) F_S(ds_k) \\
&\qquad + \int\!\!\int_{\mathcal{R}_1^\pi(n)} r_k \, F_{R|S}(dr_k|s_k) F_S(ds_k) \Bigg] \\
&\quad + \int\!\!\int_{\mathcal{R}_0^\pi(n)} \int_t^\infty \lambda(n) e^{-(\alpha + \lambda(n))(a_k - t)} E\Bigg[\sum_{\{i : A_i \in (a_k, \mathcal{T}^\pi]\}} e^{-\alpha(A_i - a_k)} \times \\
&\qquad [R_i D^\pi(N^\pi(A_i), R_i, S_i) - p(N^\pi(A_i)) (1 - D^\pi(N^\pi(A_i), R_i, S_i))]
\end{aligned}
$$

$$- \int_{a_k}^{\mathcal{T}^\pi} e^{-\alpha(\tau - a_k)} c(N^\pi(\tau))\, d\tau \;+\; e^{-\alpha(\mathcal{T}^\pi - a_k)} v\left(N^\pi\left(\mathcal{T}^{\pi+}\right)\right) \Bigg|$$

$$\left. N^\pi(a_k^+) = n, A_k = a_k, R_k = r_k, S_k = s_k \right] da_k\, F_{R|S}(dr_k|s_k) F_S(ds_k)$$

$$+ \iint_{\mathcal{R}_1^\pi(n)} \int_t^\infty \lambda(n) e^{-(\alpha+\lambda(n))(a_k - t)} E\left[\sum_{\{i: A_i \in (a_k, \mathcal{T}^\pi]\}} e^{-\alpha(A_i - a_k)} \times \right.$$
$$[R_i D^\pi(N^\pi(A_i), R_i, S_i) - p(N^\pi(A_i))\,(1 - D^\pi(N^\pi(A_i), R_i, S_i))]$$
$$- \int_{a_k}^{\mathcal{T}^\pi} e^{-\alpha(\tau - a_k)} c(N^\pi(\tau))\, d\tau \;+\; e^{-\alpha(\mathcal{T}^\pi - a_k)} v\left(N^\pi\left(\mathcal{T}^{\pi+}\right)\right) \Bigg|$$

$$\left. N^\pi(a_k^+) = n - s_k, A_k = a_k, R_k = r_k, S_k = s_k \right] da_k\, F_{R|S}(dr_k|s_k) F_S(ds_k)$$

From Equation (4.6), independence of $\{A_i\}$ and $\{(R_i, S_i)\}$, and the memoryless arrival process, it follows that for $\pi \in \Pi_{MDP}^{SD}$, if $I^\pi(n) = 1$, then

$$\iint_{\mathcal{R}_0^\pi(n)} \int_t^\infty \lambda(n) e^{-(\alpha+\lambda(n))(a_k - t)} E\left[\sum_{\{i: A_i \in (a_k, \mathcal{T}^\pi]\}} e^{-\alpha(A_i - a_k)} \times \right.$$
$$[R_i D^\pi(N^\pi(A_i), R_i, S_i) - p(N^\pi(A_i))\,(1 - D^\pi(N^\pi(A_i), R_i, S_i))]$$
$$- \int_{a_k}^{\mathcal{T}^\pi} e^{-\alpha(\tau - a_k)} c(N^\pi(\tau))\, d\tau \;+\; e^{-\alpha(\mathcal{T}^\pi - a_k)} v\left(N^\pi\left(\mathcal{T}^{\pi+}\right)\right) \Bigg|$$

$$\left. N^\pi(a_k^+) = n, A_k = a_k, R_k = r_k, S_k = s_k \right] da_k\, F_{R|S}(dr_k|s_k) F_S(ds_k)$$

$$\begin{aligned}
&= \iint_{\mathcal{R}_0^\pi(n)} \int_t^\infty \lambda(n) e^{-(\alpha+\lambda(n))(a_k - t)} V^\pi(n, a_k)\, da_k\, F_{R|S}(dr_k|s_k) F_S(ds_k) \\
&= \iint_{\mathcal{R}_0^\pi(n)} \int_t^\infty \lambda(n) e^{-(\alpha+\lambda(n))(a_k - t)} V^\pi(n)\, da_k\, F_{R|S}(dr_k|s_k) F_S(ds_k) \\
&= \frac{\lambda(n)}{\alpha + \lambda(n)} V^\pi(n) \iint_{\mathcal{R}_0^\pi(n)} F_{R|S}(dr_k|s_k) F_S(ds_k)
\end{aligned}$$

Similarly,

$$\iint_{\mathcal{R}_1^\pi(n)} \int_t^\infty \lambda(n) e^{-(\alpha+\lambda(n))(a_k - t)} E\left[\sum_{\{i: A_i \in (a_k, \mathcal{T}^\pi]\}} e^{-\alpha(A_i - a_k)} \times \right.$$
$$[R_i D^\pi(N^\pi(A_i), R_i, S_i) - p(N^\pi(A_i))\,(1 - D^\pi(N^\pi(A_i), R_i, S_i))]$$
$$- \int_{a_k}^{\mathcal{T}^\pi} e^{-\alpha(\tau - a_k)} c(N^\pi(\tau))\, d\tau \;+\; e^{-\alpha(\mathcal{T}^\pi - a_k)} v\left(N^\pi\left(\mathcal{T}^{\pi+}\right)\right) \Bigg|$$

$$\left. N^\pi(a_k^+) = n - s_k, A_k = a_k, R_k = r_k, S_k = s_k \right] da_k\, F_{R|S}(dr_k) F_S(ds_k)$$

$$= \frac{\lambda(n)}{\alpha+\lambda(n)} \int\int_{\mathcal{R}_1^\pi(n)} V^\pi(n-s_k)\, F_{R|S}(dr_k|s_k)F_S(ds_k)$$

Therefore, if $I^\pi(n)=1$, then

$$\begin{aligned}
V^\pi(n) &= -\frac{1}{\alpha+\lambda(n)}c(n) + \frac{\lambda(n)}{\alpha+\lambda(n)}\left[-p(n)\int\int_{\mathcal{R}_0^\pi(n)} F_{R|S}(dr|s)F_S(ds)\right. \\
&\qquad \left. + \int\int_{\mathcal{R}_1^\pi(n)} r\, F_{R|S}(dr|s)F_S(ds)\right] \\
&\quad + \frac{\lambda(n)}{\alpha+\lambda(n)}\left[V^\pi(n)\int\int_{\mathcal{R}_0^\pi(n)} F_{R|S}(dr|s)F_S(ds)\right. \\
&\qquad \left. + \int\int_{\mathcal{R}_1^\pi(n)} V^\pi(n-s)\, F_{R|S}(dr|s)F_S(ds)\right] \\
&= \frac{\lambda(n)}{\alpha+\lambda(n)}\left\{[V^\pi(n)-p(n)]\int\int_{\mathcal{R}_0^\pi(n)} F_{R|S}(dr|s)F_S(ds)\right. \\
&\qquad \left. + \int\int_{\mathcal{R}_1^\pi(n)} [r+V^\pi(n-1)]\, F_{R|S}(dr|s)F_S(ds)\right\} - \frac{1}{\alpha+\lambda(n)}c(n) \\
&= \frac{\lambda(n)}{\alpha+\lambda(n)}\left\{[V^\pi(n)-p(n)]\left[1-\int\int_{\mathcal{R}_1^\pi(n)} F_{R|S}(dr|s)F_S(ds)\right]\right. \\
&\qquad \left. + \int\int_{\mathcal{R}_1^\pi(n)} [r+V^\pi(n-s)]\, F_{R|S}(dr|s)F_S(ds)\right\} - \frac{1}{\alpha+\lambda(n)}c(n) \qquad (5.1) \\
&= \frac{\lambda(n)}{\alpha+\lambda(n)}\int\int_{\mathcal{R}_1^\pi(n)} \{r-[V^\pi(n)-V^\pi(n-s)-p(n)]\}\, F_{R|S}(dr|s)F_S(ds) \\
&\quad + \frac{\lambda(n)}{\alpha+\lambda(n)}V^\pi(n) - \frac{\lambda(n)}{\alpha+\lambda(n)}p(n) - \frac{1}{\alpha+\lambda(n)}c(n) \qquad (5.2)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad \alpha V^\pi(n) &= \lambda(n)\int\int_{\mathcal{R}_1^\pi(n)} \{r-[V^\pi(n)-V^\pi(n-s)-p(n)]\}\, F_{R|S}(dr|s)F_S(ds) \\
&\quad - \lambda(n)p(n) - c(n) \qquad (5.3)
\end{aligned}$$

It also follows that if $I^\pi(n)=1$, then

$$V^\pi(n) = \frac{\lambda(n)\int\int_{\mathcal{R}_1^\pi(n)} [r+V^\pi(n-1)+p(n)]\, F_{R|S}(dr|s)F_S(ds) - \lambda(n)p(n) - c(n)}{\alpha + \lambda(n)\int\int_{\mathcal{R}_1^\pi(n)} F_{R|S}(dr|s)F_S(ds)}$$

5.3 The Optimal Value Function

Yushkevich and Feinberg [187] (Theorem 2) showed for an MDP with a countable state space, a measurable action space, and an infinite horizon, that if $\alpha > 0$ and

$\sup_{n\in[0,N_0]}\sup_{\pi\in\Pi^{SD}_{MDP}}\xi(n,\pi(n))<\infty$, then V^* is the unique bounded solution of

$$\begin{aligned}\alpha V^*(n) \quad = \quad & \sup_{(D(n,\cdot,\cdot),I(n))}\left\{\left[\lambda(n)\int\!\!\int_{\mathcal{R}_1(n)}\{r-[V^*(n)-V^*(n-s)-p(n)]\}\,F_{R|S}(dr|s)F_S(ds)\right.\right.\\ & \left.\left. -\ \lambda(n)p(n)\ -\ c(n)\right]I(n)\ +\ \alpha v(n)\,(1-I(n))\right\} \qquad (5.4)\end{aligned}$$

and for any $\varepsilon>0$ there exists a stationary deterministic policy $\pi\in\Pi^{SD}_{MDP}$ that is ε-optimal among all history-dependent deterministic policies $\pi\in\Pi^{HD}_{MDP}$, i.e., $V^*(n)\equiv\sup_{\pi\in\Pi^{SD}_{MDP}}V^\pi(n)=\sup_{\pi\in\Pi^{HD}_{MDP}}V^\pi(n)$.

To obtain a countable state space, assume that the item size distribution is discrete with probability mass function f_S such that $\inf\{s\in(0,\infty):f_S(s)>0\}>0$. Then at most a finite number of items with positive sizes can be accepted, because $N_0<\infty$. Thus the set of feasible combinations of positive item sizes accepted is a subset of a finite cartesian product of the countable set $\{s\in(0,\infty):f_S(s)>0\}$. Therefore, the state space can be restricted to a countable set, because any finite cartesian product of countable sets is countable.

5.4 An Optimal Policy for the MDP and DSKP

In this section it is shown that there exists a stationary threshold policy with the simple reward threshold $x^*(n,s)=V^*(n)-V^*(n-s)-p(n)$ for all $s\le n$, that is optimal among all history-dependent deterministic policies for both the DSKP and the MDP.

5.4.1 Preliminary Results

First some preliminary results are established. For any $n\in[0,N_0]$, let $\mathcal{Y}$ denote the set of all Borel-measurable functions $\zeta:\{(n,s):n\in[0,N_0],s\in[0,n]\}\mapsto\Re$, that are F_S-integrable, i.e., $\int_0^n|\zeta(n,s)|\,F_S(ds)<\infty$. In order to derive some structural characteristics of π^* and V^*, consider the function $g:[0,N_0]\times\Re\times\mathcal{Y}\mapsto\Re$ defined by

$$\begin{aligned}g(n,y,\zeta) \quad &\equiv \quad \int_0^n\int_{y-\zeta(n,s)}^\infty\{r-[y-\zeta(n,s)]\}\,F_{R|S}(dr|s)F_S(ds)\\ &= \quad \int_0^n\int_{y-\zeta(n,s)}^\infty\left[1-F_{R|S}(r|s)\right]\,drF_S(ds)\end{aligned}$$

Lemma 5

1. *For all* $n \in [0, N_0]$ *and* $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ *satisfies the following Lipschitz condition with respect to* y:

$$|g(n, y_2, \zeta) - g(n, y_1, \zeta)| \quad \leq \quad |y_2 - y_1| \quad \text{for all} \;\; y_1, y_2 \in \Re$$

2. *For all* $n \in [0, N_0]$ *and* $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ *is absolutely continuous in* y *on* $\Re$.

3. *For all* $n \in [0, N_0]$ *and* $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ *is nonincreasing in* y *on* $\Re$, *and strictly decreasing in* y *on* $\{y \in \Re : g(n, y, \zeta) > 0\}$.

4. *For any* $n \in [0, N_0]$ *such that* $\int_0^n F_S(ds) > 0$, *and* $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ *is unbounded above, i.e., for any* $M > 0$, *there exists a* y_1 *such that* $g(n, y_1, \zeta) > M$.

5. *For any* $n \in [0, N_0]$, $\zeta \in \mathcal{Y}$, *and* $\varepsilon > 0$, *there exists a* y_2 *such that* $g(n, y_2, \zeta) < \varepsilon$.

6. *For any Borel set* $B \subset \Re \times [0, n]$,

$$\begin{aligned} & \int\int_B \{r - [y - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \\ \leq \quad & \int_0^n \int_{y-\zeta(n,s)}^{\infty} \{r - [y - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \quad \equiv \quad g(n, y, \zeta) \end{aligned}$$

The proof of Lemma 5 is given in Appendix B.

5.4.2 An Optimal Threshold Policy for the MDP and DSKP

With this result, an optimal policy π^* can now be obtained. An even more intuitive characterization of an optimal policy will be given in Section 5.5. For any $V_1, V_2 : [0, N_0] \mapsto \Re$, let

$$\|V_2 - V_1\|_\infty \quad \equiv \quad \sup_{n \in [0, N_0]} |V_2(n) - V_1(n)|$$

Theorem 1 *If the item size distribution is discrete such that* $\inf \{s \in \Re : F_S(s) > 0\} > 0$, *then the following stationary deterministic threshold policy* π^* *is an optimal policy among all history-dependent deterministic policies for the DSKP and the MDP. An optimal stopping rule is*

$$I^*(n) \quad = \quad \begin{cases} 1 & \text{if } \; \lambda(n) g(n, V^*(n), \zeta^*) - \lambda(n) p(n) - c(n) \geq \alpha v(n) \\ 0 & \text{if } \; \lambda(n) g(n, V^*(n), \zeta^*) - \lambda(n) p(n) - c(n) < \alpha v(n) \end{cases}$$

where $\zeta^*(n,s) \equiv V^*(n-s) + p(n)$. *An optimal acceptance rule is*

$$D^*(n,r,s) \quad = \quad \begin{cases} 1 & \text{if } s \le n \text{ and } r \ge V^*(n) - V^*(n-s) - p(n) \\ 0 & \text{if } s > n \text{ or } r < V^*(n) - V^*(n-s) - p(n) \end{cases}$$

That is, an optimal threshold $x^*(n,s)$ *for* $s \le n$ *is*

$$x^*(n,s) \quad = \quad V^*(n) - V^*(n-s) - p(n)$$

Proof: From Equation (5.4) and Lemma 5,

$$\begin{aligned} \alpha V^*(n) \quad &= \quad \max\left\{ \sup_{\mathcal{R}_1(n)\in\mathcal{B}_n} \lambda(n) \iint_{\mathcal{R}_1(n)} \{r - [V^*(n) - V^*(n-s) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \right. \\ &\qquad \left. -\ \lambda(n)p(n)\ -\ c(n),\ \ \alpha v(n) \right\} \\ &= \quad \max\left\{ \lambda(n) \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} \{r - [V^*(n) - V^*(n-s) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \right. \\ &\qquad \left. -\ \lambda(n)p(n)\ -\ c(n),\ \ \alpha v(n) \right\} \\ &= \quad \max\{\lambda(n)g(n, V^*(n), \zeta^*)\ -\ \lambda(n)p(n)\ -\ c(n),\ \ \alpha v(n)\} \end{aligned} \tag{5.5}$$

Thus, if $\lambda(n)g(n,V^*(n),\zeta^*) - \lambda(n)p(n) - c(n) < \alpha v(n)$, then $V^*(n) = v(n)$. Also, then $I^*(n) = 0$ and hence $V^{\pi^*}(n) = v(n) = V^*(n)$. Else, if $\lambda(n)g(n,V^*(n),\zeta^*) - \lambda(n)p(n) - c(n) \ge \alpha v(n)$, then

$$\begin{aligned} \alpha V^*(n) \quad &= \quad \lambda(n) \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} \{r - [V^*(n) - V^*(n-s) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\ &\qquad -\ \lambda(n)p(n)\ -\ c(n) \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad V^*(n) \quad &= \quad \frac{\lambda(n)}{\alpha+\lambda(n)} \left\{ [V^*(n) - p(n)] \left[1 - \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} F_{R|S}(dr|s)F_S(ds) \right] \right. \\ &\qquad \left. + \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} [r + V^*(n-s)]\ F_{R|S}(dr|s)F_S(ds) \right\} \\ &\qquad - \frac{1}{\alpha+\lambda(n)}\ c(n) \end{aligned}$$

Also, then $I^*(n) = 1$ and $x^*(n,s) = V^*(n) - V^*(n-s) - p(n)$ for all $s \le n$. Hence, from Equation (5.1),

$$V^{\pi^*}(n) \quad = \quad \frac{\lambda(n)}{\alpha+\lambda(n)} \left\{ \left[V^{\pi^*}(n) - p(n) \right] \left[1 - \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} F_{R|S}(dr|s)F_S(ds) \right] \right.$$

$$
\begin{aligned}
& \quad + \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} \left[r + V^{\pi^*}(n-s) \right] F_{R|S}(dr|s) F_S(ds) \Bigg\} \\
& - \frac{1}{\alpha + \lambda(n)} \, c(n)
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \quad & \left| V^*(n) - V^{\pi^*}(n) \right| \\
= \quad & \left| \frac{\lambda(n)}{\alpha + \lambda(n)} \left\{ \left[V^*(n) - V^{\pi^*}(n) \right] \left[1 - \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} F_{R|S}(dr|s) F_S(ds) \right] \right. \right. \\
& \left. \left. + \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} \left[V^*(n-s) - V^{\pi^*}(n-s) \right] F_{R|S}(dr|s) F_S(ds) \right\} \right| \\
\leq \quad & \frac{\lambda(n)}{\alpha + \lambda(n)} \left\{ \left| V^*(n) - V^{\pi^*}(n) \right| \left[1 - \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} F_{R|S}(dr|s) F_S(ds) \right] \right. \\
& \left. + \int_0^n \int_{V^*(n)-V^*(n-s)-p(n)}^{\infty} \left| V^*(n-s) - V^{\pi^*}(n-s) \right| F_{R|S}(dr|s) F_S(ds) \right\} \\
\leq \quad & \frac{\lambda(n)}{\alpha + \lambda(n)} \left\| V^* - V^{\pi^*} \right\|_\infty
\end{aligned}
$$

Thus for all $n \in [0, N_0]$,

$$
\begin{aligned}
\left| V^*(n) - V^{\pi^*}(n) \right| & \leq \frac{\lambda(n)}{\alpha + \lambda(n)} \left\| V^* - V^{\pi^*} \right\|_\infty \\
& \leq \frac{\overline{\lambda}}{\alpha + \overline{\lambda}} \left\| V^* - V^{\pi^*} \right\|_\infty \\
\Rightarrow \quad \left\| V^* - V^{\pi^*} \right\|_\infty & \leq \frac{\overline{\lambda}}{\alpha + \overline{\lambda}} \left\| V^* - V^{\pi^*} \right\|_\infty
\end{aligned}
$$

which can only hold if $\left\| V^* - V^{\pi^*} \right\|_\infty = 0$, because $\overline{\lambda}/(\alpha + \overline{\lambda}) < 1$. Therefore $V^{\pi^*} = V^*$, and policy π^* is optimal among all $\pi \in \Pi_{MDP}^{SD}$. From Yushkevich and Feinberg [187] Theorem 2, for any $\varepsilon > 0$, there is a $\pi \in \Pi_{MDP}^{SD}$ that is ε-optimal among all $\pi \in \Pi_{MDP}^{HD}$. Hence, π^* is optimal for MDP among all $\pi \in \Pi_{MDP}^{HD}$. From Lemma 3, $V_{MDP}^* \geq V_{DSKP}^*$. But π^* is admissible for DSKP ($\pi^* \in \Pi_{DSKP}^{HD}$), and $V_{DSKP}^{\pi^*} = V_{MDP}^{\pi^*} = V_{MDP}^* \geq V_{DSKP}^*$. Therefore, π^* is optimal for DSKP among all $\pi \in \Pi_{DSKP}^{HD}$. □

An optimal policy π^* is not a unique optimal policy, because $D^*(n, \cdot, \cdot)$ can be modified on any set with $F_{R,S}$-measure 0, without changing the expected value V^{π^*}. Hence, there exist optimal policies which are not threshold policies.

5.5 A Recursion for the Optimal Expected Value

In this section it is shown that under conditions that are typically found in applications, the optimal expected value V^* is given by a simple recursion, which gives an optimal policy π^* with an intuitive structure and an efficient algorithm for computing V^* and π^*. First some preliminary results are established in Section 5.5.1. Thereafter the recursive policies $\{\psi(n)\}$ are defined in Section 5.5.2. Finally it is shown in Section 5.5.3 that the recursive policies give an optimal policy, through the intuitive relationship $V^*(n) = \max\left\{v(n), V^{\psi(n)}(n)\right\}$. In Section 5.6 an efficient algorithm for the DSKP based on these results are given.

5.5.1 Preliminary Results

For any $n \in [0, N_0]$, $\zeta \in \mathcal{Y}$, and $y_3 \in \Re$, consider the following equation

$$\begin{aligned}\alpha y &= \lambda(n) \int_0^n \int_{y-\zeta(n,s)}^{\infty} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds) \ - \ y_3 \\ &= \lambda(n) g(n, y, \zeta) \ - \ y_3 \end{aligned} \tag{5.6}$$

Lemma 6 *For any $n \in [0, N_0]$, $\zeta \in \mathcal{Y}$, and $y_3 \in \Re$, if $\alpha > 0$, or if $\alpha = 0$ and $y_3 > 0$ and $\int_0^n F_S(ds) > 0$, then Equation (5.6) has a unique solution y.*

Proof:

Case 1: $\alpha > 0$:

Then αy is strictly increasing in y, and takes on all values in $\Re$. Also, from Lemma 5, $\lambda(n)g(n, y, \zeta) - y_3$ is nonincreasing and continuous in y. Therefore, from the intermediate value theorem, there exists a unique value y such that $\alpha y = \lambda(n)g(n, y, \zeta) - y_3$.

Case 2: $\alpha = 0$:

From Lemma 5, if $\int_0^n F_S(ds) > 0$, then for any $y_3 > 0$, there exists a y_1 such that $\lambda(n)g(n, y_1, \zeta) > y_3$, and a y_2 such that $\lambda(n)g(n, y_2, \zeta) < y_3$. Hence, from the continuity of $g(n, y, \zeta)$ in y and the intermediate value theorem, there exists at least one value y such that $\lambda(n)g(n, y, \zeta) = y_3 > 0$. Also, from Lemma 5, $g(n, y, \zeta)$ is strictly decreasing in y

at any y such that $g(n, y, \zeta) > 0$, and nonincreasing everywhere. Therefore, there exists a unique value y such that $\lambda(n)g(n, y, \zeta) - y_3 = 0 = \alpha y$. $\square$

For any $n \in [0, N_0]$, consider the function $h : [0, N_0] \times \Re \times \Re \times \mathcal{Y} \mapsto \Re$ defined by

$$h(n, y_n, y, \zeta) \equiv \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} \{r - [y - \zeta(n, s)]\} F_{R|S}(dr|s) F_S(ds)$$

Lemma 7

1. *For all $n \in [0, N_0]$, $y_n \in \Re$, and $\zeta \in \mathcal{Y}$, $h(n, y_n, y, \zeta)$ satisfies the following Lipschitz condition with respect to y:*

$$|h(n, y_n, y_2, \zeta) - h(n, y_n, y_1, \zeta)| \leq |y_2 - y_1| \quad \textit{for all} \quad y_1, y_2 \in \Re$$

2. *For all $n \in [0, N_0]$, $y_n \in \Re$, and $\zeta \in \mathcal{Y}$, $h(n, y_n, y, \zeta)$ is absolutely continuous in y on $\Re$.*

3. *For all $n \in [0, N_0]$, $y_n \in \Re$, and $\zeta \in \mathcal{Y}$, $h(n, y_n, y, \zeta)$ is nonincreasing in y on $\Re$, and strictly decreasing in y if $\int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s) F_S(ds) > 0$.*

4. *For any $n \in [0, N_0]$, $y_n \in \Re$, and $\zeta \in \mathcal{Y}$ such that $\int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s) F_S(ds) > 0$, $h(n, y_n, y, \zeta)$ is unbounded above and below, i.e., for any $M > 0$, there exists a y_1 such that $h(n, y_n, y_1, \zeta) > M$, and a y_2 such that $h(n, y_n, y_2, \zeta) < -M$.*

The proof of Lemma 7 is given in Appendix C.

For any $n \in [0, N_0]$, $y_n \in \Re$, $\zeta \in \mathcal{Y}$, and $y_3 \in \Re$, consider the following equation

$$\begin{aligned} \alpha y &= \lambda(n) \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} \{r - [y - \zeta(n, s)]\} F_{R|S}(dr|s) F_S(ds) - y_3 \\ &= \lambda(n) h(n, y_n, y, \zeta) - y_3 \end{aligned} \tag{5.7}$$

Lemma 8 *For any $n \in [0, N_0]$, $y_n \in \Re$, $\zeta \in \mathcal{Y}$, and $y_3 \in \Re$, if $\alpha > 0$, or if $\alpha = 0$ and $\int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s) F_S(ds) > 0$, then Equation (5.7) has a unique solution y.*

Proof:

Case 1: $\alpha > 0$:

Then αy is strictly increasing in y, and takes on all values in $\Re$. Also, from Lemma 7,

$\lambda(n)h(n, y_n, y, \zeta) - y_3$ is nonincreasing and continuous in y. Therefore, from the intermediate value theorem, there exists a unique value y such that $\alpha y = \lambda(n)h(n, y_n, y, \zeta) - y_3$.

Case 2: $\alpha = 0$:

From Lemma 7, if $\int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds) > 0$, then for any y_3 there exists a y_1 such that $\lambda(n)h(n, y_n, y_1, \zeta) > y_3$, and a y_2 such that $\lambda(n)h(n, y_n, y_2, \zeta) < y_3$. Hence, from the continuity of $h(n, y_n, y, \zeta)$ in y and the intermediate value theorem, there exists at least one value y such that $\lambda(n)h(n, y_n, y, \zeta) = y_3$. Also, from Lemma 7, if $\int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds) > 0$, then $h(n, y_n, y, \zeta)$ is strictly decreasing in y. Therefore, there exists a unique value y such that $\lambda(n)h(n, y_n, y, \zeta) - y_3 = 0 = \alpha y$. □

5.5.2 Definition of the Recursive Policies

Assume that the size distribution F_S is such that $F_S(0) = 0$, and the possible values of the remaining capacity are $0 = n_0 < n_1 < n_2 < \ldots < N_0$. This is the case if item sizes are measured in multiples of a unit size, for example if all item sizes are positive integers. As the positive numbers represented in a computer satisfy this property, and solutions are usually computed numerically, this assumption is not very restrictive. Another important special case that satisfies this assumption, is the case where all items have the same size. Also assume that $\alpha > 0$ or $\lambda(n)p(n) + c(n) > 0$ for all $n \in \{n_0, n_1, \ldots, N_0\}$. Without loss of generality, to simplify notation, assume that $n_k = k$ for all k.

Inductively define the sequence of threshold policies $\{\psi(n)\}$ as follows. $D^{\psi(0)}(0) = 0$, $I^{\psi(0)}(0) = 1$. From Equation (5.3), $\alpha V^{\psi(0)}(0) = -\lambda(n)p(n) - c(0)$, where $V^{\psi(0)}(0) \equiv -\infty$ if $\alpha = 0$. Let $\psi(n-s)$ and $V^{\psi(n-s)}(n') < \infty$ be defined for all $s \in \{1, \ldots, n\}$ and all $n' \in \{0, 1, \ldots, n-s\}$. Note that because $\max\{v(n-s), V^{\psi(n-s)}(n-s)\}$ assumes only a finite number of values for $n \in \{0, 1, \ldots, N_0\}$ and $s \in \{1, \ldots, n\}$, the function $\zeta : \{(n, s) : n \in \{0, 1, \ldots, N_0\}, s \in \{1, \ldots, n\}\}$ defined by $\zeta(n, s) \equiv \max\left\{v(n-s), V^{\psi(n-s)}(n-s)\right\} + p(n)$ is bounded, hence $\zeta \in \mathcal{Y}$. Let $\hat{V}(n)$ be the unique solution to

$$\begin{aligned}
\alpha \hat{V}(n) &= \lambda(n) \int_0^n \int_{\hat{V}(n) - \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n)}^{\infty} \Big\{ r - \Big[\hat{V}(n) - \\
&\qquad\qquad \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n) \Big] \Big\} F_{R|S}(dr|s)F_S(ds) \\
&\quad - \lambda(n)p(n) - c(n) \\
&= \lambda(n)\, g\left(n, \hat{V}(n), \zeta\right) - \lambda(n)p(n) - c(n) \qquad (5.8)
\end{aligned}$$

which exists by Lemma 6. Let $I^{\psi(n)}(n) = 1$, and $x^{\psi(n)}(n,s) = \hat{V}(n) - \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n) = \hat{V}(n) - \zeta(n,s)$ for all $s \leq n$. Let

$$I^{\psi(n)}(n-s) = \begin{cases} 1 & \text{if } V^{\psi(n-s)}(n-s) > v(n-s) \\ 0 & \text{if } V^{\psi(n-s)}(n-s) \leq v(n-s) \end{cases}$$

and $D^{\psi(n)}(n-s,\cdot,\cdot) = D^{\psi(n-s)}(n-s,\cdot,\cdot)$ for all $s \in \{1,\ldots,n\}$. Then it follows by induction that $V^{\psi(n)}(n-s) = \max\{v(n-s), V^{\psi(n-s)}(n-s)\}$ for all $s \in \{1,\ldots,n\}$. It follows from Equation (5.3) that $V^{\psi(n)}(n)$ satisfies

$$\begin{aligned}
\alpha V^{\psi(n)}(n) &= \lambda(n) \int_0^n \int_{x^{\psi(n)}(n,s)}^{\infty} \left\{ r - \left[V^{\psi(n)}(n) - V^{\psi(n)}(n-s) - p(n) \right] \right\} F_{R|S}(dr|s) F_S(ds) \\
&\quad - \lambda(n)p(n) - c(n) \\
&= \lambda(n) \int_0^n \int_{\hat{V}(n)-\max\{v(n-s),V^{\psi(n-s)}(n-s)\}-p(n)}^{\infty} \left\{ r - \left[V^{\psi(n)}(n) - \right.\right. \\
&\qquad\qquad \left.\left. \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n) \right] \right\} F_{R|S}(dr|s) F_S(ds) \\
&\quad - \lambda(n)p(n) - c(n) \\
&= \lambda(n)\, h\left(n, \hat{V}(n), V^{\psi(n)}(n), \zeta\right) - \lambda(n)p(n) - c(n) \qquad (5.9)
\end{aligned}$$

From Equation (5.8), Equation (5.9) has a solution $V^{\psi(n)}(n) = \hat{V}(n)$, and it follows from Lemma 8 that this is the unique solution of Equation (5.9). Therefore,

$$\begin{aligned}
\alpha V^{\psi(n)}(n) &= \lambda(n) \int_0^n \int_{V^{\psi(n)}(n)-\max\{v(n-s),V^{\psi(n-s)}(n-s)\}-p(n)}^{\infty} \left\{ r - \left[V^{\psi(n)}(n) - \right.\right. \\
&\quad \left.\left. \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n) \right] \right\} F_{R|S}(dr|s) F_S(ds) - \lambda(n)p(n) - c(n) \\
&= \lambda(n)\, g(n, V^{\psi(n)}(n), \zeta) - \lambda(n)p(n) - c(n) \qquad (5.10)
\end{aligned}$$

5.5.3 Optimality Results for the Recursive Policies

By definition, $V^*(n) \geq \max\{v(n), V^{\psi(n)}(n)\}$ for all n. Theorem 2 shows that $V^*(n) = \max\{v(n), V^{\psi(n)}(n)\}$ for all n. Therefore an optimal policy is as follows. For each n, if $V^{\psi(n)}(n) > v(n)$, then continue (i.e., $I^*(n) = 1$), using threshold $x^*(n,s) = x^{\psi(n)}(n,s) = V^{\psi(n)}(n) - \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n) = V^*(n) - V^*(n-s) - p(n)$ for all $s \leq n$; else stop (i.e., $I^*(n) = 0$), and collect $v(n)$. This result is useful, not only because it gives a clear, intuitive characterization of an optimal policy and the optimal expected value, but also because it provides an efficient method for computing the optimal expected value V^* and optimal threshold x^*.

Theorem 2 *The optimal expected value V^* satisfies $V^*(n) = \max\{v(n), V^{\psi(n)}(n)\}$ for all $n \in \{0, 1, \ldots, N_0\}$.*

Proof: By induction on n. For $n = 0$, $V^*(0) = \max\{v(0), V^{\psi(0)}(0)\}$. Suppose $V^*(n-s) = \max\{v(n-s), V^{\psi(n-s)}(n-s)\}$ for all $s \in \{1, \ldots, n\}$. Hence, $V^{\psi(n)}(n-s) = \max\{v(n-s), V^{\psi(n-s)}(n-s)\} = V^*(n-s)$ for all $s \in \{1, \ldots, n\}$.

Case 1: $V^\pi(n) > v(n)$ for some policy $\pi \in \Pi^{SD}_{MDP}$:

Consider any such policy π. Then $I^\pi(n) = 1$, and $V^\pi(n-s) \leq V^*(n-s) = V^{\psi(n)}(n-s)$ for all $s \in \{1, \ldots, n\}$. It is shown by contradiction that $V^\pi(n) \leq V^{\psi(n)}(n)$. Suppose $V^\pi(n) > V^{\psi(n)}(n)$. Then $V^\pi(n) - V^\pi(n-s) - p(n) > V^{\psi(n)}(n) - V^{\psi(n)}(n-s) - p(n)$ for all $s \in \{1, \ldots, n\}$. From Equation (5.3) and Lemma 5

$$
\begin{aligned}
\alpha V^\pi(n) &= \lambda(n) \int\!\!\int_{\mathcal{R}_1^\pi(n)} \{r - [V^\pi(n) - V^\pi(n-s) - p(n)]\}\ F_{R|S}(dr|s) F_S(ds) \\
&\quad - \lambda(n) p(n) \ - \ c(n) \\
&\leq \lambda(n) \int_0^n \int_{V^\pi(n) - V^\pi(n-s) - p(n)}^\infty \{r - [V^\pi(n) - V^\pi(n-s) - p(n)]\}\ F_{R|S}(dr|s) F_S(ds) \\
&\quad - \lambda(n) p(n) \ - \ c(n) \\
&\leq \lambda(n) \int_0^n \int_{V^\pi(n) - V^\pi(n-s) - p(n)}^\infty \Big\{r - \Big[V^{\psi(n)}(n) \\
&\qquad\qquad - V^{\psi(n)}(n-s) - p(n)\Big]\Big\}\ F_{R|S}(dr|s) F_S(ds) \\
&\quad - \lambda(n) p(n) \ - \ c(n) \\
&\leq \lambda(n) \int_0^n \int_{V^{\psi(n)}(n) - V^{\psi(n)}(n-s) - p(n)}^\infty \Big\{r - \Big[V^{\psi(n)}(n) \\
&\qquad\qquad - V^{\psi(n)}(n-s) - p(n)\Big]\Big\}\ F_{R|S}(dr|s) F_S(ds) \\
&\quad - \lambda(n) p(n) \ - \ c(n) \\
&= \alpha V^{\psi(n)}(n)
\end{aligned}
$$

which contradicts the assumption. Therefore, $V^\pi(n) \leq V^{\psi(n)}(n)$ for all $\pi \in \Pi^{SD}_{MDP}$, and $V^*(n) = V^{\psi(n)}(n) = \max\{v(n), V^{\psi(n)}(n)\}$.

Case 2: $V^\pi(n) \leq v(n)$ for every policy $\pi \in \Pi^{SD}_{MDP}$:

Then $V^*(n) = v(n) = \max\{v(n), V^{\psi(n)}(n)\}$. □

Corollary 1 *If F_S is such that $F_S(0) = 0$, and the possible values of the remaining capacity are $0 = n_0 < n_1 < n_2 < \ldots < N_0$, then the following stationary deterministic threshold* *policy π^* is an optimal policy among all history-dependent deterministic policies for the DSKP and the MDP. An optimal stopping rule is*

$$I^*(n) = \begin{cases} 1 & \text{if } V^{\psi(n)}(n) > v(n) \\ 0 & \text{if } V^{\psi(n)}(n) \leq v(n) \end{cases}$$

An optimal acceptance rule is

$$D^*(n, r, s) = \begin{cases} 1 & \text{if } s \leq n \text{ and } r \geq V^{\psi(n)}(n) - \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n) \\ 0 & \text{if } s > n \text{ or } r < V^{\psi(n)}(n) - \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n) \end{cases}$$

That is, an optimal threshold $x^*(n)$ for $n > 0$ is

$$x^*(n, s) = V^{\psi(n)}(n) - \max\{v(n-s), V^{\psi(n-s)}(n-s)\} - p(n)$$

5.6 An Efficient Algorithm for the Infinite Horizon DSKP

In this section it is assumed that the size distribution F_S is such that $F_S(0) = 0$, and the possible values of the remaining capacity are $0 = n_0 < n_1 < n_2 < \ldots < N_0$. There are a number of algorithms for solving infinite horizon Markov decision processes, including value iteration (successive approximation), policy iteration, and modified policy iteration. As their names suggest, these algorithms are iterative; they repeatedly compute approximations to the optimal value function V^* and approximate optimal policies. In this section an efficient algorithm is proposed for the infinite horizon DSKP that computes the optimal value function V^* and optimal policy π^* by processing each state n only once, and not repeatedly as the iterative algorithms mentioned above.

As a by-product of the algorithm, an optimal initial capacity can be chosen. Suppose the initial capacity $\overline{N}$ can be chosen from a set $\mathcal{N}$ of available capacities, $\mathcal{N} \subseteq \{0, \ldots, N_0\}$. This is typical when an initial size is chosen for a ship, a truck, or a batch processor, from a set of sizes available in the market. Once in operation, the option also exists to stop waiting and transport or process the items, even if the capacity of the vehicle or processor is not exhausted. An optimal initial capacity $\overline{N}^*$ is given by

$$\overline{N}^* \in \arg\max_{N \in \mathcal{N}} V^*(N)$$

Algorithm Infinite-Horizon-Knapsack computes the optimal value function V^*, an optimal stopping rule I^*, an optimal threshold x^*, and an optimal initial capacity $\overline{N}^*$ in $\Theta(N_0)$ time, if solving Equation (5.10) is counted as an operation for each value of n.

Algorithm Infinite-Horizon-Knapsack

compute $V^{\psi(0)}(0)$ from Equation (5.3);
if $V^{\psi(0)}(0) > v(0)$ **then**
  $V^*(0) = V^{\psi(0)}(0)$; $I^*(0) = 1$;
else
  $V^*(0) = v(0)$; $I^*(0) = 0$;
endif;
$\overline{N}^* = 0$;
for $n = 1$ **to** N_0
  solve Equation (5.10) for $V^{\psi(n)}(n)$;
  if $V^{\psi(n)}(n) > v(n)$ **then**
    $V^*(n) = V^{\psi(n)}(n)$; $I^*(n) = 1$; $x^*(n,s) = V^*(n) - V^*(n-s) - p(n)$;
  else
    $V^*(n) = v(n)$; $I^*(n) = 0$;
  endif;
  if $V^*(n) > V^*(\overline{N}^*)$ **and** $n \in \mathcal{N}$ **then**
    $\overline{N}^* = n$;
  endif;
endfor;

6. The Finite Horizon DSKP

In this chapter the DSKP with a finite deadline is studied. It has many similarities with the DSKP with no deadline, such as the existence of an optimal threshold policy. There are also many differences; for example, in the case with a finite deadline, time plays a fundamental role in the decision process, in that the optimal value and optimal policy are functions of time.

First an equation is derived in Section 6.1 which gives the value of a policy $\pi \in \Pi_{MDP}^{MD}$. It is shown that the equation has a unique solution, and other useful properties of the equation are derived. In Section 6.2 it is shown that the optimality equation for the MDP has a unique solution, and with it an optimal policy for the MDP is found. Under specific conditions, the optimal policy for the MDP is shown to be admissible and optimal for the DSKP in Section 6.3. For computational purposes, a discretization method is studied in Section 6.4, where it is shown that the discrete approximations converge uniformly to the optimal solution V^*. Examples are discussed in Section 6.5. A number of structural characteristics of the optimal value function and optimal policy are established in Section 6.6. It is shown that under specific conditions the optimal value and optimal policy are given by a recursion similar to that of the infinite horizon case. A number of monotonicity and convexity properties of the optimal value function and the optimal threshold are established, for the case where all items have the same size, and the case where items have random sizes.

6.1 The Value of a Policy

In Section 6.1.1 a differential equation is derived, the solution of which, if it exists, is the value function V_{MDP}^{π} of a policy $\pi \in \Pi_{MDP}^{MD}$. In Section 6.1.2 it is shown that the equation has a unique solution V_{MDP}^{π}. A number of continuity results that are used later are derived in Section 6.1.3.

6.1.1 Derivation of the Bellman-Hamilton-Jacobi Equation

A differential equation satisfied by the expected value function V^π_{MDP} under a policy $\pi \in \Pi^{MD}_{MDP}$ can be derived intuitively as follows. If $I^\pi(n,t) = 1$, then by conditioning on whether an arrival takes place in the next Δt time units, and on the size s and reward r of the item if there is an arrival, it follows that

$$\begin{aligned}
V^\pi_{MDP}(n,t) \;=\; (1-\alpha\Delta t)\Bigg\{\lambda(n,t)\Delta t\Bigg[&\int\!\!\int_{\mathcal{R}^\pi_1(n,t)} [r + V^\pi_{MDP}(n-s,t+\Delta t)]\; F_{R|S}(dr|s)F_S(ds) \\
&+ \int\!\!\int_{\mathcal{R}^\pi_0(n,t)} [V^\pi_{MDP}(n,t+\Delta t) - p(n,t)]\; F_{R|S}(dr|s)F_S(ds)\Bigg] \\
&+ (1-\lambda(n,t)\Delta t)\, V^\pi_{MDP}(n,t+\Delta t) \;-\; c(n,t)\Delta t\Bigg\} \;+\; o(\Delta t)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow\quad & \frac{V^\pi_{MDP}(n,t) - V^\pi_{MDP}(n,t+\Delta t)}{\Delta t} \\
&= (1-\alpha\Delta t)\,\lambda(n,t)\Bigg[\int\!\!\int_{\mathcal{R}^\pi_1(n,t)} [r + V^\pi_{MDP}(n-s,t+\Delta t)]\; F_{R|S}(dr|s)F_S(ds) \\
&\qquad + [V^\pi_{MDP}(n,t+\Delta t) - p(n,t)]\int\!\!\int_{\mathcal{R}^\pi_0(n,t)} F_{R|S}(dr|s)F_S(ds)\Bigg] \\
&\quad + (-\alpha - \lambda(n,t) + \alpha\lambda(n,t)\Delta t)\, V^\pi_{MDP}(n,t+\Delta t) \;-\; (1-\alpha\Delta t)\, c(n,t) \;+\; \frac{o(\Delta t)}{\Delta t}
\end{aligned}$$

where $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$, from the corresponding property for the Poisson process, and the assumption that $\int_0^{N_0}\int_{-\infty}^{\infty} |r|\, F_{R|S}(dr|s)F_S(ds) < \infty$. Letting $\Delta t \to 0$,

$$\begin{aligned}
\frac{\partial V^\pi_{MDP}}{\partial t}(n,t) &= -\lambda(n,t)\Bigg[\int\!\!\int_{\mathcal{R}^\pi_1(n,t)} [r + V^\pi_{MDP}(n-s,t)]\; F_{R|S}(dr|s)F_S(ds) \\
&\qquad + [V^\pi_{MDP}(n,t) - p(n,t)]\int\!\!\int_{\mathcal{R}^\pi_0(n,t)} F_{R|S}(dr|s)F_S(ds)\Bigg] \\
&\quad - (-\alpha - \lambda(n,t))\, V^\pi_{MDP}(n,t) \;+\; c\,(n,t) \\
&= -\lambda(n,t)\Bigg[\int\!\!\int_{\mathcal{R}^\pi_1(n,t)} [r + V^\pi_{MDP}(n-s,t)]\; F_{R|S}(dr|s)F_S(ds) \\
&\qquad + [V^\pi_{MDP}(n,t) - p(n,t)]\left[1 - \int\!\!\int_{\mathcal{R}^\pi_1(n,t)} F_{R|S}(dr|s)F_S(ds)\right]\Bigg] \\
&\quad + (\alpha + \lambda(n,t))\, V^\pi_{MDP}(n,t) \;+\; c\,(n,t) \\
&= -\lambda(n,t)\int\!\!\int_{\mathcal{R}^\pi_1(n,t)} \{r - [V^\pi_{MDP}(n,t) \\
&\qquad\qquad - V^\pi_{MDP}(n-s,t) - p(n,t)]\}\; F_{R|S}(dr|s)F_S(ds) \\
&\quad + \alpha V^\pi_{MDP}(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t)
\end{aligned}$$

If $I^\pi(n,t) = 0$, then

$$\begin{aligned}
V^\pi_{MDP}(n,t) &= (1-\alpha\Delta t)\, V^\pi_{MDP}(n,t+\Delta t) \ + \ \alpha\Delta t v(n) \\
\Rightarrow \quad \frac{V^\pi_{MDP}(n,t) - V^\pi_{MDP}(n,t+\Delta t)}{\Delta t} &= -\alpha V^\pi_{MDP}(n,t+\Delta t) \ + \ \alpha v(n) \\
\Rightarrow \quad \frac{\partial V^\pi_{MDP}}{\partial t}(n,t) &= \alpha V^\pi_{MDP}(n,t) \ - \ \alpha v(n)
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{\partial V^\pi_{MDP}}{\partial t}(n,t) = {} & \Bigg[-\lambda(n,t) \iint_{\mathcal{R}^\pi_1(n,t)} \{ r - [V^\pi_{MDP}(n,t) \\
& \qquad - V^\pi_{MDP}(n-s,t) - p(n,t)]\} \ F_{R|S}(dr|s) F_S(ds) \\
& + \ \alpha V^\pi_{MDP}(n,t) \ + \ \lambda(n,t)p(n,t) \ + \ c(n,t) \Bigg] \, I^\pi(n,t) \\
& + \ \Big[\alpha V^\pi_{MDP}(n,t) \ - \ \alpha v(n) \Big] \, (1 - I^\pi(n,t)) \qquad (6.1)
\end{aligned}$$

The boundary condition is $V^\pi_{MDP}(n,T) = v(n)$ for all $n \in [0, N_0]$.

This differential equation is usually called the Bellman equation. It is similar to that derived in Pliska [120] for a Markov decision process, where the existence of a unique absolutely continuous solution for each policy π under special conditions is shown, in Boel and Varaiya [21] for more general controlled jump processes, and in Brémaud [23], where this equation is called the Hamilton-Jacobi equation. From Boel and Varaiya [21] Theorem 5.6, if Equation (6.1) has a solution that is absolutely continuous in t and that satisfies the boundary condition, then the solution is the expected value function V^π_{MDP} of policy $\pi \in \Pi^{MD}_{MDP}$.

6.1.2 Solution of the Bellman-Hamilton-Jacobi Equation

In this section it is shown that Equation (6.1) has a solution that is the expected value function V^π_{MDP} of policy $\pi \in \Pi^{MD}_{MDP}$. At the same time a number of structural properties of Equation (6.1) are established that are also used later.

Let $\mathcal{V}$ denote the set of bounded Borel-measurable functions $V : [0, N_0] \times [0,T] \mapsto \Re$. Let $V_1, V_2 \in \mathcal{V}$. Define the following two metrics on $\mathcal{V}$:

$$\begin{aligned}
d_t(V_1,V_2) &\equiv \sup_{n\in[0,N_0]} |V_2(n,t) - V_1(n,t)| \\
d_{\mathcal{V}}(V_1,V_2) &\equiv \sup_{(n,t)\in[0,N_0]\times[0,T]} |V_2(n,t) - V_1(n,t)|
\end{aligned}$$

For any given policy $\pi \in \Pi^{MD}_{MDP}$, consider the differential equation

$$\begin{aligned}
\frac{\partial V}{\partial t}(n,t) &= \Bigg[-\lambda(n,t) \int\!\!\int_{\mathcal{R}_1^\pi(n,t)} \{r - [V(n,t) - V(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&\qquad + \; \alpha V(n,t) \; + \; \lambda(n,t) p(n,t) \; + \; c(n,t) \Bigg] I^\pi(n,t) \\
&\quad + \Big[\alpha V(n,t) \; - \; \alpha v(n) \Big] (1 - I^\pi(n,t)) \\
&\equiv f^\pi(n,t,V)
\end{aligned}$$

where $f^\pi : [0,N_0] \times [0,T] \times \mathcal{V} \mapsto \Re$. Because it is assumed that λ, p, c and v are bounded, and $\int_0^{N_0} \int_{-\infty}^{\infty} |r| \; F_{R|S}(dr|s) F_S(ds) < \infty$, it follows that for any given $V \in \mathcal{V}$, $f^\pi(\cdot,\cdot,V)$ is bounded on $[0,N_0] \times [0,T]$. Hence, for any $\pi \in \Pi^{MD}_{MDP}$ and any $V \in \mathcal{V}$, $f^\pi(\cdot,\cdot,V) \in \mathcal{V}$. The boundary condition is $V(n,T) = v(n)$ for all $n \in [0,N_0]$.

Lemma 9 *For any $\pi \in \Pi^{MD}_{MDP}$, f^π satisfies the following Lipschitz condition with respect to $V_1, V_2 \in \mathcal{V}$:*

$$|f^\pi(n,t,V_1) - f^\pi(n,t,V_2)| \;\; \leq \;\; \left(2\overline{\lambda} + \alpha\right) d_t(V_1,V_2) \;\; \leq \;\; \left(2\overline{\lambda} + \alpha\right) d_{\mathcal{V}}(V_1,V_2)$$

Proof: If $I^\pi(n,t) = 1$, then

$$\begin{aligned}
& |f^\pi(n,t,V_1) - f^\pi(n,t,V_2)| \\
= & \Bigg| -\lambda(n,t) \int\!\!\int_{\mathcal{R}_1^\pi(n,t)} \{r - [V_1(n,t) - V_1(n-s,t) - p]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \lambda(n,t) \int\!\!\int_{\mathcal{R}_1^\pi(n,t)} \{r - [V_2(n,t) - V_2(n-s,t) - p]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \; \alpha V_1(n,t) \; - \; \alpha V_2(n,t) \Bigg| \\
= & \Bigg| \lambda(n,t) \int\!\!\int_{\mathcal{R}_1^\pi(n,t)} \{[V_1(n,t) - V_2(n,t)] + [V_2(n-s,t) - V_1(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \; \alpha [V_1(n,t) - V_2(n,t)] \Bigg| \\
\leq & \; \overline{\lambda} \int\!\!\int_{\mathcal{R}_1^\pi(n,t)} 2d_t\left(V_1,V_2\right) \; F_{R|S}(dr|s) F_S(ds) \\
& + \alpha d_t\left(V_1,V_2\right) \\
\leq & \; \left(2\overline{\lambda} + \alpha\right) d_t\left(V_1,V_2\right) \\
\leq & \; \left(2\overline{\lambda} + \alpha\right) d_{\mathcal{V}}\left(V_1,V_2\right)
\end{aligned}$$

If $I^\pi(n,t) = 0$, then

$$\begin{aligned} |f^\pi(n,t,V_1) - f^\pi(n,t,V_2)| &= |\alpha V_1(n,t) - \alpha V_2(n,t)| \\ &\leq \alpha d_t(V_1,V_2) \\ &\leq \alpha d_{\mathcal{V}}(V_1,V_2) \end{aligned} \qquad \square$$

For any policy $\pi \in \Pi^{MD}_{MDP}$, consider the mapping $\mathcal{F}^\pi : \mathcal{V} \mapsto \mathcal{V}$ defined by

$$\mathcal{F}^\pi V(n,t) \equiv V(n,T) - \int_t^T f^\pi(n,\tau,V)\, d\tau$$

It was shown that for any given $V \in \mathcal{V}$, $f^\pi(\cdot,\cdot,V)$ is bounded on $[0,N_0]\times[0,T]$. Hence, for any given $V \in \mathcal{V}$, $\mathcal{F}^\pi V$ is bounded on $[0,N_0]\times[0,T]$, and thus $\mathcal{F}^\pi V \in \mathcal{V}$ and $d_{\mathcal{V}}(V,\mathcal{F}^\pi V) < \infty$. Recursively define $\mathcal{F}^{\pi,m+1}V \equiv \mathcal{F}^\pi\mathcal{F}^{\pi,m}V$, $m = 0,1,\ldots$, where $\mathcal{F}^{\pi,0}V \equiv V$. It follows by induction that for any $V \in \mathcal{V}$, $\mathcal{F}^{\pi,m}V \in \mathcal{V}$ for all $m = 0,1,\ldots$.

Lemma 10 *For any $\pi \in \Pi^{MD}_{MDP}$, $\mathcal{F}^{\pi,m}V$ satisfies the same boundary condition as V at T, i.e., $\mathcal{F}^{\pi,m}V(n,T) = V(n,T)$ for all $n \in [0,N_0]$, for all $m = 0,1,\ldots$.*

Proof: By induction on m. For $m = 0$, $\mathcal{F}^{\pi,0}V(n,T) \equiv V(n,T)$. Suppose $\mathcal{F}^{\pi,m}V(n,T) = V(n,T)$. Then $\mathcal{F}^{\pi,m+1}V(n,T) \equiv \mathcal{F}^\pi\mathcal{F}^{\pi,m}V(n,T) \equiv \mathcal{F}^{\pi,m}V(n,T) - \int_T^T f^\pi(n,\tau,\mathcal{F}^{\pi,m}V)\, d\tau = V(n,T)$, for all $n \in [0,N_0]$. $\square$

Corollary 2 *For any $\pi \in \Pi^{MD}_{MDP}$, and any bounded Borel-measurable function $v : [0,N_0] \mapsto \Re$, $\mathcal{F}^{\pi,m} : \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\} \mapsto \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\}$, for all $m = 0,1,\ldots$.*

Lemma 11 *For any $\pi \in \Pi^{MD}_{MDP}$, if $V_1, V_2 \in \mathcal{V}$ satisfy the same boundary condition at T, $V_1(n,T) = V_2(n,T)$ for all $n \in [0,N_0]$, then $\mathcal{F}^{\pi,m}$ satisfies the following inequality for all $t \in [0,T]$, for all $m = 0,1,\ldots$*

$$\begin{aligned} d_t(\mathcal{F}^{\pi,m}V_1,\mathcal{F}^{\pi,m}V_2) &\equiv \sup_{n\in[0,N_0]} |\mathcal{F}^{\pi,m}V_1(n,t) - \mathcal{F}^{\pi,m}V_2(n,t)| \\ &\leq \frac{\left(2\overline{\lambda}+\alpha\right)^m (T-t)^m}{m!} d_{\mathcal{V}}(V_1,V_2) \end{aligned}$$

Proof: By induction on m. For $m = 0$, for all $n \in [0, N_0]$ and all $t \in [0, T]$:

$$\begin{aligned}\left|\mathcal{F}^{\pi,0}V_1(n,t) - \mathcal{F}^{\pi,0}V_2(n,t)\right| &\equiv |V_1(n,t) - V_2(n,t)| \\ &\leq d_t(V_1, V_2) \\ &\leq d_{\mathcal{V}}(V_1, V_2)\end{aligned}$$

Suppose, for all $t \in [0, T]$,

$$d_t\left(\mathcal{F}^{\pi,m}V_1, \mathcal{F}^{\pi,m}V_2\right) \leq \frac{\left(2\overline{\lambda} + \alpha\right)^m (T-t)^m}{m!} d_{\mathcal{V}}(V_1, V_2)$$

Then, for all $n \in [0, N]$ and all $t \in [0, T]$:

$$\begin{aligned}\left|\mathcal{F}^{\pi,m+1}V_1(n,t) - \mathcal{F}^{\pi,m+1}V_2(n,t)\right| &\equiv \left|\mathcal{F}^{\pi}\mathcal{F}^{\pi,m}V_1(n,t) - \mathcal{F}^{\pi}\mathcal{F}^{\pi,m}V_2(n,t)\right| \\ &= \left|\left[\mathcal{F}^{\pi,m}V_1(n,T) - \int_t^T f^{\pi}\left(n, \tau, \mathcal{F}^{\pi,m}V_1\right)\, d\tau\right]\right. \\ &\quad \left. - \left[\mathcal{F}^{\pi,m}V_2(n,T) - \int_t^T f^{\pi}\left(n, \tau, \mathcal{F}^{\pi,m}V_2\right)\, d\tau\right]\right| \\ &= \left|\int_t^T \left[f^{\pi}\left(n,\tau,\mathcal{F}^{\pi,m}V_2\right) - f^{\pi}\left(n,\tau,\mathcal{F}^{\pi,m}V_1\right)\right] d\tau\right| \\ &\leq \int_t^T \left|f^{\pi}\left(n,\tau,\mathcal{F}^{\pi,m}V_2\right) - f^{\pi}\left(n,\tau,\mathcal{F}^{\pi,m}V_1\right)\right| d\tau \\ &\leq \int_t^T \left(2\overline{\lambda}+\alpha\right) d_\tau\left(\mathcal{F}^{\pi,m}V_1, \mathcal{F}^{\pi,m}V_2\right) d\tau \\ &\leq \int_t^T \left(2\overline{\lambda}+\alpha\right) \frac{\left(2\overline{\lambda}+\alpha\right)^m (T-t)^m}{m!}\, d_{\mathcal{V}}(V_1,V_2)\, d\tau \\ &= \frac{\left(2\overline{\lambda}+\alpha\right)^{m+1}}{m!}\, d_{\mathcal{V}}(V_1,V_2) \left[-\frac{(T-\tau)^{m+1}}{m+1}\right]_t^T \\ &= \frac{\left(2\overline{\lambda}+\alpha\right)^{m+1}(T-t)^{m+1}}{(m+1)!}\, d_{\mathcal{V}}(V_1,V_2)\end{aligned}$$

where the second inequality follows from Lemma 9, and the last inequality follows from the induction hypothesis. □

Corollary 3 *For any $\pi \in \Pi^{MD}_{MDP}$, if $V_1, V_2 \in \mathcal{V}$ satisfy the same boundary condition at T, $V_1(n,T) = V_2(n,T)$ for all $n \in [0, N_0]$, then $\mathcal{F}^{\pi,m}$ satisfies the following Lipschitz condition, for all $m = 0, 1, \ldots$*

$$d_{\mathcal{V}}\left(\mathcal{F}^{\pi,m}V_1, \mathcal{F}^{\pi,m}V_2\right) \leq \frac{\left[\left(2\overline{\lambda}+\alpha\right)T\right]^m}{m!}\, d_{\mathcal{V}}(V_1, V_2)$$

Proof:

$$\begin{aligned} d_{\mathcal{V}}\left(\mathcal{F}^{\pi,m}V_1, \mathcal{F}^{\pi,m}V_2\right) &\equiv \sup_{(n,t)\in[0,N_0]\times[0,T]} \left|\mathcal{F}^{\pi,m}V_1(n,t) - \mathcal{F}^{\pi,m}V_2(n,t)\right| \\ &= \sup_{t\in[0,T]} d_t\left(\mathcal{F}^{\pi,m}V_1, \mathcal{F}^{\pi,m}V_2\right) \\ &\leq \sup_{t\in[0,T]} \frac{\left(2\overline{\lambda}+\alpha\right)^m (T-t)^m}{m!} d_{\mathcal{V}}(V_1,V_2) \\ &= \frac{\left[\left(2\overline{\lambda}+\alpha\right)T\right]^m}{m!} d_{\mathcal{V}}(V_1,V_2) \qquad \square \end{aligned}$$

Lemma 12 *For any* $\pi \in \Pi^{MD}_{MDP}$*, and any bounded Borel-measurable function* $v : [0, N_0] \mapsto \Re$*,* $\mathcal{F}^\pi$ *has a unique fixed point* $\hat{V}^\pi \in \mathcal{V}$ *that satisfies the following boundary condition at* T*:* $\hat{V}^\pi(n,T) = v(n)$ *for all* $n \in [0, N_0]$*.*

Proof: From Corollary 2, for any bounded Borel-measurable function $v : [0, N_0] \mapsto \Re$, $\mathcal{F}^{\pi,m} : \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\} \mapsto \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\}$, for all $m = 0, 1, \ldots$. Because $\mathcal{V}$ with metric $d_{\mathcal{V}}$ is a complete metric space, any Cauchy sequence $\{V_m\}$ in $\mathcal{V}$ converges to some $\hat{V} \in \mathcal{V}$. If each V_m satisfies $V_m(n,T) = v(n)$ for all $n \in [0, N_0]$, then $\hat{V}(n,T) = v(n)$ for all $n \in [0, N_0]$. Hence, $\{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\}$ with metric $d_{\mathcal{V}}$ is a complete metric space. From Corollary 3, for any $V_1, V_2 \in \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\}$,

$$d_{\mathcal{V}}\left(\mathcal{F}^{\pi,m}V_1, \mathcal{F}^{\pi,m}V_2\right) \leq \frac{\left[\left(2\overline{\lambda}+\alpha\right)T\right]^m}{m!} d_{\mathcal{V}}(V_1,V_2)$$

Noting that $\left[\left(2\overline{\lambda}+\alpha\right)T\right]^m/m! \to 0$ as $m \to \infty$, it follows that there exists an M such that $\left[\left(2\overline{\lambda}+\alpha\right)T\right]^M/M! < 1$. Thus $\mathcal{F}^\pi : \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\} \mapsto \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\}$ is an M-stage contraction mapping. It follows from the Banach fixed point theorem that for any bounded Borel-measurable function $v : [0, N_0] \mapsto \Re$, $\mathcal{F}^\pi$ has a unique fixed point $\hat{V}^\pi \in \{V \in \mathcal{V} : V(n,T) = v(n)\ \forall\ n \in [0,N_0]\}$. $\square$

Theorem 3 *The expected value function* V^π_{MDP} *of policy* $\pi \in \Pi^{MD}_{MDP}$ *is the unique solution of the differential equation*

$$\frac{\partial V^\pi_{MDP}}{\partial t}(n,t) = \left[-\lambda \int\int_{\mathcal{R}^\pi_1(n,t)} \{r - [V^\pi_{MDP}(n,t) - V^\pi_{MDP}(n-s,t) - p]\}\ F_{R|S}(dr|s)F_S(ds)\right.$$

$$+ \;\alpha V^{\pi}_{MDP}(n,t) \;+\; \lambda p \;+\; c(n)\Big]\, I^{\pi}(n,t)$$
$$+ \;\Big[\alpha V^{\pi}_{MDP}(n,t) \;-\; \alpha v(n)\Big]\,(1 - I^{\pi}(n,t))$$

with boundary condition $V^{\pi}_{MDP}(n,T) = v(n)$ *for all* $n \in [0, N_0]$*, that is absolutely continuous on* $[0,T]$*.*

Proof: From Lemma 12, $\mathcal{F}^{\pi}$ has a unique fixed point $\hat{V}^{\pi} \in \mathcal{V}$ that satisfies $\hat{V}^{\pi}(n,T) = v(n)$ for all $n \in [0, N_0]$. Hence,

$$\begin{aligned}
\hat{V}^{\pi}(n,t) &= \mathcal{F}^{\pi}\hat{V}^{\pi}(n,t) \;\equiv\; \hat{V}^{\pi}(n,T) - \int_t^T f^{\pi}(n,\tau,\hat{V}^{\pi})\,d\tau \\
\Rightarrow \quad \frac{\partial \hat{V}^{\pi}}{\partial t}(n,t) &= f^{\pi}(n,t,\hat{V}^{\pi}) \qquad\qquad \text{Lebesgue measure a.e.} \\
&= \Big[-\lambda(n,t)\iint_{\mathcal{R}^{\pi}_1(n,t)} \Big\{r - \Big[\hat{V}^{\pi}(n,t) \\
&\qquad\qquad -\hat{V}^{\pi}(n-s,t) - p(n,t)\Big]\Big\}\, F_{R|S}(dr|s)F_S(ds) \\
&\quad + \;\alpha\hat{V}^{\pi}(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t)\Big]\, I^{\pi}(n,t) \\
&\quad + \;\Big[\alpha\hat{V}^{\pi}(n,t) \;-\; \alpha v(n)\Big]\,(1 - I^{\pi}(n,t)) \qquad (6.2)
\end{aligned}$$

From Boel and Varaiya [21] Theorem 5.6, if Equation (6.2) has a solution $\hat{V}^{\pi} \in \mathcal{V}$ that is absolutely continuous in t such that $\hat{V}^{\pi}(n,T) = v(n)$ for all $n \in [0, N_0]$, then $\hat{V}^{\pi} = V^{\pi}_{MDP}$, the expected value function of policy $\pi \in \Pi^{MD}_{MDP}$. □

Corollary 4 *For any* $\pi \in \Pi^{MD}_{MDP}$*, and any function* $V \in \mathcal{V}$ *such that* $V(n,T) = v(n)$ *for all* $n \in [0, N_0]$*, the sequence* $\{\mathcal{F}^{\pi,m}V\}_{m=0}^{\infty}$ *converges uniformly to* $V^{\pi}_{MDP} \in \mathcal{V}$*.*

The iterative solution method suggested by Corollary 4 usually is slow to converge, and numerical methods such as finite difference approximations are commonly used to solve the differential equation (6.2). The uniform convergence of one such method is shown in Section 6.4.

6.1.3 Continuity Results

Some continuity results that are used later, are established in this section.

Lemma 13 *For any $\pi \in \Pi^{MD}_{MDP}$ and any $V \in \mathcal{V}$, $\{\mathcal{F}^{\pi}V(n,\cdot)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$.*

Proof: For any $\pi \in \Pi^{MD}_{MDP}$ and any $V \in \mathcal{V}$, $f^{\pi}(\cdot,\cdot,V)$ is bounded on $[0,N_0]\times[0,T]$. Thus $\sup_{n\in[0,N_0]}|f^{\pi}(n,t,V)|$ is integrable on $[0,T]$. Hence, for any $\varepsilon > 0$, there exists a $\delta >$ such that

$$\int_{t_1}^{t_2} \sup_{n\in[0,N_0]} |f^{\pi}(n,\tau,V)|\,d\tau \quad < \quad \varepsilon$$

for all $[t_1,t_2] \subset [0,T]$ such that $|t_2 - t_1| < \delta$. Then, for any $n \in [0,N_0]$,

$$\begin{aligned}
|\mathcal{F}^{\pi}V(n,t_2) - \mathcal{F}^{\pi}V(n,t_1)| &= \left|\int_{t_1}^{T} f^{\pi}(n,\tau,V)\,d\tau \;-\; \int_{t_2}^{T} f^{\pi}(n,\tau,V)\,d\tau\right| \\
&= \left|\int_{t_1}^{t_2} f^{\pi}(n,\tau,V)\,d\tau\right| \\
&\leq \int_{t_1}^{t_2} |f^{\pi}(n,\tau,V)|\,d\tau \\
&\leq \int_{t_1}^{t_2} \sup_{n\in[0,N_0]} |f^{\pi}(n,\tau,V)|\,d\tau \quad < \quad \varepsilon
\end{aligned}$$

Therefore $\{\mathcal{F}^{\pi}V(n,\cdot)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$. □

Corollary 5 *For any $\pi \in \Pi^{MD}_{MDP}$, $\{V^{\pi}(n,\cdot)\}_{n\in[0,N_0]} = \{\mathcal{F}^{\pi}V^{\pi}(n,\cdot)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$.*

6.2 The Optimal Value and an Optimal Policy for the MDP

In Section 6.2.1 the optimality equation for the MDP is given, the solution of which, if it exists, is the optimal value function V^*_{MDP}. The fundamental property that leads to the optimality of a reward threshold is also exhibited. In Section 6.2.2 it is shown that the optimality equation has a unique solution V^*_{MDP}. A number of continuity results that are used later are derived in Section 6.2.3.

6.2.1 The Optimality Equation

From Boel and Varaiya [21] Theorem 5.6, if the following differential equation has a solution $\hat{V}^* \in \mathcal{V}$ that is absolutely continuous in t, and there is a policy $\pi^* \in \Pi^{MD}_{MDP}$ that attains the supremum, then $\hat{V}^* = V^*_{MDP}$, the optimal value function for the MDP, and

π^* is an optimal policy for the MDP among all history-dependent deterministic policies $\pi \in \Pi_{MDP}^{HD}$.

$$\begin{aligned} -\frac{\partial V}{\partial t}(n,t) &= \sup_{(D(n,t,\cdot,\cdot),I(n,t))} \left\{ \left[\lambda(n,t) \int\int_{\mathcal{R}_1(n,t)} \{r - [V(n,t) \right.\right. \\ &\qquad -V(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s)F_S(ds) \\ &\qquad \left. - \; \alpha V(n,t) \; - \; \lambda(n,t)p(n,t) \; - \; c(n,t) \right] I(n,t) \\ &\qquad \left. + \; \left[-\alpha V(n,t) \; + \; \alpha v(n) \right] (1 - I(n,t)) \right\} \end{aligned}$$

with boundary condition $V(n,T) = v(n)$ for all $n \in [0, N_0]$.

From Lemma 5,

$$\begin{aligned} &\sup_{D(n,t,\cdot,\cdot)} \lambda(n,t) \int\int_{\mathcal{R}_1(n,t)} \{r - [V(n,t) - V(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s)F_S(ds) \\ = \;& \lambda(n,t) \int_0^n \int_{V(n,t)-V(n-s,t)-p(n,t)}^{\infty} \{r - [V(n,t) - V(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s)F_S(ds) \end{aligned}$$

Therefore, consider the differential equation

$$\begin{aligned} \frac{\partial V}{\partial t}(n,t) &= \min \left\{ -\lambda(n,t) \int_0^n \int_{V(n,t)-V(n-s,t)-p(n,t)}^{\infty} \{r - [V(n,t) \right. \\ &\qquad -V(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s)F_S(ds) \\ &\qquad \left. + \; \alpha V(n,t) \; + \; \lambda(n,t)p(n,t) \; + \; c(n,t), \;\; \alpha V(n,t) \; - \; \alpha v(n) \right\} \\ &= \min \Big\{ -\lambda(n,t)\Upsilon(n,t,V) \; + \; \alpha V(n,t) \; + \; \lambda(n,t)p(n,t) \; + \; c(n,t), \\ &\qquad \alpha V(n,t) \; - \; \alpha v(n) \Big\} \\ &\equiv f(n,t,V) \end{aligned}$$

where $\Upsilon : [0, N_0] \times [0, T] \times \mathcal{V} \mapsto \Re$ is defined by

$$\begin{aligned} \Upsilon(n,t,V) \;\; \equiv \;\; & \int_0^n \int_{V(n,t)-V(n-s,t)-p(n,t)}^{\infty} \{r - [V(n,t) \\ & \qquad -V(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s)F_S(ds) \end{aligned}$$

and $f : [0, N_0] \times [0, T] \times \mathcal{V} \mapsto \Re$. The boundary condition is $V(n,T) = v(n)$ for all $n \in [0, N_0]$. Note that $\Upsilon \geq 0$. Also, because it is assumed that λ, p, c and v are bounded, and $\int_0^{N_0} \int_{-\infty}^{\infty} |r| \; F_{R|S}(dr|s)F_S(ds) < \infty$, it follows that for any given $V \in \mathcal{V}$, $\Upsilon(\cdot,\cdot,V)$ and $f(\cdot,\cdot,V)$ are bounded on $[0, N_0] \times [0, T]$. Hence, for any $V \in \mathcal{V}$, $\Upsilon(\cdot,\cdot,V) \in \mathcal{V}$ and $f(\cdot,\cdot,V) \in \mathcal{V}$.

6.2.2 Solution of the Optimality Equation

In this section it is shown that Equation (6.3) has a solution that is the optimal value function V^*_{MDP}. The corresponding policy π^* that attains the minimum in Equation (6.3) is an optimal policy for the MDP.

Lemma 14 *The function Υ satisfies the following Lipschitz condition with respect to $V_1, V_2 \in \mathcal{V}$:*

$$|\Upsilon(n,t,V_2) - \Upsilon(n,t,V_1)| \;\leq\; 2d_t(V_1,V_2) \;\leq\; 2d_{\mathcal{V}}(V_1,V_2)$$

The proof of Lemma 14 is given in Appendix D.

Lemma 15 *The function f satisfies the following Lipschitz condition with respect to $V_1, V_2 \in \mathcal{V}$:*

$$|f(n,t,V_1) - f(n,t,V_2)| \;\leq\; \left(2\overline{\lambda} + \alpha\right) d_t(V_1,V_2) \;\leq\; \left(2\overline{\lambda} + \alpha\right) d_{\mathcal{V}}(V_1,V_2)$$

The proof of Lemma 15 is given in Appendix E.

Consider the mapping $\mathcal{F} : \mathcal{V} \mapsto \mathcal{V}$ defined by

$$\mathcal{F}V(n,t) \;\equiv\; V(n,T) - \int_t^T f(n,\tau,V)\,d\tau$$

The analysis proceeds similar to that in Section 6.1 to obtain the optimal expected value V^*_{MDP} given in Theorem 4 and the optimal policy π^* given in Theorem 5.

Theorem 4 *The optimal expected value function V^*_{MDP} for the MDP is the unique absolutely continuous solution of the differential equation*

$$\begin{aligned}\frac{\partial V^*_{MDP}}{\partial t}(n,t) \;=\; \min\Bigg\{ & -\lambda(n,t)\int_0^n \int_{V^*_{MDP}(n,t)-V^*_{MDP}(n-s,t)-p(n,t)}^{\infty} \{r - [V^*_{MDP}(n,t) \\ & \qquad - V^*_{MDP}(n-s,t) - p(n,t)]\}\; F_{R|S}(dr|s)F_S(ds) \\ & + \alpha V^*_{MDP}(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t), \\ & \alpha V^*_{MDP}(n,t) \;-\; \alpha v(n) \Bigg\} \end{aligned} \tag{6.3}$$

with boundary condition $V^*_{MDP}(n,T) = v(n)$ for all $n \in [0, N_0]$.

Theorem 5 *The following memoryless deterministic threshold policy $\pi^* \in \Pi^{MD}_{MDP}$ is optimal for the MDP among all history-dependent deterministic policies $\pi \in \Pi^{HD}_{MDP}$. An optimal stopping rule is*

$$I^*(n,t) \quad = \quad \begin{cases} 1 & \text{if } \lambda(n,t)\Upsilon(n,t,V^*_{MDP}) - \lambda(n,t)p(n,t) - c(n,t) \ \geq \ \alpha v(n) \\ 0 & \text{if } \lambda(n,t)\Upsilon(n,t,V^*_{MDP}) - \lambda(n,t)p(n,t) - c(n,t) \ < \ \alpha v(n) \end{cases}$$

An optimal acceptance rule is

$$D^*(n,t,r,s) \quad = \quad \begin{cases} 1 & \text{if } s \leq n \text{ and } r \geq V^*_{MDP}(n,t) - V^*_{MDP}(n-s,t) - p(n,t) \\ 0 & \text{if } s > n \text{ or } r < V^*_{MDP}(n,t) - V^*_{MDP}(n-s,t) - p(n,t) \end{cases}$$

That is, an optimal threshold x^* for $s \leq n$ is

$$x^*(n,t,s) \quad = \quad V^*_{MDP}(n,t) - V^*_{MDP}(n-s,t) - p(n,t)$$

6.2.3 Continuity Results

In this section a number of continuity properties that are used later, are established.

Lemma 16 $\{V^*_{MDP}(n,\cdot)\}_{n\in[0,N_0]} = \{\mathcal{F}V^*_{MDP}(n,\cdot)\}_{n\in[0,N_0]}$ *is an equicontinuous family of functions on* $[0,T]$.

The proof of Lemma 16 is similar to that of Lemma 13 and Corollary 5.

Lemma 17 *If* $\{p(n,\cdot)\}_{n\in[0,N_0]}$ *and* $\{V(n,\cdot)\}_{n\in[0,N_0]}$ *are equicontinuous families of functions on* $[0,T]$, *then* $\{\Upsilon(n,\cdot,V)\}_{n\in[0,N_0]}$ *is an equicontinuous family of functions on* $[0,T]$.

The proof of Lemma 17 is given in Appendix F.

Corollary 6 now follows from Corollary 5 and Lemmas 16 and 17.

Corollary 6 *If* $\{p(n,\cdot)\}_{n\in[0,N_0]}$ *is an equicontinuous family of functions on* $[0,T]$, *then for any* $\pi \in \Pi^{MD}_{MDP}$, $\{\Upsilon(n,\cdot,V^\pi)\}_{n\in[0,N_0]}$, *including* $\{\Upsilon(n,\cdot,V^*_{MDP})\}_{n\in[0,N_0]}$, *is an equicontinuous family of functions on* $[0,T]$.

Lemma 18 *If* $\{\lambda(n,\cdot)\}_{n\in[0,N_0]}$, $\{p(n,\cdot)\}_{n\in[0,N_0]}$, $\{c(n,\cdot)\}_{n\in[0,N_0]}$, *and* $\{V(n,\cdot)\}_{n\in[0,N_0]}$ *are equicontinuous families of functions on* $[0,T]$, *then* $\{f(n,\cdot,V)\}_{n\in[0,N_0]}$ *is an equicontinuous family of functions on* $[0,T]$.

Proof: From Lemma 17, if $\{p(n,\cdot)\}_{n\in[0,N_0]}$ and $\{V(n,\cdot)\}_{n\in[0,N_0]}$ are equicontinuous families of functions on $[0,T]$, then $\{\Upsilon(n,\cdot,V)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$. Then it follows from

$$\begin{aligned} f(n,t,V) \;\equiv\; \min\Big\{ & -\lambda(n,t)\Upsilon(n,t,V) \;+\; \alpha V(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t), \\ & \alpha V(n,t) \;-\; \alpha v(n)\Big\} \end{aligned}$$

that $\{f(n,\cdot,V)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$. □

Corollary 7 now follows from Corollary 5 and Lemmas 16 and 18.

Corollary 7 *If $\{\lambda(n,\cdot)\}_{n\in[0,N_0]}$, $\{p(n,\cdot)\}_{n\in[0,N_0]}$, and $\{c(n,\cdot)\}_{n\in[0,N_0]}$ are equicontinuous families of functions on $[0,T]$, then for any $\pi \in \Pi_{MDP}^{MD}$, $\{f(n,\cdot,V^\pi)\}_{n\in[0,N_0]}$, including $\{f(n,\cdot,V^*_{MDP})\}_{n\in[0,N_0]}$, is an equicontinuous family of functions on $[0,T]$.*

Lemma 19 *If $\lambda(n,\cdot)$, $p(n,\cdot)$, and $c(n,\cdot)$ are continuous on $[0,T]$, then $\partial V^*_{MDP}(n,\cdot)/\partial t = f(n,\cdot,V^*_{MDP})$ everywhere on $[0,T]$.*

Proof: Similar to Lemma 17 and Corollary 6, if $p(n,\cdot)$ is continuous on $[0,T]$, then $\Upsilon(n,\cdot,V^*_{MDP})$ is continuous on $[0,T]$. From Lemma 16, $V^*_{MDP}(n,\cdot)$ is continuous on $[0,T]$. It follows that $f(n,\cdot,V^*_{MDP})$ is continuous on $[0,T]$. From $V^*_{MDP}(n,t) = \mathcal{F}V^*_{MDP}(n,t) = v(n) - \int_t^T f(n,\tau,V^*_{MDP})\,d\tau$, it follows that $\partial V^*_{MDP}(n,\cdot)/\partial t = f(n,\cdot,V^*_{MDP})$ everywhere on $[0,T]$ (as opposed to "almost everywhere with respect to Lebesgue measure"). □

Let $\mathcal{C}$ denote the set of continuous functions $V : [0,N_0]\times[0,T] \mapsto \Re$. Define the metrics $d_{\mathcal{C}}$ and d_t and the functions Υ, f and $\mathcal{F}$ on $\mathcal{C}$, similar to those for the metric space $\mathcal{V}$ above.

Lemma 20 *If F_S is continuous on $[0,N_0]$ and $p \in \mathcal{C}$, then for any $V \in \mathcal{C}$, $\Upsilon(\cdot,\cdot,V) \in \mathcal{C}$.*

The proof of Lemma 20 is given in Appendix G.

Lemma 21 *If F_S is continuous on $[0,N_0]$ and $\lambda, p, c \in \mathcal{C}$, then for any $V \in \mathcal{C}$, $f(\cdot,\cdot,V) \in \mathcal{C}$.*

Proof: The result follows from the definition of f,

$$\begin{aligned} f(n,t,V) \;\equiv\; \min\Big\{ & -\lambda(n,t)\Upsilon(n,t,V) \;+\; \alpha V(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t), \\ & \alpha V(n,t) \;-\; \alpha v(n)\Big\} \end{aligned}$$

and the facts that the product, sum, and minimum of continuous functions are continuous. □

Lemma 22 *If F_S is continuous on $[0, N_0]$ and $\lambda, p, c \in \mathcal{C}$, then for any $V \in \mathcal{C}$, $\mathcal{F}V \in \mathcal{C}$, i.e., $\mathcal{F} : \mathcal{C} \mapsto \mathcal{C}$.*

Proof: Since $V \in \mathcal{C}$, V is uniformly continuous on $[0, N_0] \times [0, T]$. Hence, for any $\varepsilon > 0$ there exists a $\delta_1 > 0$ such that $|V(n_1, t_1) - V(n_2, t_2)| < \varepsilon$ for all $(n_1, t_1), (n_2, t_2) \in [0, N_0] \times [0, T]$ with $d_{\Re^2}((n_1, t_1), (n_2, t_2)) < \delta_1$.

From Lemma 21, $f(\cdot, \cdot, V)$ is uniformly continuous on $[0, N_0] \times [0, T]$. Hence, for any $\varepsilon > 0$ there exists a $\delta_2 > 0$ such that $|f(n_1, t_1, V) - f(n_2, t_2, V)| < \varepsilon / T$ for all $(n_1, t_1), (n_2, t_2) \in [0, N_0] \times [0, T]$ with $d_{\Re^2}((n_1, t_1), (n_2, t_2)) < \delta_2$.

Also, the continuity of $f(\cdot, \cdot, V)$ on $[0, N_0] \times [0, T]$ implies that $|f(n, \cdot, V)|$ is integrable on $[0, T]$ for any $n \in [0, N_0]$. Hence, for any $\varepsilon > 0$ there exists a $\delta_3 > 0$ such that $\int_{t_1}^{t_2} |f(n, \tau, V)| \, d\tau < \varepsilon$ for all $t_1, t_2 \in [0, T]$ with $|t_2 - t_1| < \delta_3$.

Let $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. For any $(n_1, t_1), (n_2, t_2) \in [0, N_0] \times [0, T]$ such that $d_{\Re^2}((n_1, t_1), (n_2, t_2)) < \delta$ and, without loss of generality, $t_1 \leq t_2$,

$$
\begin{aligned}
|\mathcal{F}V(n_1, t_1) - \mathcal{F}V(n_2, t_2)| &= \left| \left[V(n_1, T) - \int_{t_1}^{T} f(n_1, \tau, V) \, d\tau \right] \right. \\
&\quad \left. - \left[V(n_2, T) - \int_{t_2}^{T} f(n_2, \tau, V) \, d\tau \right] \right| \\
&\leq |V(n_1, T) - V(n_2, T)| + \int_{t_1}^{t_2} |f(n_1, \tau, V)| \, d\tau \\
&\quad + \int_{t_2}^{T} |f(n_2, \tau, V) - f(n_1, \tau, V)| \, d\tau \\
&< \varepsilon + \varepsilon + \int_{t_2}^{T} \frac{\varepsilon}{T} d\tau \leq 3\varepsilon
\end{aligned}
$$

Hence, for any $V \in \mathcal{C}$, $\mathcal{F}V$ is uniformly continuous on $[0, N_0] \times [0, T]$. □

The analysis proceeds similar to that in Section 6.1 to show that $\mathcal{F} : \mathcal{C} \mapsto \mathcal{C}$ is a contraction mapping in the complete metric space $\mathcal{C}$, and to obtain the result of Proposition 1.

Proposition 1 *If F_S is continuous on $[0, N_0]$ and $\lambda, p, c \in \mathcal{C}$, then for any continuous function $v : [0, N_0] \mapsto \Re$, and any function $V \in \mathcal{C}$ such that $V(n, T) = v(n)$ for all $n \in$*

$[0, N_0]$, *the sequence* $\{\mathcal{F}^m V\}_{m=0}^{\infty}$ *converges uniformly to the unique fixed point* $V^*_{MDP} \in \mathcal{C}$ *of* $\mathcal{F}$ *that satisfies the boundary condition* $V^*_{MDP}(n,T) = v(n)$ *for all* $n \in [0, N_0]$.

6.3 The Optimal Value and an Optimal Policy for the DSKP

An optimal policy $\pi^* \in \Pi^{MD}_{MDP}$ for the MDP is not necessarily admissible for the DSKP, because under policy π^* the process may be switched *on* and *off* multiple times. In this section it is shown that if λ, p and c are independent of t, then policy $\pi^* \in \Pi^{MD}_{MDP}$ given in Theorem 5 switches *off* only once and then remains switched *off*, i.e., $I^*(n,\cdot) \in \mathcal{I}_s$ for each $n \in [0, N_0]$ (Lemma 23), and therefore policy π^* is admissible and optimal for the DSKP among all policies $\pi \in \Pi^{HD}_{DSKP}$ (Theorem 6).

Assume that λ, p and c are independent of t. Then it follows from Corollary 7 that $\{f(n,\cdot,V^*_{MDP})\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$, and from Lemma 19 that $\partial V^*_{MDP}(n,t)/\partial t = f(n,t,V^*_{MDP})$ everywhere on $[0,N_0]\times[0,T]$. Proposition 2 also follows from these assumptions.

Proposition 2 *If* λ, p, *and* c *are independent of* t, *then for any* $n \in [0, N_0]$, *and* $T \in (0, \infty]$, $V^*(n,t)$ *is nonincreasing in* t *on* $[0,T]$.

The proof of Proposition 2 is given in Appendix H.

Consider the threshold policy $\pi^* \in \Pi^{MD}_{MDP}$ for the MDP given in Theorem 5. Lemma 23 shows that policy π^* is admissible for the DSKP.

Lemma 23 *For each* $n \in [0, N_0]$, $I^*(n,\cdot)$ *is a unit step function of the form*

$$I^*(n,t) = \begin{cases} 1 & \text{if } t \in (0, \tau^*(n)] \\ 0 & \text{if } t \in (\tau^*(n), T] \end{cases}$$

where $\tau^*(n) \in [0,T]$.

Proof:

Case 1: $\alpha > 0$:

Let $t_1 \equiv \sup\{t \in [0,T] : V^*_{MDP}(n,t) > v(n)\}$. From Proposition 2, $V^*_{MDP}(n,t)$ is nonincreasing in t, hence $V^*_{MDP}(n,t) > v(n)$ for all $t \in [0, t_1)$, and $V^*_{MDP}(n,t) = v(n)$ for all $t \in [t_1, T]$. Thus, for all $t \in [0, t_1)$, $\alpha V^*_{MDP}(n,t) - \alpha v(n) > 0$. But from Proposition 2, $\partial V^*_{MDP}(n,t)/\partial t \le 0$, hence $\partial V^*_{MDP}(n,t)/\partial t = -\lambda(n)\Upsilon(n,t,V^*_{MDP}) + \alpha V^*_{MDP}(n,t) + \lambda(n)p(n) + c(n) \le 0 < \alpha V^*_{MDP}(n,t) - \alpha v(n)$. Thus $\lambda(n)\Upsilon(n,t,V^*_{MDP}) - \lambda(n)p(n) - c(n) > \alpha v(n)$, and $I^*(n,t) = 1$ for all $t \in [0,t_1)$. If $t_1 = T$, then $\lambda(n)\Upsilon(n,T,V^*_{MDP}) - \lambda(n)p(n) - c(n) \ge \alpha v(n)$, from the continuity of $\Upsilon(n,t,V^*_{MDP})$ in t, hence $I^*(n,T) = 1$ and $\tau^*(n) = T$. If $t_1 < T$, then for $t \in [t_1, T)$, $\partial V^*_{MDP}(n,t)/\partial t = 0 = \alpha V^*(n,t) - \alpha v(n) \le -\lambda(n)\Upsilon(n,t,V^*_{MDP}) + \alpha V^*_{MDP}(n,t) + \lambda(n)p(n) + c(n)$, and $\lambda(n)\Upsilon(n,t_1,V^*_{MDP}) - \lambda(n)p(n) - c(n) = \alpha v(n)$ from the continuity of $\Upsilon(n,t,V^*_{MDP})$ in t. Because $V^*_{MDP}(n,t)$ is nonincreasing in t, for any $\tau_1 \le \tau_2$, $\tau_1, \tau_2 \in [t_1, T]$, $V^*_{MDP}(n-s,\tau_1) \ge V^*_{MDP}(n-s,\tau_2)$ for all $s \in [0,n]$, and hence

$$
\begin{aligned}
\Upsilon(n,\tau_1,V^*_{MDP}) &\equiv \int_0^n \int_{V^*_{MDP}(n,\tau_1)-V^*_{MDP}(n-s,\tau_1)-p(n)}^{\infty} \{r - [V^*_{MDP}(n,\tau_1) \\
&\qquad - V^*_{MDP}(n-s,\tau_1) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\
&= \int_0^n \int_{v(n)-V^*_{MDP}(n-s,\tau_1)-p(n)}^{\infty} \{r - [v(n) \\
&\qquad - V^*_{MDP}(n-s,\tau_1) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\
&\ge \int_0^n \int_{v(n)-V^*_{MDP}(n-s,\tau_2)-p(n)}^{\infty} \{r - [v(n) \\
&\qquad - V^*_{MDP}(n-s,\tau_1) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\
&\ge \int_0^n \int_{v(n)-V^*_{MDP}(n-s,\tau_2)-p(n)}^{\infty} \{r - [v(n) \\
&\qquad - V^*_{MDP}(n-s,\tau_2) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\
&= \int_0^n \int_{V^*_{MDP}(n,\tau_2)-V^*_{MDP}(n-s,\tau_2)-p(n)}^{\infty} \{r - [V^*_{MDP}(n,\tau_2) \\
&\qquad - V^*_{MDP}(n-s,\tau_2) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\
&= \Upsilon(n,\tau_2,V^*_{MDP})
\end{aligned}
$$

Thus $\Upsilon(n,t,V^*_{MDP})$ is nonincreasing in t on $[t_1,T]$. Let $t_2 \equiv \sup\{t \in [t_1,T] : \lambda(n)\Upsilon(n,t,V^*_{MDP}) - \lambda(n)p(n) - c(n) = \alpha v(n)\}$. Then $\lambda(n)\Upsilon(n,t,V^*_{MDP}) - \lambda(n)p(n) - c(n) \ge \alpha v(n)$ and $I^*(n,t) = 1$ for all $t \in [0,t_2]$, and $\lambda(n)\Upsilon(n,t,V^*_{MDP}) - \lambda(n)p(n) - c(n) < \alpha v(n)$ and $I^*(n,t) = 0$ for all $t \in (t_2,T]$. Therefore, $\tau^*(n) = t_2$, and the result holds.

Case 2: $\alpha = 0$:

By contradiction. Suppose there exists an $n \in [0,N_0]$ and $0 < t_a < t_c \le T$ such that $I^*(n,t_a) = 0$ and $I^*(n,t_c) = 1$, i.e., $\lambda(n)\Upsilon(n,t_a,V^*_{MDP}) - \lambda(n)p(n) - c(n) < 0$ and $\lambda(n)\Upsilon(n,t_c,V^*_{MDP}) - \lambda(n)p(n) - c(n) \ge 0$. Let $t_b \equiv \inf\{t \in (t_a,t_c] : \lambda(n)\Upsilon(n,t,V^*_{MDP})$

$-\lambda(n)p(n) - c(n) \geq 0\}$. From the continuity of $\Upsilon(n,t,V^*_{MDP})$ in t, it follows that $t_b > t_a$, and $\lambda(n)\Upsilon(n,t,V^*_{MDP}) - \lambda(n)p(n) - c(n) < 0$ for all $t \in [t_a, t_b)$, and $\lambda(n)\Upsilon(n,t_b,V^*_{MDP}) - \lambda(n)p(n) - c(n) = 0$. Then $\partial V^*_{MDP}(n,t)/\partial t = 0$ for all $t \in [t_a, t_b]$, and $V^*_{MDP}(n,t)$ is absolutely continuous in t, hence $V^*_{MDP}(n,t_a) = V^*_{MDP}(n,t_b)$. Because $V^*_{MDP}(n,t)$ is nonincreasing in t, $V^*_{MDP}(n-s,t_a) \geq V^*_{MDP}(n-s,t_b)$ for all $s \in [0,n]$. Thus, $V^*_{MDP}(n,t_a) - V^*_{MDP}(n-s,t_a) \leq V^*_{MDP}(n,t_b) - V^*_{MDP}(n-s,t_b)$ for all $s \in [0,n]$, and hence

$$
\begin{aligned}
& \lambda(n)\Upsilon(n,t_a,V^*_{MDP}) \;-\; \lambda(n)p(n) \;-\; c(n) \\
\equiv\;& \lambda(n)\int_0^n \int_{V^*_{MDP}(n,t_a)-V^*_{MDP}(n-s,t_a)-p(n)}^{\infty} \{r - [V^*_{MDP}(n,t_a) \\
&\qquad -V^*_{MDP}(n-s,t_a) - p(n)]\}\, F_{R|S}(dr|s)F_S(ds) \;-\; \lambda(n)p(n) \;-\; c(n) \\
\geq\;& \lambda(n)\int_0^n \int_{V^*_{MDP}(n,t_b)-V^*_{MDP}(n-s,t_b)-p(n)}^{\infty} \{r - [V^*_{MDP}(n,t_a) \\
&\qquad -V^*_{MDP}(n-s,t_a) - p(n)]\}\, F_{R|S}(dr|s)F_S(ds) \;-\; \lambda(n)p(n) \;-\; c(n) \\
\geq\;& \lambda(n)\int_0^n \int_{V^*_{MDP}(n,t_b)-V^*_{MDP}(n-s,t_b)-p(n)}^{\infty} \{r - [V^*_{MDP}(n,t_b) \\
&\qquad -V^*_{MDP}(n-s,t_b) - p(n)]\}\, F_{R|S}(dr|s)F_S(ds) \;-\; \lambda(n)p(n) \;-\; c(n) \\
\equiv\;& \lambda(n)\Upsilon(n,t_b,V^*_{MDP}) \;-\; \lambda(n)p(n) \;-\; c(n) \;=\; 0
\end{aligned}
$$

But this contradicts the assumption that $\lambda(n)\Upsilon(n,t_a,V^*_{MDP}) - \lambda(n)p(n) - c(n) < 0$. Therefore, $\lambda(n)\Upsilon(n,t,V^*_{MDP}) - \lambda(n)p(n) - c(n) \geq 0$ and $I^*(n,t) = 1$ for all $t \in (0,\tau^*(n)]$, and $\lambda(n)\Upsilon(n,t,V^*_{MDP}) - \lambda(n)p(n) - c(n) < 0$ and $I^*(n,t) = 0$ for all $t \in (\tau^*(n), T]$ for some $\tau^*(n) \in [0,T]$. □

Theorem 6 *The memoryless deterministic threshold policy π^* is an optimal policy among all history-dependent deterministic policies for the DSKP.*

Proof: From Theorem 5, π^* is an optimal policy for the MDP among all $\pi \in \Pi^{HD}_{MDP}$. From Lemma 23, π^* satisfies $I^*(n,\cdot) \in \mathcal{I}_s$ for all $n \in [0,N_0]$; hence π^* is admissible for the DSKP ($\pi^* \in \Pi^{MD}_{DSKP} \subset \Pi^{HD}_{DSKP}$), and $V^{\pi^*}_{DSKP} = V^{\pi^*}_{MDP}$. From Lemma 3, $V^*_{MDP} \geq V^*_{DSKP}$. Therefore, $V^{\pi^*}_{DSKP} = V^{\pi^*}_{MDP} = V^*_{MDP} \geq V^*_{DSKP}$, and π^* is optimal for the DSKP among all $\pi \in \Pi^{HD}_{DSKP}$. □

Note. From here on the subscripts of V^*_{DSKP} and V^*_{MDP} are dropped, V^* denoting both V^*_{DSKP} and V^*_{MDP}, unless noted otherwise.

6.4 Convergence of Discrete-Time Approximations

To solve the DSKP, we need to compute a solution to the optimality equation (6.3). Usually a numerical method is used to solve such a differential equation. A desired property for a numerical approximation scheme is that it should converge to the optimal solution as the discretization becomes finer. Theorem 7 shows that the discretization method described in this section converges uniformly to the optimal value function V^*. This convergence result is also used to establish monotonicity and convexity results in Section 6.6.3.3.

In this section, assume that $\{\lambda(n,\cdot)\}_{n\in[0,N_0]}$, $\{p(n,\cdot)\}_{n\in[0,N_0]}$, and $\{c(n,\cdot)\}_{n\in[0,N_0]}$ are equicontinuous families of functions on $[0,T]$. Then it follows from Corollary 7 that $\{f(n,\cdot,V^*)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$.

Construct a discrete-time approximation to the differential equation (6.3) as follows: For $m = 1, 2, \ldots$, let $\Delta t = T/m$, and let $\{t_i = iT/m\}_{i=0}^{m}$ be a partition of $[0,T]$. Let V_m denote the solution of the following difference equation:

$$\begin{aligned}
V_m(n,t_{i-1}) &= V_m(n,t_i) \\
&\quad + \left(\frac{T}{m}\right)\max\Bigg\{\lambda(n,t_i)\int_0^n\int_{V_m(n,t_i)-V_m(n-s,t_i)-p(n,t_i)}^{\infty}\{r-[V_m(n,t_i) \\
&\qquad\qquad -V_m(n-s,t_i)-p(n,t_i)]\}\ F_{R|S}(dr|s)F_S(ds) \\
&\qquad\qquad - \alpha V_m(n,t_i) - \lambda(n,t_i)p(n,t_i) - c(n,t_i), -\alpha V_m(n,t_i) + \alpha v(n)\Bigg\} \\
&\equiv V_m(n,t_i) - (\Delta t)\, f(n,t_i,V_m) \qquad (6.4)
\end{aligned}$$

The boundary conditions are $V_m(n,t_m) = V_m(n,T) = v(n)$ for all $n \in [0,N_0]$, for $m = 1,2,\ldots$. This discrete-time approximation corresponds to Euler's method for solving differential equations.

Theorem 7 *The sequence of solutions $\{V_m\}_{m=1}^{\infty}$ of the sequence of difference equations* (6.4) *converges uniformly to the solution $V^* \in \mathcal{V}$ of the differential equation* (6.3) *on $[0,N_0]\times[0,T]$, as $m\to\infty$, with boundary condition $V^*(n,T) = v(n)$ for all $n\in[0,N_0]$.*

Proof: Let the discretization error be denoted by

$$e_m(n,t_i) \equiv V^*(n,t_i) - V_m(n,t_i) \qquad (6.5)$$

Let the local discretization error, obtained by using V^* instead of V_m, be denoted by

$$\hat{e}_m(n, t_{i-1}) \equiv V^*(n, t_{i-1}) - [V^*(n, t_i) - (\Delta t) f(n, t_i, V^*)] \tag{6.6}$$

Then it follows from Equations (6.4), (6.5) and (6.6) that

$$\begin{aligned} V^*(n, t_{i-1}) - V_m(n, t_{i-1}) &= V^*(n, t_i) - V_m(n, t_i) - \Delta t [f(n, t_i, V^*) - f(n, t_i, V_m)] \\ &\quad + \hat{e}_m(n, t_{i-1}) \end{aligned}$$

$$\Rightarrow \quad e_m(n, t_{i-1}) = e_m(n, t_i) - \Delta t [f(n, t_i, V^*) - f(n, t_i, V_m)] + \hat{e}_m(n, t_{i-1}) \tag{6.7}$$

Because $\partial V^*(n,t)/\partial t = f(n, t, V^*)$ everywhere on $[0, N_0] \times [0, T]$, it follows from the mean value theorem applied to $V^*(n, \cdot)$ on the interval $[t_{i-1}, t_i]$ that

$$\begin{aligned} V^*(n, t_i) - V^*(n, t_{i-1}) &= (\Delta t) f(n, t_{i-1} + \theta \Delta t, V^*) \quad \text{for some } \theta \in (0,1) \\ \Rightarrow \quad \hat{e}_m(n, t_{i-1}) &= \Delta t [f(n, t_i, V^*) - f(n, t_{i-1} + \theta \Delta t, V^*)] \end{aligned} \tag{6.8}$$

From Corollary 7, $\{f(n, \cdot, V^*)\}_{n \in [0, N_0]}$ is an equicontinuous family of functions on $[0, T]$. Hence, for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(n, t_1, V^*) - f(n, t_2, V^*)| < \varepsilon$ for all $t_1, t_2 \in [0, T]$ with $|t_2 - t_1| < \delta$, for all $n \in [0, N_0]$. Thus, if $\Delta t < \delta$, then $|t_i - (t_{i-1} + \theta \Delta t)| < \Delta t < \delta$, and

$$\sup_{n \in [0, N_0]} |f(n, t_i, V^*) - f(n, t_{i-1} + \theta \Delta t, V^*)| < \varepsilon$$

for all $i \in \{1, 2, \ldots, m\}$. Let

$$\overline{e}_m \equiv \max_{i \in \{1, 2, \ldots, m\}} \sup_{n \in [0, N_0]} |f(n, t_i, V^*) - f(n, t_{i-1} + \theta \Delta t, V^*)|$$

Therefore, $\overline{e}_m \to 0$ as $\Delta t \to 0$, and thus as $m \to \infty$. Note that

$$\sup_{n \in [0, N_0]} |e_m(n, t_i)| = \sup_{n \in [0, N_0]} |V^*(n, t_i) - V_m(n, t_i)| = d_{t_i}(V^*, V_m)$$

and from Lemma 15 that

$$\sup_{n \in [0, N_0]} |f(n, t_i, V^*) - f(n, t_i, V_m)| \leq \left(2\overline{\lambda} + \alpha\right) d_{t_i}(V^*, V_m)$$

and from Equation (6.8) that

$$\begin{aligned}
\sup_{n\in[0,N_0]} \left|\hat{e}_m\left(n,t_{i-1}\right)\right| &= \Delta t \sup_{n\in[0,N_0]} \left|f\left(n,t_i,V^*\right) - f\left(n,t_{i-1}+\theta\Delta t,V^*\right)\right| \\
&\le \Delta t \max_{i\in\{1,2,\ldots,m\}} \sup_{n\in[0,N_0]} \left|f\left(n,t_i,V^*\right) - f\left(n,t_{i-1}+\theta\Delta t,V^*\right)\right| \\
&= (\Delta t)\,\overline{e}_m
\end{aligned}$$

Then it follows from Equation (6.7) that

$$\left|e_m\left(n,t_{i-1}\right)\right| \le \left|e_m\left(n,t_i\right)\right| + \Delta t\left|f\left(n,t_i,V^*\right) - f\left(n,t_i,V_m\right)\right| + \left|\hat{e}_m\left(n,t_{i-1}\right)\right|$$

$$\begin{aligned}
\Rightarrow \quad \sup_{n\in[0,N_0]} \left|e_m\left(n,t_{i-1}\right)\right| &\le \sup_{n\in[0,N_0]} \left|e_m\left(n,t_i\right)\right| + (\Delta t) \sup_{n\in[0,N_0]} \left|f\left(n,t_i,V^*\right) - f\left(n,t_i,V_m\right)\right| \\
&\quad + \sup_{n\in[0,N_0]} \left|\hat{e}_m\left(n,t_{i-1}\right)\right|
\end{aligned}$$

$$\Rightarrow \quad d_{t_{i-1}}\left(V^*,V_m\right) \le d_{t_i}\left(V^*,V_m\right) + \Delta t\left(2\overline{\lambda}+\alpha\right) d_{t_i}\left(V^*,V_m\right) + (\Delta t)\,\overline{e}_m \tag{6.9}$$

for all $i \in \{1,2,\ldots,m\}$, for all $m = 1,2,\ldots$. Next it is shown by (reverse) induction on i that the following inequality holds for $i = 0,1,\ldots,m$, for all $m = 1,2,\ldots$:

$$\begin{aligned}
d_{t_i}\left(V^*,V_m\right) &\le \overline{e}_m \frac{\left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-i} - 1}{2\overline{\lambda}+\alpha} + \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-i} d_T\left(V^*,V_m\right) \\
&\le \overline{e}_m \frac{e^{\left(2\overline{\lambda}+\alpha\right)T} - 1}{2\overline{\lambda}+\alpha} + e^{\left(2\overline{\lambda}+\alpha\right)T} d_T\left(V^*,V_m\right)
\end{aligned}$$

For $i = m$:

$$\begin{aligned}
d_{t_m}\left(V^*,V_m\right) &= d_T\left(V^*,V_m\right) \\
&= \overline{e}_m \frac{\left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-m} - 1}{2\overline{\lambda}+\alpha} + \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-m} d_T\left(V^*,V_m\right)
\end{aligned}$$

As induction hypothesis, suppose

$$d_{t_i}\left(V^*,V_m\right) \le \overline{e}_m \frac{\left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-i} - 1}{2\overline{\lambda}+\alpha} + \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-i} d_T\left(V^*,V_m\right)$$

Then it follows from Equation (6.9) that

$$\begin{aligned}
d_{t_{i-1}}\left(V^*,V_m\right) &\le \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right] d_{t_i}\left(V^*,V_m\right) + (\Delta t)\,\overline{e}_m \\
&\le \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right] \Bigg[\overline{e}_m \frac{\left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-i} - 1}{2\overline{\lambda}+\alpha}
\end{aligned}$$

$$
\begin{aligned}
& + \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-i} d_T\left(V^*,V_m\right)\Bigg] \;+\; (\Delta t)\,\overline{e}_m \\
= \;& \overline{e}_m \frac{\left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-(i-1)}-1}{2\overline{\lambda}+\alpha} \;+\; \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-(i-1)} d_T\left(V^*,V_m\right)
\end{aligned}
$$

This establishes the first inequality. For the second inequality,

$$
\begin{aligned}
1+\left(2\overline{\lambda}+\alpha\right)\Delta t &\leq e^{\left(2\overline{\lambda}+\alpha\right)\Delta t} \\
\Rightarrow \quad \left[1+\left(2\overline{\lambda}+\alpha\right)\Delta t\right]^{m-i} &\leq e^{\left(2\overline{\lambda}+\alpha\right)(m-i)\Delta t} \;\leq\; e^{\left(2\overline{\lambda}+\alpha\right)m\Delta t} \;=\; e^{\left(2\overline{\lambda}+\alpha\right)T} \\
\Rightarrow \quad d_{t_i}\left(V^*,V_m\right) &\leq \overline{e}_m \frac{e^{\left(2\overline{\lambda}+\alpha\right)T}-1}{2\overline{\lambda}+\alpha} \;+\; e^{\left(2\overline{\lambda}+\alpha\right)T} d_T\left(V^*,V_m\right)
\end{aligned}
$$

for all $i \in \{0,1,\ldots,m\}$. Noting that $V_m(n,T) = V^*(n,T) = v(n)$ for all $n \in [0,N_0]$, it follows that $d_T\left(V^*,V_m\right) = 0$ for all $m = 1,2,\ldots$. It was also shown that $\overline{e}_m \to 0$ as $m \to \infty$. Therefore, $d_{t_i}\left(V^*,V_m\right) \to 0$ as $m \to \infty$, uniformly for all $i \in \{0,1,\ldots,m\}$. Thus, for any $\varepsilon > 0$ there exists an $M_1 > 0$ such that $d_{t_i}\left(V^*,V_m\right) < \varepsilon$ for all $i \in \{0,1,\ldots,m\}$, for all $m \geq M_1$. From Lemma 16, $\{V^*(n,\cdot)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$. Hence, for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $|V^*(n,\tau_1) - V^*(n,\tau_2)| < \varepsilon$ for all $\tau_1,\tau_2 \in [0,T]$ with $|\tau_2-\tau_1| < \delta$, for all $n \in [0,N_0]$. Choose $M_2 > T/\delta$, which implies that $\Delta t \equiv T/m < \delta$ for all $m \geq M_2$. Let $M = \max\{M_1,M_2\}$. Complete $\{V_m(n,t_i)\}_{i=0}^m$ on $[0,T]$ for every $n \in [0,N_0]$ by linear interpolation, i.e., for any $t \in [0,T]$, $t \in [t_{i-1},t_i]$ for some $i \in \{1,\ldots,m\}$, and let

$$
V_m(n,t) \;\equiv\; V_m(n,t_{i-1}) + \left[V_m(n,t_i) - V_m(n,t_{i-1})\right]\frac{t-t_{i-1}}{t_i-t_{i-1}}
$$

Then for all $m \geq M$ and all $(n,t) \in [0,N_0]\times[0,T]$,

$$
\begin{aligned}
\left|V^*(n,t) - V_m(n,t_{i-1})\right| &\leq \left|V^*(n,t) - V^*(n,t_{i-1})\right| + \left|V^*(n,t_{i-1}) - V_m(n,t_{i-1})\right| \\
&< \varepsilon + d_{t_{i-1}}\left(V^*,V_m\right) \;<\; 2\varepsilon
\end{aligned}
$$

$$
\begin{aligned}
\left|V^*(n,t) - V_m(n,t_i)\right| &\leq \left|V^*(n,t) - V^*(n,t_i)\right| + \left|V^*(n,t_i) - V_m(n,t_i)\right| \\
&< \varepsilon + d_{t_i}\left(V^*,V_m\right) \;<\; 2\varepsilon
\end{aligned}
$$

Then, because $V_m(n,t)$ is between $V_m(n,t_{i-1})$ and $V_m(n,t_i)$, it follows that

$$
\left|V^*(n,t) - V_m(n,t)\right| \;<\; 2\varepsilon
$$

for all $m \geq M$, for all $(n,t) \in [0, N_0] \times [0,T]$. Therefore, $\{V_m\}_{m=1}^{\infty}$ converges uniformly to V^* on $[0, N_0] \times [0,T]$. □

In computations, a refinement of the method described in this section was used, which corresponds to Heun's version of the second order Runge-Kutta method.

6.5 Examples

Closed-form solutions for the optimal expected value $V^*(n,t)$ and the optimal threshold $x^*(n,t,s)$ can be obtained for some distributions $F_{R,S}$. Suppose all items have unit size, and the rewards are exponentially distributed with mean $1/\mu$. (Kincaid and Darling [97], Stadje [171] and Gallego and Van Ryzin [75] considered examples of a pricing problem where the maximum price a customer is willing to pay, or the arrival rate of buying customers, is exponentially distributed.) Let $\alpha = 0$, $p = 0$ and $c = 0$. Then

$$\frac{\partial V^*(n,t)}{\partial t} = -\frac{\lambda}{\mu} e^{-\mu[V^*(n,t) - V^*(n-1,t)]}$$

It can be shown by induction on n that

$$V^*(n,t) = \frac{1}{\mu} \ln \left[\sum_{i=0}^{n} \frac{\lambda^i (T-t)^i}{i!} e^{\mu v(n-i)} \right]$$

If v is constant, then

$$V^*(n,t) = \frac{1}{\mu} \ln \left[\sum_{i=0}^{n} \frac{\lambda^i (T-t)^i}{i!} \right] + v$$

It is interesting to note that, due to the continuity of the ln function,

$$\begin{aligned}
\lim_{n\to\infty} V^*(n,t) &= \frac{1}{\mu} \ln \left[\lim_{n\to\infty} \sum_{i=0}^{n} \frac{\lambda^i (T-t)^i}{i!} \right] + v \\
&= \frac{\lambda}{\mu}(T-t) + v \\
\lim_{n\to\infty} x^*(n,t) &= 0
\end{aligned}$$

This result is intuitive, because if the remaining capacity is very large, it is optimal to accept all arrivals, and the optimal threshold $x^*(n,t) = 0$ for all t. From Wald's equation the expected value is the expected number of arrivals in the remaining time, $\lambda(T-t)$, times the expected reward per arrival, $1/\mu$, plus the terminal value v.

Consider another example with all items having unit size, and with exponentially distributed rewards with mean $1/\mu$. Let $\alpha = 0$. If $\lambda(0)p(0) + c(0) \geq 0$, then $V^*(0,t) = v(0)$ for all t; otherwise $V^*(0,t) = v(0) - [\lambda(0)p(0) + c(0)](T-t)$. Suppose $p = 0$, $c(0) \geq 0$, v is constant, and $\lambda(1)/\mu \geq c(1) > 0$. Then $\lambda(1)\Upsilon(1,\cdot,0) = \lambda(1)/\mu \geq c(1)$. Hence, as shown in Corollary 8, it is optimal to continue if $n = 1$ for all $t < T$. Then it follows from Equation (6.3) that

$$\frac{\partial V^*}{\partial t}(1,t) = -\frac{\lambda(1)}{\mu} e^{-\mu[V^*(1,t)-v]} + c(1)$$

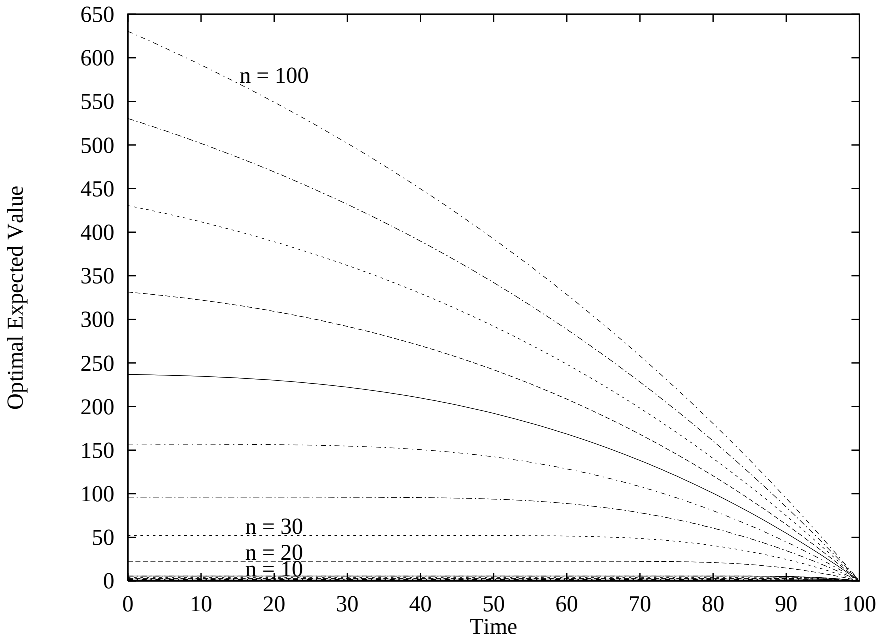
The solution is

$$V^*(1,t) = \frac{1}{\mu} \ln \left[e^{-\mu c(1)(T-t)} + \frac{\lambda(1)}{\mu c(1)} \left[1 - e^{-\mu c(1)(T-t)} \right] \right] + v$$

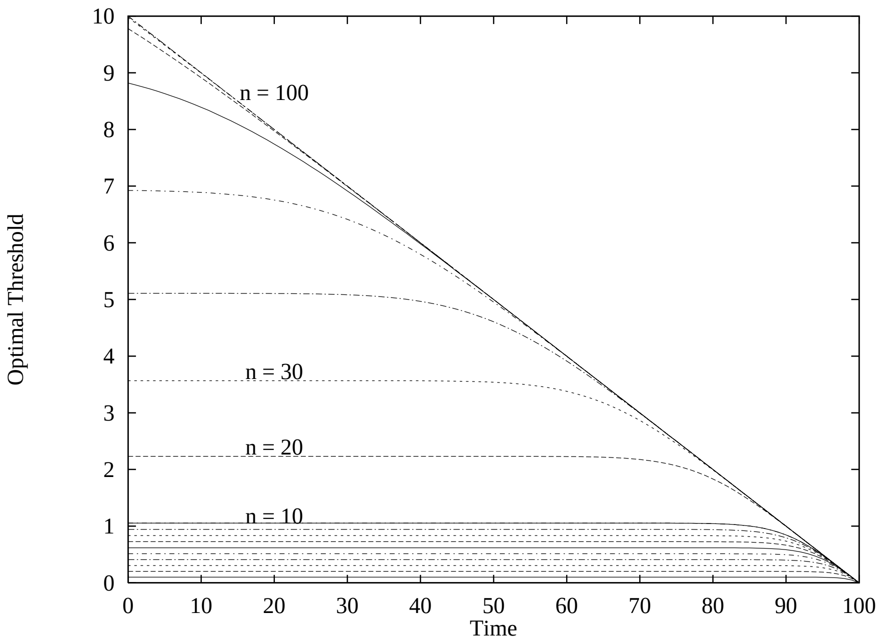
Solutions for $n > 1$ were obtained numerically. Computation times were less than a second on a Sun Sparc 2 workstation. Figure 6.1(a) shows $V^*(n,t)$ as a function of t for different values of n, for arrival rate $\lambda = 1$, deadline $T = 100$, mean reward $1/\mu = 10$, penalty $p = 0$, holding cost per unit time $c(n) = 10 - n/10$, terminal value $v = 0$, and discount rate $\alpha = 0$. Figure 6.1(b) shows $x^*(n,t)$ versus t for different n. Note that the optimal expected value is decreasing and concave in time, and increasing in the remaining capacity. The optimal threshold is decreasing in time, and increasing in the remaining capacity. Monotonicity and convexity results are discussed in more detail in Section 6.6.3.

Figure 6.2(a) shows $V^*(n,t)$ as a function of t for different values of n, for an example with all items having unit size, arrival rate $\lambda = 1$, deadline $T = 100$, exponentially distributed rewards with mean $1/\mu = 10$, penalty $p = 0$, constant holding cost per unit time $c = 5$, terminal value $v = 0$, and discount rate $\alpha = 0$. Figure 6.2(b) shows $x^*(n,t)$ versus t for different n. The optimal expected value is again decreasing and concave in time, and increasing in remaining capacity. The optimal threshold is decreasing in time, and decreasing in remaining capacity. The shape of the optimal expected value curves are similar to those in Figure 6.1(a) for a decreasing holding cost. However, the optimal threshold curves are very different for the different holding cost structures. This suggests that an optimal policy is quite sensitive with respect to the cost structure.

Consider an example with λ constant, $\alpha = 0$, $p = 0$, $c = 0$, and $v = 0$, where all items have unit size and rewards follow a uniform distribution $u(a,b)$, with $a > 0$. Then, because $V^*(n,T) = v(n) = 0$ for all n, and $V^*(n,t)$ is continuous in t for all n, there exists a time

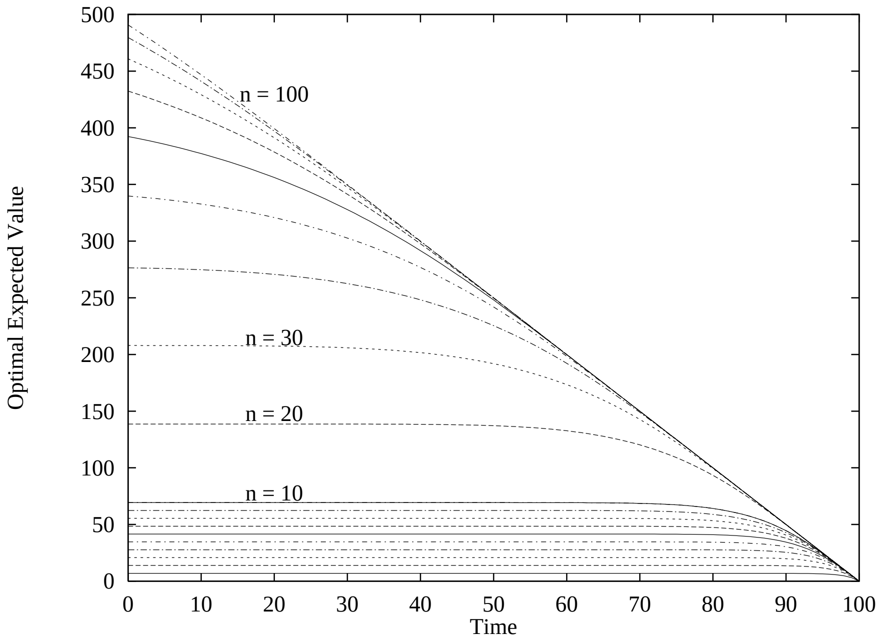


(a) Optimal Expected Value versus Time for Different Remaining Capacities n

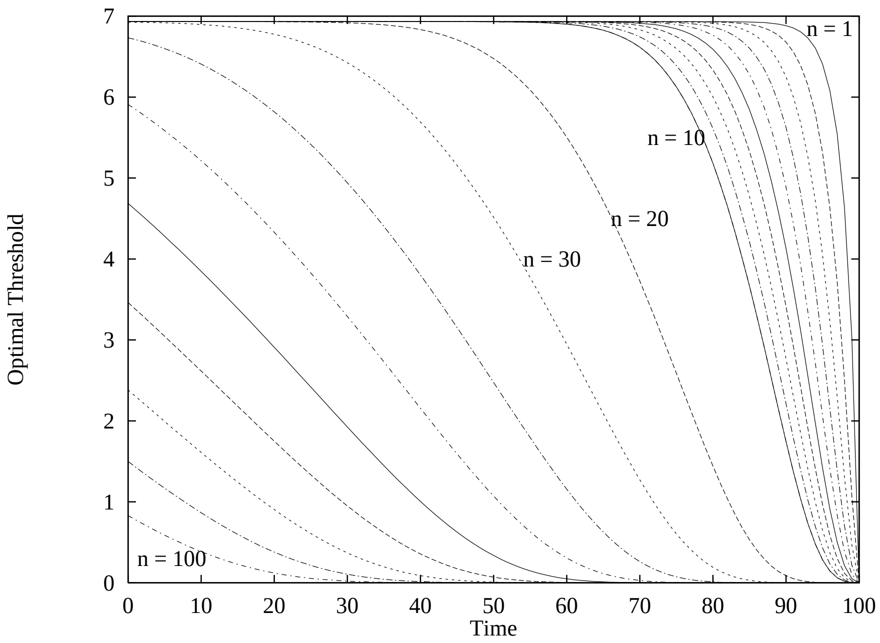


(b) Optimal Threshold versus Time for Different Remaining Capacities n

Figure 6.1 Example with Exponential Rewards and Variable Holding Cost Per Unit Time



(a) Optimal Expected Value versus Time for Different Remaining Capacities n



(b) Optimal Threshold versus Time for Different Remaining Capacities n

Figure 6.2 Example with Exponential Rewards and Constant Holding Cost Per Unit Time interval $(t_n, T]$ during which $x^*(n,t) = V^*(n,t) - V^*(n-1,t) < a$. During this time interval all arrivals are accepted, as long as the remaining capacity is $n > 0$, and any threshold less than a is optimal. Hence the following differential equation holds just before the deadline, with the usual boundary condition:

$$\frac{\partial V^*(n,t)}{\partial t} = -\lambda\left[\frac{a+b}{2} + V^*(n-1,t) - V^*(n,t)\right] \qquad t_n < t < T$$

With remaining capacity $n = 1$,

$$\frac{\partial V^*(1,t)}{\partial t} = -\lambda\left[\frac{a+b}{2} - V^*(1,t)\right] < 0 \qquad t_1 < t < T$$

The solution is

$$V^*(1,t) = x^*(1,t) = \frac{a+b}{2}\left[1 - e^{-\lambda(T-t)}\right] \qquad t_1 \leq t \leq T \quad (6.10)$$

The same solution could have been obtained by noting that, because any arrival will be accepted, the expected reward of an accepted arrival is $(a+b)/2$, and because arrivals occur according to a Poisson process, the probability of at least one arrival in the remaining time $(T-t)$ is $1 - e^{-\lambda(T-t)}$.

The optimal threshold $x^*(n,t)$ decreases with time, as shown in greater generality in Theorem 9, and $(a+b)\left[1 - e^{-\lambda(T-t)}\right]/2 \to (a+b)/2 > a$ as $(T-t) \to \infty$. Hence, from the continuity of $x^*(n,t)$ in t, for sufficiently large T, there exists a point in time $t_1 < T$, such that $x^*(1,t_1) = V^*(1,t_1) = a$, and $a < x^*(1,t) < b$ for $t < t_1$. Time t_1 is given by

$$t_1 = T - \frac{1}{\lambda}\ln\left(\frac{b+a}{b-a}\right) \quad (6.11)$$

For $t < t_n$, $a < x^*(n,t) < b$, and the following differential equation holds:

$$\frac{\partial V^*(n,t)}{\partial t} = -\frac{\lambda}{2(b-a)}\left[b - \left(V^*(n,t) - V^*(n-1,t)\right)\right]^2 \qquad t < t_n \quad (6.12)$$

with boundary condition $x^*(n,t_n) = V^*(n,t_n) - V^*(n-1,t_n) = a$. For $n = 1$,

$$\frac{\partial V^*(1,t)}{\partial t} = -\frac{\lambda}{2(b-a)}\left[b - V^*(1,t)\right]^2$$

$$\Rightarrow \quad V^*(1,t) = x^*(1,t) = \frac{\lambda b(t_1 - t) + 2a}{\lambda(t_1 - t) + 2} \qquad t \leq t_1 \quad (6.13)$$

where t_1 is given by Equation (6.11). Note that $V^*(1,t) = x^*(1,t) \to b$ as $t_1 - t \to \infty$, which is an intuitive result.

For remaining capacity $n = 2$, there exists a time interval $(t_2, T]$ when $x^*(2,t) = V^*(2,t) - V^*(1,t) < a$, and it follows that

$$\frac{\partial V^*(2,t)}{\partial t} = -\lambda \left[\frac{a+b}{2} + V^*(1,t) - V^*(2,t)\right] \qquad t_2 < t < T \quad (6.14)$$

It is shown in Theorem 9 that $x^*(n+1,t) < x^*(n,t)$, hence $(t_1, T] \subset (t_2, T]$. For $t \in [t_1, T]$, $V^*(1,t)$ is given by Equation (6.10). Substituting into Equation (6.14), and solving, gives the optimal expected value

$$V^*(2,t) = \frac{a+b}{2}\left\{2 - e^{-\lambda(T-t)}\left[2 + \lambda(T-t)\right]\right\} \qquad t_1 \le t \le T$$

and the optimal threshold

$$x^*(2,t) = V^*(2,t) - V^*(1,t) = \frac{a+b}{2}\left\{1 - e^{-\lambda(T-t)}\left[1 + \lambda(T-t)\right]\right\} \qquad t_1 \le t \le T$$

It is insightful to note that at t_1, where $V^*(1,t_1) = x^*(1,t_1) = a$,

$$V^*(2,t_1) = 2a - \frac{b-a}{2}\ln\left(\frac{b+a}{b-a}\right)$$

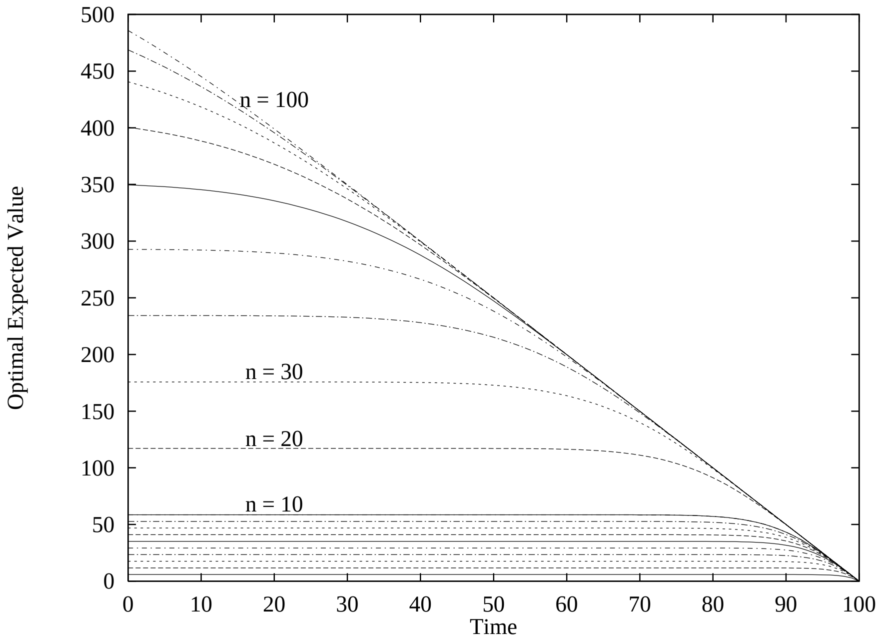
$$x^*(2,t_1) = V^*(2,t_1) - V^*(1,t_1) = a - \frac{b-a}{2}\ln\left(\frac{b+a}{b-a}\right) < a$$

This confirms that $(t_1, T] \subset (t_2, T]$, as established above. For $t \in [t_2, t_1]$, Equation (6.13) has to be substituted into Equation (6.14) to solve for $V^*(2,t)$ and $x^*(2,t)$. Further calculations were intricate, and numerical methods were used to compute the solutions recursively.

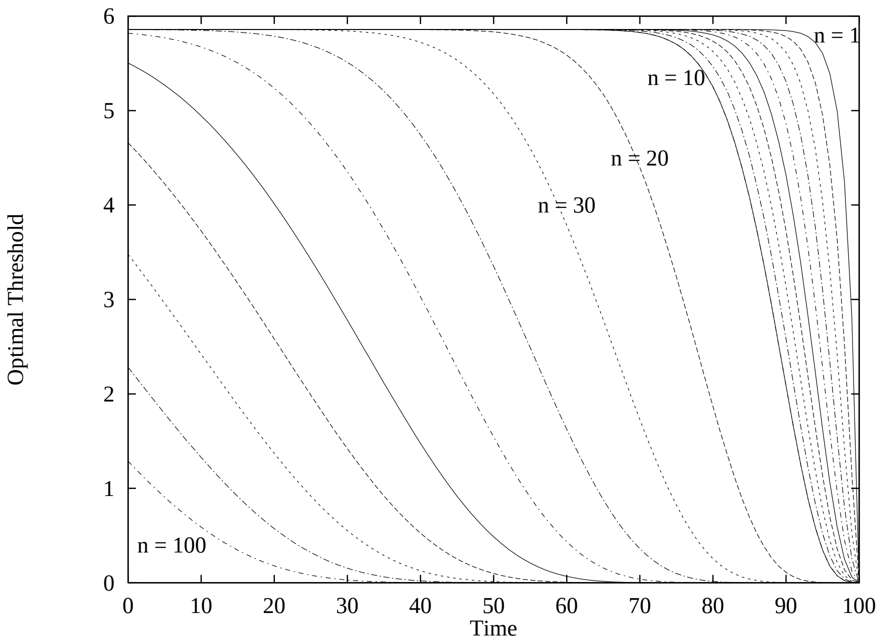
Figure 6.3(a) shows $V^*(n,t)$ as a function of t for different values of n, for an example with all items having unit size, arrival rate $\lambda = 1$, deadline $T = 100$, uniform reward distribution $u(0,20)$, penalty $p = 0$, constant holding cost per unit time $c = 5$, terminal value $v = 0$, and discount rate $\alpha = 0$. Figure 6.3(b) shows $x^*(n,t)$ versus t for different n. The curves are similar to those of Figure 6.2(b) for the case of exponential rewards. This and other experimentation suggest that an optimal policy is not very sensitive with respect to the reward distribution, but quite sensitive with respect to the cost structure, as observed above.

6.6 Structural Characteristics

Conditions under which intuitive structural characteristics of the optimal expected value V^* and optimal threshold x^* hold, are established in this section. Section 6.6.1 discusses the



(a) Optimal Expected Value versus Time for Different Remaining Capacities n



(b) Optimal Threshold versus Time for Different Remaining Capacities n

Figure 6.3 Example with Uniform Rewards and Constant Holding Cost Per Unit Time fundamental roles played by the rate at which rewards are earned, the rate at which cost is incurred, and their relative magnitudes. A recursive form for the optimal expected value and an optimal policy is given in Section 6.6.2. that holds under typical conditions. This characterization is useful because it gives a simple, intuitive recipe for following an optimal policy, and it simplifies computation of V^* and x^*. Thereafter some intuitive monotonicity and convexity properties are shown to hold under certain conditions. Also interesting are the cases where counter-intuitive properties are exhibited, some of which are illustrated by example.

6.6.1 A Fundamental Relationship

The properties of V^* depend to a large extent on the relative magnitudes of $\lambda(n)\Upsilon(n,\cdot,0)$ and $\lambda(n)p(n)+c(n)+\alpha v(n)$. An example of this is given in Proposition 4. The importance of these quantities makes intuitive sense, by noting that

$$\begin{aligned}
\lambda\left[\Upsilon(n,\cdot,0)-p(n)\right] &= \lambda(n)\left[\int_0^n\int_{-p(n)}^{\infty}[r+p(n)]\,F_{R|S}(dr|s)F_S(ds)-p(n)\right]\\
&= \lambda(n)\left[\int_0^n\int_{-p(n)}^{\infty} r\,F_{R|S}(dr|s)F_S(ds)\right.\\
&\qquad \left. -\,p(n)\left(1-\int_0^n\int_{-p(n)}^{\infty}F_{R|S}(dr|s)F_S(ds)\right)\right]\\
&= \lambda(n)\left(P\left[\{S\le n\}\cap\{R\ge -p(n)\}\right]E\left[R\mid\{S\le n\}\cap\{R\ge -p(n)\}\right]\right.\\
&\qquad \left. -\,p(n)P\left[\{S>n\}\cup\{R<-p(n)\}\right]\right)
\end{aligned}$$

and by interpreting $\lambda(n)P\left[\{S\le n\}\cap\{R\ge -p(n)\}\right]E\left[R\mid\{S\le n\}\cap\{R\ge -p(n)\}\right]$ as the effective reward rate while one continues to wait, and comparing it with

$$\lambda(n)p(n)P\left[\{S>n\}\cup\{R<-p(n)\}\right]+c(n)+\alpha v(n)$$

the rate at which (opportunity) cost is incurred while one continues to wait.

Proposition 4 gives conditions that are sufficient to make it optimal to stop at any time when the remaining capacity is n. Not surprisingly, one of the conditions is that the rate at which rewards are earned be less than the rate at which cost is incurred. First, Proposition 3 establishes the intuitive result that if λ is constant, $p(n)$ and $c(n)$ are nonincreasing, and $v(n)$ is nondecreasing, then $V^*(n,t)$ is nondecreasing in n. These are conditions that often hold in applications. It is typical for the holding cost to increase as the number of accepted customers increases, and for the terminal value to decrease (for example, for the dispatching and transportation cost to increase) as the final number of customers increases.

Proposition 3 *If λ is constant, $p(n)$ and $c(n)$ are nonincreasing, and $v(n)$ is nondecreasing, then for any $T \in (0, \infty]$, and $t \in [0, T]$, $V^*(n, t)$ is nondecreasing in n on $[0, N_0]$.*

The proof of Proposition 3 is given in Appendix I. The reasons the stated conditions are needed for Proposition 3 to hold, are obvious, and thus no counterexamples are given to show that the conditions cannot be dropped.

Proposition 4 *If λ is constant, $p(n)$ and $c(n)$ are nonincreasing, $v(n)$ is nondecreasing, and $\lambda\Upsilon(n, \cdot, 0) \leq \lambda p(n) + c(n) + \alpha v(n)$ for an n, then $V^*(n, t) = v(n)$ for all $t \in [0, T]$.*

Proof: From Proposition 3, if λ is constant, $p(n)$ and $c(n)$ are nonincreasing, and $v(n)$ is nondecreasing, then $V^*(n, t)$ is nondecreasing in n. Hence, $V^*(n, t) - V^*(n - s, t) \geq 0$ for all $s \in [0, n]$ and all $t \in [0, T]$. Thus

$$\begin{aligned}
& \lambda \int_0^n \int_{V^*(n,t)-V^*(n-s,t)-p(n)}^{\infty} \{r - [V^*(n,t) - V^*(n-s,t) - p(n)]\} \, F_{R|S}(dr|s) F_S(ds) \\
& - \alpha V^*(n,t) \; - \; \lambda p(n) \; - \; c(n) \\
\leq \; & \lambda \int_0^n \int_{V^*(n,t)-V^*(n-s,t)-p(n)}^{\infty} [r + p(n)] \; F_{R|S}(dr|s) F_S(ds) \; - \; \alpha V^*(n,t) \; - \; \lambda p(n) \; - \; c(n) \\
\leq \; & \lambda \int_0^n \int_{-p(n)}^{\infty} [r + p(n)] \; F_{R|S}(dr|s) F_S(ds) \; - \; \alpha V^*(n,t) \; - \; \lambda p(n) \; - \; c(n) \\
= \; & \lambda \Upsilon(n, \cdot, 0) \; - \; \alpha V^*(n,t) \; - \; \lambda p(n) \; - \; c(n) \\
\leq \; & -\alpha V^*(n,t) \; + \; \alpha v(n)
\end{aligned}$$

Therefore, $\partial V^*(n, t)/\partial t = \alpha V^*(n, t) - \alpha v(n)$, and $V^*(n, t) = v(n)$ for all $t \in [0, T]$. □

Counterexamples where Proposition 4 Does Not Hold

It is shown by way of counterexamples that the assumptions made in Proposition 4 cannot be dropped. The first assumption is that λ is constant. Figure 6.4 shows $V^*(n, t)$ as a function of t for different n, for an example where all items have size 1, rewards are exponentially distributed with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda(n) = 0.1 - 0.01n$ for $n \in \{0, 1, \ldots, 10\}$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = 0.5$, the terminal value $v(n) = 0$, and the discount rate $\alpha = 0$. Note that for $n = 6$, $V^*(6,t) > v(6) = 0$ for all $t < 69$, although $\lambda(6)\Upsilon(6,t,0) = \lambda(6)/\mu = 0.4 < \lambda(6)p(6) + c(6) + \alpha v(6) = 0.5$ for all $t \in [0,T]$. Therefore, Proposition 4 does not hold if λ is an arbitrary decreasing function of n.

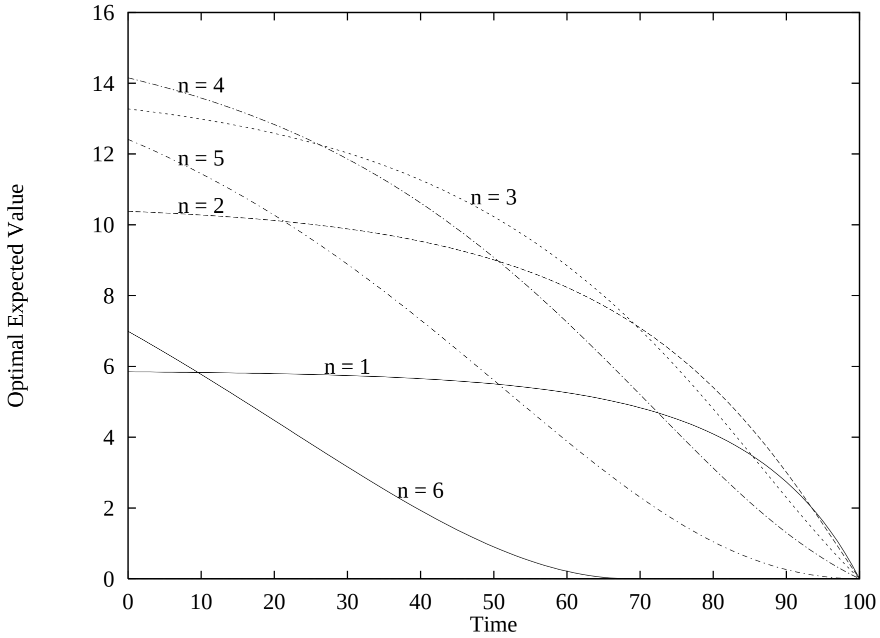


Figure 6.4 Example to show that Proposition 4 does not hold if λ is an arbitrary decreasing function of n

Figure 6.5 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are normally distributed with mean $\mu = -1$ and standard deviation $\sigma = 1$, the deadline $T = 100$, items arrive at rate $\lambda(n) = 0.01n$, the penalty $p(n) = 1$, the holding cost per unit time $c(n) = -0.05$, the terminal value $v(n) = 0$, and the discount rate $\alpha = 0$. Note that for $n = 9$, $V^*(9,t) > v(9) = 0$ for all $t < 79$, although $\lambda(9)\Upsilon(9,t,0) = 0.036 < \lambda(9)p(9) + c(9) + \alpha v(9) = 0.04$ for all $t \in [0,T]$. A similar result holds for $n = 10$. Therefore, Proposition 4 does not hold if λ is an arbitrary increasing function of n.

The next assumption is that $p(n)$ is nonincreasing. Figure 6.6 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are normally distributed with mean $\mu = 0$ and standard deviation $\sigma = 5$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 2.5n$, the holding cost per unit time $c(n) = 0.1$, the terminal value $v(n) = 0$, and the discount rate $\alpha = 0$. Note that for $n = 10$, $V^*(10,t) > v(10) = 0$ for all $t < 100$, although $\lambda\Upsilon(5,t,0) = 1.286 < \lambda p(5) + c(5) + \alpha v(5) = 1.35$ for all

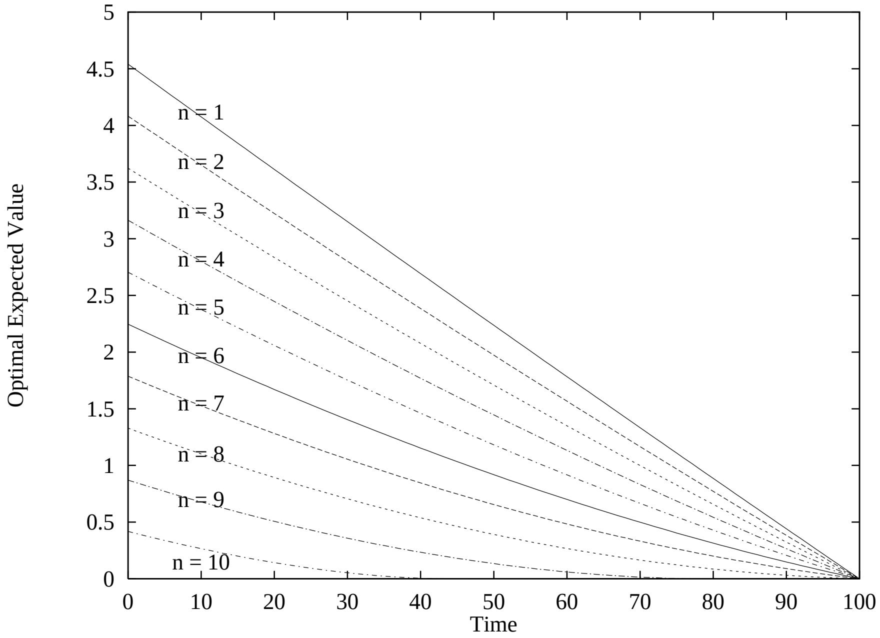


Figure 6.5 Example to show that Proposition 4 does not hold if λ is an arbitrary increasing function of n

$t \in [0, T]$. Similar results hold for $n \in \{6, \ldots, 10\}$. Therefore, Proposition 4 does not hold if p is an arbitrary increasing function of n.

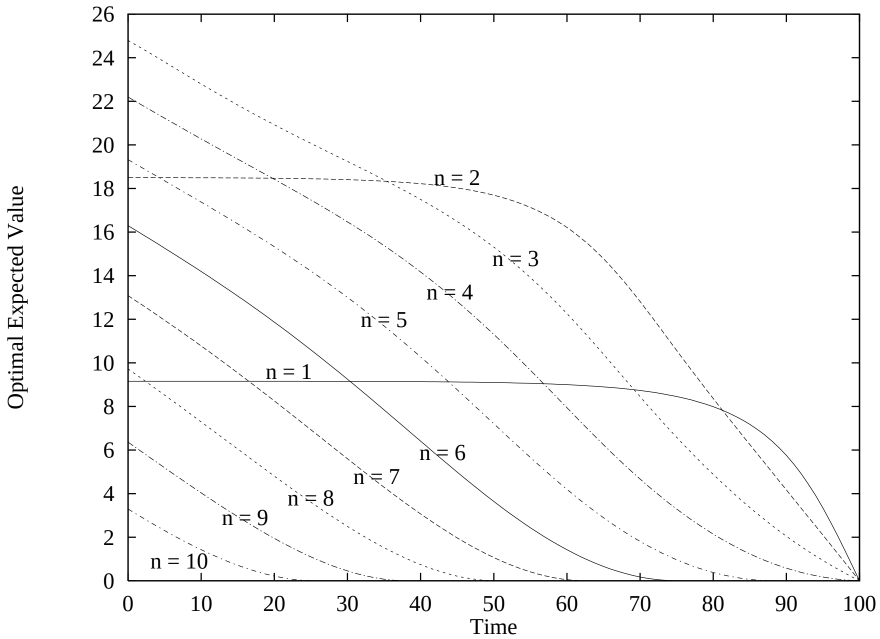


Figure 6.6 Example to show that Proposition 4 does not hold if p is an arbitrary increasing function of n

The next assumption is that $c(n)$ is nonincreasing. Figure 6.7 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are exponentially distributed with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda = 0.5$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = n$, the terminal value $v(n) = n$, and the discount rate $\alpha = 0$. Note that for $n = 9$, $V^*(9,t) > v(9) = 9$ for all $t < 69$, although $\lambda\Upsilon(9,t,0) = 5.0 < \lambda p(9) + c(9) + \alpha v(9) = 9.0$ for all $t \in [0,T]$. Similar results hold for $n \in \{6,7,8\}$. Therefore, Proposition 4 does not hold if c is an arbitrary increasing function of n.

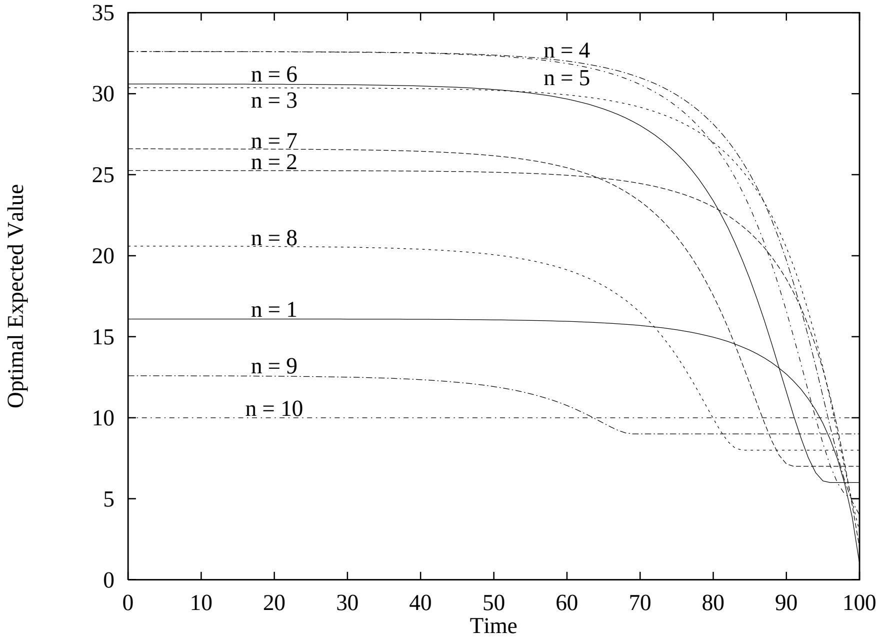


Figure 6.7 Example to show that Proposition 4 does not hold if c is an arbitrary increasing function of n

The next assumption is that $v(n)$ is nondecreasing. Figure 6.8 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are exponentially distributed with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = 1.1$, the terminal value $v(n) = 20 - 2n$, and the discount rate $\alpha = 0$. Note that for $n = 10$, $V^*(10,t) > v(10) = 0$ for all $t < 100$, although $\lambda\Upsilon(10,t,0) = 1.0 < \lambda p(10) + c(10) + \alpha v(10) = 1.1$ for all $t \in [0,T]$. Similar results hold for $n \in \{1,2,\ldots,9\}$. Therefore, Proposition 4 does not hold if v is an arbitrary decreasing function of n.

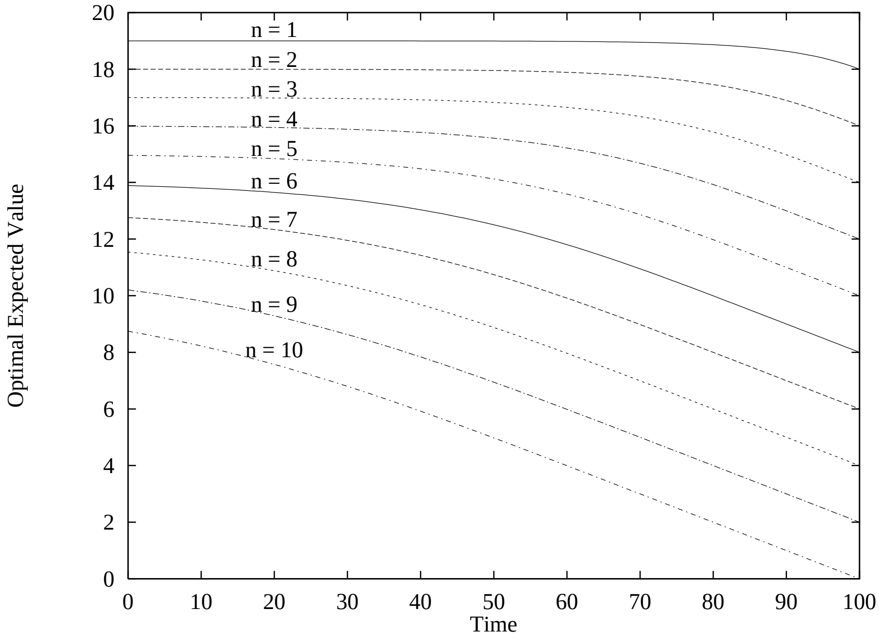


Figure 6.8 Example to show that Proposition 4 does not hold if v is an arbitrary decreasing function of n

6.6.2 A Recursion for the Optimal Expected Value

Similar to the infinite horizon case, a recursive form for the optimal expected value V^* is derived. The conditions under which $V^*(n,t) = \max\{v(n), V^{\psi(n)}(n,t)\}$ for the finite horizon case are more restrictive than those for the infinite horizon case. Examples are given to show that these conditions cannot be dropped.

6.6.2.1 Definition of Recursive Policies

Again assume that the size distribution F_S is such that $F_S(0) = 0$, and the possible values of the remaining capacity are $0 = n_0 < n_1 < n_2 < \ldots < N_0$. Without loss of generality, to simplify notation, assume that $n_k = k$ for all k.

Inductively define the sequence of threshold policies $\{\psi(n)\}_{n=0}^{N_0}$. $D^{\psi(0)}(0,t) = 0$, $I^{\psi(0)}(0,t) = 1$ for all $t \in (0,T]$. Then

$$V^{\psi(0)}(0,t) \quad = \quad e^{-\alpha(T-t)}v(0) - \frac{\lambda(0)p(0)+c(0)}{\alpha}\left(1 - e^{-\alpha(T-t)}\right)$$

if $\alpha > 0$, and $V^{\psi(0)}(0,t) = -[\lambda(0)p(0) + c(0)](T-t) + v(0)$ if $\alpha = 0$. Let $\psi(n-s)$ and $V^{\psi(n-s)}(n',\cdot)$ be defined, with $V^{\psi(n-s)}(n',\cdot)$ absolutely continuous on $[0,T]$, for all

$s \in \{1, \ldots, n\}$ and all $n' \in \{0, 1, \ldots, n-s\}$. Because n and s assume only a finite number of values $n \in \{0, 1, \ldots, N_0\}$ and $s \in \{1, \ldots, n\}$, and each $V^{\psi(n-s)}(n-s, \cdot)$ is continuous and thus bounded on $[0, T]$, the function $\zeta : \{(n, t, s) : n \in \{0, 1, \ldots, N_0\}, t \in [0, T], s \in \{1, \ldots, n\}\} \mapsto \Re$ defined by $\zeta(n, t, s) \equiv \max\left\{v(n-s), V^{\psi(n-s)}(n-s, t)\right\} + p(n)$ is bounded. Let $\hat{V}(n, \cdot)$ satisfy

$$\begin{aligned}
\frac{\partial \hat{V}}{\partial t}(n, t) &= -\lambda(n) \int_0^n \int_{\hat{V}(n,t) - \max\{v(n-s), V^{\psi(n-s)}(n-s,t)\} - p(n)}^{\infty} \Big\{ r - \Big[\hat{V}(n, t) \\
&\qquad - \max\{v(n-s), V^{\psi(n-s)}(n-s, t)\} - p(n) \Big] \Big\} F_{R|S}(dr|s) F_S(ds) \\
&\quad + \alpha \hat{V}(n, t) + \lambda(n) p(n) + c(n) \\
&= -\lambda(n)\, g\left(n, t, \hat{V}(n, t), \zeta\right) + \alpha \hat{V}(n, t) + \lambda(n) p(n) + c(n) \qquad (6.15)
\end{aligned}$$

for $t \in (0, T)$ with boundary condition $\hat{V}(n, T) = v(n)$. The right hand side of Equation (6.15) satisfies a Lipschitz condition with respect to $\hat{V}$ with constant $\overline{\lambda} + \alpha$. Then the analysis proceeds similar to that in Section 6.1 to show that Equation (6.15) has a unique solution $\hat{V}(n, \cdot)$ that is absolutely continuous on $[0, T]$. Let $I^{\psi(n)}(n, t) = 1$, and let $x^{\psi(n)}(n, t, s) = \hat{V}(n, t) - \max\{v(n-s), V^{\psi(n-s)}(n-s, t)\} - p(n) = \hat{V}(n, t) - \zeta(n, t, s)$ for all $s \in \{1, \ldots, n\}$ and all $t \in (0, T]$. Let

$$I^{\psi(n)}(n-s, t) = \begin{cases} 1 & \text{if } V^{\psi(n-s)}(n-s, t) > v(n-s) \\ 0 & \text{if } V^{\psi(n-s)}(n-s, t) \leq v(n-s) \end{cases}$$

and $D^{\psi(n)}(n-s, t, \cdot, \cdot) = D^{\psi(n-s)}(n-s, t, \cdot, \cdot)$ for all $s \in \{1, \ldots, n\}$ and all $t \in (0, T]$. From the definition of $\{\psi(n)\}$, for all $s \in \{1, \ldots, n\}$ and all $s' \in \{1, \ldots, n-s\}$, $D^{\psi(n)}(n-s-s', t, \cdot, \cdot) = D^{\psi(n-s-s')}(n-s-s', t, \cdot, \cdot) = D^{\psi(n-s)}(n-s-s', t, \cdot, \cdot)$, and $I^{\psi(n)}(n-s-s', t) = I^{\psi(n-s)}(n-s-s', t)$ for all $t \in (0, T]$. Hence $V^{\psi(n)}(n-s-s', t) = V^{\psi(n-s)}(n-s-s', t)$ for all $t \in [0, T]$. Then it follows from the definition of $\{\psi(n)\}$, Theorem 3 and Equation (6.1) that $V^{\psi(n)}(n, t)$ is the unique solution of

$$\begin{aligned}
\frac{\partial V^{\psi(n)}}{\partial t}(n, t) &= -\lambda(n) \int_0^n \int_{x^{\psi(n)}(n,t,s)}^{\infty} \Big\{ r - \Big[V^{\psi(n)}(n, t) \\
&\qquad - V^{\psi(n)}(n-s, t) - p(n) \Big] \Big\} F_{R|S}(dr|s) F_S(ds) \\
&\quad + \alpha V^{\psi(n)}(n, t) + \lambda(n) p(n) + c(n) \\
&= -\lambda(n)\, h\left(n, t, \hat{V}(n, t), V^{\psi(n)}(n, t), \zeta\right) + \alpha V^{\psi(n)}(n, t) + \lambda(n) p(n) + c(n) \quad (6.16)
\end{aligned}$$

with boundary condition $V^{\psi(n)}(n,T)=v(n)$, and for any $s\in\{1,\ldots,n\}$, $V^{\psi(n)}(n-s,t)$ is the unique solution of

$$\begin{aligned}
\frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) &= \Bigg[-\lambda(n-s)\int\!\!\int_{\mathcal{R}_1^{\psi(n)}(n-s,t)}\Big\{r-\Big[V^{\psi(n)}(n-s,t)\\
&\qquad -V^{\psi(n)}(n-s-s',t)-p(n-s)\Big]\Big\}\,F_{R|S}(dr|s')F_S(ds')\\
&\qquad +\alpha V^{\psi(n)}(n-s,t)\;+\;\lambda(n-s)p(n-s)\;+\;c(n-s)\Bigg]\,I^{\psi(n)}(n-s,t)\\
&\qquad +\Big[\alpha V^{\psi(n)}(n-s,t)-\alpha v(n-s)\Big]\Big[1-I^{\psi(n)}(n-s,t)\Big]\\
&= \Bigg[-\lambda(n-s)\int_0^{n-s}\int_{x^{\psi(n-s)}(n-s,t,s')}^{\infty}\Big\{r-\Big[V^{\psi(n)}(n-s,t)\\
&\qquad -V^{\psi(n-s)}(n-s-s',t)-p(n-s)\Big]\Big\}\,F_{R|S}(dr|s')F_S(ds')\\
&\qquad +\alpha V^{\psi(n)}(n-s,t)\;+\;\lambda(n-s)p(n-s)\;+\;c(n-s)\Bigg]\times\\
&\qquad I_{\{V^{\psi(n-s)}(n-s,t)>v(n-s)\}}\\
&\qquad +\Big[\alpha V^{\psi(n)}(n-s,t)-\alpha v(n-s)\Big]\,I_{\{V^{\psi(n-s)}(n-s,t)\le v(n-s)\}}
\end{aligned}\tag{6.17}$$

with boundary condition $V^{\psi(n)}(n-s,T)=v(n-s)$, that is absolutely continuous on $[0,T]$. Because the right-hand-side of Equation (6.16) is continuous on $[0,T]$, $V^{\psi(n)}(n,t)$ satisfies the differential equation everywhere on $[0,T]$. The same applies to Equation (6.17) on any interval $(t_0,t_2)\subseteq(0,T)$ on which $I^{\psi(n)}(n-s,t)$ remains constant.

Lemma 24 *For all $n>0$, all $t\in[0,T]$, and all $s\in\{1,\ldots,n\}$,*

$$V^{\psi(n)}(n-s,t)=\max\Big\{v(n-s),V^{\psi(n-s)}(n-s,t)\Big\}$$

The proof of Lemma 24 is given in Appendix J.

It follows from Equation (6.16) and Lemma 24 that $V^{\psi(n)}(n,\cdot)$ is the unique solution of

$$\begin{aligned}
\frac{\partial V^{\psi(n)}(n,t)}{\partial t} &= -\lambda(n)\int_0^n\int_{\hat V(n,t)-\max\{v(n-s),V^{\psi(n-s)}(n-s,t)\}-p(n)}^{\infty}\Big\{r-\Big[V^{\psi(n)}(n,t)\\
&\qquad -\max\{v(n-s),V^{\psi(n-s)}(n-s,t)\}-p(n)\Big]\Big\}\,F_{R|S}(dr|s)F_S(ds)\\
&\qquad +\alpha V^{\psi(n)}(n,t)\;+\;\lambda(n)p(n)\;+\;c(n)\\
&= -\lambda(n)\,g\left(n,t,\hat V(n,t),\zeta\right)\;+\;\alpha V^{\psi(n)}(n,t)\;+\;\lambda(n)p(n)\;+\;c(n)
\end{aligned}\tag{6.18}$$

with boundary condition $V^{\psi(n)}(n,T) = v(n)$, that is absolutely continuous on $[0,T]$. From Equation (6.15), Equation (6.18) has a unique solution $V^{\psi(n)}(n,\cdot) = \hat{V}(n,\cdot)$. Therefore,

$$\begin{aligned}\frac{\partial V^{\psi(n)}(n,t)}{\partial t} &= -\lambda(n)\int_0^n \int_{V^{\psi(n)}(n,t)-V^{\psi(n)}(n-s,t)-p(n)}^{\infty} \Big\{r - \Big[V^{\psi(n)}(n,t) \\ &\qquad\qquad -V^{\psi(n)}(n-s,t) - p(n)\Big]\Big\} F_{R|S}(dr|s)F_S(ds) \\ &\quad + \alpha V^{\psi(n)}(n,t) \;+\; \lambda(n)p(n) \;+\; c(n) \\ &= -\lambda(n)\,\Upsilon\left(n,t,V^{\psi(n)}\right) \;+\; \alpha V^{\psi(n)}(n,t) \;+\; \lambda(n)p(n) \;+\; c(n) \qquad (6.19)\end{aligned}$$

6.6.2.2 Optimality of Continuing

In this section it is shown that it is optimal to continue until the deadline if $v(n)$ is nonincreasing and $\lambda(n)\Upsilon(n,\cdot,0) > \lambda(n)p(n) + c(n) + \alpha v(n)$. The second condition makes sense in light of the discussion of Section 6.6.1, because it requires that the rate at which rewards are earned be greater than the rate at which costs are incurred. The condition that $v(n)$ be nonincreasing is needed because if $v(n)$ were increasing too steeply, then the loss in terminal value when accepting an item would be too much to justify continuing to wait for arrivals.

Proposition 5 *If $v(n)$ is nonincreasing and $\lambda(n)\Upsilon(n,\cdot,0) > \lambda(n)p(n) + c(n) + \alpha v(n)$ for an $n > 0$, then $V^{\psi(n)}(n,t) > v(n)$ for all $t \in [0,T)$.*

Proof: By contradiction. Suppose there exists $t_1 \in [0,T)$ such that $V^{\psi(n)}(n,t_1) \le v(n)$. Then $x^{\psi(n)}(n,t_1,s) = V^{\psi(n)}(n,t_1) - \max\{v(n-s), V^{\psi(n-s)}(n-s,t_1)\} - p(n) \le -p(n)$. Thus

$$\begin{aligned}&\lambda(n)\Upsilon\left(n,t_1,V^{\psi(n)}\right) \;-\; \alpha V^{\psi(n)}(n,t_1) \;-\; \lambda(n)p(n) \;-\; c(n) \\ =\;& \lambda(n)\int_0^n \int_{V^{\psi(n)}(n,t_1)-V^{\psi(n)}(n-s,t_1)-p(n)}^{\infty} \Big\{r - \Big[V^{\psi(n)}(n,t_1) \\ &\qquad\qquad -V^{\psi(n)}(n-s,t_1) - p(n)\Big]\Big\} F_{R|S}(dr|s)F_S(ds) \\ &-\; \alpha V^{\psi(n)}(n,t_1) \;-\; \lambda(n)p(n) \;-\; c(n) \\ \ge\;& \lambda(n)\int_0^n \int_{-p(n)}^{\infty} \Big\{r - \Big[V^{\psi(n)}(n,t_1) - V^{\psi(n)}(n-s,t_1) - p(n)\Big]\Big\} F_{R|S}(dr|s)F_S(ds) \\ &-\; \alpha V^{\psi(n)}(n,t_1) \;-\; \lambda(n)p(n) \;-\; c(n) \\ \ge\;& \lambda(n)\int_0^n \int_{-p(n)}^{\infty} [r + p(n)]\, F_{R|S}(dr|s)F_S(ds) \;-\; \alpha V^{\psi(n)}(n,t_1) \;-\; \lambda(n)p(n) \;-\; c(n) \\ =\;& \lambda(n)\Upsilon(n,\cdot,0) \;-\; \alpha V^{\psi(n)}(n,t_1) \;-\; \lambda(n)p(n) \;-\; c(n) \\ \ge\;& \lambda(n)\Upsilon(n,\cdot,0) \;-\; \alpha v(n) \;-\; \lambda(n)p(n) \;-\; c(n) \;>\; 0\end{aligned}$$

Then, from the continuity of $\Upsilon(n,t,V^{\psi(n)})$ and $V^{\psi(n)}(n,t)$ in t, there exists a neighborhood $(t_0,t_2) \subseteq (0,T)$ of t_1 such that $-\lambda(n)\Upsilon(n,t,V^{\psi(n)}) + \alpha V^{\psi(n)}(n,t) + \lambda(n)p(n) + c(n) < 0$ for all $t \in (t_0,t_2)$. Then from Equation (6.19)

$$\frac{\partial V^{\psi(n)}}{\partial t}(n,t) = -\lambda(n)\Upsilon(n,t,V^{\psi(n)}) + \alpha V^{\psi(n)}(n,t) + \lambda(n)p(n) + c(n) < 0$$

for all $t \in (t_0,t_2)$. By repeating the argument at t_2, it follows that $V^{\psi(n)}(n,t)$ is strictly decreasing on $[t_1,T]$. This implies that $V^{\psi(n)}(n,T) < V^{\psi(n)}(n,t_1) \leq v(n)$, which violates the boundary condition $V^{\psi(n)}(n,T) = v(n)$. Therefore, $V^{\psi(n)}(n,t) > v(n)$ for all $t \in [0,T)$. $\square$

Corollary 8 *If $v(n)$ is nonincreasing and $\lambda(n)\Upsilon(n,\cdot,0) > \lambda(n)p(n) + c(n) + \alpha v(n)$ for an $n > 0$, then $V^*(n,t) > v(n)$, and it is optimal to continue (i.e., $I^*(n,t) = 1$) for all $t \in [0,T)$.*

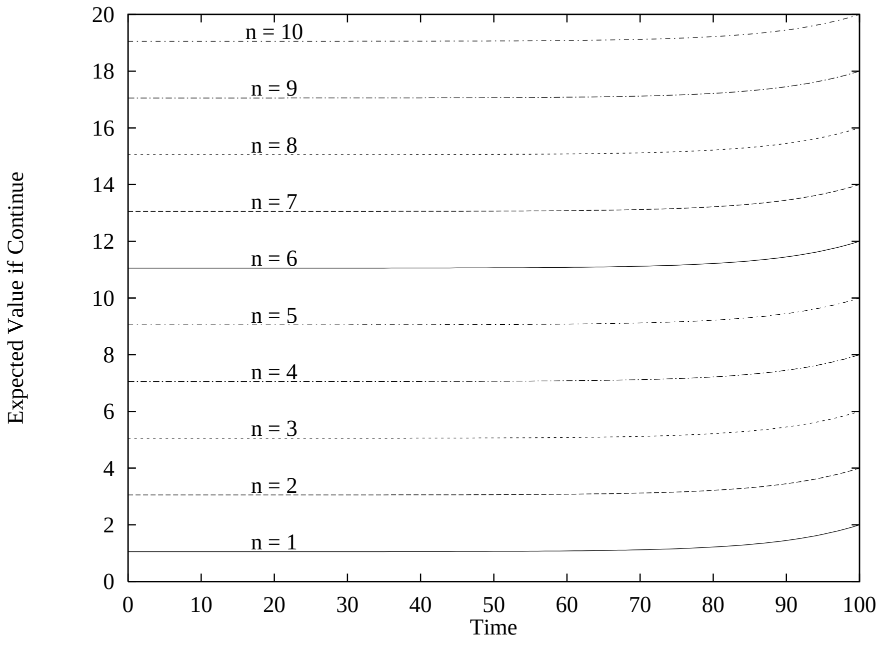


Figure 6.9 Example to show that Corollary 8 does not hold if v is an arbitrary increasing function of n

It is shown by way of a counterexample that the assumption that $v(n)$ is nonincreasing in Proposition 5 cannot be dropped. Figure 6.9 shows $V^{\psi(n)}(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are exponentially distributed with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda(n) = 0.1$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = 0.9$, the terminal value $v(n) = 2n$, and the discount rate $\alpha = 0$. Note that for all $n > 0$, $V^{\psi(n)}(n,t) < v(n)$ for all $t < T$, although $\lambda(n)\Upsilon(n,t,0) = \lambda(n)/\mu = 1.0 > \lambda(n)p(n) + c(n) + \alpha v(n) = 0.9$ for all $t \in [0,T]$. Therefore, Proposition 5 does not hold if v is an arbitrary increasing function of n. For the same example, for all n, $V^*(n,t) = v(n)$ for all $t \in [0,T]$, although $\lambda(n)\Upsilon(n,t,0) > \lambda(n)p(n) + c(n) + \alpha v(n)$ for all $t \in [0,T]$. Therefore, Corollary 8 also does not hold if v is an arbitrary increasing function of n.

6.6.2.3 Optimality Results for Recursive Policies

By definition, $V^*(n,t) \geq \max\{v(n), V^{\psi(n)}(n,t)\}$ for all n and t. As noted before, it is typical in applications for $c(n)$ to be nonincreasing. In most cases p does not depend on n. It is also common that v does not vary much with n, for example, the dispatching cost of a vehicle or a batch processor does not depend very much on the number of items. Also, in many applications the item arrival rate λ does not depend on n. It is shown in Theorem 8 that if $p(n)$ and $c(n)$ are nonincreasing and λ and v are constant, then similar to the infinite horizon case, $V^*(n,t) = \max\{v, V^{\psi(n)}(n,t)\}$ for all n and t.

Theorem 8 *If $p(n)$ and $c(n)$ are nonincreasing and λ and v are constant, then $V^*(n,t) = \max\{v, V^{\psi(n)}(n,t)\}$ for all n and t.*

Proof: By induction on n. For $n = 0$, if $-\lambda p(0) - c(0) \geq \alpha v$, then $I^*(0,t) = 1$ for all $t \in (0,T]$. Therefore, $V^*(0,t) = V^{\psi(0)}(0,t) \geq v$ for all $t \in [0,T]$. Else, if $-\lambda p(0) - c(0) < \alpha v$, then $I^*(0,t) = 0$ for all $t \in (0,T]$. Then $V^*(0,t) = v \geq V^{\psi(0)}(0,t)$ for all $t \in [0,T]$.

As induction hypothesis on n, suppose $V^*(n-s,t) = \max\{v, V^{\psi(n-s)}(n-s,t)\}$ for all $s \in \{1,\ldots,n\}$ and all $t \in [0,T]$. For $n > 0$, consider the following two cases.

Case 1: $\lambda\Upsilon(n,\cdot,0) \leq \lambda p(n) + c(n) + \alpha v$:

Then, from Proposition 4, $V^*(n,t) = v$ for all $t \in [0,T]$, and because $V^{\psi(n)}(n,t) \leq V^*(n,t)$, $V^*(n,t) = \max\{v, V^{\psi(n)}(n,t)\}$ for all $t \in [0,T]$.

Case 2: $\lambda\Upsilon(n,\cdot,0) > \lambda p(n) + c(n) + \alpha v$:

Then, from Corollary 8, $V^*(n,t) > v$ and $I^*(n,t) = 1$ for all $t \in [0,T)$. Then, from

Equation (6.3), $V^*(n,\cdot)$ is the unique solution of

$$\begin{aligned}\frac{\partial V^*}{\partial t}(n,t) &= -\lambda \int_0^n \int_{V^*(n,t)-V^*(n-s,t)-p(n)}^{\infty} \{r - [V^*(n,t) \\ &\qquad\qquad -V^*(n-s,t) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\ &\quad + \alpha V^*(n,t) \ + \ \lambda p(n) \ + \ c(n)\end{aligned}$$

with boundary condition $V^*(n,T) = v(n)$, that is absolutely continuous on $[0,T]$. From Equation (6.18), $V^{\psi(n)}(n,\cdot)$ is the unique solution of

$$\begin{aligned}\frac{\partial V^{\psi(n)}}{\partial t}(n,t) &= -\lambda \int_0^n \int_{V^{\psi(n)}(n,t)-\max\{v,V^{\psi(n-s)}(n-s,t)\}-p(n)}^{\infty} \Big\{r - \Big[V^{\psi(n)}(n,t) \\ &\qquad\qquad - \max\{v, V^{\psi(n-s)}(n-s,t)\} - p(n)\Big]\Big\}\ F_{R|S}(dr|s)F_S(ds) \\ &\quad + \alpha V^{\psi(n)}(n,t) \ + \ \lambda p(n) \ + \ c(n) \\ &= -\lambda \int_0^n \int_{V^{\psi(n)}(n,t)-V^*(n-s,t)-p(n)}^{\infty} \Big\{r - \Big[V^{\psi(n)}(n,t) \\ &\qquad\qquad -V^*(n-s,t) - p(n)]\}\ F_{R|S}(dr|s)F_S(ds) \\ &\quad + \alpha V^{\psi(n)}(n,t) \ + \ \lambda p(n) \ + \ c(n)\end{aligned}$$

with boundary condition $V^{\psi(n)}(n,T) = v(n)$, that is absolutely continuous on $[0,T]$. Hence, $V^*(n,\cdot)$ and $V^{\psi(n)}(n,\cdot)$ satisfy the same differential equation with the same boundary condition. Therefore, $V^*(n,t) = V^{\psi(n)}(n,t) \geq v$, and $V^*(n,t) = \max\{v, V^{\psi(n)}(n,t)\}$ for all $t \in [0,T]$. □

If the conditions of Theorem 8 hold, then an optimal policy π^* has the following convenient form. If $-\lambda p(0) - c(0) > \alpha v$, then $\lambda\Upsilon(n,\cdot,0) > \lambda p(n) + c(n) + \alpha v$ for all $n \in \{0,1,\ldots,N_0\}$, because $\lambda\Upsilon(n,\cdot,0) \geq 0 > \lambda p(0) + c(0) + \alpha v \geq \lambda p(n) + c(n) + \alpha v$. Then from the proof of Theorem 8, $V^*(n,t) = V^{\psi(n)}(n,t)$ and $I^*(n,t) = 1$ for all n and t. Else, if $-\lambda p(0) - c(0) \leq \alpha v$, then let $n^* = \max\{n \in \{0,1,\ldots,N_0\} : \lambda\Upsilon(n,\cdot,0) \leq \lambda p(n) + c(n) + \alpha v\}$. Then for all $n \leq n^*$, $\lambda\Upsilon(n,\cdot,0) \leq \lambda\Upsilon(n^*,\cdot,0) \leq \lambda p(n^*) + c(n^*) + \alpha v \leq \lambda p(n) + c(n) + \alpha v$. Then $V^*(n,t) = v$ and $I^*(n,t) = 0$ for all $n \leq n^*$ and all $t \in (0,T]$. Also, for all $n > n^*$, $\lambda\Upsilon(n,\cdot,0) > \lambda p(n) + c(n) + \alpha v$. Thus $V^*(n,t) = V^{\psi(n)}(n,t)$ and $I^*(n,t) = 1$ for all $n > n^*$ and all $t \in (0,T]$. Hence, as long as $t < T$ and $N^*(t) > n^*$, it is optimal to continue, using threshold $x^*(n,t,s) = x^{\psi(n)}(n,t,s) = V^{\psi(n)}(n,t) - \max\{v, V^{\psi(n-s)}(n-s,t)\} - p(n) = V^*(n,t) - V^*(n-s,t) - p(n)$. It is optimal to stop and collect v as soon as $N^*(t)$ becomes less than or equal to n^*. It is interesting that n^* does not depend on t, i.e., it is optimal to continue if $N^*(t) > n^*$ for all $t < T$, and it is optimal to stop if $N^*(t) \leq n^*$ for all t. In other words, closeness to the deadline does not affect the optimal stopping decision. This result characterizes an optimal policy and the optimal expected value in a simple, intuitive way. It also leads to an easy method for computing the optimal expected value V^* and optimal threshold x^*, by simply computing $V^{\psi(n)}(n,t)$ for $n = 0, 1, \ldots, N_0$.

Counterexamples where Theorem 8 Does Not Hold

It is shown by way of counterexamples that the assumptions made in Theorem 8 cannot be dropped. Consider the assumption that λ is constant. Figure 6.10 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are exponentially distributed with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda(n) = 0.1 - 0.01n$ for $n \in \{0, 1, \ldots, 10\}$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = 0.6$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that for $n = 5$, $V^*(5,t) > \max\{v, V^{\psi(5)}(5,t)\}$ for all $t < 73$. Therefore, Theorem 8 does not hold if λ is an arbitrary decreasing function of n.

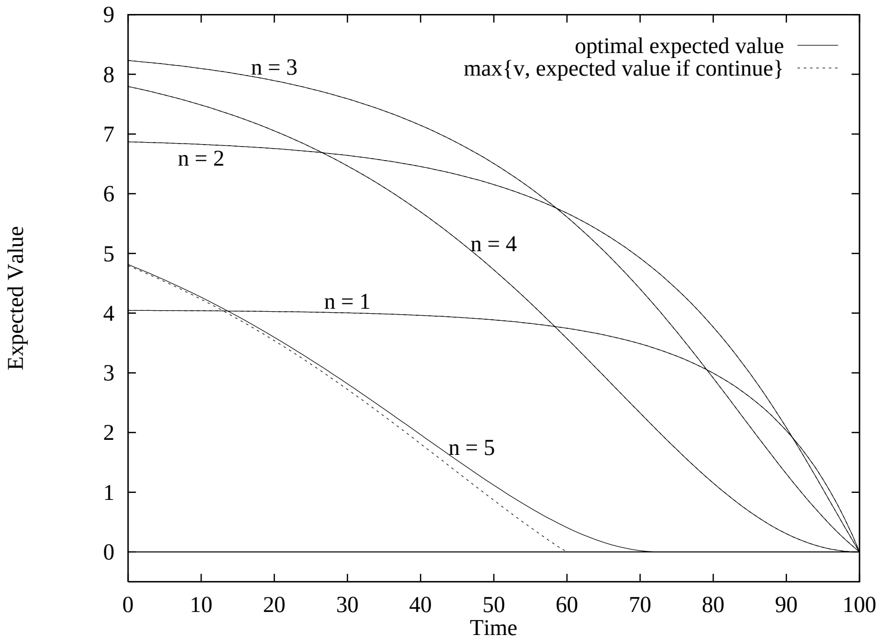


Figure 6.10 Example to show that Theorem 8 does not hold if λ is an arbitrary decreasing function of n

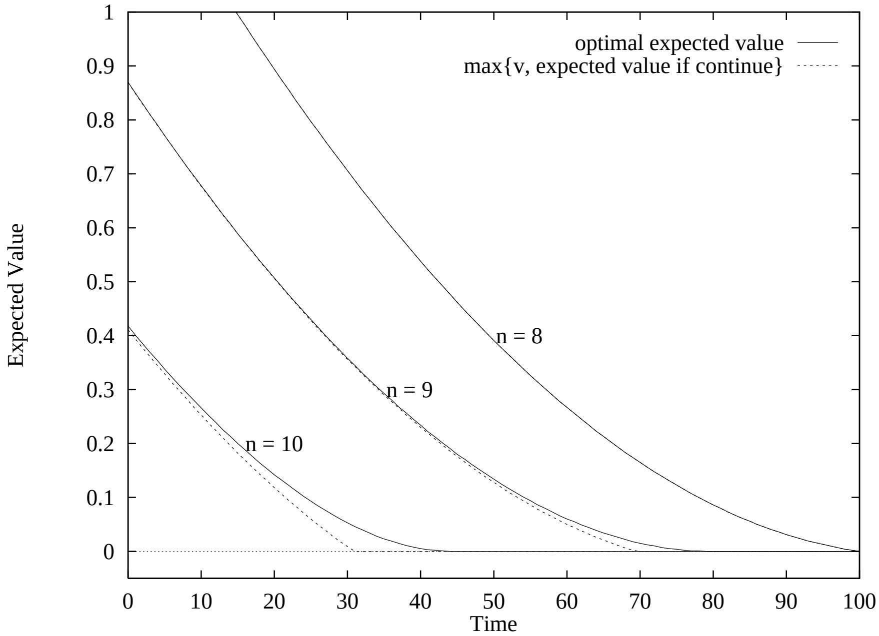


Figure 6.11 Example to show that Theorem 8 does not hold if λ is an arbitrary increasing function of n

Figure 6.11 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are normally distributed with mean $\mu = -1$ and standard deviation $\sigma = 1$, the deadline $T = 100$, items arrive at rate $\lambda(n) = 0.01n$, the penalty $p(n) = 1$, the holding cost per unit time $c(n) = -0.05$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that for $n = 9$, $V^*(9,t) > \max\{v, V^{\psi(9)}(9,t)\}$ for all $t < 79$, and for $n = 10$, $V^*(10,t) > \max\{v, V^{\psi(10)}(10,t)\}$ for all $t < 44$, Therefore, Theorem 8 does not hold if λ is an arbitrary increasing function of n.

Next consider the assumption that $p(n)$ is nonincreasing. Figure 6.12 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are normally distributed with mean $\mu = 0$ and standard deviation $\sigma = 5$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 2.5n$, the holding cost per unit time $c(n) = 0.1$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that for $n = 5$, $V^*(5,t) > \max\{v, V^{\psi(5)}(5,t)\}$ for all $t < 88$. Similar results hold for $n \in \{6, \ldots, 10\}$. Therefore, Theorem 8 does not hold if p is an arbitrary increasing function of n.

Next consider the assumption that $c(n)$ is nonincreasing. Figure 6.13 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are

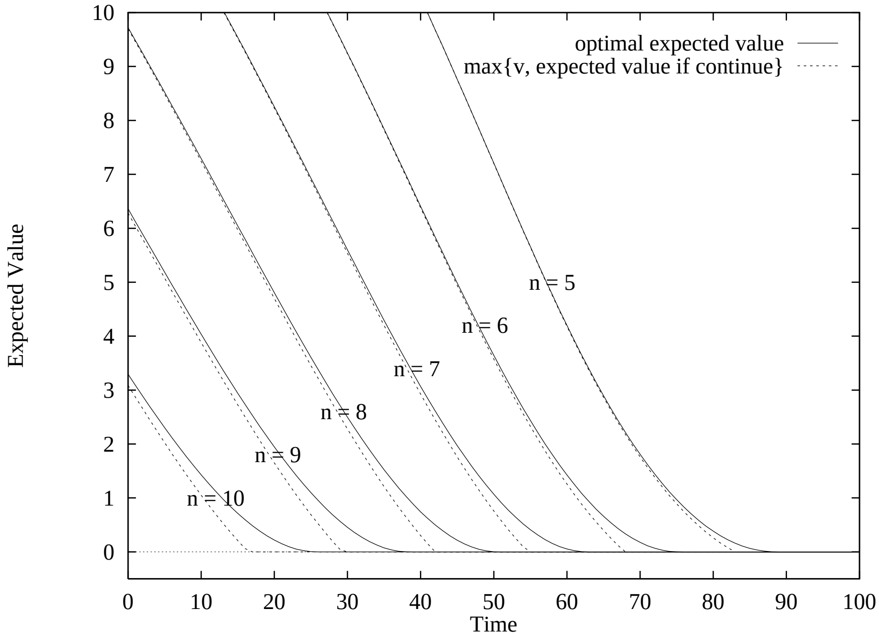


Figure 6.12 Example to show that Theorem 8 does not hold if p is an arbitrary increasing function of n

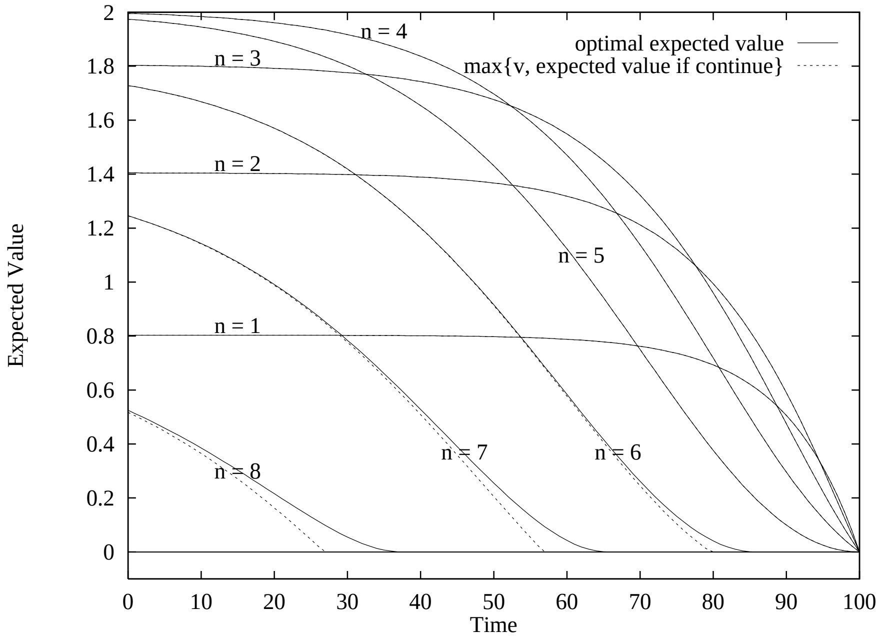


Figure 6.13 Example to show that Theorem 8 does not hold if c is an arbitrary increasing function of n

exponentially distributed with mean $1/\mu = 100$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = 9.9{+}0.02n$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that for $n = 6$, $V^*(6,t) > \max\{v, V^{\psi(6)}(6,t)\}$ for all $t < 86$. Similar results hold for $n = 7$ and $n = 8$. Therefore, Theorem 8 does not hold if c is an arbitrary increasing function of n.

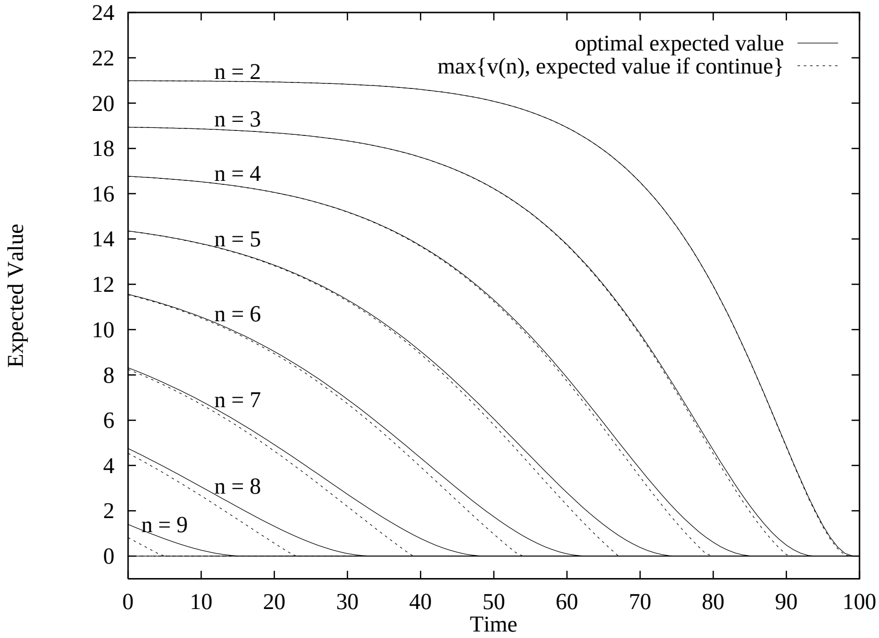


Figure 6.14 Example to show that Theorem 8 does not hold if v is an arbitrary decreasing function of n

Next consider the assumption that v is constant. Figure 6.14 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are exponentially distributed with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = 1.2$, the terminal value $v(0) = 25$; $v(n) = 0$ if $n > 0$, and the discount rate $\alpha = 0$. Note that for $n = 9$, $V^*(9,t) > \max\{v(9), V^{\psi(9)}(9,t)\}$ for all $t < 15$. Similar results hold for $n \in \{2,\ldots,8\}$. Therefore, Theorem 8 does not hold if v is an arbitrary decreasing function of n.

Figure 6.15 shows $V^*(n,t)$ as a function of t for different n, for an example where all items have size 1, rewards are exponentially distributed with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 0$, the holding cost per unit time $c(n) = 0.5$, the terminal value $v(n) = 25$ if $n > 7$; $v(n) = 0$ if $n \leq 7$, and the discount rate $\alpha = 0$. Note that for $n = 8$, $V^*(8,t) > \max\{v(8), V^{\psi(8)}(8,t)\}$ for all $t < 63$. Similar

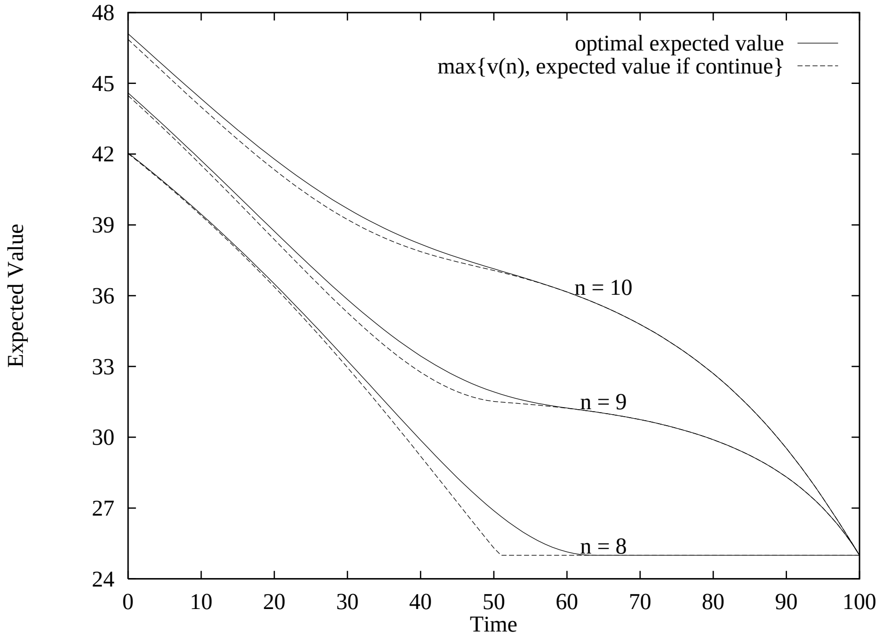


Figure 6.15 Example to show that Theorem 8 does not hold if v is an arbitrary increasing function of n

results hold for $n = 9$ and $n = 10$. Therefore, Theorem 8 does not hold if v is an arbitrary increasing function of n.

6.6.3 Monotonicity and Convexity Properties

It was shown in Proposition 2 that $V^*(n,t)$ is nonincreasing in t. In Proposition 3 it was shown that if λ is constant, $p(n)$ and $c(n)$ are nonincreasing, and $v(n)$ is nondecreasing, then $V^*(n,t)$ is nondecreasing in n. It is intuitive to expect that an optimal policy becomes more lenient in accepting items as the deadline approaches, in other words the optimal threshold is decreasing in time for any given remaining capacity and size of item. This property implies that the optimal expected value is concave in time. It is also intuitive to expect that an optimal policy becomes more selective in accepting items as the remaining capacity decreases, in other words the optimal threshold is decreasing in the remaining capacity for any given time and size of item. This property in turn is equivalent to the optimal expected value being concave in the remaining capacity.

Conditions under which these properties hold are established in this section. Examples are also given of cases where these properties do not hold.

6.6.3.1 General Results

A number of monotonicity and convexity properties for V^* and x^* that hold in general are derived in this section. In Section 6.6.3.2 further results are derived for the special case with all items having the same size. Thereafter the results are extended to the case where items have random sizes.

Lemma 25 *If p is constant, then $V^*(n,t)$ is concave in n if and only if $x^*(n,t,s)$ is nonincreasing in n for any $s \in [0,n]$.*

Proof: $V^*(n,t)$ is concave in n if and only if for each $n_1 \leq n_2$ and any $s \in [0, n_1]$,

$$\begin{aligned} V^*(n_1,t) - V^*(n_1-s,t) &\geq V^*(n_2,t) - V^*(n_2-s,t) \\ \Leftrightarrow \quad V^*(n_1,t) - V^*(n_1-s,t) - p &\geq V^*(n_2,t) - V^*(n_2-s,t) - p \\ \Leftrightarrow \quad x^*(n_1,t,s) &\geq x^*(n_2,t,s) \qquad \square \end{aligned}$$

Lemma 26 *If λ is constant, $p(n)$ and $c(n)$ are nonincreasing, $\alpha = 0$, and $x^*(n,t,s)$ is nonincreasing in n for all $s \in [0,n]$, then $\partial V^*(n,t)/\partial t$ is nonincreasing in n.*

Proof: $\partial V^*(n,t)/\partial t = \min\{-\lambda\Upsilon(n,t,V^*) + \lambda p(n) + c(n), 0\}$. For any $n_1 \leq n_2$, if $\partial V^*(n_1,t)/\partial t = 0$, then $\partial V^*(n_2,t)/\partial t \leq 0 = \partial V^*(n_1,t)/\partial t$. If $\partial V^*(n_1,t)/\partial t = -\lambda\Upsilon(n_1,t,V^*) + \lambda p(n_1) + c(n_1)$, then

$$\begin{aligned} \frac{\partial V^*}{\partial t}(n_2,t) &\leq -\lambda \int_0^{n_2}\int_{x^*(n_2,t,s)}^{\infty} [r - x^*(n_2,t,s)]\, F_{R|S}(dr|s)F_S(ds) \;+\; \lambda p(n_2) \;+\; c(n_2) \\ &\leq -\lambda \int_0^{n_1}\int_{x^*(n_1,t,s)}^{\infty} [r - x^*(n_2,t,s)]\, F_{R|S}(dr|s)F_S(ds) \\ &\quad - \lambda \int_{n_1}^{n_2}\int_{x^*(n_2,t,s)}^{\infty} [r - x^*(n_2,t,s)]\, F_{R|S}(dr|s)F_S(ds) \;+\; \lambda p(n_1) \;+\; c(n_1) \\ &\leq -\lambda \int_0^{n_1}\int_{x^*(n_1,t,s)}^{\infty} [r - x^*(n_1,t,s)]\, F_{R|S}(dr|s)F_S(ds) \;+\; \lambda p(n_1) \;+\; c(n_1) \\ &= \frac{\partial V^*}{\partial t}(n_1,t) \qquad \square \end{aligned}$$

Lemma 27 *$\partial V^*(n,t)/\partial t$ is nonincreasing in n if and only if $x^*(n,t,s)$ is nonincreasing in t for any $s \in [0,n]$.*

Proof: $\partial V^*(n,t)/\partial t$ is nonincreasing in n if and only if for each $n \in [0, N_0]$ and any $s \in [0, n]$,

$$\begin{aligned} \frac{\partial V^*(n,t)}{\partial t} &\leq \frac{\partial V^*(n-s,t)}{\partial t} \\ \Leftrightarrow \quad \frac{\partial x^*(n,t,s)}{\partial t} &= \frac{\partial\left[V^*(n,t) - V^*(n-s,t) - p(n)\right]}{\partial t} \leq 0 \end{aligned}$$ □

Lemma 28 *If it is optimal to continue (i.e., $I^*(n,t) = 1$) for all $t \in (0,T)$, and $x^*(n,t,s)$ is nonincreasing in t for all $s \in [0,n]$, then $V^*(n,t)$ is concave in t on $[0,T]$.*

Proof: For any $t_1 \leq t_2$, $x^*(n,t_1,s) \geq x^*(n,t_2,s)$ for all $s \in [0,n]$, and $V^*(n,t_1) \geq V^*(n,t_2)$. Then, because $I^*(n,t) = 1$,

$$\begin{aligned} \frac{\partial V^*}{\partial t}(n,t_1) &= -\lambda(n) \int_0^n \int_{x^*(n,t_1,s)}^{\infty} \left[r - x^*(n,t_1,s)\right] F_{R|S}(dr|s) F_S(ds) \\ &\quad + \alpha V^*(n,t_1) \; + \; \lambda(n)p(n) \; + \; c(n) \\ &\geq -\lambda(n) \int_0^n \int_{x^*(n,t_1,s)}^{\infty} \left[r - x^*(n,t_2,s)\right] F_{R|S}(dr|s) F_S(ds) \\ &\quad + \alpha V^*(n,t_2) \; + \; \lambda(n)p(n) \; + \; c(n) \\ &\geq -\lambda(n) \int_0^n \int_{x^*(n,t_2,s)}^{\infty} \left[r - x^*(n,t_2,s)\right] F_{R|S}(dr|s) F_S(ds) \\ &\quad + \alpha V^*(n,t_2) \; + \; \lambda(n)p(n) \; + \; c(n) \\ &= \frac{\partial V^*}{\partial t}(n,t_2) \end{aligned}$$

Thus, $\partial V^*(n,t)/\partial t$ is nonincreasing in t. Because $V^*(n,t)$ is continuous in t, $V^*(n,t)$ is concave in t if and only if $\partial V^*(n,t)/\partial t$ is nonincreasing in t, from Royden [153] Propositions 5.17 and 5.18. □

6.6.3.2 The Case With All Items Having The Same Size

The important special case where all items have the same size is studied in this section. The case where items have random sizes is discussed in Section 6.6.3.3. In transportation applications, this case occurs for example when ships are loaded with standard sized containers, when airline tickets for a specific flight are sold (when there are different types of seats, for example first class and coach class seats, these can be modeled as separate knapsacks, because the number of each class of seats cannot be changed with short notice), or when rail freight is collected in units of car loads. This case also applies when standard sized jobs are accepted for a batch processor, and when an inventory of similar assets are sold, for example when tickets are sold for a performance or a sports event. The case where all items have the same size exhibits special structure that does not hold in general if items have random sizes.

Additional Results

Without loss of generality, let all items have unit size. In this case an optimal threshold $x^* : \{1, \ldots, N_0\} \times (0, T] \mapsto \Re$ is given by $x^*(n,t) = V^*(n,t) - V^*(n-1,t) - p(n)$.

Let the function $\Upsilon_1 : \Re \mapsto \Re$ be defined by $\Upsilon_1(x) \equiv \int_x^\infty (r-x)\, F_R(dr)$. Note that $\Upsilon_1 \geq 0$. Also, $\Upsilon(n,t,V^*) = \Upsilon_1(x^*(n,t))$, and $\Upsilon(n,\cdot,0) = \Upsilon_1(-p(n))$ for all $n \in \{1, \ldots, N_0\}$.

Lemma 29

1. Υ_1 *satisfies the following Lipschitz condition:*

$$|\Upsilon_1(x_1) - \Upsilon_1(x_2)| \quad \leq \quad |x_2 - x_1| \quad \textit{for all} \;\; x_1, x_2 \in \Re$$

2. Υ_1 *is absolutely continuous on* $\Re$.

3. Υ_1 *is nonincreasing on* $\Re$.

The proof of Lemma 29 is given in Appendix K.

In Lemma 26 it was shown that if λ is constant, $p(n)$ and $c(n)$ are nonincreasing, $\alpha = 0$, and $x^*(n,t)$ is nonincreasing in n, then $\partial V^*(n,t)/\partial t$ is nonincreasing in n. Lemma 30 shows that the converse holds under modified conditions.

Lemma 30 *If λ is constant, $p(n)$ and $c(n)$ are nondecreasing, $\alpha = 0$, it is optimal to continue (i.e., $I^*(n,t) = 1$) for all $n > 0$, and $\partial V^*(n,t)/\partial t$ is nonincreasing in n, then $x^*(n,t)$ is nonincreasing in n.*

Proof: For any $0 < n_1 \leq n_2$, since $\alpha = 0$, it is optimal to continue, and Υ_1 is nonincreasing from Lemma 29,

$$\begin{aligned} \frac{\partial V^*}{\partial t}(n_1,t) \quad &\geq \quad \frac{\partial V^*}{\partial t}(n_2,t) \\ \Leftrightarrow \quad -\lambda\Upsilon_1(x^*(n_1,t)) \; + \; \lambda p(n_1) \; + \; c(n_1) \quad &\geq \quad -\lambda\Upsilon_1(x^*(n_2,t)) \; + \; \lambda p(n_2) \; + \; c(n_2) \\ \Rightarrow \quad x^*(n_1,t) \quad &\geq \quad x^*(n_2,t) \end{aligned}$$

□

In Lemma 28 it was shown that if it is optimal to continue (i.e., $I^*(n,t) = 1$) for all $t \in (0,T)$, and $x^*(n,t)$ is nonincreasing in t, then $V^*(n,t)$ is concave in t on $[0,T]$. Lemma 31 shows that the converse holds if $\alpha = 0$.

Lemma 31 *If $\alpha = 0$, it is optimal to continue (i.e., $I^*(n,t) = 1$) for all $t \in (0,T)$, and $V^*(n,t)$ is concave in t, then $x^*(n,t)$ is nonincreasing in t.*

Proof: Because $V^*(n,t)$ is continuous in t, $V^*(n,t)$ is concave in t if and only if $\partial V^*(n,t)/\partial t$ is nonincreasing in t. For any $t_1 \le t_2$, because $I^*(n,t) = 1$ and $\alpha = 0$,

$$\begin{aligned} \frac{\partial V^*}{\partial t}(n,t_1) &\geq \frac{\partial V^*}{\partial t}(n,t_2) \\ \Leftrightarrow \quad -\lambda(n)\Upsilon_1 x^*(n,t_1) + \lambda(n)p(n) + c(n) &\geq -\lambda(n)\Upsilon_1 x^*(n,t_2) + \lambda(n)p(n) + c(n) \\ \Leftrightarrow \quad x^*(n,t_1) &\geq x^*(n,t_2) \end{aligned}$$

The last result follows because Υ_1 is nonincreasing from Lemma 29. $\square$

Equivalent Intuitive Properties

Now we are ready to give the conditions under which the intuitive monotonicity and convexity properties hold. This is the subject of Theorem 9. Thereafter some examples are given of cases where these properties do not hold.

Theorem 9 *If λ, p, and c are constant, $v(n)$ is concave nonincreasing, $\alpha = 0$, and $\lambda\Upsilon_1(-p) \ge \lambda p + c$, then the following conditions hold.*

(i) $V^*(n+1,t) - V^*(n,t) \le V^*(n,t) - V^*(n-1,t)$ *for all* $n \in \{1,\ldots,N_0 - 1\}$ *and all* $t \in [0,T]$ (the optimal expected value is concave in remaining capacity).

(ii) $x^*(n+1,t) \le x^*(n,t)$ *for all* $n \in \{1,\ldots,N_0 - 1\}$ *and all* $t \in [0,T]$ (the optimal threshold is nonincreasing in remaining capacity).

(iii) $\partial V^*(n,t)/\partial t \le \partial V^*(n-1,t)/\partial t$ *for all* $n \in \{1,\ldots,N_0\}$ *and all* $t \in (0,T)$ (the marginal optimal expected value of remaining time $-\partial V^*(n,t)/\partial t$ is nondecreasing in remaining capacity).

(iv) $\partial x^*(n,t)/\partial t \le 0$ *for all* $n \in \{1,\ldots,N_0\}$ *and all* $t \in (0,T)$ (the optimal threshold is nonincreasing in time).

(v) $\partial V^*(n,t_2)/\partial t \leq \partial V^*(n,t_1)/\partial t$ *for all* $n \in \{0,\ldots,N_0\}$ *and all* $0 < t_1 \leq t_2 < T$ ($\partial V^*(n,t)/\partial t$ is nonincreasing in time, or the optimal expected value is concave in time).

Proof: From Corollary 8, because $v(n)$ is nonincreasing and $\lambda\Upsilon(\cdot,\cdot,0) = \lambda\Upsilon_1(-p) \geq \lambda p + c$, it follows that $I^*(n,t) = 1$ and $\partial V^*(n,t)/\partial t = -\lambda\Upsilon(n,t,V^*) + \lambda p + c = -\lambda\Upsilon_1(x^*(n,t)) + \lambda p + c$ for all $n \in \{1,\ldots,N_0\}$ and all $t \in (0,T)$. First it is shown that all the conditions are equivalent, and then it will be shown that (iii) and (v) hold.

(i) $\Leftrightarrow$ (ii): From Lemma 25.

(ii) $\Leftrightarrow$ (iii): From Lemmas 26 and 30.

(iii) $\Leftrightarrow$ (iv): From Lemma 27.

(iv) $\Leftrightarrow$ (v): From Lemmas 28 and 31.

For $n > 0$, $\partial V^*(n,t)/\partial t = -\lambda\Upsilon_1(V^*(n,t) - V^*(n-1,t) - p) + \lambda p + c$ for all $t \in (0,T)$. From the continuity of $V^*(n,t)$ in t for all n, $V^*(n,t) \to V^*(n,T) = v(n)$ as $t \to T$. Hence, from the continuity of Υ_1, $\partial V^*(n,t)/\partial t \to -\lambda\Upsilon_1(v(n) - v(n-1) - p) + \lambda p + c$ as $t \to T$ for all $n > 0$.

It is shown by induction on n that (iii) and (v) hold. For $n = 0$, if $\lambda p + c \leq 0$, then $\partial V^*(0,t)/\partial t = \lambda p + c$ for all $t \in (0,T)$. Else, if $\lambda p + c > 0$, then $\partial V^*(0,t)/\partial t = 0$ for all $t \in (0,T)$. Hence, $\partial V^*(0,t)/\partial t = \min\{0, \lambda p + c\}$ for all $t \in (0,T)$. For $n = 1$, it is shown by contradiction that (iii) holds. Suppose there exists a $t_1 \in (0,T)$ such that $\partial V^*(1,t_1)/\partial t > \partial V^*(0,t_1)/\partial t$. From the continuity of $\partial V^*(n,t)/\partial t$ in t, there exists a neighborhood $(t_0,t_2) \subseteq (0,T)$ of t_1 such that $\partial V^*(1,t)/\partial t > \partial V^*(0,t)/\partial t$ for all $t \in (t_0,t_2)$. Then for all $t \in (t_1,t_2]$

$$\begin{aligned}
\int_{t_1}^{t} \frac{\partial V^*}{\partial \tau}(1,\tau)\,d\tau &> \int_{t_1}^{t} \frac{\partial V^*}{\partial \tau}(0,\tau)\,d\tau \\
\Rightarrow \quad V^*(1,t) - V^*(1,t_1) &> V^*(0,t) - V^*(0,t_1) \\
\Rightarrow \quad V^*(1,t) - V^*(0,t) &> V^*(1,t_1) - V^*(0,t_1) \\
\Rightarrow \quad -\lambda\Upsilon_1(V^*(1,t) - V^*(0,t) - p) + \lambda p + c &\geq -\lambda\Upsilon_1(V^*(1,t_1) - V^*(0,t_1) - p) + \lambda p + c \\
\Rightarrow \quad \frac{\partial V^*}{\partial t}(1,t) &\geq \frac{\partial V^*}{\partial t}(1,t_1)
\end{aligned}$$

By repeating the argument at t_2, it follows that $\partial V^*(1,t)/\partial t$ is nondecreasing on $[t_1,T)$. But $\partial V^*(1,t)/\partial t \to -\lambda\Upsilon_1(v(1) - v(0) - p) + \lambda p + c$ as $t \to T$, and $\Upsilon_1 \geq 0 \Rightarrow -\lambda\Upsilon_1(v(1) - v(0) -$

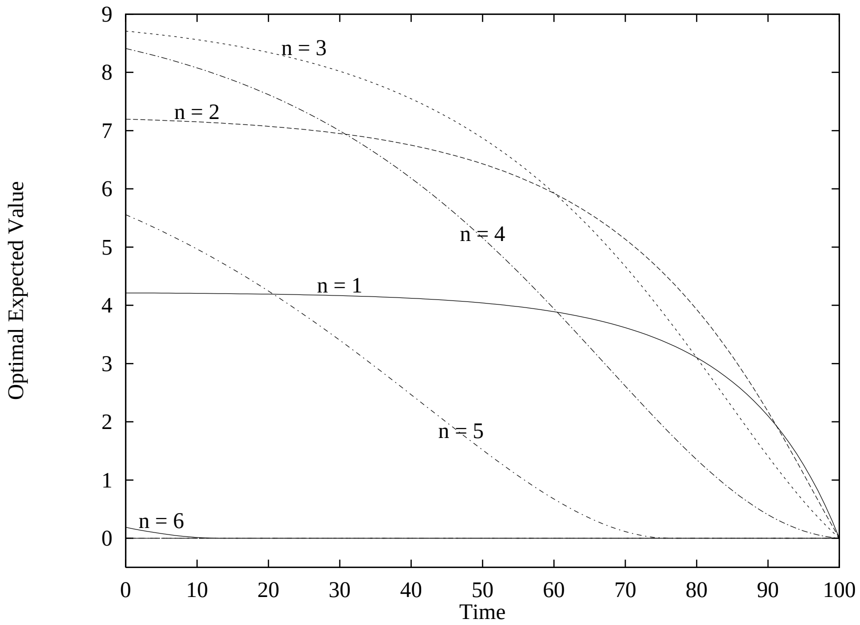
$p) + \lambda p + c \leq \lambda p + c$. Also, $v(1) \leq v(0)$ because $v(n)$ is nonincreasing, and thus from the fact that Υ_1 is nonincreasing, $-\lambda\Upsilon_1(v(1) - v(0) - p) + \lambda p + c \leq -\lambda\Upsilon_1(-p) + \lambda p + c \leq 0$, which follows from the assumptions. Hence $\lim_{t\to T} \partial V^*(1,t)/\partial t \leq \min\{0, \lambda p + c\} = \partial V^*(0,t)/\partial t$, which contradicts $\partial V^*(1,t_1)/\partial t > \partial V^*(0,t_1)/\partial t$, $\partial V^*(1,t)/\partial t$ nondecreasing on $[t_1, T)$, and $\partial V^*(0,t)/\partial t$ constant on $(0,T)$. Therefore, $\partial V^*(1,t)/\partial t \leq \partial V^*(0,t)/\partial t$, $\partial x^*(1,t)/\partial t \leq 0$, and $\partial V^*(1,t)/\partial t$ is nonincreasing on $(0,T)$.

As induction hypothesis on n, suppose that $\partial V^*(n,t)/\partial t$ is nonincreasing on $(0,T)$. Similar to the case for $n = 1$, it is shown by contradiction that (iii) holds. Suppose there exists a $t_1 \in (0,T)$ such that $\partial V^*(n+1,t_1)/\partial t > \partial V^*(n,t_1)/\partial t$. In the same way as for $n = 1$, it follows that $\partial V^*(n+1,t)/\partial t$ is nondecreasing on $[t_1, T)$. But $v(n+1) - v(n) \leq v(n) - v(n-1)$ because $v(n)$ is concave, and thus from the fact that Υ_1 is nonincreasing, $\lim_{t\to T} \partial V^*(n+1,t)/\partial t = -\lambda\Upsilon_1(v(n+1) - v(n) - p) + \lambda p + c \leq -\lambda\Upsilon_1(v(n) - v(n-1) - p) + \lambda p + c = \lim_{t\to T} \partial V^*(n,t)/\partial t$, which contradicts $\partial V^*(n+1,t_1)/\partial t > \partial V^*(n,t_1)/\partial t$, $\partial V^*(n+1,t)/\partial t$ nondecreasing on $[t_1,T)$, and $\partial V^*(n,t)/\partial t$ nonincreasing on $(0,T)$. Therefore, $\partial V^*(n+1,t)/\partial t \leq \partial V^*(n,t)/\partial t$, $\partial x^*(n+1,t)/\partial t \leq 0$, and $\partial V^*(n+1,t)/\partial t$ is nonincreasing on $(0,T)$. □

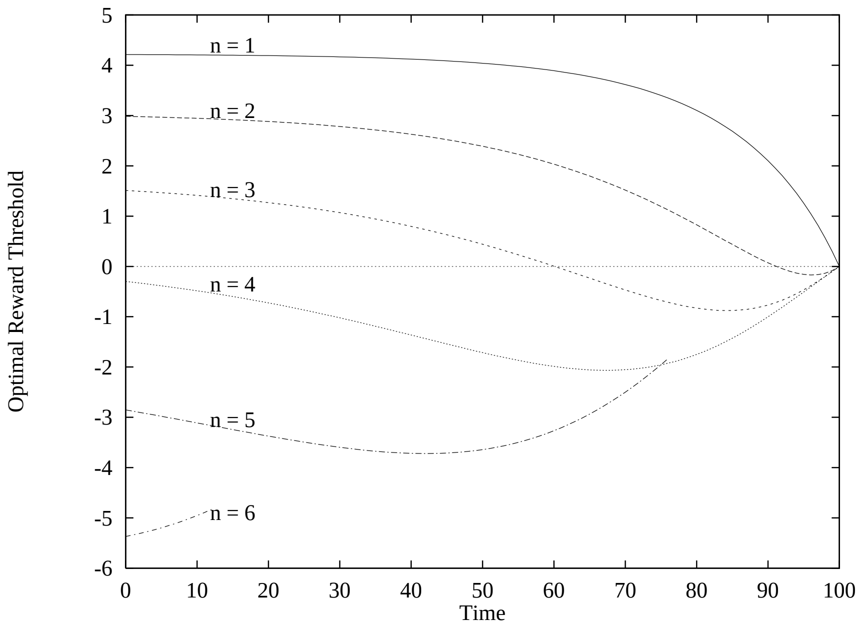
Counterexamples where the Monotonicity and Convexity Properties Do Not Hold

It is shown by way of counterexamples that the assumptions made in Theorem 9 cannot be dropped. Consider the assumption that λ is constant. Figure 6.16 shows $V^*(n,t)$ and $x^*(n,t)$ as functions of t for different n, for an example with exponentially distributed rewards with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda(n) = 0.1 - 0.01n$ for $n \in \{0, 1, \ldots, 10\}$, the penalty $p = 0$, the holding cost per unit time $c = 0.59$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that for $n = 4$, $V^*(4,t)$ is not concave in t on $[0,T]$, and $x^*(4,t)$ is increasing in t for $t > 67$, although $\lambda(4)\Upsilon_1(-p) = 0.06/\mu = 0.6 > \lambda(4)p + c = 0.59$. Therefore, Theorem 9 does not hold if λ is an arbitrary decreasing function of n.

Figure 6.17 shows $V^*(n,t)$ and $x^*(n,t)$ as functions of t for different n, for an example with normally distributed rewards with mean $\mu = -1$ and standard deviation $\sigma = 1$, the deadline $T = 100$, items arrive at rate $\lambda(n) = 0.01n$, the penalty $p = 1$, the holding

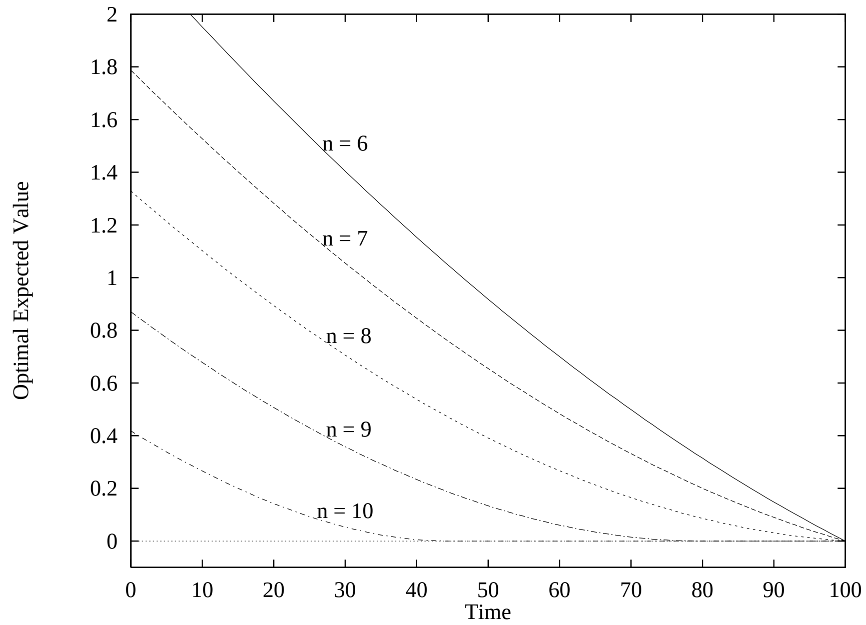


(a) Optimal Expected Value versus Time for Different Remaining Capacities n

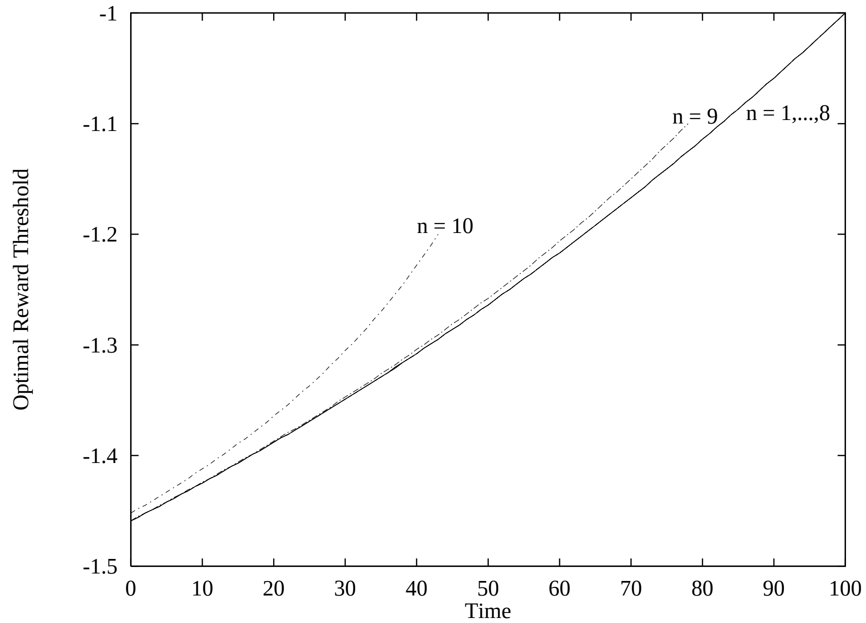


(b) Optimal Threshold versus Time for Different Remaining Capacities n

Figure 6.16 Example to show that Theorem 9 does not hold if λ is an arbitrary decreasing function of n



(a) Optimal Expected Value versus Time for Different Remaining Capacities n



(b) Optimal Threshold versus Time for Different Remaining Capacities n

Figure 6.17 Example to show that Theorem 9 does not hold if λ is an arbitrary increasing function of n

cost per unit time $c = -0.05$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that for $n = 6$, $V^*(6,t)$ is not concave in t, and $x^*(6,t)$ is increasing in t, although $\lambda(6)\Upsilon_1(-p) = 0.024 > \lambda(6)p + c = 0.01$. Therefore, Theorem 9 does not hold if λ is an arbitrary increasing function of n.

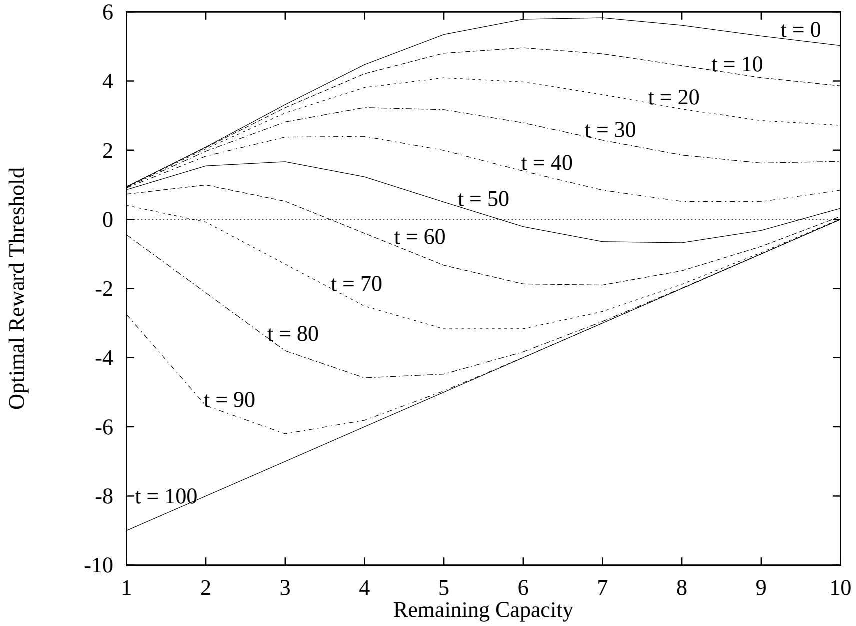
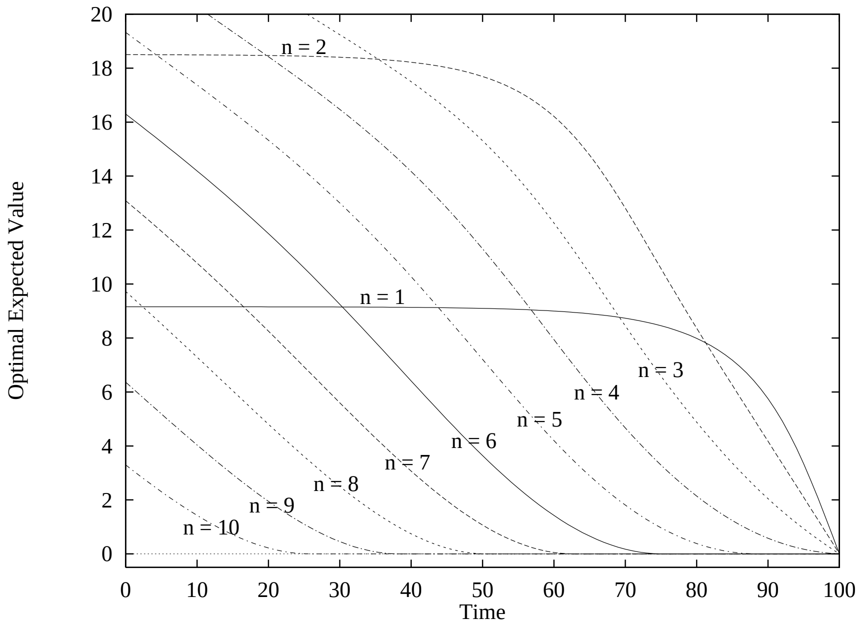


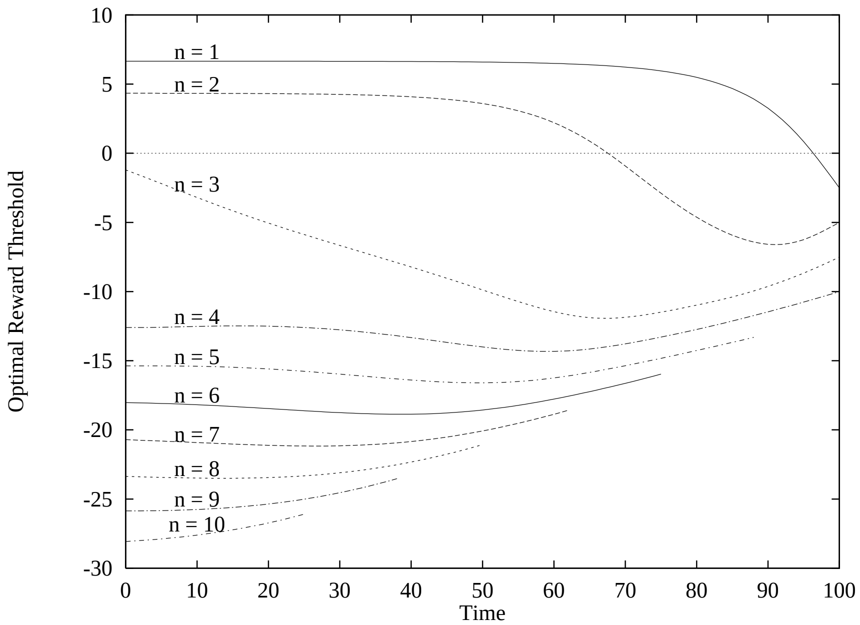
Figure 6.18 Example to show that Theorem 9 does not hold if p is an arbitrary decreasing function of n

Next consider the assumption that p is constant. Figure 6.18 shows $x^*(n,t)$ as a function of n for different t, for an example with exponentially distributed rewards with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 10 - n$, the holding cost per unit time $c = 0.01$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that $x^*(n,t)$ is not decreasing in n, although $\lambda\Upsilon_1(-p(n)) = \lambda/\mu + \lambda p(n) = 1.0 + \lambda p(n) > \lambda p(n) + c = \lambda p(n) + 0.01$. Therefore, Theorem 9 does not hold if p is an arbitrary decreasing function of n.

Figure 6.19 shows $V^*(n,t)$ and $x^*(n,t)$ as functions of t for different n, for an example with normally distributed rewards with mean $\mu = 0$ and standard deviation $\sigma = 5$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p(n) = 2.5n$, the holding cost per unit time $c = 0.1$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that for $n = 3$, $V^*(3,t)$ is not concave in t on $[0,T]$, and $x^*(3,t)$ is increasing in t for $t > 67$,



(a) Optimal Expected Value versus Time for Different Remaining Capacities n



(b) Optimal Reward Threshold versus Time for Different Remaining Capacities n

Figure 6.19 Example to show that Theorem 9 does not hold if p is an arbitrary increasing function of n

although $\lambda\Upsilon_1(-p(3)) = 1.02 > \lambda p(3) + c = 0.85$. Therefore, Theorem 9 does not hold if p is an arbitrary increasing function of n.

Next consider the assumption that c is constant. Figure 6.20 shows $V^*(n,t)$ and $x^*(n,t)$ as functions of n for different t, for an example with exponentially distributed rewards with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p = 0$, the holding cost per unit time $c(n) = 1.0 - 0.1n$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that $V^*(n,t)$ is not concave in n, and $x^*(n,t)$ is not decreasing in n, although $\lambda\Upsilon_1(-p) = \lambda/\mu = 1.0 > \lambda p + c(n) = c(n)$. Therefore, Theorem 9 does not hold if c is an arbitrary decreasing function of n.

Figure 6.21 shows $V^*(n,t)$ and $x^*(n,t)$ as functions of t for different n, for an example with exponentially distributed rewards with mean $1/\mu = 10$, the deadline $T = 100$, items arrive at rate $\lambda = 0.1$, the penalty $p = 0$, the holding cost per unit time $c(n) = 0.1n$, the terminal value $v = 0$, and the discount rate $\alpha = 0$. Note that $V^*(n,t)$ is not concave in t, and $x^*(n,t)$ is not decreasing in t, although $\lambda\Upsilon_1(-p) = \lambda/\mu = 1.0 > \lambda p + c(n) = c(n)$, for all $n < 10$. Therefore, Theorem 9 does not hold if c is an arbitrary increasing function of n.

It is clear that it is necessary that $v(n)$ is concave for $V^*(n,t)$ to be concave in n for all $t \in [0,T]$, because $V^*(n,T) = v(n)$. It does not seem to be necessary for $v(n)$ to be nonincreasing, or for α to be 0.

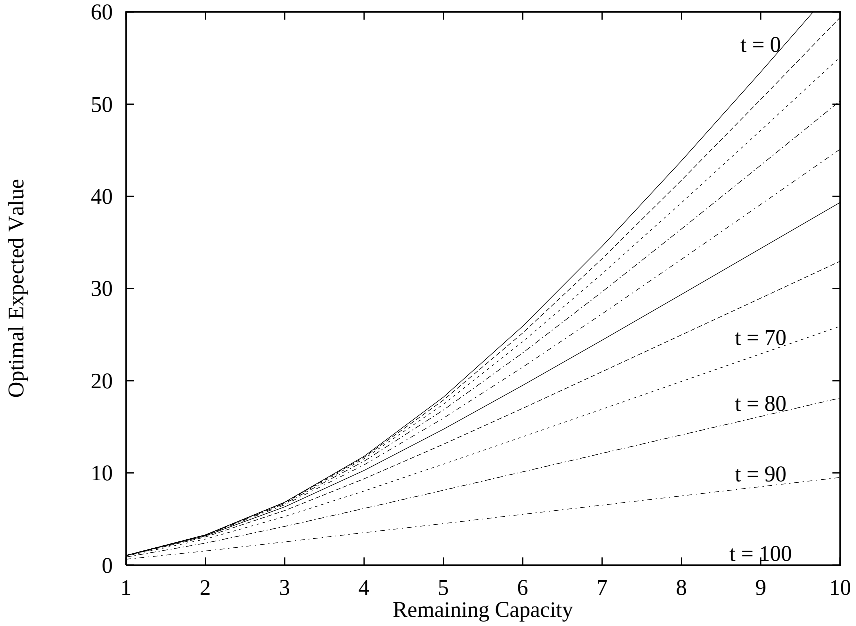
6.6.3.3 The Case With Items Having Random Sizes

In this section monotonicity and convexity properties are established for the case where items have random sizes. The conditions under which these properties hold are more restricted than those for the case where all items have the same size. First it is shown that under conditions that typically hold in applications, the optimal reward threshold x^* is equivalent to an optimal size threshold z^*.

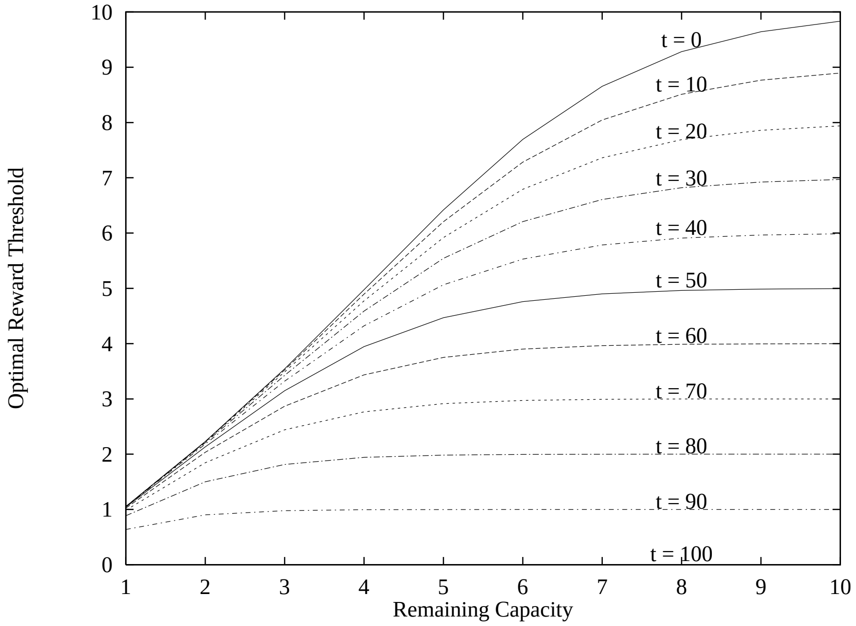
An Equivalent Optimal Size Threshold

Theorems 5 and 6 established that threshold policy π^* with reward threshold $x^*(n,t,s) \equiv V^*(n,t) - V^*(n-s,t) - p(n)$, and acceptance rule

$$D^*(n,t,r,s) \equiv \begin{cases} 1 & \text{if } s \leq n \text{ and } r \geq x^*(n,t,s) \\ 0 & \text{if } s > n \text{ or } r < x^*(n,t,s) \end{cases}$$

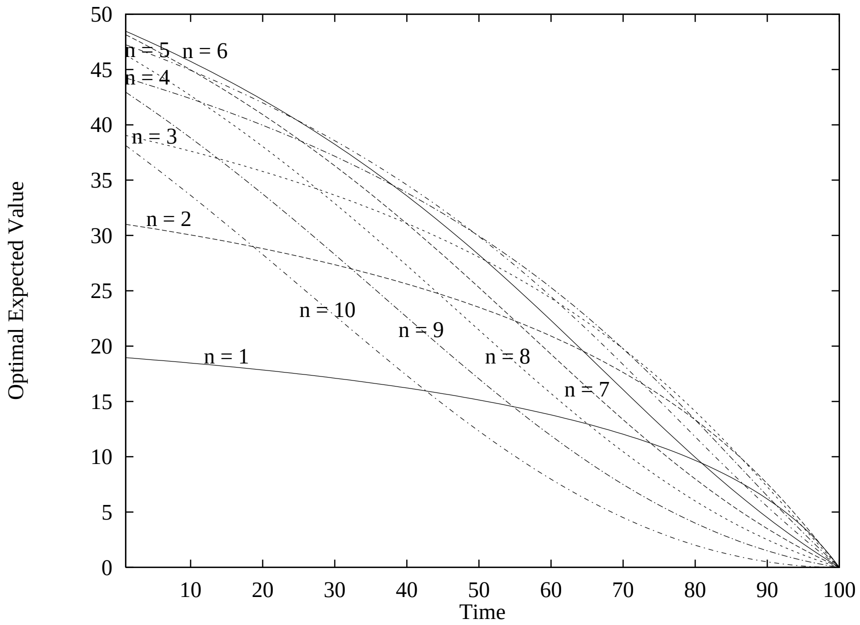


(a) Optimal Expected Value versus Remaining Capacity at Different Times t

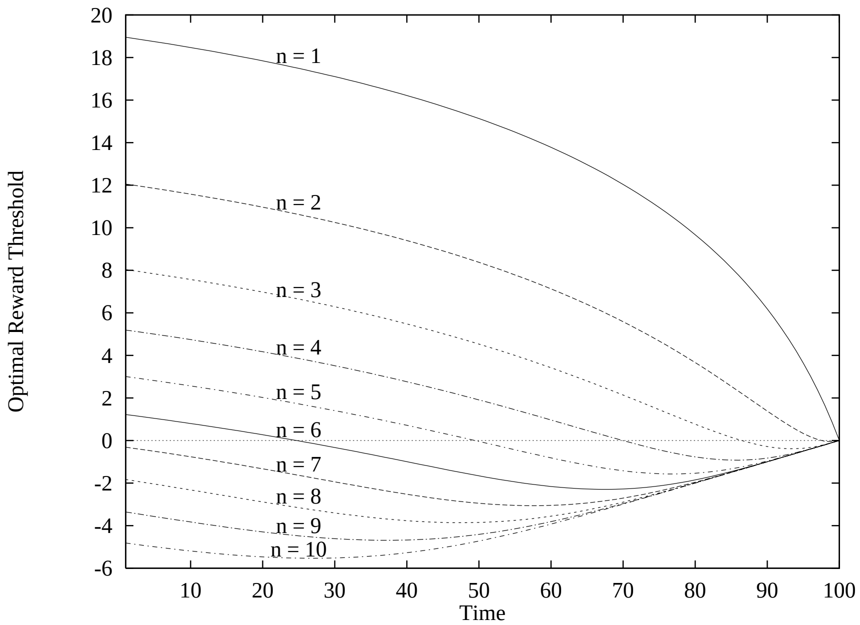


(b) Optimal Reward Threshold versus Remaining Capacity at Different Times t

Figure 6.20 Example to show that Theorem 9 does not hold if c is an arbitrary decreasing function of n



(a) Optimal Expected Value versus Remaining Capacity at Different Times t



(b) Optimal Reward Threshold versus Remaining Capacity at Different Times t

Figure 6.21 Example to show that Theorem 9 does not hold if c is an arbitrary increasing function of n

is an optimal policy for the MDP and the DSKP. It is intuitive to expect that there should be an optimal policy with size threshold

$$z^*(n,t,r) \equiv \sup\{s \in [0,n] : r \geq V^*(n,t) - V^*(n-s,t) - p(n)\}$$

(where $\sup \emptyset \equiv -\infty$), and acceptance rule

$$D^*(n,t,r,s) \equiv \begin{cases} 1 & \text{if } s \leq z^*(n,t,r) \\ 0 & \text{if } s > z^*(n,t,r) \end{cases}$$

Lemma 32 shows that if $V^*(n,t)$ is nondecreasing and right-continuous in n, then the acceptance rule with size threshold $z^*(n,t,r)$ is equivalent to the acceptance rule with reward threshold $x^*(n,t,s)$. Proposition 3 established that if λ is constant, $p(n)$ and $c(n)$ are nonincreasing, and $v(n)$ is nondecreasing, then $V^*(n,t)$ is nondecreasing in n. These conditions usually hold in applications. Two cases where $V^*(n,t)$ is right-continuous in n, are as follows.

1. $V^*(n,t)$ is continuous on $[0,N_0] \times [0,T]$. Proposition 1 established that if F_S and v are continuous on $[0,N_0]$ and $\lambda, p, c \in \mathcal{C}$, then $V^* \in \mathcal{C}$, i.e., $V^*(n,t)$ is continuous on $[0,N_0] \times [0,T]$.
2. $V^*(n,t)$ is upper-semicontinuous in n. Then it follows from $V^*(n,t)$ being nondecreasing in n that $V^*(n,t)$ is right-continuous in n. Pliska [120] showed that under very general conditions, $V^*(n,t)$ is upper-semicontinuous in n.

Lemma 32 *If $V^*(n,t)$ is nondecreasing and right-continuous in n, then the acceptance rule with size threshold $z^*(n,t,r)$ is equivalent to the acceptance rule with reward threshold $x^*(n,t,s)$.*

Proof: If $V^*(n,t)$ is nondecreasing in n, then $V^*(n-s,t)$ is nonincreasing in s, thus $x^*(n,t,s) \equiv V^*(n,t) - V^*(n-s,t) - p(n)$ is nondecreasing in s. Also, if $V^*(n,t)$ is right-continuous in n, then $V^*(n-s,t)$ is left-continuous in s, thus $x^*(n,t,s) \equiv V^*(n,t) - V^*(n-s,t) - p(n)$ is left-continuous in s.

For any $(n,t) \in [0,N_0] \times (0,T]$, if $(r_1,s_1) \in \{(r,s) \in \Re \times \Re_+ : s \leq n \text{ and } r \geq x^*(n,t,s)\}$, then $s_1 \leq \sup\{s \in [0,n] : r_1 \geq V^*(n,t) - V^*(n-s,t) - p(n)\} \equiv z^*(n,t,r_1)$.

Otherwise, suppose $(r_1,s_1) \in \{(r,s) \in \Re \times \Re_+ : s > n \text{ or } r < x^*(n,t,s)\}$. Then, if $s_1 > n$, it follows that $s_1 > z^*(n,t,r_1)$ because $z^*(n,t,r_1) \leq n$. Else, if $r_1 < x^*(n,t,s_1)$ and

$s_1 = 0$, then it follows from the fact that $x^*(n,t,s)$ is nondecreasing in s that $r_1 < x^*(n,t,s)$ for all $s \in [0,n]$. Then $z^*(n,t,r_1) \equiv \sup \emptyset \equiv -\infty$, and thus $s_1 > z^*(n,t,r_1)$. Else, if $r_1 < x^*(n,t,s_1)$ and $s_1 \in (0,n]$, then it follows from the left-continuity of $x^*(n,t,s)$ in s that there exists a left-neighborhood $(s_0,s_1]$ of s_1 such that $r_1 < x^*(n,t,s)$ for all $s \in (s_0,s_1]$. Then it follows from the fact that $x^*(n,t,s)$ is nondecreasing in s that $r_1 < x^*(n,t,s)$ for all $s \in (s_0,n]$. Thus $s_1 > s_0 \geq \sup\{s \in [0,n] : r_1 \geq V^*(n,t) - V^*(n-s,t) - p(n)\} \equiv z^*(n,t,r_1)$.

Therefore, the acceptance sets and the rejection sets of the two acceptance rules coincide. □

It follows that the order of integration in Equation (6.3) can be interchanged, so that

$$\begin{aligned}
\frac{\partial V^*}{\partial t}(n,t) &= \min\Bigg\{-\lambda \int_0^n \int_{V^*(n,t)-V^*(n-s,t)-p(n)}^{\infty} \{r - [V^*(n,t) \\
&\qquad\qquad -V^*(n-s,t) - p(n)]\}\, F_{R|S}(dr|s) F_S(ds) \\
&\qquad + \alpha V^*(n,t) \ + \ \lambda p(n) \ + \ c(n), \\
&\qquad \alpha V^*(n,t) \ - \ \alpha v(n)\Bigg\} \\
&= \min\Bigg\{-\lambda \int_{-\infty}^{\infty} \int_0^{z^*(n,t,r)} \{r - [V^*(n,t) \\
&\qquad\qquad -V^*(n-s,t) - p(n)]\}\, F_{S|R}(ds|r) F_R(dr) \\
&\qquad + \alpha V^*(n,t) \ + \ \lambda p(n) \ + \ c(n), \\
&\qquad \alpha V^*(n,t) \ - \ \alpha v(n)\Bigg\} \\
&= \min\Bigg\{-\lambda \int_{-\infty}^{\infty} \Bigg[\int_0^{z^*(n,t,r)} [r + V^*(n-s,t)]\, F_{S|R}(ds|r) \\
&\qquad\qquad + \int_{z^*(n,t,r)}^{\infty} [V^*(n,t) - p(n)]\, F_{S|R}(ds|r)\Bigg] F_R(dr) \\
&\qquad + \ (\alpha+\lambda)\, V^*(n,t) \ + \ c(n), \\
&\qquad \alpha V^*(n,t) \ - \ \alpha v(n)\Bigg\}
\end{aligned} \tag{6.20}$$

Convexity Results for the Discrete-Time Case

It was shown by Papastavrou *et al.* [117] for the discrete-time case with $p = 0$, $c = 0$, $v = 0$ and $\alpha = 0$, that if $F_{S|R}(s|r)$ is concave in s on $[0, N_0]$ for all $r \geq 0$, then the monotonicity and convexity properties of Theorem 9 hold. These results are extended to the discrete-time approximation of Section 6.4, and to the continuous-time case.

Similar to Section 6.4, construct a discrete-time approximation to the differential equation (6.20) as follows: For $m = 1, 2, \ldots$, let $\Delta t = T/m$, and let $\{t_i = iT/m\}_{i=0}^{m}$ be a partition of $[0, T]$. Let V_m denote the solution of the following difference equation:

$$\begin{aligned}
V_m(n, t_{i-1}) &= V_m(n, t_i) - (\Delta t) f(n, t_i, V_m) \\
&= \max\left\{ \lambda \Delta t \int_{-\infty}^{\infty} \left[\int_0^{z_m(n,t_i,r)} [r + V_m(n-s, t_i)]\, F_{S|R}(ds|r) \right.\right. \\
&\qquad \left. + \int_{z_m(n,t_i,r)}^{\infty} [V_m(n, t_i) - p(n)]\, F_{S|R}(ds|r) \right] F_R(dr) \\
&\qquad + [1 - (\alpha + \lambda)\Delta t]\, V_m(n, t_i) - c(n)\Delta t, \\
&\qquad \left. (1 - \alpha\Delta t)\, V_m(n, t_i) + \alpha v(n)\Delta t \right\}
\end{aligned} \tag{6.21}$$

The boundary conditions are $V_m(n, t_m) = V_m(n, T) = v(n)$ for all $n \in [0, N_0]$, for all $m > (\alpha+\lambda)T$. The size threshold is $z_m(n, t_i, r) \equiv \sup\{s \in [0, n] : r \geq V_m(n, t_i) - V_m(n-s, t_i) - p(n)\}$. Complete $\{V_m(n, t_i)\}_{i=0}^{m}$ on $[0, T]$ for every $n \in [0, N_0]$ by linear interpolation, i.e., for any $t \in [t_{i-1}, t_i]$, let

$$V_m(n, t) \equiv \frac{t_i - t}{t_i - t_{i-1}} V_m(n, t_{i-1}) + \frac{t - t_{i-1}}{t_i - t_{i-1}} V_m(n, t_i) \tag{6.22}$$

Lemma 33 *If λ is constant, $p = 0$, $c = 0$, $v(n)$ is concave nondecreasing and right-continuous, $\alpha = 0$, and $F_{S|R}(s|r)$ is concave in s on $[0, N_0]$ for F_R-almost all $r \geq 0$, then $V_m(n, t)$ is concave in n for all $t \in [0, T]$ and all $m > \lambda T$.*

Proof: If $p = 0$, $c = 0$, and $\alpha = 0$, then because $\Upsilon \geq 0$ it follows that $\lambda(n)\Upsilon(n, t, V^*) - \lambda(n)p(n) - c(n) \geq \alpha v(n)$, and $I^*(n, t) = 1$ for all $(n, t) \in [0, N_0] \times [0, T]$, i.e., it is always optimal to continue. Also, from Proposition 3, because λ is constant, $p(n)$ and $c(n)$ are nonincreasing, and $v(n)$ is nondecreasing, it follows that $V^*(n, t)$ is nondecreasing in n for all $t \in [0, T]$. Hence, $x^*(n, t, s) \equiv V^*(n, t) - V^*(n-s, t) - p(n) \geq 0$ for all $(n, t) \in [0, N_0] \times [0, T]$ and all $s \in [0, n]$. Thus attention can be restricted to nonnegative rewards. Also, because $F_{S|R}(s|r)$ is concave in s on $[0, N_0]$ for F_R-almost all $r \geq 0$, it follows that F_S is concave on $[0, N_0]$, and hence F_S is absolutely continuous on $[0, N_0]$ and F_S has a nonincreasing density

f_S on $[0, N_0]$. It also follows from $v(n)$ being concave nondecreasing and right-continuous, that v is continuous on $[0, N_0]$. Then, from Proposition 1, because F_S and v are continuous on $[0, N_0]$ and $\lambda, p, c \in \mathcal{C}$, it follows that $V^*(n,t)$ is continuous on $[0, N_0] \times [0, T]$. From Proposition 32, because $V^*(n,t)$ is nondecreasing and right-continuous in n, the acceptance rule with size threshold $z^*(n,t,r)$ is equivalent to the acceptance rule with reward threshold $x^*(n,t,s)$.

It is shown by (reverse) induction on i that $V_m(n,t_i)$ is concave in n on $[0, N_0]$ for all $i \in \{0, 1, \ldots, m\}$ and all $m > \lambda T$. For $i = m$, $V_m(n, t_m) = V_m(n, T) = v(n)$ is concave in n on $[0, N_0]$. As induction hypothesis on i, suppose $V_m(n, t_i)$ is concave in n on $[0, N_0]$. From [117], if $F_{S|R}(s|r)$ is concave in s, then $\int_0^{z_m(n,t_i,r)}[r + V_m(n-s, t_i)]\, F_{S|R}(ds|r) + \int_{z_m(n,t_i,r)}^{\infty} V_m(n,t_i)\, F_{S|R}(ds|r)$ is concave in n. Thus

$$\lambda\Delta t \left\{\int_0^{\infty} \left[\int_0^{z_m(n,t_i,r)} [r + V_m(n-s,t_i)]\, F_{S|R}(ds|r) + \int_{z_m(n,t_i,r)}^{\infty} V_m(n,t_i)\, F_{S|R}(ds|r)\right] F_R(dr)\right\}$$

is concave in n. For all $m > \lambda T$, $1 - \lambda \Delta t = 1 - \lambda T/m > 0$. From the induction hypothesis, $V_m(n, t_i)$ is concave in n. Hence

$$\begin{aligned} V_m(n, t_{i-1}) \;=\; & \lambda \Delta t \int_0^{\infty} \left[\int_0^{z_m(n,t_i,r)} [r + V_m(n-s,t_i)]\; F_{S|R}(ds|r)\right. \\ & \qquad \left. + \int_{z_m(n,t_i,r)}^{\infty} V_m(n,t_i)\, F_{S|R}(ds|r)\right] F_R(dr) \\ & + (1 - \lambda\Delta t)\, V_m(n, t_i) \end{aligned}$$

is concave in n. Thus $V_m(n,t_i)$ is concave in n for all $i \in \{0, 1, \ldots, m\}$, for all $m > \lambda T$. Therefore, $V_m(n,t)$ is concave in n on $[0, N_0]$ for all $t \in [0, T]$, for all $m > \lambda T$, because $V_m(n,t)$ is a nonnegative linear (convex) combination of $V_m(n, t_{i-1})$ and $V_m(n, t_i)$, where $t \in [t_{i-1}, t_i]$. □

Monotonicity and Convexity Properties for the Continuous-Time Case

The convexity properties for the discrete-time case, derived from the results in Papastavrou *et al.* [117], are now extended to the continuous-time case. This, together with the results established in Section 6.6.3.1, give a number of monotonicity and convexity properties for the case where items have random sizes, which hold under specific conditions.

Lemma 34 *If λ is constant, $p = 0$, $c = 0$, $v(n)$ is concave nondecreasing and right-continuous, $\alpha = 0$, and $F_{S|R}(s|r)$ is concave in s on $[0, N_0]$ for F_R-almost all $r \geq 0$, then $V^*(n,t)$ is concave in n for all $t \in [0,T]$.*

Proof: It follows from Lemma 33 that $V_m(n,t)$ is concave in n on $[0, N_0]$ for all $t \in [0,T]$, for all $m > \lambda T$. It also follows from Theorem 7 that the sequence of discrete-time approximations $\{V_m(n,t)\}_{m=1}^{\infty}$ converges uniformly to the continuous-time optimal expected value $V^*(n,t)$. Therefore, from Theorem 10.8 of [145] (or Proposition 16.5(b) of [37]) it follows that $V^*(n,t)$ is concave in n on $[0, N_0]$ for all $t \in [0,T]$. □

Lemma 35 *If p is constant and $V^*(n,t)$ is concave in n, then $z^*(n,t,r)$ is nondecreasing in n for all $r \in \Re$.*

Proof: $V^*(n,t)$ is concave in n if and only if for any $0 \leq n_1 \leq n_2 \leq N_0$ and any $s \in [0, n_1]$,

$$\begin{aligned}
& V^*(n_1,t) - V^*(n_1 - s, t) - p \quad \geq \quad V^*(n_2,t) - V^*(n_2 - s, t) - p \\
\Rightarrow \quad & \text{for any } r \in \Re,\ \{s \in [0,n_1] : r \geq V^*(n_1,t) - V^*(n_1 - s,t) - p\} \\
& \subseteq \{s \in [0,n_1] : r \geq V^*(n_2,t) - V^*(n_2 - s,t) - p\} \\
& \subseteq \{s \in [0,n_2] : r \geq V^*(n_2,t) - V^*(n_2 - s,t) - p\} \\
\Rightarrow \quad & \sup\{s \in [0,n_1] : r \geq V^*(n_1,t) - V^*(n_1 - s,t) - p\} \\
& \leq \sup\{s \in [0,n_2] : r \geq V^*(n_2,t) - V^*(n_2 - s,t) - p\} \\
\Leftrightarrow \quad & z^*(n_1,t,r) \leq z^*(n_2,t,r) \qquad \text{for all } r \in \Re
\end{aligned}$$

□

Lemma 36 *If λ and p are constant, $c(n)$ is nonincreasing, $\alpha = 0$, and $V^*(n,t)$ is concave in n, then $z^*(n,r,t)$ is nondecreasing in t for all $n \in [0, N_0]$ and all $r \in \Re$.*

Proof: From Lemma 25, if p is constant, then $V^*(n,t)$ is concave in n if and only if $x^*(n,t,s)$ is nonincreasing in n for any $s \in [0,n]$. From Lemma 26, if λ is constant, $p(n)$ and $c(n)$ are nonincreasing, $\alpha = 0$, and $x^*(n,t,s)$ is nonincreasing in n for all $s \in [0,n]$, then $\partial V^*(n,t)/\partial t$ is nonincreasing in n. Thus, for any $n \in [0, N_0]$ and any $s \in [0,n]$,

$$\begin{aligned}
& \frac{\partial V^*}{\partial t}(n,t) \quad \leq \quad \frac{\partial V^*}{\partial t}(n-s,t) \\
\Leftrightarrow \quad & \frac{\partial\left[V^*(n,t) - V^*(n-s,t) - p\right]}{\partial t} \quad \leq \quad 0
\end{aligned}$$

which implies that for any $0 \leq t_1 \leq t_2 \leq T$ and any $r \in \Re$,

$$\begin{aligned}
& \{s \in [0,n] : r \geq V^*(n,t_1) - V^*(n-s,t_1) - p\} \\
\subseteq \;& \{s \in [0,n] : r \geq V^*(n,t_2) - V^*(n-s,t_2) - p\} \\
\Rightarrow \;& \sup\{s \in [0,n] : r \geq V^*(n,t_1) - V^*(n-s,t_1) - p\} \\
\leq \;& \sup\{s \in [0,n] : r \geq V^*(n,t_2) - V^*(n-s,t_2) - p\} \\
\Leftrightarrow \;& z^*(n,t_1,r) \leq z^*(n,t_2,r) \qquad \square
\end{aligned}$$

Theorem 10 *If λ is constant, $p = 0$, $c = 0$, $v(n)$ is concave nondecreasing and right-continuous, $\alpha = 0$, and $F_{S|R}(s|r)$ is concave in s on $[0, N_0]$ for F_R-almost all $r \geq 0$, then the following conditions hold.*

(i) $V^*(n_1,t) - V^*(n_1-s,t) \geq V^*(n_2,t) - V^*(n_2-s,t)$ *for all* $0 \leq n_1 \leq n_2 \leq N_0$, *all* $s \in [0,n_1]$, *and all* $t \in [0,T]$ (the optimal expected value is concave in remaining capacity).

(ii) $x^*(n_1,t,s) \geq x^*(n_2,t,s)$ *for all* $0 \leq n_1 \leq n_2 \leq N_0$, *all* $s \in [0,n_1]$, *and all* $t \in [0,T]$ (the optimal reward threshold is nonincreasing in remaining capacity).

(iii) $\partial V^*(n_1,t)/\partial t \geq \partial V^*(n_2,t)/\partial t$ *for all* $0 \leq n_1 \leq n_2 \leq N_0$ *and all* $t \in (0,T)$ (the marginal optimal expected value of remaining time $-\partial V^*(n,t)/\partial t$ is nondecreasing in remaining capacity).

(iv) $\partial x^*(n,t,s)/\partial t \leq 0$ *for all* $n \in [0,N_0]$, *all* $s \in [0,n]$, *and all* $t \in (0,T)$ (the optimal reward threshold is nonincreasing in time).

(v) $\partial V^*(n,t_1)/\partial t \geq \partial V^*(n,t_2)/\partial t$ *for all* $n \in [0,N_0]$ *and all* $0 < t_1 \leq t_2 < T$ ($\partial V^*(n,t)/\partial t$ is nonincreasing in time, or the optimal expected value is concave in time).

(vi) $z^*(n_1,t,r) \leq z^*(n_2,t,r)$ *for all* $0 \leq n_1 \leq n_2 \leq N_0$, *all* $t \in [0,T]$, *and all* $r \in \Re$ (the optimal size threshold is nondecreasing in remaining capacity).

(vii) $z^*(n,t_1,r) \leq z^*(n,t_2,r)$ *for all* $n \in [0,N_0]$, *all* $0 \leq t_1 \leq t_2 \leq T$, *and all* $r \in \Re$ (the optimal size threshold is nondecreasing in time).

Proof: From Lemma 34, $V^*(n,t)$ is concave in n, thus (i) holds.

From Lemma 25, (i) $\Leftrightarrow$ (ii).

From Lemma 26, (ii) $\Rightarrow$ (iii).

From Lemma 27, (iii) $\Leftrightarrow$ (iv).

From Lemma 28, (iv) $\Rightarrow$ (v).

From Lemma 35, (i) $\Rightarrow$ (vi).

From Lemma 36, (i) $\Rightarrow$ (vii). $\square$

Part II

THE DYNAMIC AND STOCHASTIC DISTRIBUTION PROBLEM

7. Introduction

Consider a carrier that operates a network of terminals and a fleet of vehicles. Customers call in with requests for the transportation of loads. Each request involves a batch of loads that have to be transported between different origins and destinations. A price is negotiated for transporting the loads, and the carrier has the option to reject a request if the negotiated price is not attractive enough. Accepted loads are collected, either by a fleet of pickup and delivery vehicles, or by the customers dropping the loads off, and consolidated at their origin terminals. If the loads are to be picked up and delivered, vehicle tours have to be formed to deliver loads and to pick up loads at customers' locations, often with time window constraints. This problem, the Vehicle Routing Problem (VRP), has been studied extensively. These pickup and delivery operations are not part of this study. Instead, the focus is on the line-haul operations.

At any point in time there are a number of loads at each terminal, waiting to be transported to their destination terminals. There are also a number of vehicles at each terminal, as well as vehicles en route between different terminals. The dispatcher has to decide when to dispatch vehicles, to which destinations to send them, and which loads to put on the vehicles. It is assumed that loads are shipped directly from their origin terminals to their destination terminals. For an interesting study of the role of terminals and transhipments in one-to-many and many-to-many freight distribution operations, see Daganzo [48, 51]. In his terminology, the operations investigated in this study have a maximum of two transhipments, one close to the origin and one close to the destination.

A number of costs have to be taken into account. If a request is accepted, costs are incurred for picking up, handling, and delivering the loads, and for the administrative activities involved. It is assumed that these costs can be assigned directly to the request, and that they can therefore be subtracted from the negotiated price to give the net reward of the request. If a request is rejected, some customer goodwill and possible future sales are lost, and this is represented by a penalty for rejecting requests. There is a holding cost per unit time, that is a function of the number of loads and vehicles at each terminal, as well as the number of vehicles en route between each pair of terminals. The holding cost represents actual costs for keeping an inventory of waiting loads and vehicles, as well as an incentive for delivering loads as quickly as possible to obtain good customer service. There is also a dispatching and transportation cost that is incurred for each shipment as a whole, that depends on the origin-destination pair of the shipment and the number of loads in the shipment. The objective is to find a policy for accepting transportation requests and for dispatching vehicles that maximizes the expected value (rewards minus costs) of these distribution operations. We call this problem the *Dynamic and Stochastic Distribution Problem* (DSDP).

Chapter DSDP Literature Review reviews the literature on research related to the DSDP. The DSDP is defined in Chapter 9. A Markov decision process model is developed in Chapter 10. Equations are derived for the expected value of a policy, and some fundamental properties of these equations are established. It is shown that it is sufficient to consider only policies that prescribe the dispatching of at most one vehicle for every state of the process. It is also shown that an optimal acceptance rule is given by a simple threshold rule. A number of algorithms for computing the optimal expected value and an optimal policy are studied in Chapter 11. It is shown that the classical algorithms for Markov decision processes converge when applied to the DSDP, in spite of the instantaneous transitions that occur with the dispatching of vehicles. An algorithm that exploits the structural results obtained for the DSDP is proposed and analyzed. Computational tests are conducted to compare the algorithms, and a numerical example is given.

8. Literature Review

In this chapter the research related to the DSDP is reviewed. Much of the research on transportation systems design and distribution operations is similar in nature to the DSDP, and therefore a brief overview of the literature on several classes of problems in this area is given. These problems are classified into Vehicle Routing Problems, Vehicle and Crew Scheduling Problems, Dynamic Vehicle Allocation Problems, and Service Network Design Problems. There are many problems that overlap between and combine the different classes, which indicates the rich variety of problems encountered in the areas of transportation and logistics systems. The operations research literature related to vehicle routing and scheduling alone is extensive; for example, a 1983 review paper by Bodin *et al.*, titled "Routing and Scheduling of Vehicles and Crews" [20], contains 699 references.

Also related to the DSDP is the research on the optimal control of bulk service queues. The literature in this area is also reviewed, especially the aspects of it that are related to transportation problems.

8.1 Vehicle Routing Problems

The Vehicle Routing Problem (VRP) is possibly the most widely studied problem in the area of freight distribution. Instants of the VRP with time window constraints have been solved to optimality, using a set partitioning formulation, which was solved using column generation with the shortest path problem with time window constraints as subproblem [65, 63, 62]. Dynamic and/or stochastic versions of the VRP have also been studied [138, 87, 68, 38, 16, 15, 17, 22, 100]. Closely related are the Pickup and Delivery problem, as well as the Dial-a-Ride problem. Solution procedures for deterministic versions have been proposed in [46, 135, 136, 137, 160, 161, 89, 7, 70, 182], and dynamic and stochastic versions have been studied in [53, 175]. VRP's with additional considerations, such as backhauls, depot locations, inventory, equipment constraints, time-dependent travel costs, and weight as well as volume constraints, have been studied in [47, 49, 50, 101, 102, 140, 169, 36, 118, 83, 82, 141, 99, 105, 107, 130, 180, 34, 33].

8.2 Vehicle and Crew Scheduling Problems

The assignment of crews and vehicles to loads in a deterministic and static problem setting can be done efficiently by solving an assignment problem. In the deterministic and dynamic case the vehicle and crew scheduling problem (VCS) is formulated as a network flow problem on a time expanded network, where every node (state) represents a location-time combination. For an overview, see Powell [128]. Vehicle and crew scheduling problems for public transit and airlines have been studied widely, for example in [20, 8, 9, 71, 64, 166, 18, 86, 142]. Most formulations are deterministic.

8.3 Dynamic Vehicle Allocation Problems

Freight flows are usually not balanced, i.e., the amount of freight that flows into an area in a time period is not equal to the amount of freight that flows from the area. This phenomenon is even more pronounced if the differences in transportation modes and equipment requirements and the variance in freight flows over time are taken into account. Therefore, the vehicles that are used to transport freight into an area are not always used to transport freight out of the area. Dynamic Vehicle Allocation (DVA) problems attempt to find the most efficient ways to relocate empty vehicles from areas with surplus vehicles to areas with shortages of vehicles. Empty Freight Car Relocation problems for railways were investigated in [90, 59, 78, 40, 2]. The DVA problem for truckload trucking was studied in [125, 126, 127, 73, 128].

8.4 Service Network Design Problems

The design of a service network involves questions regarding the location of end-of-line terminals and consolidation centers or breakbulks; the design of vehicle tours and transfers to move freight from origins to destinations; and the choice of operating policy. An operating policy can be stated in terms of the frequencies of vehicle movements between nodes in the network, or in terms of dispatching rules and rules for the relocation of empty vehicles. A rule can involve the solution of one or more optimization problems.

Distribution Network Design (DND) problems are very hard, and an iterative decomposition approach is usually followed to find solutions. The DND problem is decomposed into its natural subproblems (e.g., location of terminals and breakbulks, design of vehicle tours and transfers, and design of operating policy), and the subproblems are iteratively solved until a satisfactory solution is obtained or some type of convergence takes place. Even the subproblems are hard, and are usually solved with optimization based heuristics. The design of so-called load planning networks, which involves choosing the routes and frequency of service between pairs of terminals, was studied by Powell [124], Powell and Sheffi [131], and Leung, Magnanti, and Singhal [104]. Other DND problems were studied in [106, 44, 72, 43, 41, 40, 42, 77, 76, 31, 30, 119]. Most of these models are deterministic; where queueing delays are regarded as important, they are incorporated into the problem by imposing fixed delays. No stochastic DND formulations, that take the variances in interarrival times, sizes, travel and handling times into account, have been encountered in the literature.

Four basic strategies for routing shipments from origins to destinations through terminals were investigated by Hall [81]. Hall [84] also evaluated a number of configurations for local area freight networks. His complete network topology corresponds to the type of system modeled by the DSDP. Soumis, Sauvé, and Le Beau [170] considered the design of vehicle tours to distribute a homogeneous product from multiple supply points to multiple demand points, similar to the transportation problem, but with vehicle movements explicitly taken into account. Distribution problems that incorporate transportation and inventory costs were studied in Burns *et al.* [29], Daganzo [48, 51], and Benjamin [13]. The abovementioned problems (VRP, VCS, DVA) can all be classified as operating policy problems. Other operating policy problems were studied in [39, 79, 80, 52, 54, 93, 178, 94, 67, 91, 98, 10, 32, 66, 88, 121, 146].

8.5 Control of Bulk Service Queues

The general bulk service rule, proposed by Neuts [115], seems to be the first control policy for bulk service queues studied in the literature. Under this policy a bulk server is kept idle as long as there is not a sufficient number of customers in the queue, and service commences on the customers in the system up to the capacity of the server, as soon as the number of customers in the queue reaches a threshold level.

Deb and Serfozo [58] showed that under certain conditions a bulk service rule with a specific threshold is an optimal control policy. In their system customers arrive according to a Poisson process with rate λ, the batch server has a fixed capacity $Q \leq \infty$, and service times are i.i.d. and independent of everything else. There is a cost for holding customers in the system, and a fixed and variable cost for serving a batch of customers. They show that under certain conditions there exists an $M \leq Q$ such that the following bulk service rule minimizes the expected discounted cost: when the server is idle and a customer arrives, or at a service completion, if there are x customers in the system, do not serve if $x < M$, and serve $\min\{x, Q\}$ customers if $x \geq M$. They also show that under certain conditions a similar bulk service rule minimizes the expected (long run) average cost.

Deb [55] extended these results to the case of a bulk server that is controlled by switching the server on and off. A cost is charged for switching the server on or off additional to the service and holding costs. He shows that under certain conditions, there exists an $m \in \{1, \ldots, Q\}$ and an $M \in \{m, \ldots\}$, such that the following policy minimizes the expected discounted cost: when the server is off and a customer arrives, if there are x customers in the system, leave the server off if $x < M$, and serve $\min\{x, Q\}$ customers if $x \geq M$; and at a service completion, when the server is on, if there are x customers in the system, serve $\min\{x, Q\}$ customers if $x \geq m$, and switch the server off if $x < m$. He also showed that a similar policy minimizes the expected average cost under certain conditions.

Deb [56] and Deb and Schmidt [57] showed that such a bulk service policy is also optimal for a two-terminal shuttle. Customers arrive at terminals 1 and 2 according to Poisson processes with rates λ_1 and λ_2 respectively, the batch server has a fixed capacity $Q \leq \infty$, and service times are i.i.d. and independent of everything else. There is a linear cost for holding customers in the system, and a fixed and variable cost for serving a batch of customers. Let x_1 and x_2 denote the number of customers waiting at terminals 1 and 2 respectively. They show that under certain conditions there exist two monotone decreasing functions $G_1 \leq Q, G_2 \leq Q$ such that the following bulk service rule minimizes the expected discounted cost: if the shuttle is at terminal 1 (2), then wait for the next arrival if $x_1 < G_1(x_2)$ ($x_2 < G_2(x_1)$), and dispatch the shuttle with $\min\{x_1, Q\}$ ($\min\{x_2, Q\}$) customers if $x_1 \geq G_1(x_2)$ ($x_2 \geq G_2(x_1)$). They also show that a similar policy minimizes the expected average cost.

Control strategies for bulk arrival, bulk service queues were also studied by Powell. The following strategies were considered: (1) No control; (2) Cancel service if the queue has fewer than m customers; (3) Cancel service if the excess capacity of the server would be more than s (this strategy differs from that of (2) if the server's capacity is not fixed); (4) Keep the server idle if there are fewer than m customers waiting until there are at least m customers in the queue (this is the bulk service rule also studied by Neuts, Deb, Serfozo and Schmidt); (5) Keep the server idle if there are fewer than m customers waiting and the server has been kept idle for less than time T. If there are at least m customers in the queue before the server has been kept idle for time T, then serve the customers up to the capacity of the server. If the server is kept idle for time T, then cancel the service. In [122] Powell derived transforms for the distributions of the number of customers in the queue at service completions under the strategies above. Expressions for the moments of these distributions can then be derived. The mean and approximate variance of the waiting time distribution were also related to the moments of the queue length distribution. The probability q_i that there are i customers in the queue at service completion can be found for $i \in \{0, 1, \ldots, C-1\}$, where C is the maximum server capacity, by solving a system of C linear equations, or by the polynomial expansion algorithm proposed by Powell and Humblet in [129]. However, computing these probabilities and moments requires solving for the zeroes of a function in the complex plane, which is a drawback of this approach. To overcome this drawback, Powell developed approximate closed-form formulas for the mean and variance of the queue length and waiting time in [123]. He also proposed an iterative convolutions algorithm to compute the p.m.f. of the queue length at service completions, and compared this algorithm to a matrix manipulation algorithm proposed by Neuts [116]. The queue length distribution can then be used to evaluate certain performance measures.

Powell and Simão [132, 164, 165] also studied the queue length distribution for transient bulk arrival, bulk service queues under more general operating conditions using discrete time simulation.

9. Problem Definition

A verbal definition of the DSDP is given in Section 9.1. This is followed by a mathematical formulation of the DSDP in Section 9.2. A Markov decision process model of the DSDP is developed in Chapter 10.

9.1 Problem Description

Transportation requests arrive according to a stochastic process in time. Each customer's request consists of a batch of loads that have to be transported between a number of origins and destinations. Loads of different customers may be consolidated into shipments to be carried by vehicles. Loads are classified into load classes according to attributes that determine which loads can be transported together in the same shipment. For example, loads with different origin-destination pairs, or load types that should not be transported together, such as foods and chemical products, are classified into different load classes. The loads in a customer's batch may thus belong to a number of load classes. To simplify notation, load classes will be determined by the origin-destination pairs of the loads. Because the distribution of loads between terminals is studied, and the pickup and delivery of loads at customers' sites is considered as a separate problem, the origin-destination pair of a load is given by a pair of different terminals. It is assumed that the composition of a batch of loads can be represented by integer numbers of unit sized loads; the unit sizes of different load classes may be different. The composition of a batch of loads is the number of loads of each load class in the batch. The composition of a customer's batch is unknown before the request is received, and becomes known upon arrival of the request. The distribution of batch compositions is known. Each request has an associated reward that depends on the composition of the batch. The reward of a request is unknown before arrival, and becomes known upon arrival. The conditional distribution of the rewards given the batch composition is known. Each request can be accepted or rejected, subject to the constraint that a request cannot be accepted if the loads will cause the capacity of one of the terminals to be exceeded. If a request is accepted, the reward associated with the request is received, and if the request is rejected, a penalty is incurred. Once a request is rejected, it cannot be recalled.

There is a holding cost per unit time that depends on the number of loads of each load class at each terminal, the number of vehicles at each terminal, and the number of vehicles en route between each pair of terminals. If a vehicle is available at a terminal, the vehicle can be loaded with one of the load classes at that terminal, subject to the capacity constraint of the vehicle, and be dispatched at any time. A dispatching and transportation cost is incurred each time one or more vehicles are dispatched, that depends on the origin-destination pairs of the vehicles and the number of loads on the vehicles. The objective is to find a policy for accepting requests and for dispatching vehicles that maximizes the expected total discounted value (rewards minus costs) over an infinite time horizon, among the class of nonanticipatory policies.

9.2 Dynamic and Stochastic Distribution Problem Formulation

Let $\mathcal{K}$ be the set of terminals, and let K be the number of terminals. Let the arrival times of the transportation requests be a point process on $(0, \infty]$, with arrival times denoted by $\{A_i\}_{i=1}^{\infty}$. Let $\{B_i\}_{i=1}^{\infty}$ denote the sequence of batch compositions of the arriving requests, where $B_i \in \mathcal{Z}_{+}^{K \times K}$, i.e., each B_i is a $K \times K$ matrix of nonnegative integers. $B_i(k, l)$ denotes the number of loads in batch B_i to be transported from terminal k to terminal l, i.e., the number of loads of class (k, l) in batch B_i. Let $s(k, l)$ denote the size of a load of class (k, l). Let R_i denote the reward of request i. Let the vehicle arrival times at the terminals be a point process on $(0, \infty]$, with arrival times denoted by $\{U_i\}_{i=1}^{\infty}$. Let W_i be the origin-destination pair of the vehicle trip that arrives at time U_i.

Let $\overline{M}$ be the number of vehicles in the fleet, and let $\mathcal{M} \equiv \{0, 1, \ldots, \overline{M}\}$. Let $\overline{N}_v$ be the vehicle capacity, and let $\overline{N}(k)$ be the capacity of terminal k. Let $N_0(k, l)$ denote the initial number of loads at terminal k that have to be transported to terminal l. Let $M_0(k)$ denote the initial number of vehicles at terminal k, and let $\tilde{M}_0(k, l)$ denote the initial number of vehicles en route from terminal k to terminal l. Then

$$\overline{M} \;=\; \sum_{k\in\mathcal{K}} M_0(k) \;+\; \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}} \tilde{M}_0(k,l)$$

Let D_i denote the decision whether to accept or reject request i, defined as follows:

$$D_i \;\equiv\; \begin{cases} 1 & \text{if arrival } i \text{ is accepted} \\ 0 & \text{if arrival } i \text{ is rejected} \end{cases}$$

Let $\{T_i\}_{i=1}^{\infty}$ denote the sequence of vehicle dispatching times. Let L_i be the list of origin-destination pairs of the shipments dispatched at time T_i, i.e., $L_i = \{(k_{i,j}, l_{i,j})\}_{j=1}^{\ell_i}$, where ℓ_i is the number of shipments dispatched at time T_i. Let Q_i be the list of numbers of loads in the shipments dispatched at time T_i, i.e., $Q_i = \{q_{i,j}\}_{j=1}^{\ell_i}$.

For any $t \in [0,\infty)$, let $\mathcal{H}_t$ be the history of the process $\{A_i, B_i, R_i, U_i, W_i\}$ up to time t (i.e., the σ-algebra generated by $\{(A_i, B_i, R_i) : A_i \le t\}$ and $\{(U_i, W_i) : U_i \le t\}$), denoted by $\mathcal{H}_t \equiv \sigma\left(\{(A_i, B_i, R_i) : A_i \le t\},\, \{(U_i, W_i) : U_i \le t\}\right)$. Let $\mathcal{H}_\infty \equiv \sigma\left(\{(A_i, B_i, R_i, U_i, W_i)\}_{i=1}^{\infty}\right)$, and let $\mathcal{H}_{\tau_i} \equiv \{E \in \mathcal{H}_\infty : E \cap \{\tau_i \le t\} \in \mathcal{H}_t \;\forall\; t \in [0,\infty]\}$. Let $\mathcal{A} \equiv \{\{A_i\}_{i=1}^{\infty} : 0 < A_1 < A_2 < \ldots \le \infty\}$, let $\mathcal{B} \equiv \{\{B_i\}_{i=1}^{\infty} : B_i \in \mathcal{Z}_+^{K\times K}\}$, and let $\mathcal{R} \equiv \{\{R_i\}_{i=1}^{\infty} : R_i \in \Re\}$. Let $\mathcal{U} \equiv \{\{U_i\}_{i=1}^{\infty} : 0 < U_1 < U_2 < \ldots \le \infty\}$, and let $\mathcal{W} \equiv \{\{W_i\}_{i=1}^{\infty} : W_i \in \mathcal{K}\times\mathcal{K}\}$. Let $\mathcal{D} \equiv \{\{D_i\}_{i=1}^{\infty} : D_i \in \{0,1\}\}$, let $\mathcal{T} \equiv \{\{T_i\}_{i=1}^{\infty} : 0 < T_1 < T_2 < \ldots \le \infty\}$, let $\mathcal{L} \equiv \{\{L_i\}_{i=1}^{\infty} : L_i = \{(k_{i,j}, l_{i,j})\}_{j=1}^{\ell_i}, (k_{i,j}, l_{i,j}) \in \mathcal{K}\times\mathcal{K}, k_{i,j} \ne l_{i,j}, \ell_i \in \{1,\ldots,\overline{M}\}\}$, and let $\mathcal{Q} \equiv \{\{Q_i\}_{i=1}^{\infty} : Q_i = \{q_{i,j}\}_{j=1}^{\ell_i}, q_{i,j} \in \mathcal{Z}_+, \ell_i \in \{1,\ldots,\overline{M}\}\}$.

The set of history-dependent deterministic policies Π^{HD} for the DSDP is the set of Borel-measurable functions $\pi : \mathcal{A}\times\mathcal{B}\times\mathcal{R}\times\mathcal{U}\times\mathcal{W} \mapsto \mathcal{D}\times\mathcal{T}\times\mathcal{L}\times\mathcal{Q}$ which satisfy the conditions

$$D_i^\pi \text{ is } \mathcal{H}_{A_i} \text{ measurable, i.e., } \{D_i^\pi = 1\} \in \mathcal{H}_{A_i} \text{ for all } i \in \{1,2,\ldots\}$$

$$T_i^\pi \text{ is an } \mathcal{H}_t \text{ stopping time, i.e., } \{T_i^\pi \le t\} \in \mathcal{H}_t \text{ for all } t \in [0,\infty], \text{ for all } i \in \{1,2,\ldots\}$$

$$L_i^\pi \text{ is } \mathcal{H}_{T_i^\pi} \text{ measurable, i.e., } \{L_i^\pi = \{(k_{i,j}^\pi, l_{i,j}^\pi)\}_{j=1}^{\ell_i^\pi}\} \in \mathcal{H}_{T_i^\pi} \text{ for all sequences } \{(k_{i,j}^\pi, l_{i,j}^\pi)\}_{j=1}^{\ell_i^\pi},$$
$$(k_{i,j}^\pi, l_{i,j}^\pi) \in \mathcal{K}\times\mathcal{K}, k_{i,j}^\pi \ne l_{i,j}^\pi, \ell_i^\pi \in \{1,\ldots,\overline{M}\}, \text{ for all } i \in \{1,2,\ldots\}$$

$$Q_i^\pi \text{ is } \mathcal{H}_{T_i^\pi} \text{ measurable, i.e., } \{Q_i^\pi = \{q_{i,j}^\pi\}_{j=1}^{\ell_i^\pi}\} \in \mathcal{H}_{T_i^\pi} \text{ for all sequences } \{q_{i,j}^\pi\}_{j=1}^{\ell_i^\pi},$$
$$q_{i,j}^\pi \in \mathcal{Z}_+, \ell_i^\pi \in \{1,\ldots,\overline{M}\}, \text{ for all } i \in \{1,2,\ldots\}$$

$$s(k_{i,j}^\pi, l_{i,j}^\pi) q_{i,j}^\pi \;\le\; \overline{N}_v \quad \text{for all } (k_{i,j}^\pi, l_{i,j}^\pi) \in L_i^\pi, q_{i,j}^\pi \in Q_i^\pi, \text{ for all } i \in \{1,2,\ldots\}$$

$$\begin{aligned}
N^\pi(k,l,t) &\geq 0 \quad \text{for all } (k,l) \in \mathcal{K}\times\mathcal{K}, \text{ for all } t \in [0,\infty) \\
\textstyle\sum_{l\in\mathcal{K}} s(k,l) N^\pi(k,l,t) &\leq \overline{N}(k) \quad \text{for all } k \in \mathcal{K}, \text{ for all } t \in [0,\infty) \\
M^\pi(k,t) &\geq 0 \text{ for all } k \in \mathcal{K}, \text{ for all } t \in [0,\infty) \\
\tilde{M}^\pi(k,l,t) &\geq 0 \quad \text{for all } (k,l) \in \mathcal{K}\times\mathcal{K}, \text{ for all } t \in [0,\infty)
\end{aligned}$$

where $(\{D_i^\pi\}, \{T_i^\pi\}, \{L_i^\pi\}, \{Q_i^\pi\}) \equiv \pi(\{A_i\}, \{B_i\}, \{R_i\}, \{U_i\}, \{W_i\})$,

$$N^\pi(k,l,t) \equiv N_0(k,l) + \sum_{\{i\,:\,A_i<t\}} B_i(k,l) D_i^\pi - \sum_{\{i\,:\,T_i^\pi<t\}} \sum_{\{j\,:\,(k_{i,j}^\pi, l_{i,j}^\pi)=(k,l)\}} q_{i,j}^\pi$$

is the number of loads at terminal k at time t that have to be transported to terminal l,

$$M^\pi(k,t) \equiv M_0(k) + \sum_{\{i\,:\,U_i\leq t\}} \sum_{l\in\mathcal{K}} I_{\{W_i=(l,k)\}} - \sum_{\{i\,:\,T_i^\pi<t\}} \left|\left\{j : k_{i,j}^\pi = k\right\}\right|$$

is the number of vehicles at terminal k at time t, and

$$\tilde{M}^\pi(k,l,t) \equiv \tilde{M}_0(k,l) + \sum_{\{i\,:\,T_i^\pi<t\}} \left|\left\{j : (k_{i,j}^\pi, l_{i,j}^\pi) = (k,l)\right\}\right| - \sum_{\{i\,:\,U_i\leq t\}} I_{\{W_i=(k,l)\}}$$

is the number of vehicles en route between terminal k and terminal l at time t. Note that

$$\sum_{k\in\mathcal{K}} M^\pi(k,t) + \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}} \tilde{M}^\pi(k,l,t) = \sum_{k\in\mathcal{K}} M_0(k) + \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}} \tilde{M}_0(k,l) = \overline{M}$$

for all t.

Let $c(m,\tilde{m},n)$ denote the holding cost per unit time while the number of vehicles at each terminal is $m = (m(1),\ldots,m(K)) \in \mathcal{M}^K$, the number of vehicles en route between each pair of terminals is $\tilde{m} = (\tilde{m}(1,1), \tilde{m}(1,2), \ldots, \tilde{m}(K,K)) \in \mathcal{M}^{K\times K}$, and the number of loads of each load class waiting to be transported is $n = (n(1,1), n(1,2), \ldots, n(K,K)) \in \mathcal{Z}_+^{K\times K}$. Let p denote the penalty that is incurred if a request is rejected. Let $\varphi(L,Q)$ denote the dispatching/transportation cost that is incurred if vehicles are dispatched with origin-destination pairs given by the list $L = \{(k_1,l_1),\ldots,(k_\ell,l_\ell)\}$ and numbers of loads given by the list $Q = \{q_1,\ldots,q_\ell\}$. Let $\alpha > 0$ be the discount rate.

Let V^π denote the expected total discounted value under policy $\pi \in \Pi^{HD}$, i.e.,

$$\begin{aligned}
V^\pi \equiv E\Bigg[&\sum_{i=1}^\infty e^{-\alpha A_i}\left[R_i D_i^\pi - p\,(1-D_i^\pi)\right] \\
&- \int_0^\infty e^{-\alpha\tau} c(M^\pi(\tau), \tilde{M}^\pi(\tau), N^\pi(\tau))\,d\tau \\
&- \sum_{i=1}^\infty e^{-\alpha T_i^\pi}\varphi(L_i^\pi, Q_i^\pi)\,\Bigg|\, M^\pi(0) = M_0, \tilde{M}^\pi(0) = \tilde{M}_0, N^\pi(0) = N_0\Bigg]
\end{aligned}$$

The objective is to find the optimal expected value V^*, i.e.,

$$V^* \equiv \sup_{\pi \in \Pi^{HD}} V^\pi$$

and to find an optimal policy $\pi^* \in \Pi^{HD}$ that achieves V^*, if such a policy exists.

10. A Markov Decision Process Model

In this chapter a Markov decision process model of the DSDP is developed. The assumptions are stated in Section 10.1. Section 10.2 gives the Markov decision process formulation of the DSDP. Expressions for the value V^π of a policy π are derived in Section 10.3. Two fundamental structural properties for the further development of the model and the solution procedures are established in Section 10.4. The optimality equation, that gives the optimal expected value V^* and an optimal policy π^*, is obtained in Section 10.5.

10.1 Assumptions

Assume that requests arrive according to a Poisson process $\{A_i\}_{i=1}^\infty$ on $(0, \infty)$, with rate $\lambda \in (0, \infty)$. Further assume that vehicles arrive at the terminals according to Poisson processes $\{U_i(k,l)\}_{i=1}^\infty$ on $(0, \infty)$, with rates $\tilde{\mu}(\tilde{m})(k,l) \in [0, \infty)$ while the number of vehicles en route between each pair of terminals is given by $\tilde{m} \in \mathcal{M}^{K \times K}$. Let $\mu(\tilde{m}) \equiv \sum_{k \in \mathcal{K}} \sum_{l \in \mathcal{K}} \tilde{\mu}(\tilde{m})(k,l)$ be the total vehicle arrival rate while the number of vehicles en route between each pair of terminals is given by $\tilde{m}$. Assume that $\{B_i\}_{i=1}^\infty$ is an i.i.d. sequence, independent of $\{A_i\}_{i=1}^\infty$, and that $\{R_i\}_{i=1}^\infty$ is an i.i.d. sequence, independent of $\{A_i\}_{i=1}^\infty$. Also assume that, given $\tilde{m}$, the Poisson processes $\{U_i(k,l)\}_{i=1}^\infty$ are independent of each other and of $\{A_i\}_{i=1}^\infty$, $\{B_i\}_{i=1}^\infty$, and $\{R_i\}_{i=1}^\infty$. Let F_B denote the probability distribution of B, with probability mass function f_B. Assume that $F_B(0) = 0$. Let $F_{R|B}$ denote the conditional probability distribution of R given B. Assume that $E[R \mid B = b] < \infty$ for all b such that $f_B(b) > 0$ and $\sum_{l \in \mathcal{K}} s(k,l) b(k,l) \leq \overline{N}(k)$ for all $k \in \mathcal{K}$.

10.2 Markov Decision Process Formulation

With the assumptions of Section 10.1, the DSDP can be formulated as a continuous time Markov decision process, with the following elements.

1. State space $\mathcal{X} \equiv \mathcal{M}^K \times \mathcal{M}^{K\times K} \times \mathcal{Z}_+^{K\times K}$. The first component m indicates the number of vehicles at each terminal k, the second component $\tilde{m}$ the number of vehicles en route between each pair of terminals (k,l), and the third component n the number of loads of each class (k,l) to be transported. Because of the fixed number $\overline{M}$ of vehicles, and the capacity constraint $\overline{N}(k)$ on each terminal $k \in \mathcal{K}$, not all the states in $\mathcal{X}$ as defined above are feasible.

2. The set of stationary deterministic policies Π^{SD} for the DSDP is defined as follows. Let $\mathcal{G}_D$ denote the set of all Borel-measurable functions $f : \Re \times \mathcal{Z}_+^{K\times K} \mapsto \{0,1\}$. Let $\mathcal{G}_L$ denote the set of all lists $\{(k_j,l_j)\}_{j=1}^{\ell}$, $(k_j,l_j) \in \mathcal{K}\times\mathcal{K}, k_j \neq l_j, \ell \in \{0,1,\ldots,\overline{M}\}$. Let $\mathcal{G}_Q$ denote the set of all lists $\{q_j\}_{j=1}^{\ell}$, $q_j \in \mathcal{Z}_+, \ell \in \{0,1,\ldots,\overline{M}\}$. Then Π^{SD} is the set of functions $\pi : \mathcal{X} \mapsto \mathcal{G}_D \times \mathcal{G}_L \times \mathcal{G}_Q$, where $\pi \equiv (D^\pi, L^\pi, Q^\pi)$, and D^π, L^π and Q^π are as follows. $D^\pi(m,\tilde{m},n,r,b)$ denotes the decision under policy π whether to accept or reject a request i with batch composition $B_i = b$ and reward $R_i = r$ if the state of the system is $(m,\tilde{m},n)$, and is defined as follows:

$$D^\pi(m,\tilde{m},n,r,b) \equiv \begin{cases} 1 & \text{if} \quad \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l)+b(k,l)] \leq \overline{N}(k) \text{ for all } k \in \mathcal{K}, \\ & \quad\;\; \text{and request } i \text{ is accepted} \\ 0 & \text{if} \quad \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l)+b(k,l)] > \overline{N}(k) \text{ for some } k \in \mathcal{K}, \\ & \quad\;\; \text{or request } i \text{ is rejected} \end{cases}$$

The constraints are to ensure that the terminal capacities $\overline{N}(k)$ are not exceeded. Let the acceptance set for policy π be denoted by $\mathcal{R}_1^\pi(m,\tilde{m},n) \equiv \{(r,b) \in \Re \times \mathcal{Z}_+^{K\times K} : D^\pi(m,\tilde{m},n,r,b) = 1\}$, and let the rejection set be denoted by $\mathcal{R}_0^\pi(m,\tilde{m},n) \equiv \{(r,b) \in \Re \times \mathcal{Z}_+^{K\times K} : D^\pi(m,\tilde{m},n,r,b) = 0\}$. $L^\pi(m,\tilde{m},n)$ denotes the list of origin-destination pairs of shipments to dispatch under policy π if the state of the system is $(m,\tilde{m},n)$. No more vehicles can be dispatched from a terminal k than are available at terminal k; therefore the constraint $m(k) - |\{j : k_j = k\}| \geq 0$ for all $k \in \mathcal{K}$, where $L^\pi(m,\tilde{m},n) = \{(k_j,l_j)\}_{j=1}^{\ell}$, is imposed. $Q^\pi(m,\tilde{m},n)$ denotes the list of numbers of loads to dispatch on each shipment under policy π if the state of the system is $(m,\tilde{m},n)$. The vehicle capacities cannot be exceeded; therefore the constraint $s(k_j,l_j)q_j \leq \overline{N}_v$ for all $j \in \{1,\ldots,\ell\}$, where $Q^\pi(m,\tilde{m},n) = \{q_j\}_{j=1}^{\ell}$, is imposed.

It was shown by Yushkevich and Feinberg [187] (Theorem 2) for an infinite horizon MDP with a countable state space and an arbitrary measurable action space, that if $\alpha > 0$ and the reward rate is uniformly bounded for all policies $\pi \in \Pi^{SD}$, then for any $\varepsilon > 0$ there exists a stationary deterministic policy $\pi \in \Pi^{SD}$ that is ε-optimal among all history-dependent deterministic policies $\pi \in \Pi^{HD}$. Hence, attention is restricted to the class of policies Π^{SD}.

3. The total transition rate is $\lambda + \mu(\tilde{m})$ if $L^{\pi}(m, \tilde{m}, n) = \varnothing$, i.e., if no vehicles are dispatched while the state is $(m, \tilde{m}, n)$. If $L^{\pi}(m, \tilde{m}, n) \neq \varnothing$, then the Markov chain associated with policy $\pi \in \Pi^{SD}$ does not visit state $(m, \tilde{m}, n)$, but skips right over it to state $(m', \tilde{m}', n') = (m - \sum_{j=1}^{\ell} e(k_j), \tilde{m} + \sum_{j=1}^{\ell} e(k_j, l_j), n - \sum_{j=1}^{\ell} e(k_j, l_j) q_j)$ if $L^{\pi}(m', \tilde{m}', n') = \varnothing$, or the first state $(m'', \tilde{m}'', n'')$ skipped to thereafter for which $L^{\pi}(m'', \tilde{m}'', n'') = \varnothing$, which has to exist due to the finite number of vehicles that can be dispatched. Here, $e(k)$ denotes the kth standard basis vector in $\Re^K$, and $e(k,l)$ denotes a $K \times K$ matrix with a 1 in position (k,l) and 0's elsewhere. Thus the total transition rate while in states visited by policy π are uniformly bounded for all policies $\pi \in \Pi^{SD}$. It is convenient to include all possible states in the discussion, whether they are visited by the Markov chain associated with a policy π or not. It is shown in Section 10.4 that these states can also be retained for computational purposes, which simplifies the algorithms developed in Chapter 11.

4. The transition probabilities are as follows:

$$P\left[(m, \tilde{m}, n+b) \mid (m, \tilde{m}, n), \pi\right] = \frac{\lambda}{\lambda + \mu(\tilde{m})} \int\int_{\mathcal{R}_1^{\pi}(m,\tilde{m},n) \cap (\Re \times \{b\})} F_{R|B}(dr|b) F_B(db) \times I_{\{L^{\pi}(m,\tilde{m},n) = \varnothing\}}$$

$$P\left[(m, \tilde{m}, n) \mid (m, \tilde{m}, n), \pi\right] = \frac{\lambda}{\lambda + \mu(\tilde{m})} \int\int_{\mathcal{R}_0^{\pi}(m,\tilde{m},n)} F_{R|B}(dr|b) F_B(db) \times I_{\{L^{\pi}(m,\tilde{m},n) = \varnothing\}}$$

$$P\left[(m + e(l), \tilde{m} - e(k,l), n) \mid (m, \tilde{m}, n), \pi\right] = \frac{\tilde{\mu}(\tilde{m})(k,l)}{\lambda + \mu(\tilde{m})} I_{\{L^{\pi}(m,\tilde{m},n) = \varnothing\}}$$

$$P\left[\left(m - \sum_{j=1}^{\ell} e(k_j), \tilde{m} + \sum_{j=1}^{\ell} e(k_j, l_j), n - \sum_{j=1}^{\ell} e(k_j, l_j) q_j\right) \middle| (m, \tilde{m}, n), \pi\right]$$
$$= I_{\{L^{\pi}(m,\tilde{m},n) = \{(k_j,l_j)\}_{j=1}^{\ell}, Q^{\pi}(m,\tilde{m},n) = \{q_j\}_{j=1}^{\ell}\}}$$

All other transition probabilities are 0.

5. Let $V^\pi(m, \tilde{m}, n)$ be the expected total discounted value under policy $\pi \in \Pi^{SD}$ if the state is $(m, \tilde{m}, n)$, i.e.,

$$\begin{aligned} V^\pi(m, \tilde{m}, n) \quad \equiv \quad & E\left[\sum_{i=1}^{\infty} e^{-\alpha A_i} \left[R_i D^\pi\left(M^\pi(A_i), \tilde{M}^\pi(A_i), N^\pi(A_i), R_i, B_i\right)\right.\right. \\ & \left. - p\left(1 - D^\pi\left(M^\pi(A_i), \tilde{M}^\pi(A_i), N^\pi(A_i), R_i, B_i\right)\right)\right] \\ & - \int_0^\infty e^{-\alpha\tau} c\left(M^\pi(\tau), \tilde{M}^\pi(\tau), N^\pi(\tau)\right) d\tau \\ & - \sum_{i=1}^{\infty} e^{-\alpha T_i^\pi} \varphi\left(L^\pi\left(M^\pi(T_i^\pi), \tilde{M}^\pi(T_i^\pi), N^\pi(T_i^\pi)\right),\right. \\ & \left. Q^\pi\left(M^\pi(T_i^\pi), \tilde{M}^\pi(T_i^\pi), N^\pi(T_i^\pi)\right)\right) \\ & \left. \Bigg| M^\pi(0) = m, \tilde{M}^\pi(0) = \tilde{m}, N^\pi(0) = n\right] \end{aligned} \tag{10.1}$$

Let $V^*(m, \tilde{m}, n)$ be the corresponding optimal expected value, i.e.,

$$V^*(m, \tilde{m}, n) \quad \equiv \quad \sup_{\pi \in \Pi^{SD}} V^\pi(m, \tilde{m}, n)$$

10.3 The Value of a Policy

In this section the value function V^π of a policy $\pi \in \Pi^{SD}$ is derived by distinguishing, for each state $(m, \tilde{m}, n) \in \mathcal{X}$, whether vehicles are dispatched in state $(m, \tilde{m}, n)$ ($L^\pi(m, \tilde{m}, n) = \{(k_j, l_j)\}_{j=1}^{\ell}$ and $Q^\pi(m, \tilde{m}, n) = \{q_j\}_{j=1}^{\ell}$), or not ($L^\pi(m, \tilde{m}, n) = \emptyset$).

1. If $L^\pi(m, \tilde{m}, n) = \emptyset$, then by conditioning on the time of the next transition (arrival of a request or a vehicle), and the batch composition and reward if a request arrives, or the origin-destination pair if a vehicle arrives, it follows that

$$\begin{aligned} V^\pi(m, \tilde{m}, n) \quad = \quad & \int_0^\infty (\lambda + \mu(\tilde{m}))\, e^{-(\lambda+\mu(\tilde{m}))t} \left\{ e^{-\alpha t} \left[\frac{\lambda}{\lambda + \mu(\tilde{m})} \times \right.\right. \\ & \left\{ \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} [r + V^\pi(m, \tilde{m}, n + b)]\ F_{R|B}(dr|b) F_B(db) \right. \\ & \left. + \int\!\!\int_{\mathcal{R}_0^\pi(m,\tilde{m},n)} [V^\pi(m, \tilde{m}, n) - p]\ F_{R|B}(dr|b) F_B(db) \right\} \\ & \left. + \sum_{k \in \mathcal{K}} \sum_{l \in \mathcal{K}} \frac{\tilde{\mu}(\tilde{m})(k, l)}{\lambda + \mu(\tilde{m})}\, V^\pi(m + e(l), \tilde{m} - e(k, l), n) \right] \end{aligned}$$

$$
\begin{aligned}
& - \int_0^t e^{-\alpha\tau} c(m,\tilde{m},n)\, d\tau \Big\} \, dt \\
= \; & \frac{\lambda}{\lambda+\mu(\tilde{m})+\alpha} \left\{ \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} [r + V^\pi(m,\tilde{m},n+b)] \; F_{R|B}(dr|b) F_B(db) \right. \\
& \left. + \; [V^\pi(m,\tilde{m},n) - p] \left[1 - \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} F_{R|B}(dr|b) F_B(db) \right] \right\} \\
& + \sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{K}} \frac{\tilde{\mu}(\tilde{m})(k,l)}{\lambda+\mu(\tilde{m})+\alpha} \, V^\pi(m+e(l), \tilde{m}-e(k,l), n) \\
& - \frac{\lambda+\mu(\tilde{m})}{\alpha} \, c(m,\tilde{m},n) \int_0^\infty \left(e^{-(\lambda+\mu(\tilde{m}))t} - e^{-(\lambda+\mu(\tilde{m})+\alpha)t} \right) dt \\
= \; & \frac{\lambda}{\lambda+\mu(\tilde{m})+\alpha} \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} \{ r - [V^\pi(m,\tilde{m},n) - V^\pi(m,\tilde{m},n+b) - p] \} \\
& \qquad F_{R|B}(dr|b) F_B(db) \\
& + \sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{K}} \frac{\tilde{\mu}(\tilde{m})(k,l)}{\lambda+\mu(\tilde{m})+\alpha} \, V^\pi(m+e(l), \tilde{m}-e(k,l), n) \\
& + \frac{\lambda}{\lambda+\mu(\tilde{m})+\alpha} \, V^\pi(m,\tilde{m},n) \; - \; \frac{\lambda p + c(m,\tilde{m},n)}{\lambda+\mu(\tilde{m})+\alpha}
\end{aligned}
\tag{10.2}
$$

It also follows that

$$
\begin{aligned}
V^\pi(m,\tilde{m},n) \; = \; & \frac{\lambda}{\mu(\tilde{m})+\alpha} \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} \{ r - [V^\pi(m,\tilde{m},n) - V^\pi(m,\tilde{m},n+b) - p] \} \\
& \qquad F_{R|B}(dr|b) F_B(db) \\
& + \sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{K}} \frac{\tilde{\mu}(\tilde{m})(k,l)}{\mu(\tilde{m})+\alpha} V^\pi(m+e(l), \tilde{m}-e(k,l), n) \; - \; \frac{\lambda p + c(m,\tilde{m},n)}{\mu(\tilde{m})+\alpha}
\end{aligned}
\tag{10.3}
$$

and

$$
V^\pi(m,\tilde{m},n) \; = \; \frac{\left\{ \lambda \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} [r + V^\pi(m,\tilde{m},n+b) + p] \; F_{R|B}(dr|b) F_B(db) \; + \; \sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{K}} \tilde{\mu}(\tilde{m})(k,l)\, V^\pi(m+e(l), \tilde{m}-e(k,l), n) \; - \; \lambda p \; - \; c(m,\tilde{m},n) \right\}}{\left\{ \lambda \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} F_{R|B}(dr|b) F_B(db) \; + \; \mu(\tilde{m}) \; + \; \alpha \right\}}
\tag{10.4}
$$

2. If $L^\pi(m,\tilde{m},n) = \{(k_j,l_j)\}_{j=1}^{\ell}$, and $Q^\pi(m,\tilde{m},n) = \{q_j\}_{j=1}^{\ell}$, then

$$
\begin{aligned}
V^\pi(m,\tilde{m},n) \; = \; & V^\pi\left(m - \sum_{j=1}^{\ell} e(k_j), \tilde{m} + \sum_{j=1}^{\ell} e(k_j,l_j), n - \sum_{j=1}^{\ell} e(k_j,l_j) q_j \right) \\
& - \; \varphi\left(\{(k_j,l_j,q_j)\}_{j=1}^{\ell} \right)
\end{aligned}
$$

Thus the value function V^π is given by a functional equation of the following form:

$$
V^\pi \; = \; r^\pi \; + \; P^\pi V^\pi \tag{10.5}
$$

where P^π and r^π are given by:

1. If $L^\pi(m, \tilde{m}, n) = \emptyset$:

$$
\begin{aligned}
P^\pi\left[(m,\tilde{m},n),(m,\tilde{m},n)\right] &= \frac{\lambda}{\lambda+\mu(\tilde{m})+\alpha}\int\!\!\int_{\mathcal{R}_0^\pi(m,\tilde{m},n)} F_{R|B}(dr|b)F_B(db) \\
P^\pi\left[(m,\tilde{m},n),(m,\tilde{m},n+b)\right] &= \frac{\lambda}{\lambda+\mu(\tilde{m})+\alpha}\int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)\cap(\Re\times\{b\})} F_{R|B}(dr|b)F_B(db) \\
P^\pi\left[(m,\tilde{m},n),(m+e(l),\tilde{m}-e(k,l),n)\right] &= \frac{\tilde{\mu}(\tilde{m})(k,l)}{\lambda+\mu(\tilde{m})+\alpha}
\end{aligned}
$$

All other values $P^\pi\left[(m,\tilde{m},n),\cdot\right]$ are 0. Note that

$$
\begin{aligned}
\sum_{y\in\mathcal{X}} P^\pi\left[(m,\tilde{m},n),y\right] &= \frac{\lambda}{\lambda+\mu(\tilde{m})+\alpha}\left[\int\!\!\int_{\mathcal{R}_0^\pi(m,\tilde{m},n)} F_{R|B}(dr|b)F_B(db)\right. \\
&\qquad \left. + \int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} F_{R|B}(dr|b)F_B(db)\right] \\
&\quad + \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}}\frac{\tilde{\mu}(\tilde{m})(k,l)}{\lambda+\mu(\tilde{m})+\alpha} \\
&= \frac{\lambda+\mu(\tilde{m})}{\lambda+\mu(\tilde{m})+\alpha} \quad < \quad 1
\end{aligned}
$$

$$
\begin{aligned}
r^\pi(m,\tilde{m},n) &= \frac{\lambda}{\lambda+\mu(\tilde{m})+\alpha}\left[\int\!\!\int_{\mathcal{R}_1^\pi(m,\tilde{m},n)} r\,F_{R|B}(dr|b)F_B(db)\right. \\
&\qquad \left. - p\int\!\!\int_{\mathcal{R}_0^\pi(m,\tilde{m},n)} F_{R|B}(dr|b)F_B(db)\right] \\
&\quad - \frac{c(m,\tilde{m},n)}{\lambda+\mu(\tilde{m})+\alpha}
\end{aligned}
$$

2. If $L^\pi(m,\tilde{m},n) = \{(k_j,l_j)\}_{j=1}^{\ell}$ and $Q^\pi(m,\tilde{m},n) = \{q_j\}_{j=1}^{\ell}$:

$$
P^\pi\left[(m,\tilde{m},n),\left(m-\sum_{j=1}^{\ell}e(k_j),\tilde{m}+\sum_{j=1}^{\ell}e(k_j,l_j),n-\sum_{j=1}^{\ell}e(k_j,l_j)q_j\right)\right] = 1
$$

All other values $P^\pi\left[(m,\tilde{m},n),\cdot\right]$ are 0.

$$
r^\pi(m,\tilde{m},n) = \varphi\left(\{(k_j,l_j,q_j)\}_{j=1}^{\ell}\right)
$$

10.4 Structural Properties

Two fundamental results are established in this section. In Section 10.4.1 it is shown that for any $\pi \in \Pi^{SD}$, P^π in Equation (10.5) has spectral radius $\rho(P^\pi) < 1$. This property is essential for the convergence and monotonicity results that are established later. In Section 10.4.2 it is shown that attention can be restricted to policies that prescribe the dispatching of at most one vehicle in each state. This property simplifies the exposition and the solution procedures significantly.

10.4.1 The Spectral Radius of the MDP

It follows from Equation (10.5) that the continuous time Markov decision process is equivalent to a discrete time Markov decision process, with the condition that under policy π discounting takes place in state $(m, \tilde{m}, n)$ only if $L^{\pi}(m, \tilde{m}, n) = \emptyset$. The equivalence is in the sense that for any policy $\pi \in \Pi^{SD}$, the value function V^{π} is the same for both processes. The equivalence between continuous and discrete time Markov decision processes with uniformly bounded transition rates was shown by Serfozo [159]. Discounted Markov decision processes studied in the literature usually have a functional equation of the form $V^{\pi} = r^{\pi} + \beta P^{\pi} V^{\pi}$, where $0 < \beta < 1$ and P^{π} is a stochastic matrix. As a consequence, for any $\pi \in \Pi^{SD}$, $\|\beta P^{\pi}\|_{\infty} \equiv \sup_i \sum_j \beta P^{\pi}(i,j) = \beta < 1$. It follows that for any $\pi \in \Pi^{SD}$, βP^{π} has spectral radius $\rho\left(\beta P^{\pi}\right) \leq \|\beta P^{\pi}\|_{\infty} = \beta < 1$. This property plays a fundamental role in establishing several convergence results for discounted MDP's. However, in the model for the distribution problem, transitions occur instantaneously when vehicles are dispatched. Consequently the model has a functional equation of the form $V^{\pi} = r^{\pi} + P^{\pi} V^{\pi}$ (with no $\beta \in (0,1)$). In this case P^{π} is a substochastic matrix, where some rows may have a row sum equal to 1. Thus $\|P^{\pi}\|_{\infty} = 1$, and the usual convergence results for MDP's do not follow immediately. First it has to be established that for any $\pi \in \Pi^{SD}$, P^{π} has spectral radius $\rho\left(P^{\pi}\right) < 1$. Let $\overline{\mu} \equiv \max_{\tilde{m}} \mu(\tilde{m})$.

Lemma 37 *For any policy $\pi \in \Pi^{SD}$, P^{π} has spectral radius*

$$\rho\left(P^{\pi}\right) \quad \leq \quad \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha}\right)^{\frac{1}{\overline{M}+1}} \quad < \quad 1$$

Proof: Because $\mathcal{X}$ is finite,

$$\rho\left(P^{\pi}\right) \quad = \quad \max_{\{\text{eigenvalues } \nu \text{ of } P^{\pi}\}} |\nu|$$

Let ν be an eigenvalue of P^π with $|\nu| = \rho(P^\pi)$, and corresponding eigenvector $v \neq 0$. Let Δ^π be a diagonal matrix with diagonal entries

$$\begin{aligned}
\Delta^\pi[(m,\tilde{m},n),(m,\tilde{m},n)] &= \frac{1}{\sum_{y\in\mathcal{X}} P^\pi[(m,\tilde{m},n),y]} \\
&= \begin{cases} \frac{\lambda+\mu(\tilde{m})+\alpha}{\lambda+\mu(\tilde{m})} > 1 & \text{if } L^\pi(m,\tilde{m},n) = \emptyset \\ 1 & \text{if } L^\pi(m,\tilde{m},n) \neq \emptyset \end{cases}
\end{aligned}$$

Then

$$\Delta^\pi P^\pi v = \nu \Delta^\pi v$$

and for any $(m,\tilde{m},n) \in \mathcal{X}$

$$\begin{aligned}
\sum_{y\in\mathcal{X}} (\Delta^\pi P^\pi)[(m,\tilde{m},n),y] &= \begin{cases} \frac{\lambda+\mu(\tilde{m})+\alpha}{\lambda+\mu(\tilde{m})} \sum_{y\in\mathcal{X}} P^\pi[(m,\tilde{m},n),y] = 1 \\ \qquad\qquad \text{if } L^\pi(m,\tilde{m},n) = \emptyset \\ 1 \qquad\qquad \text{if } L^\pi(m,\tilde{m},n) \neq \emptyset \end{cases} \\
\Rightarrow \quad \|\Delta^\pi P^\pi\|_\infty &\equiv \sup_{x\in\mathcal{X}} \sum_{y\in\mathcal{X}} |(\Delta^\pi P^\pi)(x,y)| = 1
\end{aligned}$$

Also

$$\|\nu\Delta^\pi v\|_\infty = |\nu|\,\|\Delta^\pi v\|_\infty$$

and

$$\begin{aligned}
\|\nu\Delta^\pi v\|_\infty = \|\Delta^\pi P^\pi v\|_\infty &\leq \|\Delta^\pi P^\pi\|_\infty \|v\|_\infty = \|v\|_\infty \\
\Rightarrow \quad |\nu|\,\|\Delta^\pi v\|_\infty &\leq \|v\|_\infty \qquad (10.6)
\end{aligned}$$

Let

$$|v(y_0)| = \max_{x\in\mathcal{X}} |v(x)| \equiv \|v\|_\infty$$

If $L^\pi(y_0) = \emptyset$, then

$$\Delta^\pi(y_0,y_0) = \frac{\lambda+\mu(\tilde{m}(y_0))+\alpha}{\lambda+\mu(\tilde{m}(y_0))} \geq \frac{\lambda+\overline{\mu}+\alpha}{\lambda+\overline{\mu}}$$

and

$$\begin{aligned}
|(\Delta^\pi v)(y_0)| &= \left|\frac{\lambda+\mu(\tilde{m}(y_0))+\alpha}{\lambda+\mu(\tilde{m}(y_0))} v(y_0)\right| \geq \frac{\lambda+\overline{\mu}+\alpha}{\lambda+\overline{\mu}}|v(y_0)| = \frac{\lambda+\overline{\mu}+\alpha}{\lambda+\overline{\mu}}\|v\|_\infty \\
\Rightarrow \quad \|\Delta^\pi v\|_\infty &\geq \frac{\lambda+\overline{\mu}+\alpha}{\lambda+\overline{\mu}}\|v\|_\infty > 0 \\
\Rightarrow \quad |\nu|\,\|\Delta^\pi v\|_\infty &\leq \|v\|_\infty \leq \frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha}\|\Delta^\pi v\|_\infty \\
\Rightarrow \quad \rho(P^\pi) = |\nu| &\leq \frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha} \leq \left(\frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha}\right)^{\frac{1}{M+1}}
\end{aligned}$$

and the result has been established.

Else, if $L^\pi(y_0) \neq \emptyset$, then $P^\pi(y_0, y_1) = 1$ for some $y_1 \in \mathcal{X}$, $y_1 \neq y_0$, and y_1 has at least one vehicle less than y_0 at at least one terminal and no more at any terminal, i.e., $m(y_1)(k) \leq m(y_0)(k)$ for all $k \in \mathcal{K}$, and $m(y_1)(k_0) < m(y_0)(k_0)$ for at least one $k_0 \in \mathcal{K}$. Also,

$$\begin{aligned} v(y_1) &= (P^\pi v)(y_0) &&= \nu v(y_0) \\ \Rightarrow \quad |v(y_1)| &= |\nu|\,|v(y_0)| &&= |\nu|\,\|v\|_\infty \end{aligned}$$

As induction hypothesis, suppose that a sequence $y_0, y_1, \ldots, y_i$ has been obtained, where each pair y_j, y_{j+1} satisfies $m(y_{j+1})(k) \leq m(y_j)(k)$ for all $k \in \mathcal{K}$, and $m(y_{j+1})(k_j) < m(y_j)(k_j)$ for at least one $k_j \in \mathcal{K}$. Also, suppose that $L^\pi(y_{i-1}) \neq \emptyset$, and $|v(y_i)| = |\nu|^i \|v\|_\infty$. Note that this requires that $i \leq \overline{M}$, because there are only $\overline{M}$ vehicles in the fleet, and thus if a state $y_{\overline{M}}$ is reached, it has to hold that $L^\pi(y_{\overline{M}}) = \emptyset$.

If $L^\pi(y_i) = \emptyset$, then

$$\Delta^\pi(y_i, y_i) = \frac{\lambda + \mu(\tilde{m}(y_i)) + \alpha}{\lambda + \mu(\tilde{m}(y_i))} \geq \frac{\lambda + \overline{\mu} + \alpha}{\lambda + \overline{\mu}}$$

and

$$\begin{aligned} |(\Delta^\pi v)(y_i)| &= \left| \frac{\lambda + \mu(\tilde{m}(y_i)) + \alpha}{\lambda + \mu(\tilde{m}(y_i))} v(y_i) \right| \geq \frac{\lambda + \overline{\mu} + \alpha}{\lambda + \overline{\mu}} |v(y_i)| = \frac{\lambda + \overline{\mu} + \alpha}{\lambda + \overline{\mu}} |\nu|^i \|v\|_\infty \\ \Rightarrow \quad \|\Delta^\pi v\|_\infty &\geq \frac{\lambda + \overline{\mu} + \alpha}{\lambda + \overline{\mu}} |\nu|^i \|v\|_\infty > 0 \\ \Rightarrow \quad |\nu|^{i+1} \|\Delta^\pi v\|_\infty &\leq |\nu|^i \|v\|_\infty \leq \frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \|\Delta^\pi v\|_\infty \\ \Rightarrow \quad \rho(P^\pi) = |\nu| &\leq \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha}\right)^{\frac{1}{i+1}} \leq \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha}\right)^{\frac{1}{\overline{M}+1}} \end{aligned}$$

and the result has been established.

Else, if $L^\pi(y_i) \neq \emptyset$, then $P^\pi(y_i, y_{i+1}) = 1$ for some $y_{i+1} \in \mathcal{X}$, $y_{i+1} \neq y_i$, and $m(y_{i+1})(k) \leq m(y_i)(k)$ for all $k \in \mathcal{K}$, and $m(y_{i+1})(k_i) < m(y_i)(k_i)$ for at least one $k_i \in \mathcal{K}$. Also,

$$\begin{aligned} v(y_{i+1}) &= (P^\pi v)(y_i) &&= \nu v(y_i) \\ \Rightarrow \quad |v(y_{i+1})| &= |\nu|\,|v(y_i)| &&= |\nu|^{i+1} \|v\|_\infty \end{aligned}$$

and the induction has been established. □

Corollary 9 *For any policy $\pi \in \Pi^{SD}$, $(I - P^\pi)^{-1}$ exists and satisfies*

$$(I - P^\pi)^{-1} = \sum_{i=0}^{\infty} (P^\pi)^i$$

Corollary 10 *For any policy $\pi \in \Pi^{SD}$, the expected value under policy π is given by*

$$V^\pi = (I - P^\pi)^{-1} r^\pi$$

Corollary 11 *For any policy $\pi \in \Pi^{SD}$, and $V_1, V_2 : \mathcal{X} \mapsto \Re$ such that $V_2 \geq V_1$, it holds that*

$$(I - P^\pi)^{-1} V_2 \geq (I - P^\pi)^{-1} V_1$$

10.4.2 Dispatching at Most One Vehicle in Each State

Suppose the dispatching and transportation cost φ is additive over the vehicles that are dispatched. That is, for any $\{(k_j, l_j)\}_{j=1}^{\ell}$ and $\{q_j\}_{j=1}^{\ell}$

$$\varphi\left(\{(k_j, l_j)\}_{j=1}^{\ell}, \{q_j\}_{j=1}^{\ell}\right) = \sum_{j=1}^{\ell} \varphi(k_j, l_j, q_j)$$

This seems to be a reasonable assumption to make in practice. With this assumption, attention can be restricted to policies that consider the dispatching of only a single vehicle in any state. This result considerably simplifies the computation of optimal policies and optimal expected values.

Lemma 38 *For any policy $\pi_1 \in \Pi^{SD}$, there exists a policy $\pi \in \Pi^{SD}$ such that π prescribes the dispatching of at most one vehicle in each state, and $V^\pi \geq V^{\pi_1}$.*

Proof: Suppose $L^{\pi_1}(m, \tilde{m}, n) = \{(k_j, l_j)\}_{j=1}^{\ell}$, and $Q^{\pi_1}(m, \tilde{m}, n) = \{q_j\}_{j=1}^{\ell}$ with $\ell > 1$. Let $m_i = m - \sum_{j=1}^{i} e(k_j)$, $\tilde{m}_i = \tilde{m} + \sum_{j=1}^{i} e(k_j, l_j)$, and $n_i = n - \sum_{j=1}^{i} e(k_j, l_j) q_j$. Then

$$V^{\pi_1}(m, \tilde{m}, n) = V^{\pi_1}(m_\ell, \tilde{m}_\ell, n_\ell) - \sum_{j=1}^{\ell} \varphi(k_j, l_j, q_j)$$

Case 1: $V^{\pi_1}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) \geq V^{\pi_1}(m_\ell, \tilde{m}_\ell, n_\ell) - \varphi(k_\ell, l_\ell, q_\ell)$:

Consider policy π_2 that is the same as policy π_1 except that the dispatching of the ℓth vehicle is not prescribed when the state is $(m, \tilde{m}, n)$. That is, $L^{\pi_2}(m, \tilde{m}, n) = \{(k_j, l_j)\}_{j=1}^{\ell-1}$ and $Q^{\pi_2}(m, \tilde{m}, n) = \{q_j\}_{j=1}^{\ell-1}$. Then

$$\begin{aligned} r^{\pi_2}(m, \tilde{m}, n) + (P^{\pi_2} V^{\pi_1})(m, \tilde{m}, n) &= V^{\pi_1}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) - \sum_{j=1}^{\ell-1} \varphi(k_j, l_j, q_j) \\ &\geq V^{\pi_1}(m_\ell, \tilde{m}_\ell, n_\ell) - \sum_{j=1}^{\ell} \varphi(k_j, l_j, q_j) \\ &= V^{\pi_1}(m, \tilde{m}, n) \end{aligned}$$

and $r^{\pi_2}(x) + (P^{\pi_2} V^{\pi_1})(x) = V^{\pi_1}(x)$ for all other states $x \in \mathcal{X}$. Hence $r^{\pi_2} \geq (I - P^{\pi_2}) V^{\pi_1}$. It follows from Corollary 10 and the monotonicity property of Corollary 11 that

$$V^{\pi_2} = (I - P^{\pi_2})^{-1} r^{\pi_2} \geq V^{\pi_1}$$

Hence, policy π_1 is replaced by policy π_2 which prescribes the dispatching of one less vehicle in state $(m, \tilde{m}, n)$, and which has value function $V^{\pi_2} \geq V^{\pi_1}$.

Case 2: $V^{\pi_1}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) < V^{\pi_1}(m_\ell, \tilde{m}_\ell, n_\ell) - \varphi(k_\ell, l_\ell, q_\ell)$:

Consider policy π_2 that is the same as policy π_1 except that the dispatching of one vehicle is prescribed when the state is $(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1})$, namely, $L^{\pi_2}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) = \{(k_\ell, l_\ell)\}$ and $Q^{\pi_2}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) = \{q_\ell\}$. Then

$$\begin{aligned} r^{\pi_2}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) + (P^{\pi_2} V^{\pi_1})(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) &= V^{\pi_1}(m_\ell, \tilde{m}_\ell, n_\ell) - \varphi(k_\ell, l_\ell, q_\ell) \\ &> V^{\pi_1}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) \end{aligned}$$

and $r^{\pi_2}(x) + (P^{\pi_2} V^{\pi_1})(x) = V^{\pi_1}(x)$ for all other states $x \in \mathcal{X}$. Hence $r^{\pi_2} \geq (I - P^{\pi_2}) V^{\pi_1}$. It follows from Corollary 10 and the monotonicity property of Corollary 11 that

$$V^{\pi_2} = (I - P^{\pi_2})^{-1} r^{\pi_2} \geq V^{\pi_1}$$

Hence, policy π_1 is replaced by policy π_2 which prescribes the dispatching of one vehicle in state $(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1})$, and which has value function $V^{\pi_2} \geq V^{\pi_1}$, that also satisfies $V^{\pi_2}(m_{\ell-1}, \tilde{m}_{\ell-1}, n_{\ell-1}) = V^{\pi_2}(m_\ell, \tilde{m}_\ell, n_\ell) - \varphi(k_\ell, l_\ell, q_\ell)$. Now case 1 can be applied to reduce the number of vehicles that are dispatched in state $(m, \tilde{m}, n)$.

The procedure can be continued until a policy $\pi \in \Pi^{SD}$ is obtained such that π prescribes the dispatching of at most one vehicle in each state, and $V^{\pi} \geq V^{\pi_1}$. $\square$

10.5 The Optimal Value and an Optimal Policy for the DSDP

Similar to the DSKP, it is intuitive to expect that a policy with reward threshold $x^*(m,\tilde{m},n,b) = V^*(m,\tilde{m},n) - V^*(m,\tilde{m},n+b) - p$ is optimal. Suppose the state is $(m,\tilde{m},n)$ and a batch b of loads arrive with associated reward r. If $\sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l)+b(k,l)] \leq \overline{N}(k)$ for all $k \in \mathcal{K}$, and the batch is accepted, then the optimal expected value is $r + V^*(m,\tilde{m},n+b)$. Else, if the batch is rejected, then the optimal expected value is $V^*(m,\tilde{m},n) - p$. Hence, it is optimal to accept the batch b if $\sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l)+b(k,l)] \leq \overline{N}(k)$ for all $k \in \mathcal{K}$, and $r \geq V^*(m,\tilde{m},n) - V^*(m,\tilde{m},n+b) - p$. This intuition is formalized in Theorem 11.

10.5.1 An Optimal Threshold Policy for the DSDP

Theorem 11

$$
\begin{aligned}
V^*(m,\tilde{m},n) \quad = \quad & \sup_{(D,L,Q)\in\Pi^{SD}} \left\{ \frac{\lambda}{\mu(\tilde{m})+\alpha} \int\!\!\int_{\mathcal{R}_1(m,\tilde{m},n)} \{r - [V^*(m,\tilde{m},n) \right. \\
& \qquad\qquad -V^*(m,\tilde{m},n+b) - p]\}\; F_{R|B}(dr|b)F_B(db) \\
& + \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}} \frac{\tilde{\mu}(\tilde{m})(k,l)}{\mu(\tilde{m})+\alpha}\, V^*\left(m+e(l),\tilde{m}-e(k,l),n\right) \;-\; \frac{\lambda p + c(m,\tilde{m},n)}{\mu(\tilde{m})+\alpha}, \\
& \left. V^*\left(m-e(k),\tilde{m}+e(k,l),n-e(k,l)q\right) \;-\; \varphi(k,l,q) \right\} \qquad (10.7) \\
= \quad & \max\left\{ \frac{\lambda}{\mu(\tilde{m})+\alpha} \sum_{\left\{b\,:\,\sum_{l\in\mathcal{K}} s(k,l)[n(k,l)+b(k,l)]\leq\overline{N}(k)\ \forall\ k\in\mathcal{K}\right\}} f_B(b)\times \right. \\
& \qquad \int_{V^*(m,\tilde{m},n)-V^*(m,\tilde{m},n+b)-p}^{\infty} \{r - [V^*(m,\tilde{m},n) \\
& \qquad\qquad -V^*(m,\tilde{m},n+b) - p]\}\; F_{R|B}(dr|b) \\
& + \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}} \frac{\tilde{\mu}(\tilde{m})(k,l)}{\mu(\tilde{m})+\alpha}\, V^*\left(m+e(l),\tilde{m}-e(k,l),n\right) \;-\; \frac{\lambda p + c(m,\tilde{m},n)}{\mu(\tilde{m})+\alpha}, \\
& \max_{\{(k,l)\in\mathcal{K}\times\mathcal{K},q\in\mathcal{Z}_+\,:\,m(k)\geq 1,k\neq l,q\leq n(k,l),s(k,l)q\leq\overline{N}_v\}} \\
& \left. \left\{ V^*\left(m-e(k),\tilde{m}+e(k,l),n-e(k,l)q\right) \;-\; \varphi(k,l,q) \right\} \right\} \qquad (10.8)
\end{aligned}
$$

Proof: The first equality follows from Equations (10.3) and (10.5), Lemma 38, and the results in Yushkevich and Feinberg [187] (Theorem 2). The second equality follows from the property that for any distribution F_R,

$$\int_y^\infty (r-y)\ F_R(dr) \quad = \quad \sup_{E\in\mathcal{E}} \int_E (r-y)\ F_R(dr)$$

where $\mathcal{E}$ is the Borel sets on $\Re$, which was established in Lemma 5. Thus the supremum in Equation (10.7) is attained by using the optimal reward threshold $x^*(m,\tilde{m},n,b) = V^*(m,\tilde{m},n) - V^*(m,\tilde{m},n+b) - p$ if $\sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l)+b(k,l)] \leq \overline{N}(k)$ for all $k \in \mathcal{K}$. Therefore, the policy that maximizes Equation (10.8) is an optimal stationary deterministic policy. From the results in Yushkevich and Feinberg [187] (Theorem 2), it follows that such a policy is also optimal among all history-dependent deterministic policies. $\square$

Corollary 12 *An optimal acceptance rule is given by the following threshold rule.*

$$D^*(m,\tilde{m},n,r,b) \quad \equiv \quad \begin{cases} 1 \quad \textit{if} & \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l)+b(k,l)] \leq \overline{N}(k) \textit{ for all } k\in\mathcal{K}, \\ & \textit{and } r \geq V^*(m,\tilde{m},n) - V^*(m,\tilde{m},n+b) - p \\ 0 \quad \textit{if} & \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l)+b(k,l)] > \overline{N}(k) \textit{ for some } k\in\mathcal{K}, \\ & \textit{or } r < V^*(m,\tilde{m},n) - V^*(m,\tilde{m},n+b) - p \end{cases} \tag{10.9}$$

10.5.2 Monotonicity Result

Proposition 6 establishes the intuitive result that if $c(m,\tilde{m},n)$ is nondecreasing in n for all m and $\tilde{m}$, and $\varphi(k,l,q)$ is nondecreasing in q for all k and l, (these conditions hold in most practical cases), then $V^*(m,\tilde{m},n)$ is nonincreasing in n for all m and $\tilde{m}$.

Proposition 6 *If $c(m,\tilde{m},n)$ is nondecreasing in n for all m and $\tilde{m}$, and $\varphi(k,l,q)$ is nondecreasing in q for all k and l, then $V^*(m,\tilde{m},n)$ is nonincreasing in n for all m and $\tilde{m}$.*

The proof of Proposition 6 is given in Appendix L.

11. Solution Methods

A number of algorithms are available for solving infinite horizon discounted Markov decision processes with a functional equation of the form $V^\pi = r^\pi + \beta P^\pi V^\pi$ for some discount factor $\beta \in (0,1)$. Examples of these include value iteration (successive approximation), policy iteration, modified policy iteration, and many variations of these algorithms. The convergence results for these algorithms rely on the discount factor $\beta \in (0,1)$, which provides the contractive property for the iteration schemes. The MDP model of the DSDP does not have a functional equation of the form $V^\pi = r^\pi + \beta P^\pi V^\pi$ for some $\beta \in (0,1)$, and thus the convergence results commonly found in the literature do not apply. In this chapter it is shown that these classical algorithms converge when applied to the DSDP. These convergence results could have been obtained more directly by using the result from Lemma 37 that for the DSDP, for any policy $\pi \in \Pi^{SD}$, P^π has spectral radius

$$\rho\left(P^\pi\right) \quad \leq \quad \left(\frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha}\right)^{\frac{1}{\overline{M}+1}} \quad < \quad 1$$

A different approach is followed in this chapter which exhibits the nature of these algorithms in a more intuitive way.

The value iteration algorithm, including a number of versions of the basic successive approximation scheme, is discussed in Section 11.1. In Section 11.2 the modified policy iteration algorithm is studied. Both value iteration and policy iteration are special cases of modified policy iteration. In Section 11.3 a more efficient algorithm, called the two-decisions algorithm, is developed. This algorithm exploits the property of the DSDP that there are two types of decisions which have to be made in the DSDP, namely acceptance decisions and dispatching decisions, and that given a value function, an optimal acceptance rule is much easier to compute than an optimal dispatching rule, thanks to the simple optimal threshold acceptance rule. The efficiency of these algorithms is compared in Section 11.4, and a numerical example is given.

11.1 Value Iteration

Jacobi and Gauss-Seidel versions of value iteration are studied. Jacobi methods use the values of the previous iteration throughout the next iteration, whereas Gauss-Seidel methods use the updated values as soon as they become available. Pre-Jacobi and pre-Gauss-Seidel methods compute the value of the current state in the next iteration using the value of the current state obtained from the previous iteration on the right-hand-side of the equation. Post-Jacobi and post-Gauss-Seidel methods "move" the value of the current state from the right-hand-side to the left-hand-side of the equation, thereby eliminating the diagonal entries from the right-hand-side.

Let $\mathcal{V}$ denote the set of value functions on the state space $\mathcal{X}$, i.e., $\mathcal{V} \equiv \{V : \mathcal{X} \mapsto \Re\}$. For any $V \in \mathcal{V}$ and any $\pi \in \Pi^{SD}$, let

$$\mathcal{M}^\pi V \equiv r^\pi + P^\pi V$$

and let

$$\mathcal{M}V \equiv \sup_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V\}$$

From Equation (10.5), $V^\pi = r^\pi + P^\pi V^\pi = \mathcal{M}^\pi V^\pi$, and thus V^π is a fixed point of $\mathcal{M}^\pi$. $\mathcal{M}$ is the pre-Jacobi value iteration operator.

11.1.1 Jacobi Value Iteration

Pre-Jacobi Value Iteration Algorithm

begin

choose $V_0 \in \mathcal{V}$ and $\varepsilon > 0$, and set $i = 0$ and $err = \varepsilon + 1$;

while $err > \varepsilon$

$$V_{i+1} = \max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_i\}; \qquad (11.1)$$

$$\pi_{i+1} \in \arg\max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_i\}; \qquad (11.2)$$

$$err = \|V_{i+1} - V_i\|_\infty;$$

$$i = i + 1;$$

endwhile;

choose policy $\pi_\varepsilon = \pi_i$ and stop;

end;

Pre-Jacobi value iteration uses Equation (10.2) to update the value function in Equation (11.1), whereas post-Jacobi value iteration uses Equation (10.4). Post-Jacobi value iteration can also be expressed by "splitting" $I - P^\pi$ into $I - P^\pi_\Delta - P^\pi_O$, where P^π_Δ is the diagonal part of P^π and P^π_O is the off-diagonal part of P^π. The update step of Equation (11.1) can then be expressed as

$$V_{i+1} \quad = \quad \max_{\pi \in \Pi^{SD}} \left\{ \left(I - P^\pi_\Delta\right)^{-1} r^\pi + \left(I - P^\pi_\Delta\right)^{-1} P^\pi_O V_i \right\}. \tag{11.3}$$

and the update step of Equation (11.2) as

$$\pi_{i+1} \quad \in \quad \arg\max_{\pi \in \Pi^{SD}} \left\{ \left(I - P^\pi_\Delta\right)^{-1} r^\pi + \left(I - P^\pi_\Delta\right)^{-1} P^\pi_O V_i \right\}$$

For any given value function V, the simple threshold rule

$$D(m, \tilde{m}, n, r, b) \quad \equiv \quad \begin{cases} 1 \quad \text{if} & \sum_{l \in \mathcal{K}} s(k,l)\left[n(k,l) + b(k,l)\right] \leq \overline{N}(k) \text{ for all } k \in \mathcal{K}, \\ & \text{and } r \geq V(m, \tilde{m}, n) - V(m, \tilde{m}, n+b) - p \\ 0 \quad \text{if} & \sum_{l \in \mathcal{K}} s(k,l)\left[n(k,l) + b(k,l)\right] > \overline{N}(k) \text{ for some } k \in \mathcal{K}, \\ & \text{or } r < V(m, \tilde{m}, n) - V(m, \tilde{m}, n+b) - p \end{cases} \tag{11.4}$$

maximizes the right-hand-side of Equation (10.2), and thus gives an optimal acceptance rule for the maximization in Equation (11.1) when performing pre-Jacobi value iteration. However, the threshold rule of Equation (11.4) does not necessarily maximize the right-hand-side of Equation (10.4), and therefore using post-Jacobi value iteration might involve a search for an optimal acceptance rule when performing the maximization in Equation (11.3).

Lemma 39

1. *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \leq V$, $V^* \leq V$.*
2. *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $V^* \geq V$.*

Proof: Suppose $\mathcal{M}V \leq V$. Then for any $\pi \in \Pi^{SD}$, and any $i \geq 1$

$$\begin{aligned} V \quad &\geq \quad r^\pi + P^\pi V \quad \geq \quad r^\pi + P^\pi\left(r^\pi + P^\pi V\right) \\ &\geq \quad \cdots \quad \geq \quad \sum_{j=0}^{i-1} \left(P^\pi\right)^j r^\pi \; + \; \left(P^\pi\right)^i V \end{aligned}$$

$$\Rightarrow \quad V - V^\pi \quad \geq \quad \left(P^\pi\right)^i V \; - \; \sum_{j=i}^{\infty} \left(P^\pi\right)^j r^\pi$$

Let $\nu_1, \ldots, \nu_z$ be the eigenvalues of P^π, with corresponding eigenvectors $v_1, \ldots, v_z$, where $z \equiv |\mathcal{X}|$. Then $(P^\pi)^i V = \sum_{k=1}^{z} c_k \nu_k^i v_k$, and

$$\begin{aligned}
\left\| (P^\pi)^i V \right\|_\infty &\leq \sum_{k=1}^{z} |c_k| \, |\nu_k|^i \, \|v_k\|_\infty \\
&\leq \sum_{k=1}^{z} |c_k| \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \right)^{\frac{i}{\overline{M}+1}} \|v_k\|_\infty \\
&= c \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \right)^{\frac{i}{\overline{M}+1}}
\end{aligned}$$

where $c = \sum_{k=1}^{z} |c_k| \, \|v_k\|_\infty$. Hence, for any $\varepsilon > 0$ there exists an N_1 such that $(P^\pi)^i V > -\varepsilon \mathbf{1}$ for all $i \geq N_1$, where $\mathbf{1}$ denotes a vector of 1's. Also, $\sum_{j=i}^{\infty} (P^\pi)^j r^\pi = \sum_{j=i}^{\infty} \sum_{k=1}^{z} d_k \nu_k^j v_k$. Thus

$$\begin{aligned}
\left\| \sum_{j=i}^{\infty} (P^\pi)^j r^\pi \right\|_\infty &\leq \sum_{j=i}^{\infty} \sum_{k=1}^{z} |d_k| \, |\nu_k|^j \, \|v_k\|_\infty \\
&\leq \sum_{j=i}^{\infty} \sum_{k=1}^{z} |d_k| \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \right)^{\frac{j}{\overline{M}+1}} \|v_k\|_\infty \\
&= d \, \frac{\left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \right)^{\frac{i}{\overline{M}+1}}}{1 - \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \right)^{\frac{1}{\overline{M}+1}}}
\end{aligned}$$

where $d = \sum_{k=1}^{z} |d_k| \, \|v_k\|_\infty$. Hence, for any $\varepsilon > 0$ there exists an N_2 such that $-\sum_{j=i}^{\infty} (P^\pi)^j r^\pi > -\varepsilon \mathbf{1}$ for all $i \geq N_2$. Therefore, for any $\pi \in \Pi^{SD}$, $V > V^\pi - 2\varepsilon \mathbf{1}$. Because $\varepsilon > 0$ is arbitrary, $V \geq V^\pi$ for any $\pi \in \Pi^{SD}$, and thus $V \geq V^*$.

Suppose $\mathcal{M}V \geq V$. Then there exists $\pi \in \Pi^{SD}$ such that $V \leq r^\pi + P^\pi V$. It follows from Corollaries 10 and 11 that $V \leq (I - P^\pi)^{-1} r^\pi = V^\pi$. Therefore $V \leq V^*$, and the result follows. □

Corollary 13 *If $\mathcal{M}\hat{V} = \hat{V}$, then $\hat{V} = V^*$.*

Lemma 40 *$\mathcal{M} : \mathcal{V} \mapsto \mathcal{V}$ is an $(\overline{M}+1)$-stage contraction mapping with constant $(\lambda+\overline{\mu})/(\lambda+\overline{\mu}+\alpha)$, i.e., for any $V_1, V_2 \in \mathcal{V}$,*

$$\left\| \mathcal{M}^{\overline{M}+1} V_2 - \mathcal{M}^{\overline{M}+1} V_1 \right\|_\infty \leq \frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \|V_2 - V_1\|_\infty$$

Proof: Let $V_1, V_2 \in \mathcal{V}$. For any $x_0 = (m, \tilde{m}, n) \in \mathcal{X}$, consider the following two cases:

Case 1. $\mathcal{M}^{\overline{M}+1}V_2(x_0) \geq \mathcal{M}^{\overline{M}+1}V_1(x_0)$:

Let $\pi_2 \in \arg\max_{\pi\in\Pi^{SD}} \left\{r^\pi + P^\pi \mathcal{M}^{\overline{M}} V_2\right\}$. Then

$$\begin{aligned}\left|\mathcal{M}^{\overline{M}+1}V_2(x_0) - \mathcal{M}^{\overline{M}+1}V_1(x_0)\right| &\leq \left[r^{\pi_2}(x_0) + \left(P^{\pi_2}\mathcal{M}^{\overline{M}}V_2\right)(x_0)\right] \\ &\quad - \left[r^{\pi_2}(x_0) + \left(P^{\pi_2}\mathcal{M}^{\overline{M}}V_1\right)(x_0)\right] \\ &= \left(P^{\pi_2}\left(\mathcal{M}^{\overline{M}}V_2 - \mathcal{M}^{\overline{M}}V_1\right)\right)(x_0)\end{aligned}$$

If $L^{\pi_2}(x_0) = \emptyset$, then

$$\left(P^{\pi_2}\left(\mathcal{M}^{\overline{M}}V_2 - \mathcal{M}^{\overline{M}}V_1\right)\right)(x_0) \leq \frac{\lambda + \mu(\tilde{m}(x_0))}{\lambda + \mu(\tilde{m}(x_0)) + \alpha}\left\|\mathcal{M}^{\overline{M}}V_2 - \mathcal{M}^{\overline{M}}V_1\right\|_\infty$$

If $L^{\pi_2}(x_0) = (k, l)$ and $Q^{\pi_2}(x_0) = q$, then let $x_1 = (m - e(k), \tilde{m} + e(k, l), n - e(k, l)q)$. Then

$$\left(P^{\pi_2}\left(\mathcal{M}^{\overline{M}}V_2 - \mathcal{M}^{\overline{M}}V_1\right)\right)(x_0) = \mathcal{M}^{\overline{M}}V_2(x_1) - \mathcal{M}^{\overline{M}}V_1(x_1) \leq \left\|\mathcal{M}^{\overline{M}}V_2 - \mathcal{M}^{\overline{M}}V_1\right\|_\infty$$

Note that in state x_1 there is one vehicle less at a terminal than in state x_0. The argument above can be repeated with state x_1. Because at most $\overline{M}$ vehicles can be dispatched, there exists a $z \in \{0, 1, \ldots, \overline{M}\}$ such that $L^{\pi_2}(x_i) \neq \emptyset$ for all $i = 0, 1, \ldots, z-1$ and $L^{\pi_2}(x_z) = \emptyset$. Therefore

$$\begin{aligned}\left|\mathcal{M}^{\overline{M}+1}V_2(x_0) - \mathcal{M}^{\overline{M}+1}V_1(x_0)\right| &\leq \mathcal{M}^{\overline{M}}V_2(x_1) - \mathcal{M}^{\overline{M}}V_1(x_1) \\ &\leq \ldots \leq \mathcal{M}^{\overline{M}-z+1}V_2(x_z) - \mathcal{M}^{\overline{M}-z+1}V_1(x_z) \\ &\leq \frac{\lambda + \mu(\tilde{m}(x_z))}{\lambda + \mu(\tilde{m}(x_z)) + \alpha}\left\|\mathcal{M}^{\overline{M}-z}V_2 - \mathcal{M}^{\overline{M}-z}V_1\right\|_\infty \\ &\leq \ldots \leq \frac{\lambda + \mu(\tilde{m}(x_z))}{\lambda + \mu(\tilde{m}(x_z)) + \alpha}\|V_2 - V_1\|_\infty \\ &\leq \frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha}\|V_2 - V_1\|_\infty\end{aligned}$$

Case 2. $\mathcal{M}^{\overline{M}+1}V_2(x_0) < \mathcal{M}^{\overline{M}+1}V_1(x_0)$:

Repeat the argument for Case 1 with policy $\pi_1 \in \arg\max_{\pi\in\Pi^{SD}} \left\{r^\pi + P^\pi\mathcal{M}^{\overline{M}}V_1\right\}$. Thus,

$$\left\|\mathcal{M}^{\overline{M}+1}V_2 - \mathcal{M}^{\overline{M}+1}V_1\right\|_\infty \leq \frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha}\|V_2 - V_1\|_\infty \qquad \square$$

Theorem 12 *$\mathcal{M}$ has a unique fixed point equal to V^*, and for any $V \in \mathcal{V}$, $\mathcal{M}^i V \to V^*$ in $\|\cdot\|_\infty$-norm as $i \to \infty$.*

Proof: From the Banach fixed point theorem, if $\mathcal{M} : \mathcal{V} \mapsto \mathcal{V}$ is an $(\overline{M}+1)$-stage contraction mapping, then $\mathcal{M}$ has a unique fixed point $\hat{V} \in \mathcal{V}$, i.e., $\mathcal{M}\hat{V} = \hat{V}$, and for any $V \in \mathcal{V}$, $\mathcal{M}^i V \to \hat{V}$ as $i \to \infty$. It follows from Corollary 13 that $\hat{V} = V^*$. □

It follows in a similar way that for any $\pi \in \Pi^{SD}$, $\mathcal{M}^\pi$ is an $(\overline{M}+1)$-stage contraction mapping, with unique fixed point V^π.

Two important monotonicity properties are established next.

Lemma 41 *$\mathcal{M}$ is nondecreasing, i.e., for any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$, $\mathcal{M}^i V_2 \geq \mathcal{M}^i V_1$ for all $i \geq 0$.*

Proof: By induction on i. For $i = 0$, $\mathcal{M}^i V_2 = V_2 \geq V_1 = \mathcal{M}^i V_1$. As induction hypothesis, suppose $\mathcal{M}^i V_2 \geq \mathcal{M}^i V_1$. For any $\pi \in \Pi^{SD}$,

$$\begin{aligned}
& & r^\pi + P^\pi \mathcal{M}^i V_2 &\geq r^\pi + P^\pi \mathcal{M}^i V_1 \\
\Rightarrow \quad \mathcal{M}\mathcal{M}^i V_2 &\equiv \sup_{\pi \in \Pi^{SD}} \left\{ r^\pi + P^\pi \mathcal{M}^i V_2 \right\} &&\geq r^\pi + P^\pi \mathcal{M}^i V_1 \\
\Rightarrow \quad \mathcal{M}^{i+1} V_2 &\geq \sup_{\pi \in \Pi^{SD}} \left\{ r^\pi + P^\pi \mathcal{M}^i V_1 \right\} &&\equiv \mathcal{M}^{i+1} V_1
\end{aligned}$$ □

Similarly, for any $\pi \in \Pi^{SD}$, $\mathcal{M}^\pi$ is nondecreasing, i.e., for any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$, $(\mathcal{M}^\pi)^i V_2 \geq (\mathcal{M}^\pi)^i V_1$ for all $i \geq 0$.

Lemma 42 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}^i V \leq \mathcal{M}^{i+1} V \leq V^*$ for all $i \geq 0$.*

Proof: If $\mathcal{M}V \geq V$, then $\mathcal{M}^{i+1} V = \mathcal{M}^i \mathcal{M} V \geq \mathcal{M}^i V$ for all $i \geq 0$. Thus $\mathcal{M}\mathcal{M}^i V \geq \mathcal{M}^i V$, which implies $\mathcal{M}^i V \leq V^*$ by Lemma 39. □

Similarly, if $\mathcal{M}V \leq V$, then $\mathcal{M}^i V \geq \mathcal{M}^{i+1} V \geq V^*$ for all $i \geq 0$.

An error bound on the value function and the policy obtained when the pre-Jacobi value iteration algorithm stops, is provided next.

Proposition 7 *If the pre-Jacobi value iteration algorithm stops when $\|V_{i+1} - V_i\|_\infty < \varepsilon$, then*

$$\|V_{i+1} - V^*\|_\infty \quad < \quad \frac{\lambda + \overline{\mu} + \overline{M}\,(\lambda + \overline{\mu} + \alpha)}{\alpha}\,\varepsilon$$

and

$$\|V^{\pi_\varepsilon} - V^*\|_\infty \quad < \quad 2\,\frac{\lambda + \overline{\mu} + \overline{M}\,(\lambda + \overline{\mu} + \alpha)}{\alpha}\,\varepsilon$$

Proof:

$$\begin{aligned}
\|V_{i+1} - V^*\|_\infty &= \lim_{j\to\infty} \|V_{i+1} - V_j\|_\infty \\
&\le \sum_{j=0}^{\infty} \left\| \mathcal{M}^{(j+1)(\overline{M}+1)} V_{i+1} - \mathcal{M}^{j(\overline{M}+1)} V_{i+1} \right\|_\infty \\
&\le \sum_{j=0}^{\infty} \left(\frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha}\right)^j \left\| \mathcal{M}^{\overline{M}+1} V_{i+1} - V_{i+1} \right\|_\infty \\
&\le \frac{\lambda+\overline{\mu}+\alpha}{\alpha} \sum_{j=0}^{\overline{M}} \left\| \mathcal{M}^{j+1} V_{i+1} - \mathcal{M}^{j} V_{i+1} \right\|_\infty \\
&= \frac{\lambda+\overline{\mu}+\alpha}{\alpha} \sum_{j=0}^{\overline{M}} \left\| \mathcal{M}^{j+1} V_{i+1} - \mathcal{M}^{j+1} V_{i} \right\|_\infty \\
&\le \frac{\lambda+\overline{\mu}+\alpha}{\alpha} \left(\frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha} + \overline{M} \right) \|V_{i+1} - V_i\|_\infty \\
&< \frac{\lambda + \overline{\mu} + \overline{M}\,(\lambda + \overline{\mu} + \alpha)}{\alpha}\,\varepsilon
\end{aligned}$$

$$\begin{aligned}
\|V^{\pi_i} - V_{i+1}\|_\infty &\le \left\| \mathcal{M}^{\pi_i} V^{\pi_i} - (\mathcal{M}^{\pi_i})^{\overline{M}+1} V_{i+1} \right\|_\infty \\
&\quad + \left\| (\mathcal{M}^{\pi_i})^{\overline{M}+1} V_{i+1} - V_{i+1} \right\|_\infty \\
&\le \left\| (\mathcal{M}^{\pi_i})^{\overline{M}+1} V^{\pi_i} - (\mathcal{M}^{\pi_i})^{\overline{M}+1} V_{i+1} \right\|_\infty \\
&\quad + \sum_{j=0}^{\overline{M}} \left\| (\mathcal{M}^{\pi_i})^{j+1} V_{i+1} - (\mathcal{M}^{\pi_i})^{j} V_{i+1} \right\|_\infty \\
&= \left\| (\mathcal{M}^{\pi_i})^{\overline{M}+1} V^{\pi_i} - (\mathcal{M}^{\pi_i})^{\overline{M}+1} V_{i+1} \right\|_\infty \\
&\quad + \sum_{j=0}^{\overline{M}} \left\| (\mathcal{M}^{\pi_i})^{j+1} V_{i+1} - (\mathcal{M}^{\pi_i})^{j+1} V_{i} \right\|_\infty \\
&\le \frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha} \|V^{\pi_i} - V_{i+1}\|_\infty \;+\; \left(\frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha} + \overline{M} \right) \|V_{i+1} - V_i\|_\infty
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad \frac{\alpha}{\lambda+\overline{\mu}+\alpha} \|V^{\pi_i} - V_{i+1}\|_\infty &\le \left(\frac{\lambda+\overline{\mu}}{\lambda+\overline{\mu}+\alpha} + \overline{M} \right) \|V_{i+1} - V_i\|_\infty \\
\Rightarrow \quad \|V^{\pi_i} - V_{i+1}\|_\infty &< \frac{\lambda + \overline{\mu} + \overline{M}\,(\lambda + \overline{\mu} + \alpha)}{\alpha}\,\varepsilon \\
\Rightarrow \quad \|V^{\pi_i} - V^*\|_\infty &\le \|V^{\pi_i} - V_{i+1}\|_\infty \;+\; \|V_{i+1} - V^*\|_\infty \\
&< 2\frac{\lambda + \overline{\mu} + \overline{M}\,(\lambda + \overline{\mu} + \alpha)}{\alpha}\,\varepsilon
\end{aligned}$$

□

11.1.2 Gauss-Seidel Value Iteration

A Gauss-Seidel value iteration algorithm uses the most up-to-date values of the value function V available. That is, with the states $x \in \mathcal{X}$ ordered $x_1, x_2, \ldots, x_{|\mathcal{X}|}$, the update step of pre-Gauss-Seidel value iteration corresponding to Equation (11.1), is

$$V_{i+1}(x_j) = \max_{a \in \Psi(x_j)} \left\{ r(x_j, a) + \sum_{k<j} P[x_k \mid x_j, a] V_{i+1}(x_k) + \sum_{k \geq j} P[x_k \mid x_j, a] V_i(x_k) \right\}$$

where $\Psi(x_j)$ denotes the set of admissible actions in state x_j. This can also be expressed by "splitting" $I - P^\pi$ into $I - P_L^\pi - P_U^\pi$, where P_L^π is the strictly lower triangular part of P^π and P_U^π is the upper triangular part of P^π. For post-Gauss-Seidel value iteration, P_L^π is the lower triangular part of P^π and P_U^π is the strictly upper triangular part of P^π. The update step in Equation (11.1) can then be expressed as

$$V_{i+1} = \max_{\pi \in \Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V_i \right\} \tag{11.5}$$

and the update step in Equation (11.2) as

$$\pi_{i+1} \in \arg\max_{\pi \in \Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V_i \right\}$$

When performing pre-Gauss-Seidel value iteration, the threshold rule of Equation (11.4) can again be used to maximize the right-hand-side of Equation (10.2), and thus to give an optimal acceptance rule for the maximization in Equation (11.5). Similar to post-Jacobi value iteration, the threshold rule of Equation (11.4) does not necessarily maximize the right-hand-side of Equation (10.4), and therefore post-Gauss-Seidel value iteration might involve a cumbersome search for an optimal acceptance rule when performing the maximization in Equation (11.5). Hence, unless noted otherwise, the rest of the analysis in this section considers pre-Gauss-Seidel value iteration.

For any $V \in \mathcal{V}$ and any $\pi \in \Pi^{SD}$, let

$$\mathcal{M}_G^\pi V \equiv (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V$$

and let

$$\mathcal{M}_{GV} V \equiv \sup_{\pi \in \Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V \right\}$$

From Equation (10.5), $V^\pi = (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V^\pi = \mathcal{M}_G^\pi V^\pi$, and thus V^π is a fixed point of $\mathcal{M}_G^\pi$.

Lemma 43 *$\mathcal{M}_{GV}$ is nondecreasing, i.e., for any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$, $\mathcal{M}_{GV}^i V_2 \geq \mathcal{M}_{GV}^i V_1$ for all $i \geq 0$.*

Proof: By induction on i. For $i = 0$, $\mathcal{M}_{GV}^i V_2 = V_2 \geq V_1 = \mathcal{M}_{GV}^i V_1$. As induction hypothesis, suppose $\mathcal{M}_{GV}^i V_2 \geq \mathcal{M}_{GV}^i V_1$. For any $\pi \in \Pi^{SD}$, $\rho(P_L^\pi) = 0 < 1$. Thus, from Corollary 9, $(I - P_L^\pi)^{-1} = \sum_{i=0}^{|\mathcal{X}|} (P_L^\pi)^i \geq 0$. Also, $P_U^\pi \geq 0$. Hence, for any $\pi \in \Pi^{SD}$,

$$\begin{aligned}
(I - P_L^\pi)^{-1} P_U^\pi \mathcal{M}_{GV}^i V_2 &\geq (I - P_L^\pi)^{-1} P_U^\pi \mathcal{M}_{GV}^i V_1 \\
\Rightarrow \quad \mathcal{M}_{GV}^{i+1} V_2 &\equiv \sup_{\pi \in \Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi \mathcal{M}_{GV}^i V_2 \right\} \\
&\geq \sup_{\pi \in \Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi \mathcal{M}_{GV}^i V_1 \right\} \\
&\equiv \mathcal{M}_{GV}^{i+1} V_1 \qquad \square
\end{aligned}$$

Similarly, for any $\pi \in \Pi^{SD}$, $\mathcal{M}_G^\pi$ is nondecreasing, i.e., for any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$, $(\mathcal{M}_G^\pi)^i V_2 \geq (\mathcal{M}_G^\pi)^i V_1$ for all $i \geq 0$.

Lemma 44 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}_{GV}^i V \geq \mathcal{M}^i V$ for all $i \geq 0$.*

Proof: By induction on i. For $i = 0$, $\mathcal{M}_{GV}^i V = V = \mathcal{M}^i V$. As induction hypothesis on i, suppose that $\mathcal{M}_{GV}^i V \geq \mathcal{M}^i V$. From Lemma 42, if $\mathcal{M}V \geq V$, then $\mathcal{M}^{i+1} V \geq \mathcal{M}^i V$ for all $i \geq 0$. It is shown by induction on j that $\mathcal{M}_{GV}^{i+1} V(x_j) \geq \mathcal{M}^{i+1} V(x_j)$ for all $x_j \in \mathcal{X}$.

$$\begin{aligned}
\mathcal{M}_{GV}^{i+1} V(x_1) &= \sup_{a \in \Psi(x_1)} \left\{ r(x_1, a) + \sum_{x_k \in \mathcal{X}} P[x_k \mid x_1, a] \mathcal{M}_{GV}^i V(x_k) \right\} \\
&= \mathcal{M}\mathcal{M}_{GV}^i V(x_1) \;\geq\; \mathcal{M}\mathcal{M}^i V(x_1) \;=\; \mathcal{M}^{i+1} V(x_1)
\end{aligned}$$

As induction hypothesis on j, suppose that $\mathcal{M}_{GV}^{i+1} V(x_k) \geq \mathcal{M}^{i+1} V(x_k)$ for all $k \in \{1, \ldots, j-1\}$. Then

$$\begin{aligned}
\mathcal{M}_{GV}^{i+1} V(x_j) &= \sup_{a \in \Psi(x_j)} \Bigg\{ r(x_j, a) + \sum_{k<j} P[x_k \mid x_j, a] \mathcal{M}_{GV}^{i+1} V(x_k) \\
&\qquad + \sum_{k \geq j} P[x_k \mid x_j, a] \mathcal{M}_{GV}^i V(x_k) \Bigg\} \\
&\geq \sup_{a \in \Psi(x_j)} \Bigg\{ r(x_j, a) + \sum_{k<j} P[x_k \mid x_j, a] \mathcal{M}^{i+1} V(x_k) \\
&\qquad + \sum_{k \geq j} P[x_k \mid x_j, a] \mathcal{M}_{GV}^i V(x_k) \Bigg\}
\end{aligned}$$

$$\begin{aligned}
&\geq \sup_{a\in\Psi(x_j)}\left\{r(x_j,a) + \sum_{k<j} P[x_k \mid x_j,a]\mathcal{M}^i V(x_k)\right. \\
&\qquad\qquad \left. + \sum_{k\geq j} P[x_k \mid x_j,a]\mathcal{M}^i V(x_k)\right\} \\
&= \mathcal{M}\mathcal{M}^i V(x_j) \quad = \quad \mathcal{M}^{i+1}V(x_j)
\end{aligned}$$

Therefore $\mathcal{M}_{GV}^{i+1}V \geq \mathcal{M}^{i+1}V$, and the induction has been established. □

Similarly, for any $V \in \mathcal{V}$ such that $\mathcal{M}V \leq V$, $\mathcal{M}_{GV}^i V \leq \mathcal{M}^i V$ for all $i \geq 0$.

Lemma 45 *V^* is a fixed point of $\mathcal{M}_{GV}$.*

Proof: From Proposition 12, $\mathcal{M}V^* = V^*$. Then from Lemma 44, $\mathcal{M}_{GV}^i V^* = \mathcal{M}^i V^* = V^*$ for all $i \geq 0$, thus $\mathcal{M}_{GV}V^* = V^*$. □

Lemma 46 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}_{GV}^i V \leq V^*$ for all $i \geq 0$.*

Proof: From Lemma 39, if $\mathcal{M}V \geq V$, then $V \leq V^*$. From Lemmas 43 and 45, $\mathcal{M}_{GV}^i V \leq \mathcal{M}_{GV}^i V^* = V^*$ for all $i \geq 0$. □

Similarly, for any $V \in \mathcal{V}$ such that $\mathcal{M}V \leq V$, $\mathcal{M}_{GV}^i V \geq V^*$ for all $i \geq 0$.

Lemma 47 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}_{GV}^i V \leq \mathcal{M}_{GV}^{i+1}V \leq V^*$ for all $i \geq 0$.*

Proof: From Lemma 44, if $\mathcal{M}V \geq V$, then $\mathcal{M}_{GV}V \geq \mathcal{M}V \geq V$. Then $\mathcal{M}_{GV}^{i+1}V = \mathcal{M}_{GV}^i \mathcal{M}_{GV}V \geq \mathcal{M}_{GV}^i V$ for all $i \geq 0$, from Lemma 43. Also, $\mathcal{M}_{GV}^i V \leq V^*$ for all $i \geq 0$, from Lemma 46. □

Similarly, for any $V \in \mathcal{V}$ such that $\mathcal{M}V \leq V$, $\mathcal{M}_{GV}^i V \geq \mathcal{M}_{GV}^{i+1}V \geq V^*$ for all $i \geq 0$.

Theorem 13 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}_{GV}^i V \uparrow V^*$ in $\|\cdot\|_\infty$-norm as $i \to \infty$.*

Proof: From Lemmas 44 and 46, if $\mathcal{M}V \geq V$, then $\mathcal{M}^i V \leq \mathcal{M}_{GV}^i V \leq V^*$ for all $i \geq 0$. From Proposition 12, $\mathcal{M}^i V \to V^*$ in $\|\cdot\|_\infty$-norm as $i \to \infty$. Therefore $\mathcal{M}_{GV}^i V \to V^*$ in $\|\cdot\|_\infty$-norm as $i \to \infty$. From Lemma 47, $\{\mathcal{M}_{GV}^i V\}_{i=0}^\infty$ is nondecreasing. □

Similarly, for any $V \in \mathcal{V}$ such that $\mathcal{M}V \leq V$, $\mathcal{M}_{GV}^i V \downarrow V^*$ in $\|\cdot\|_\infty$-norm as $i \to \infty$.

In Puterman [139], Theorem 6.3.9, it is shown that for any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$,

$$\limsup_{i \to \infty} \left(\frac{\|\mathcal{M}_{GV}^i V - V^*\|_\infty}{\|V - V^*\|_\infty} \right)^{1/i} \leq \rho\left((I - P_L^*)^{-1} P_U^* \right)$$

For pre-Gauss-Seidel value iteration, P_L^π is strictly lower triangular, thus $I - P_L^\pi$ is lower triangular with 1's on the main diagonal. Hence $(I - P_L^\pi)^{-1}$ is lower triangular with 1's on the main diagonal, and $\rho\left((I - P_L^\pi)^{-1}\right) = 1$ for all $\pi \in \Pi^{SD}$. P_U^π is upper triangular with diagonal entries

$$\begin{aligned} P_U^\pi\left[(m,\tilde{m},n),(m,\tilde{m},n)\right] &= \frac{\lambda}{\lambda + \mu(\tilde{m}) + \alpha} \int\int_{\mathcal{R}_0^\pi(m,\tilde{m},n)} F_{R|B}(dr|b) F_B(db) \\ &\leq \frac{\lambda}{\lambda + \mu(\tilde{m}) + \alpha} \leq \frac{\lambda}{\lambda + \alpha} \end{aligned}$$

if $L^\pi(m,\tilde{m},n) = \emptyset$, and $P_U^\pi[(m,\tilde{m},n),(m,\tilde{m},n)] = 0$ if $L^\pi(m,\tilde{m},n) \neq \emptyset$. Hence $\rho(P_U^\pi) \leq \lambda/(\lambda+\alpha)$ for all $\pi \in \Pi^{SD}$, and

$$\rho\left((I - P_L^*)^{-1} P_U^* \right) \leq \rho\left((I - P_L^*)^{-1} \right) \rho\left(P_U^*\right) \leq \frac{\lambda}{\lambda + \alpha}$$

Note that this is a better bound than for $\rho(P^\pi)$, because

$$\frac{\lambda}{\lambda+\alpha} \leq \frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \leq \left(\frac{\lambda + \overline{\mu}}{\lambda + \overline{\mu} + \alpha} \right)^{\frac{1}{\overline{M}+1}}$$

Therefore,

$$\limsup_{i \to \infty} \left(\frac{\|\mathcal{M}_{GV}^i V - V^*\|_\infty}{\|V - V^*\|_\infty} \right)^{1/i} \leq \frac{\lambda}{\lambda+\alpha} < 1$$

which means that for any $\varepsilon > 0$, there exists an N_ε such that

$$\left\|\mathcal{M}_{GV}^i V - V^*\right\|_\infty < \left(\frac{\lambda}{\lambda+\alpha} + \varepsilon \right)^i \|V - V^*\|_\infty$$

for all $i \geq N_\varepsilon$. Note that because $\lambda/(\lambda+\alpha) < 1$, ε can be chosen such that $\lambda/(\lambda+\alpha)+\varepsilon < 1$, to obtain convergence that is geometric at rate $\lambda/(\lambda+\alpha)+\varepsilon$.

For post-Gauss-Seidel value iteration, P_L^π is lower triangular with diagonal entries

$$\begin{aligned} P_L^\pi\left[(m,\tilde{m},n),(m,\tilde{m},n)\right] &= \frac{\lambda}{\lambda + \mu(\tilde{m}) + \alpha} \int\int_{\mathcal{R}_0^\pi(m,\tilde{m},n)} F_{R|B}(dr|b) F_B(db) \\ &\leq \frac{\lambda}{\lambda + \mu(\tilde{m}) + \alpha} \leq \frac{\lambda}{\lambda + \alpha} \end{aligned}$$

if $L^\pi(m,\tilde{m},n) = \emptyset$, and $P_L^\pi[(m,\tilde{m},n),(m,\tilde{m},n)] = 0$ if $L^\pi(m,\tilde{m},n) \neq \emptyset$. Thus $I - P_L^\pi$ is lower triangular with diagonal entries $\alpha/(\lambda+\alpha) \leq 1 - P_L^\pi[(m,\tilde{m},n),(m,\tilde{m},n)] \leq 1$. Hence

$(I-P_L^\pi)^{-1}$ is lower triangular with diagonal entries $1 \leq (1-P_L^\pi[(m,\tilde{m},n),(m,\tilde{m},n)])^{-1} \leq (\lambda+\alpha)/\alpha$, for all $\pi \in \Pi^{SD}$. P_U^π is strictly upper triangular; hence $\rho(P_U^\pi) = 0$ for all $\pi \in \Pi^{SD}$, and $\rho\left((I-P_L^*)^{-1} P_U^*\right) \leq \rho\left((I-P_L^*)^{-1}\right)\rho\left(P_U^*\right) = 0$. Therefore,

$$\limsup_{i\to\infty} \left(\frac{\|\mathcal{M}_{GV}^i V - V^*\|_\infty}{\|V-V^*\|_\infty}\right)^{1/i} = 0$$

which means that for any $\varepsilon > 0$, there exists an N_ε such that

$$\left\|\mathcal{M}_{GV}^i V - V^*\right\|_\infty < \varepsilon^i \|V-V^*\|_\infty$$

for all $i \geq N_\varepsilon$. This means that $\mathcal{M}_{GV}^i V$ converges to V^* faster than any geometric rate. However, it should be taken into account that post-Gauss-Seidel value iteration might require more computational effort to find an optimal acceptance rule than pre-Gauss-Seidel value iteration when performing the maximization in Equation (11.5).

The same error bound as in Proposition 7 applies if pre-Gauss-Seidel value iteration stops when $\|V_{i+1}-V_i\|_\infty < \varepsilon$.

Another variation on the value iteration algorithm is to use relaxation, where the value function is updated as follows for pre-Jacobi value iteration

$$V_{i+1} = \omega \sup_{\pi\in\Pi^{SD}} \{r^\pi + P^\pi V_i\} + (1-\omega)V_i$$

or as follows for Gauss-Seidel value iteration

$$V_{i+1} = \omega \sup_{\pi\in\Pi^{SD}} \left\{(I-P_L^\pi)^{-1} r^\pi + (I-P_L^\pi)^{-1} P_U^\pi V_i\right\} + (1-\omega)V_i$$

where $\omega \in (0,2)$ is a relaxation factor.

11.2 Modified Policy Iteration

Since policy iteration (and value iteration) are special cases of modified policy iteration, modified policy iteration is studied next. Jacobi and Gauss-Seidel versions can again be distinguished, depending on how the updated values are used.

11.2.1 Jacobi Modified Policy Iteration

Pre-Jacobi Modified Policy Iteration Algorithm

begin

choose $V_{0,0} \in \mathcal{V}$ such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, choose $\varepsilon > 0$,

choose a sequence of positive integers $\{z_i\}$, and set $i = 1$;

$$V_{1,1} = \max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_{0,0}\};$$

$$\pi_1 \in \arg\max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_{0,0}\};$$

$$err = \|V_{1,1} - V_{0,0}\|_\infty;$$

$$i = i + 1;$$

while $err > \varepsilon$

for $j = 1$ **to** $z_i - 1$

$$V_{i,j+1} = r^{\pi_i} + P^{\pi_i} V_{i,j}; \tag{11.6}$$

endfor;

$$V_{i+1,1} = \max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_{i,z_i}\}; \tag{11.7}$$

$$\pi_{i+1} \in \arg\max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_{i,z_i}\}; \tag{11.8}$$

$$err = \|V_{i+1,1} - V_{i,z_i}\|_\infty;$$

$$i = i + 1;$$

endwhile;

choose policy $\pi_\varepsilon = \pi_i$ and stop;

end;

Pre-Jacobi modified policy iteration uses Equation (10.2) for partial policy evaluation in Equation (11.6) and policy improvement in Equation (11.7), whereas post-Jacobi modified policy iteration uses Equation (10.4). The partial policy evaluation step of post-Jacobi modified policy iteration can be expressed as

$$V_{i,j+1} = \left\{(I - P_\Delta^{\pi_i})^{-1} r^{\pi_i} + (I - P_\Delta^{\pi_i})^{-1} P_O^{\pi_i} V_{i,j}\right\}$$

and the policy improvement step as

$$V_{i+1,1} = \max_{\pi \in \Pi^{SD}} \left\{(I - P_\Delta^{\pi})^{-1} r^{\pi} + (I - P_\Delta^{\pi})^{-1} P_O^{\pi} V_{i,z_i}\right\} \tag{11.9}$$

and

$$\pi_{i+1} \in \arg\max_{\pi \in \Pi^{SD}} \left\{(I - P_\Delta^{\pi})^{-1} r^{\pi} + (I - P_\Delta^{\pi})^{-1} P_O^{\pi} V_{i,z_i}\right\}$$

The threshold rule of Equation (11.4) maximizes the right-hand-side of Equation (10.2), and thus gives an optimal acceptance rule for the policy improvement step of pre-Jacobi modified policy iteration in Equation (11.7). However, the threshold rule of Equation (11.4) does not necessarily maximize the right-hand-side of Equation (10.4). Therefore, performing the maximization in Equation (11.9) might involve a search for an optimal acceptance rule, which makes the policy improvement step of post-Jacobi modified policy iteration more cumbersome than that of pre-Jacobi modified policy iteration. The policy evaluation steps of both should take comparable amounts of computing effort.

For any $V \in \mathcal{V}$, any $\pi_V \in \arg\max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V\}$, and any positive integer z, let

$$\mathcal{M}_{JM}^z V \equiv (\mathcal{M}^{\pi_V})^z V = \sum_{j=0}^{z-1} (P^{\pi_V})^j r^{\pi_V} + (P^{\pi_V})^z V$$

denote the pre-Jacobi modified policy iteration operator that performs one step of policy improvement and $z-1$ steps of partial policy evaluation.

Lemma 48 *For any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$, $\mathcal{M}^z V_2 \geq \mathcal{M}_{JM}^z V_1$ for all $z \geq 0$.*

Proof: By induction on z. For $z = 0$, $\mathcal{M}^z V_2 = V_2 \geq V_1 = \mathcal{M}_{JM}^z V_1$. As induction hypothesis, suppose $\mathcal{M}^z V_2 \geq \mathcal{M}_{JM}^z V_1$. Then

$$\begin{aligned}
\mathcal{M}^{z+1} V_2 &= \mathcal{M}\mathcal{M}^z V_2 \\
&\geq \mathcal{M}\mathcal{M}_{JM}^z V_1 \\
&= \sup_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi \mathcal{M}_{JM}^z V_1\} \\
&\geq r^{\pi_{V_1}} + P^{\pi_{V_1}} \mathcal{M}_{JM}^z V_1 \\
&= \mathcal{M}_{JM}^{z+1} V_1 \qquad \square
\end{aligned}$$

Lemma 49 *For any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$ and $\mathcal{M} V_2 \geq V_2$, $\mathcal{M}_{JM}^z V_2 \geq \mathcal{M} V_1$ for all $z \geq 1$.*

Proof:

$$\begin{aligned}
\mathcal{M}_{JM}^z V_2 &\equiv \sum_{j=0}^{z-1} (P^{\pi_{V_2}})^j r^{\pi_{V_2}} + (P^{\pi_{V_2}})^z V_2 \\
&= V_2 + \sum_{j=0}^{z-1} (P^{\pi_{V_2}})^j \left(r^{\pi_{V_2}} + P^{\pi_{V_2}} V_2 - V_2\right)
\end{aligned}$$

$$\begin{aligned} &= V_2 + \sum_{j=0}^{z-1} \left(P^{\pi_{V_2}}\right)^j \left(\mathcal{M}V_2 - V_2\right) \\ &\geq \mathcal{M}V_2 \quad \geq \quad \mathcal{M}V_1 \qquad \square \end{aligned}$$

Lemma 50 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}\mathcal{M}_{JM}^z V \geq \mathcal{M}_{JM}^z V$ for all $z \geq 0$.*

The proof is similar to the proof of Lemma 6.5.4 in Puterman [139].

Proof: For any $V_1, V_2 \in \mathcal{V}$ and $\pi_1 \in \arg\max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_1\}$,

$$\begin{aligned} \mathcal{M}V_1 &= r^{\pi_1} + P^{\pi_1} V_1 \\ \mathcal{M}V_2 &\geq r^{\pi_1} + P^{\pi_1} V_2 \\ \Rightarrow \quad \mathcal{M}V_2 &\geq \mathcal{M}V_1 + P^{\pi_1}\left(V_2 - V_1\right) \\ \Rightarrow \quad \mathcal{M}V_2 &\geq V_2 + \mathcal{M}V_1 - V_1 + \left(P^{\pi_1} - I\right)\left(V_2 - V_1\right) \\ \Rightarrow \quad \mathcal{M}\mathcal{M}_{JM}^z V &\geq \mathcal{M}_{JM}^z V + \mathcal{M}V - V + \left(P^{\pi_V} - I\right)\left(\mathcal{M}_{JM}^z V - V\right) \\ &= \mathcal{M}_{JM}^z V + \mathcal{M}V - V + \left(P^{\pi_V} - I\right)\left(\sum_{j=0}^{z-1}\left(P^{\pi_V}\right)^j\left(\mathcal{M}V - V\right)\right) \\ &= \mathcal{M}_{JM}^z V + \left(P^{\pi_V}\right)^z\left(\mathcal{M}V - V\right) \\ &\geq \mathcal{M}_{JM}^z V \qquad \square \end{aligned}$$

Theorem 14 *For any $V_{0,0} \in \mathcal{V}$ such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, and any sequence of positive integers $\{z_i\}$, the iterates of pre-Jacobi modified policy iteration $V_{i,z_i} \uparrow V^*$ in $\|\cdot\|_\infty$-norm as $i \to \infty$.*

Proof: Let $U_0 = W_0 = V_{0,0}$, $U_{i+1} = \mathcal{M}U_i$, $V_{i+1,z_{i+1}} = \mathcal{M}_{JM}^{z_{i+1}} V_{i,z_i}$, and $W_{i+1} = \mathcal{M}^{z_{i+1}} W_i$. It is shown by induction that $V_{i+1,z_{i+1}} \geq V_{i,z_i}$, $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$, and $U_i \leq V_{i,z_i} \leq W_i$, for all $i \geq 0$. As induction hypothesis, assume that $U_i \leq V_{i,z_i} \leq W_i$, and $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$. By assumption, the hypothesis holds for $i = 0$, (with $z_0 = 0$).

$$\begin{aligned} V_{i+1,z_{i+1}} &= \sum_{j=0}^{z_{i+1}-1} \left(P^{\pi_{i+1}}\right)^j r^{\pi_{i+1}} + \left(P^{\pi_{i+1}}\right)^{z_{i+1}} V_{i,z_i} \\ &= V_{i,z_i} + \sum_{j=0}^{z_{i+1}-1} \left(P^{\pi_{i+1}}\right)^j \left(r^{\pi_{i+1}} + P^{\pi_{i+1}} V_{i,z_i} - V_{i,z_i}\right) \\ &= V_{i,z_i} + \sum_{j=0}^{z_{i+1}-1} \left(P^{\pi_{i+1}}\right)^j \left(\mathcal{M}V_{i,z_i} - V_{i,z_i}\right) \\ &\geq V_{i,z_i} \end{aligned}$$

From Lemma 50, $\mathcal{M}V_{i+1,z_{i+1}} = \mathcal{M}\mathcal{M}_{JM}^{z_{i+1}}V_{i,z_i} \geq \mathcal{M}_{JM}^{z_{i+1}}V_{i,z_i} = V_{i+1,z_{i+1}}$. From Lemma 49, $V_{i+1,z_{i+1}} = \mathcal{M}_{JM}^{z_{i+1}}V_{i,z_i} \geq \mathcal{M}U_i = U_{i+1}$. From Lemma 48, $V_{i+1,z_{i+1}} = \mathcal{M}_{JM}^{z_{i+1}}V_{i,z_i} \leq \mathcal{M}^{z_{i+1}}W_i = W_{i+1}$. From Lemma 12, U_i and W_i converge in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. Therefore V_{i,z_i} converges in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. □

11.2.2 Gauss-Seidel Modified Policy Iteration

Similar to value iteration, a Gauss-Seidel version of modified policy iteration uses the most up-to-date values of the value function V. With the states $x \in \mathcal{X}$ ordered $x_1, x_2, \ldots, x_{|\mathcal{X}|}$, the partial policy evaluation step of pre-Gauss-Seidel modified policy iteration corresponding to Equation (11.6), is

$$V_{i,j+1}(x_k) = r(x_k, \pi_i(x_k)) + \sum_{k'<k} P[x_{k'} \mid x_k, \pi_i(x_k)]V_{i,j+1}(x_{k'}) + \sum_{k'\geq k} P[x_{k'} \mid x_k, \pi_i(x_k)]V_{i,j}(x_{k'})$$

The policy improvement step of pre-Gauss-Seidel modified policy iteration corresponding to Equation (11.7), is

$$V_{i+1,1}(x_k) = \max_{a\in\Psi(x_k)} \left\{ r(x_k, a) + \sum_{k'<k} P[x_{k'} \mid x_k, a]V_{i+1,1}(x_{k'}) + \sum_{k'\geq k} P[x_{k'} \mid x_k, a]V_{i,z_i}(x_{k'}) \right\} \tag{11.10}$$

The step of pre-Gauss-Seidel modified policy iteration corresponding to Equation (11.8) is then

$$\pi_{i+1}(x_k) \in \arg\max_{a\in\Psi(x_k)} \left\{ r(x_k, a) + \sum_{k'<k} P[x_{k'} \mid x_k, a]V_{i+1,1}(x_{k'}) + \sum_{k'\geq k} P[x_{k'} \mid x_k, a]V_{i,z_i}(x_{k'}) \right\}$$

This can again be expressed by "splitting" $I - P^\pi$ into $I - P_L^\pi - P_U^\pi$, where P_L^π is the strictly lower triangular part of P^π and P_U^π is the upper triangular part of P^π. For post-Gauss-Seidel modified policy iteration, P_L^π is the lower triangular part of P^π and P_U^π is the strictly upper triangular part of P^π. The policy improvement step can then be expressed as

$$V_{i+1,1} = \max_{\pi\in\Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V_{i,z_i} \right\} \tag{11.11}$$

and

$$\pi_{i+1} \in \arg\max_{\pi\in\Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V_{i,z_i} \right\}$$

The partial policy evaluation step can be expressed as

$$V_{i,j+1} = (I - P_L^{\pi_i})^{-1} r^{\pi_i} + (I - P_L^{\pi_i})^{-1} P_U^{\pi_i} V_{i,j}$$

When performing pre-Gauss-Seidel modified policy iteration, the threshold rule of Equation (11.4) can be used to give an optimal acceptance rule for the policy improvement step in Equation (11.10). The threshold rule of Equation (11.4) does not necessarily provide an optimal acceptance rule for the policy improvement step of post-Gauss-Seidel modified policy iteration in Equation (11.11), which might involve a search for an optimal acceptance rule. This makes the policy improvement step of post-Gauss-Seidel modified policy iteration more cumbersome than that of pre-Gauss-Seidel modified policy iteration. The policy evaluation steps of both should take comparable amounts of computing effort.

For any $V \in \mathcal{V}$, any $\pi_V \in \arg\max_{\pi \in \Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V \right\}$, and any positive integer z, let

$$\mathcal{M}_{GM}^z V \equiv (\mathcal{M}_G^{\pi_V})^z V = \sum_{j=0}^{z-1} \left[(I - P_L^{\pi_V})^{-1} P_U^{\pi_V} \right]^j (I - P_L^{\pi_V})^{-1} r^{\pi_V} + \left[(I - P_L^{\pi_V})^{-1} P_U^{\pi_V} \right]^z V$$

denote the pre-Gauss-Seidel modified policy iteration operator that performs one step of policy improvement and $z - 1$ steps of partial policy evaluation.

Lemma 51 *For any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$, $\mathcal{M}_{GM}^z V_1 \leq \mathcal{M}_{GV}^z V_2$ for all $z \geq 0$.*

Proof:

$$\begin{aligned} \mathcal{M}_{GM}^z V_1 &\equiv \left(\mathcal{M}_G^{\pi_{V_1}} \right)^z V_1 \\ &\leq \mathcal{M}_{GV}^z V_1 \\ &\leq \mathcal{M}_{GV}^z V_2 \end{aligned}$$

from Lemma 43. □

Lemma 52 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}_{GM}^z V \leq \mathcal{M}_{GM}^{z+1} V$ for all $z \geq 0$.*

Proof: $\mathcal{M}_{GM}^1 V = \mathcal{M}_{GV} V \geq \mathcal{M} V \geq V$ from Lemma 44. Then

$$\begin{aligned} \mathcal{M}_{GM}^{z+1} V \equiv (\mathcal{M}_G^{\pi_V})^{z+1} V &= (\mathcal{M}_G^{\pi_V})^z \mathcal{M}_G^{\pi_V} V \\ &= (\mathcal{M}_G^{\pi_V})^z \mathcal{M}_{GM}^1 V \\ &\geq (\mathcal{M}_G^{\pi_V})^z V \\ &\equiv \mathcal{M}_{GM}^z V \end{aligned}$$

□

Similarly, for any $V \in \mathcal{V}$ such that $\mathcal{M}V \leq V$, $\mathcal{M}_{GM}^z V \geq \mathcal{M}_{GM}^{z+1} V$ for all $z \geq 0$.

Lemma 53 *For any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$ and $\mathcal{M}V_2 \geq V_2$, $\mathcal{M}^z_{GM}V_2 \geq \mathcal{M}_{GV}V_1$ for all $z \geq 1$.*

Proof: From Lemmas 52 and 43, $\mathcal{M}^z_{GM}V_2 \geq \mathcal{M}^1_{GM}V_2 = \mathcal{M}_{GV}V_2 \geq \mathcal{M}_{GV}V_1$ for all $z \geq 1$. □

Lemma 54 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, $\mathcal{M}\mathcal{M}^z_{GM}V \geq \mathcal{M}^z_{GM}V$ for all $z \geq 0$.*

Proof: For $z = 0$, $\mathcal{M}\mathcal{M}^z_{GM}V = \mathcal{M}V \geq V = \mathcal{M}^z_{GM}V$.
For $z > 0$, and any $x_j \in \mathcal{X}$,

$$\begin{aligned}
\mathcal{M}\mathcal{M}^z_{GM}V(x_j) &\equiv \sup_{a\in\Psi(x_j)} \left\{ r(x_j, a) + \sum_{x_k\in\mathcal{X}} P[x_k \mid x_j, a]\mathcal{M}^z_{GM}V(x_k) \right\} \\
&\geq r(x_j, \pi_V(x_j)) + \sum_{x_k\in\mathcal{X}} P[x_k \mid x_j, \pi_V(x_j)]\mathcal{M}^z_{GM}V(x_k) \\
&\geq r(x_j, \pi_V(x_j)) + \sum_{k<j} P[x_k \mid x_j, \pi_V(x_j)]\mathcal{M}^z_{GM}V(x_k) \\
&\quad + \sum_{k\geq j} P[x_k \mid x_j, \pi_V(x_j)]\mathcal{M}^{z-1}_{GM}V(x_k) \\
&= \mathcal{M}^z_{GM}V(x_j)
\end{aligned}$$

where the second inequality follows from Lemma 52. □

Theorem 15 *For any $V_{0,0} \in \mathcal{V}$ such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, and any sequence of positive integers $\{z_i\}$, the iterates of pre-Gauss-Seidel modified policy iteration $V_{i,z_i} \uparrow V^*$ in $\|\cdot\|_\infty$-norm as $i \to \infty$.*

Proof: Let $U_0 = W_0 = V_{0,0}$, $U_{i+1} = \mathcal{M}_{GV}U_i$, $V_{i+1,z_{i+1}} = \mathcal{M}^{z_{i+1}}_{GM}V_{i,z_i}$, and $W_{i+1} = \mathcal{M}^{z_{i+1}}_{GV}W_i$. It is shown by induction that $V_{i+1,z_{i+1}} \geq V_{i,z_i}$, $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$, and $U_i \leq V_{i,z_i} \leq W_i$ for all $i \geq 0$. As induction hypothesis, assume that $U_i \leq V_{i,z_i} \leq W_i$, and $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$. By assumption, the hypothesis holds for $i = 0$, (with $z_0 = 0$). From Lemma 52, $V_{i+1,z_{i+1}} = \mathcal{M}^{z_{i+1}}_{GM}V_{i,z_i} \geq \mathcal{M}^0_{GM}V_{i,z_i} = V_{i,z_i}$. From Lemma 54, $\mathcal{M}V_{i+1,z_{i+1}} = \mathcal{M}\mathcal{M}^{z_{i+1}}_{GM}V_{i,z_i} \geq \mathcal{M}^{z_{i+1}}_{GM}V_{i,z_i} = V_{i+1,z_{i+1}}$. From Lemma 53, $V_{i+1,z_{i+1}} = \mathcal{M}^{z_{i+1}}_{GM}V_{i,z_i} \geq \mathcal{M}_{GV}U_i = U_{i+1}$. From

Lemma 51, $V_{i+1,z_{i+1}} = \mathcal{M}_{GM}^{z_{i+1}} V_{i,z_i} \leq \mathcal{M}_{GV}^{z_{i+1}} W_i = W_{i+1}$. From Lemma 13, U_i and W_i converge in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. Therefore V_{i,z_i} converges in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. □

The same error bound as in Proposition 7 applies if modified policy iteration stops when $\|V_{i+1,1} - V_{i,z_i}\|_\infty < \varepsilon$.

Similar to value iteration, relaxation can be used to update the values of Jacobi and Gauss-Seidel modified policy iteration.

11.3 Two-Decisions Algorithm

In this section a more efficient algorithm for the distribution problem is proposed. The algorithm distinguishes between the two types of decisions that have to be made, namely acceptance decisions and dispatching decisions. It exploits the optimal threshold result of Corollary 12, and the fact that given any value function V, an optimal threshold $V(m, \tilde{m}, n) - V(m, \tilde{m}, n + b) - p$ is very easy to compute for any batch composition b, whereas an optimal dispatching decision is harder to compute. The threshold $V(m, \tilde{m}, n) - V(m, \tilde{m}, n + b) - p$ does not necessarily give an optimal acceptance rule for post-Jacobi and post-Gauss-Seidel versions of the two-decisions algorithm. Thus these versions are unable to exploit this easily computed optimal acceptance rule, which is at the heart of the two-decisions algorithm. Therefore, only the pre-Jacobi and pre-Gauss-Seidel two-decisions algorithms will be studied.

11.3.1 Pre-Jacobi Two-Decisions Algorithm

Pre-Jacobi Two-Decisions Algorithm

begin

choose $V_{0,0} \in \mathcal{V}$ such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, choose $\varepsilon > 0$,

choose a sequence of positive integers $\{z_i\}$, and set $i = 1$;

$V_{1,1} \;=\; \max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_{0,0}\}$;

$\pi_{1,1} \;\in\; \arg\max_{\pi \in \Pi^{SD}} \{r^\pi + P^\pi V_{0,0}\}$;

$err \;=\; \|V_{1,1} - V_{0,0}\|_\infty$;

while $err > \varepsilon$

for $j = 1$ **to** $z_i - 1$

$$D^{\pi_{i,j+1}}(m,\tilde{m},n,r,b) \equiv \begin{cases} 1 & \text{if} \quad \sum_{l\in\mathcal{K}} s(k,l)\left[n(k,l)+b(k,l)\right] \le \overline{N}(k) \text{ for all } k \in \mathcal{K}, \\ & \quad\ \text{and } r \ge V_{i,j}(m,\tilde{m},n) - V_{i,j}(m,\tilde{m},n+b) - p \\ 0 & \text{if} \quad \sum_{l\in\mathcal{K}} s(k,l)\left[n(k,l)+b(k,l)\right] > \overline{N}(k) \text{ for some } k \in \mathcal{K}, \\ & \quad\ \text{or } r < V_{i,j}(m,\tilde{m},n) - V_{i,j}(m,\tilde{m},n+b) - p \end{cases}$$

$$L^{\pi_{i,j+1}}(m,\tilde{m},n) = L^{\pi_{i,j}}(m,\tilde{m},n);$$

$$Q^{\pi_{i,j+1}}(m,\tilde{m},n) = Q^{\pi_{i,j}}(m,\tilde{m},n);$$

$$V_{i,j+1} = r^{\pi_{i,j+1}} + P^{\pi_{i,j+1}} V_{i,j}; \tag{11.12}$$

endfor;

$$V_{i+1,1} = \max_{\pi\in\Pi^{SD}} \left\{ r^{\pi} + P^{\pi} V_{i,z_i} \right\}; \tag{11.13}$$

$$\pi_{i+1,1} \in \arg\max_{\pi\in\Pi^{SD}} \left\{ r^{\pi} + P^{\pi} V_{i,z_i} \right\};$$

$$err = \left\| V_{i+1,1} - V_{i,z_i} \right\|_{\infty};$$

$$i = i+1;$$

endwhile;

choose policy $\pi_\varepsilon = \pi_{i,1}$ and stop;

end;

For any $V_{i-1,z_{i-1}} \in \mathcal{V}$, and any positive integer z_i, let

$$\mathcal{M}_{JT}^{z_i} V_{i-1,z_{i-1}} \equiv \sum_{j=0}^{z_i-1} \left[\prod_{k=1}^{j} P^{\pi_{i,z_i-j+k}} \right] r^{\pi_{i,z_i-j}} + \left[\prod_{k=1}^{z_i} P^{\pi_{i,k}} \right] V_{i-1,z_{i-1}}$$

denote the operator of the pre-Jacobi two-decisions algorithm.

Lemma 55 *For any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \ge V_1$, $\mathcal{M}^z V_2 \ge \mathcal{M}_{JT}^z V_1$ for all $z \ge 0$.*

Proof: By induction on z. For $z = 0$, $\mathcal{M}^z V_2 = V_2 \ge V_1 = \mathcal{M}_{JT}^z V_1$. As induction hypothesis, suppose $\mathcal{M}^z V_2 \ge \mathcal{M}_{JT}^z V_1$. Then

$$\begin{aligned} \mathcal{M}^{z+1} V_2 &= \mathcal{M}\mathcal{M}^z V_2 \\ &\ge \mathcal{M}\mathcal{M}_{JT}^z V_1 \\ &= \sup_{\pi\in\Pi^{SD}} \left\{ r^{\pi} + P^{\pi} \mathcal{M}_{JT}^z V_1 \right\} \\ &\ge r^{\pi_{i,z+1}} + P^{\pi_{i,z+1}} \mathcal{M}_{JT}^z V_1 \\ &= \mathcal{M}_{JT}^{z+1} V_1 \end{aligned}$$

□

Lemma 56 *For any $V \in \mathcal{V}$, $\mathcal{M}_{JT}^z V \geq \mathcal{M}_{JM}^z V$ for all $z \geq 0$.*

Proof: By induction on z. For $z = 0$, $\mathcal{M}_{JT}^z V = V = \mathcal{M}_{JM}^z V$. As induction hypothesis, suppose $\mathcal{M}_{JT}^z V \geq \mathcal{M}_{JM}^z V$. Then

$$\begin{aligned}
\mathcal{M}_{JT}^{z+1} V &= r^{\pi_{i,z+1}} + P^{\pi_{i,z+1}} \mathcal{M}_{JT}^z V \\
&\geq r^{\pi_{i,1}} + P^{\pi_{i,1}} \mathcal{M}_{JT}^z V \\
&\geq r^{\pi_{i,1}} + P^{\pi_{i,1}} \mathcal{M}_{JM}^z V \\
&= \mathcal{M}_{JM}^{z+1} V
\end{aligned}$$

The first inequality follows from Lemma 5, and the second inequality follows from the induction hypothesis. □

Lemma 57 *If $V_{0,0} \in \mathcal{V}$ is such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, then the iterates of the pre-Jacobi two-decisions algorithm increase monotonically: $V_{0,0} \leq V_{1,1} \leq V_{1,2} \leq \ldots \leq V_{i,j} \leq V_{i,j+1} \leq \ldots \leq V_{i,z_i} \leq V_{i+1,1} \leq \ldots \leq V^*$.*

Proof: By induction on i and j. As induction hypothesis on i, suppose $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$. By assumption, this holds for $i = 0$ (with $z_0 = 0$). Then $V_{i+1,1} = r^{\pi_{i+1,1}} + P^{\pi_{i+1,1}} V_{i,z_i} = \mathcal{M}V_{i,z_i} \geq V_{i,z_i}$. It is shown by induction that $V_{i+1,j+1} \geq V_{i+1,j}$ for all $j \in \{1, \ldots, z_{i+1} - 1\}$. As induction hypothesis on j, suppose $V_{i+1,j} = r^{\pi_{i+1,j}} + P^{\pi_{i+1,j}} V_{i+1,j-1} \geq V_{i+1,j-1}$ (where $V_{i+1,0}$ is taken to be V_{i,z_i}). Then

$$\begin{aligned}
V_{i+1,j+1} &= r^{\pi_{i+1,j+1}} + P^{\pi_{i+1,j+1}} V_{i+1,j} \\
&\geq r^{\pi_{i+1,j}} + P^{\pi_{i+1,j}} V_{i+1,j} \\
&\geq r^{\pi_{i+1,j}} + P^{\pi_{i+1,j}} V_{i+1,j-1} \\
&= V_{i+1,j}
\end{aligned}$$

The first inequality follows from Lemma 5, and the second inequality follows from the induction hypothesis. Thus $V_{i+1,z_{i+1}} = r^{\pi_{i+1,z_{i+1}}} + P^{\pi_{i+1,z_{i+1}}} V_{i+1,z_{i+1}-1} \geq V_{i+1,z_{i+1}-1}$. Then

$$\begin{aligned}
V_{i+2,1} &= \mathcal{M}V_{i+1,z_{i+1}} \\
&\geq r^{\pi_{i+1,z_{i+1}}} + P^{\pi_{i+1,z_{i+1}}} V_{i+1,z_{i+1}} \\
&\geq r^{\pi_{i+1,z_{i+1}}} + P^{\pi_{i+1,z_{i+1}}} V_{i+1,z_{i+1}-1} \\
&= V_{i+1,z_{i+1}}
\end{aligned}$$

and the induction on i has been established. $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ implies that $V^* \geq V_{i,z_i}$ from Lemma 39; hence $V^* \geq V_{i,j}$ for each i and j. □

Corollary 14 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, and any nonnegative integer z, $\mathcal{M}\mathcal{M}_{JT}^z V \geq \mathcal{M}_{JT}^z V$.*

Theorem 16 *For any $V_{0,0} \in \mathcal{V}$ such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, and any sequence of positive integers $\{z_i\}$, the iterates $V_{i,j}$ of the pre-Jacobi two-decisions algorithm converge monotonically and in $\|\cdot\|_\infty$-norm to V^*.*

Proof: Let $U_0 = W_0 = V_{0,0}$, $U_{i+1} = \mathcal{M}U_i$, $V_{i+1,z_{i+1}} = \mathcal{M}_{JT}^{z_{i+1}} V_{i,z_i}$, and $W_{i+1} = \mathcal{M}^{z_{i+1}} W_i$. It was shown in Lemma 57 that the iterates $V_{i,j}$ increase monotonically, and in Corollary 14 that $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ for all i. It is shown by induction that $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ and $U_i \leq V_{i,z_i} \leq W_i$ for all i. By assumption, this holds for $i = 0$, (with $z_0 = 0$). As induction hypothesis, assume that $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ and $U_i \leq V_{i,z_i} \leq W_i$. From Corollary 14, $\mathcal{M}V_{i+1,z_{i+1}} = \mathcal{M}\mathcal{M}_{JT}^{z_{i+1}} V_{i,z_i} \geq \mathcal{M}_{JT}^{z_{i+1}} V_{i,z_i} = V_{i+1,z_{i+1}}$. From Lemma 56, $V_{i+1,z_{i+1}} = \mathcal{M}_{JT}^{z_{i+1}} V_{i,z_i} \geq \mathcal{M}_{JM}^{z_{i+1}} V_{i,z_i}$. From Lemma 49, $\mathcal{M}_{JM}^{z_{i+1}} V_{i,z_i} \geq \mathcal{M}U_i = U_{i+1}$. From Lemma 55, $V_{i+1,z_{i+1}} = \mathcal{M}_{JT}^{z_{i+1}} V_{i,z_i} \leq \mathcal{M}^{z_{i+1}} W_i = W_{i+1}$. Therefore, $U_{i+1} \leq V_{i+1,z_{i+1}} \leq W_{i+1}$, and the induction has been established. From Lemma 12, U_i and W_i converge in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. Therefore V_{i,z_i} converges in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. □

11.3.2 Pre-Gauss-Seidel Two-Decisions Algorithm

A pre-Gauss-Seidel version of the two-decisions algorithm uses the most up-to-date values of the value function V for computing both the acceptance thresholds and the updated values.

Pre-Gauss-Seidel Two-Decisions Algorithm

begin

choose $V_{0,0} \in \mathcal{V}$ such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, choose $\varepsilon > 0$,

choose a sequence of positive integers $\{z_i\}$, and set $i = 1$;

$$V_{1,1} \quad = \quad \max_{\pi \in \Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V_{0,0} \right\};$$

$$\pi_{1,1} \in \arg\max_{\pi\in\Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V_{0,0} \right\};$$

$$err = \|V_{1,1} - V_{0,0}\|_\infty ;$$

while $err > \varepsilon$

for $j = 1$ **to** $z_i - 1$

$$L^{\pi_{i,j+1}}(m, \tilde{m}, n) = L^{\pi_{i,j}}(m, \tilde{m}, n);$$

$$Q^{\pi_{i,j+1}}(m, \tilde{m}, n) = Q^{\pi_{i,j}}(m, \tilde{m}, n);$$

for $k = 1$ **to** $|\mathcal{X}|$

for $k' = 1$ **to** $k - 1$

$$D^{\pi_{i,j+1}}(x_k, r, b) \equiv \begin{cases} 1 \text{ if } & \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l) + b(k,l)] \le \overline{N}(k) \text{ for all } k \in \mathcal{K}, \\ & \text{and } r \ge V_{i,j}(x_k) - V_{i,j+1}(x_{k'}) - p \\ 0 \text{ if } & \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l) + b(k,l)] > \overline{N}(k) \text{ for some } k \in \mathcal{K}, \\ & \text{or } r < V_{i,j}(x_k) - V_{i,j+1}(x_{k'}) - p \end{cases}$$

where $x_k = (m, \tilde{m}, n)$ and $x_{k'} = (m, \tilde{m}, n + b)$;

endfor;

for $k' = k$ **to** $|\mathcal{X}|$

$$D^{\pi_{i,j+1}}(x_k, r, b) \equiv \begin{cases} 1 \text{ if } & \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l) + b(k,l)] \le \overline{N}(k) \text{ for all } k \in \mathcal{K}, \\ & \text{and } r \ge V_{i,j}(x_k) - V_{i,j}(x_{k'}) - p \\ 0 \text{ if } & \sum_{l\in\mathcal{K}} s(k,l)\,[n(k,l) + b(k,l)] > \overline{N}(k) \text{ for some } k \in \mathcal{K}, \\ & \text{or } r < V_{i,j}(x_k) - V_{i,j}(x_{k'}) - p \end{cases}$$

endfor;

$$\begin{aligned} V_{i,j+1}(x_k) = {} & r(x_k, \pi_{i,j+1}(x_k)) \\ & + \sum_{k'<k} P[x_{k'} \mid x_k, \pi_{i,j+1}(x_k)] V_{i,j+1}(x_{k'}) \\ & + \sum_{k'\ge k} P[x_{k'} \mid x_k, \pi_{i,j+1}(x_k)] V_{i,j}(x_{k'}); \end{aligned}$$

endfor;

endfor;

$$V_{i+1,1} = \max_{\pi\in\Pi^{SD}} \left\{ (I - P_L^\pi)^{-1} r^\pi + (I - P_L^\pi)^{-1} P_U^\pi V_{i,z_i} \right\};$$

$$\pi_{i+1,1} \in \arg\max_{\pi\in\Pi^{SD}} \left\{ (I-P_L^\pi)^{-1} r^\pi + (I-P_L^\pi)^{-1} P_U^\pi V_{i,z_i} \right\};$$
$$err = \|V_{i+1,1} - V_{i,z_i}\|_\infty;$$
$$i = i+1;$$

endwhile;

choose policy $\pi_\varepsilon = \pi_{i,1}$ and stop;

end;

For any $V_{i-1,z_{i-1}} \in \mathcal{V}$, and any positive integer z_i, let

$$\mathcal{M}_{GT}^{z_i} V_{i-1,z_{i-1}} \equiv \sum_{j=0}^{z_i-1} \left[\prod_{k=1}^{j} \left(I - P_L^{\pi_{i,z_i-j+k}}\right)^{-1} P_U^{\pi_{i,z_i-j+k}} \right] \left(I - P_L^{\pi_{i,z_i-j}}\right)^{-1} r^{\pi_{i,z_i-j}} + \left[\prod_{k=1}^{z_i} \left(I - P_L^{\pi_{i,k}}\right)^{-1} P_U^{\pi_{i,k}} \right] V_{i-1,z_{i-1}}$$

denote the operator of the pre-Gauss-Seidel two-decisions algorithm.

Lemma 58 *For any $V_1, V_2 \in \mathcal{V}$ such that $V_2 \geq V_1$, $\mathcal{M}_{GV}^z V_2 \geq \mathcal{M}_{GT}^z V_1$ for all $z \geq 0$.*

Proof: By induction on z. For $z = 0$, $\mathcal{M}_{GV}^z V_2 = V_2 \geq V_1 = \mathcal{M}_{GT}^z V_1$. As induction hypothesis, suppose $\mathcal{M}_{GV}^z V_2 \geq \mathcal{M}_{GT}^z V_1$. Then

$$\begin{aligned}
\mathcal{M}_{GV}^{z+1} V_2 &= \mathcal{M}_{GV}\mathcal{M}_{GV}^z V_2 \\
&\geq \mathcal{M}_{GV}\mathcal{M}_{GT}^z V_1 \\
&= \sup_{\pi\in\Pi^{SD}} \left\{ (I-P_L^\pi)^{-1} r^\pi + (I-P_L^\pi)^{-1} P_U^\pi \mathcal{M}_{GT}^z V_1 \right\} \\
&\geq \left\{ \left(I-P_L^{\pi_{i,z+1}}\right)^{-1} r^{\pi_{i,z+1}} + \left(I-P_L^{\pi_{i,z+1}}\right)^{-1} P_U^{\pi_{i,z+1}} \mathcal{M}_{GT}^z V_1 \right\} \\
&= \mathcal{M}_{GT}^{z+1} V_1
\end{aligned}$$

□

Lemma 59 *For any $V \in \mathcal{V}$, $\mathcal{M}_{GT}^z V \geq \mathcal{M}_{GM}^z V$ for all $z \geq 0$.*

Proof: By induction on z. For $z = 0$, $\mathcal{M}_{GT}^z V = V = \mathcal{M}_{GM}^z V$. As induction hypothesis, suppose $\mathcal{M}_{GT}^z V \geq \mathcal{M}_{GM}^z V$. Then

$$\begin{aligned}
\mathcal{M}_{GT}^{z+1} V &= \left(I-P_L^{\pi_{i,z+1}}\right)^{-1} r^{\pi_{i,z+1}} + \left(I-P_L^{\pi_{i,z+1}}\right)^{-1} P_U^{\pi_{i,z+1}} \mathcal{M}_{GT}^z V \\
&\geq \left(I-P_L^{\pi_{i,1}}\right)^{-1} r^{\pi_{i,1}} + \left(I-P_L^{\pi_{i,1}}\right)^{-1} P_U^{\pi_{i,1}} \mathcal{M}_{GT}^z V \\
&\geq \left(I-P_L^{\pi_{i,1}}\right)^{-1} r^{\pi_{i,1}} + \left(I-P_L^{\pi_{i,1}}\right)^{-1} P_U^{\pi_{i,1}} \mathcal{M}_{GM}^z V \\
&= \mathcal{M}_{GM}^{z+1} V
\end{aligned}$$

The first inequality follows from Lemma 5, and the second inequality follows from the induction hypothesis. $\square$

Lemma 60 *If $V_{0,0} \in \mathcal{V}$ is such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, then the iterates of the pre-Gauss-Seidel two-decisions algorithm increase monotonically: $V_{0,0} \leq V_{1,1} \leq V_{1,2} \leq \ldots \leq V_{i,j} \leq V_{i,j+1} \leq \ldots \leq V_{i,z_i} \leq V_{i+1,1} \leq \ldots$.*

Proof: By induction on i and j. $V_{1,1} = \mathcal{M}_{GV}V_{0,0} \geq \mathcal{M}V_{0,0} \geq V_{0,0}$ from Lemma 44. As induction hypothesis on i, suppose $V_{i+1,1} \geq V_{i,z_i}$. It is shown by induction that $V_{i+1,j+1} \geq V_{i+1,j}$ for all $j \in \{1, \ldots, z_{i+1} - 1\}$. As induction hypothesis on j, suppose $V_{i+1,j} = \left(I - P_L^{\pi_{i+1,j}}\right)^{-1} r^{\pi_{i+1,j}} + \left(I - P_L^{\pi_{i+1,j}}\right)^{-1} P_U^{\pi_{i+1,j}} V_{i+1,j-1} \geq V_{i+1,j-1}$ (where $V_{i+1,0}$ is taken to be V_{i,z_i}). Then

$$\begin{aligned}
V_{i+1,j+1} &= \left(I - P_L^{\pi_{i+1,j+1}}\right)^{-1} r^{\pi_{i+1,j+1}} + \left(I - P_L^{\pi_{i+1,j+1}}\right)^{-1} P_U^{\pi_{i+1,j+1}} V_{i+1,j} \\
&\geq \left(I - P_L^{\pi_{i+1,j}}\right)^{-1} r^{\pi_{i+1,j}} + \left(I - P_L^{\pi_{i+1,j}}\right)^{-1} P_U^{\pi_{i+1,j}} V_{i+1,j} \\
&\geq \left(I - P_L^{\pi_{i+1,j}}\right)^{-1} r^{\pi_{i+1,j}} + \left(I - P_L^{\pi_{i+1,j}}\right)^{-1} P_U^{\pi_{i+1,j}} V_{i+1,j-1} \\
&= V_{i+1,j}
\end{aligned}$$

The first inequality follows from Lemma 5, and the second inequality follows from the induction hypothesis. Thus

$$\begin{aligned}
V_{i+1,z_{i+1}} &= \left(I - P_L^{\pi_{i+1,z_{i+1}}}\right)^{-1} r^{\pi_{i+1,z_{i+1}}} + \left(I - P_L^{\pi_{i+1,z_{i+1}}}\right)^{-1} P_U^{\pi_{i+1,z_{i+1}}} V_{i+1,z_{i+1}-1} \\
&\geq V_{i+1,z_{i+1}-1}
\end{aligned}$$

Then

$$\begin{aligned}
V_{i+2,1} &= \mathcal{M}_{GV} V_{i+1,z_{i+1}} \\
&\geq \mathcal{M}_{GV} V_{i+1,z_{i+1}-1} \\
&\geq \mathcal{M}_G^{\pi_{i+1,z_{i+1}}} V_{i+1,z_{i+1}-1} \\
&= V_{i+1,z_{i+1}}
\end{aligned}$$

and the induction on i has been established. $\square$

Lemma 61 *For any $V \in \mathcal{V}$ such that $\mathcal{M}V \geq V$, and any nonnegative integer z, $\mathcal{M}\mathcal{M}_{GT}^z V \geq \mathcal{M}_{GT}^z V$.*

Proof: For $z = 0$, $\mathcal{M}\mathcal{M}_{GT}^{z}V = \mathcal{M}V \geq V = \mathcal{M}_{GT}^{z}V$.

For $z > 0$, and for any $x_j \in \mathcal{X}$,

$$\begin{aligned}
\mathcal{M}\mathcal{M}_{GT}^{z}V(x_j) &\equiv \sup_{a\in\Psi(x_j)}\left\{ r(x_j,a) \ + \sum_{x_k\in\mathcal{X}} P[x_k \mid x_j, a]\mathcal{M}_{GT}^{z}V(x_k)\right\} \\
&\geq r(x_j,\pi_{i,z}(x_j)) \ + \sum_{x_k\in\mathcal{X}} P[x_k \mid x_j, \pi_{i,z}(x_j)]\mathcal{M}_{GT}^{z}V(x_k) \\
&\geq r(x_j,\pi_{i,z}(x_j)) \ + \sum_{k<j} P[x_k \mid x_j, \pi_{i,z}(x_j)]\mathcal{M}_{GT}^{z}V(x_k) \\
&\quad + \sum_{k\geq j} P[x_k \mid x_j, \pi_{i,z}(x_j)]\mathcal{M}_{GT}^{z-1}V(x_k) \\
&= \mathcal{M}_{GT}^{z}V(x_j)
\end{aligned}$$

where the second inequality follows from Lemma 60. □

Corollary 15 *If $V_{0,0} \in \mathcal{V}$ is such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, then $V_{i,j} \leq V^*$ for all i and j.*

Proof: From Lemma 61, if $\mathcal{M}V_{0,0} \geq V_{0,0}$, then $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ for all i. From Lemma 39, $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ implies that $V^* \geq V_{i,z_i}$; hence $V^* \geq V_{i,j}$ for all i and j from Lemma 60. □

Theorem 17 *For any $V_{0,0} \in \mathcal{V}$ such that $\mathcal{M}V_{0,0} \geq V_{0,0}$, and any sequence of positive integers $\{z_i\}$, the iterates $V_{i,j}$ of the pre-Gauss-Seidel two-decisions algorithm converge monotonically and in $\|\cdot\|_\infty$-norm to V^*.*

Proof: Let $U_0 = W_0 = V_{0,0}$, $U_{i+1} = \mathcal{M}_{GV}U_i$, $V_{i+1,z_{i+1}} = \mathcal{M}_{GT}^{z_{i+1}}V_{i,z_i}$, and $W_{i+1} = \mathcal{M}_{GV}^{z_{i+1}}W_i$. It was shown in Lemma 60 that the iterates $V_{i,j}$ increase monotonically. and in Lemma 61 that $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ for all i. It is shown by induction that $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ and $U_i \leq V_{i,z_i} \leq W_i$ for all i. By assumption, this holds for $i = 0$, (with $z_0 = 0$). As induction hypothesis, assume that $\mathcal{M}V_{i,z_i} \geq V_{i,z_i}$ and $U_i \leq V_{i,z_i} \leq W_i$. From Lemma 61, $\mathcal{M}V_{i+1,z_{i+1}} = \mathcal{M}\mathcal{M}_{GT}^{z_{i+1}}V_{i,z_i} \geq \mathcal{M}_{GT}^{z_{i+1}}V_{i,z_i} = V_{i+1,z_{i+1}}$. From Lemma 59, $V_{i+1,z_{i+1}} = \mathcal{M}_{GT}^{z_{i+1}}V_{i,z_i} \geq \mathcal{M}_{GM}^{z_{i+1}}V_{i,z_i}$. From Lemma 53, $\mathcal{M}_{GM}^{z_{i+1}}V_{i,z_i} \geq \mathcal{M}_{GV}U_i = U_{i+1}$. From Lemma 58, $V_{i+1,z_{i+1}} = \mathcal{M}_{GT}^{z_{i+1}}V_{i,z_i} \leq \mathcal{M}_{GV}^{z_{i+1}}W_i = W_{i+1}$. Therefore, $U_{i+1} \leq V_{i+1,z_{i+1}} \leq W_{i+1}$, and the induction has been established. From Lemma 13, U_i and W_i converge in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. Therefore V_{i,z_i} converges in $\|\cdot\|_\infty$-norm to V^* as $i \to \infty$. □

The same error bound as in Proposition 7 applies if the two-decisions algorithm stops when $\|V_{i+1,1} - V_{i,z_i}\|_\infty < \varepsilon$.

Both the Jacobi and Gauss-Seidel versions of the two-decisions algorithm can use relaxation to update the values.

11.4 Numerical Results

11.4.1 Comparison of Algorithms

The algorithms described in Section 11 were applied to a number of distribution problems of different sizes to compare their performance. The algorithms tested were

1. Pre-Gauss-Seidel value iteration;

2. Pre-Jacobi modified policy iteration;

3. Pre-Gauss-Seidel modified policy iteration;

4. Pre-Gauss-Seidel two-decisions algorithm;

For the two modified policy iteration algorithms as well as the proposed algorithm, the sequence $\{z_i\}$ of numbers of partial policy evaluation iterations was determined as follows. A sequence $\{\eta_i\}$ of maximum number of partial policy evaluation iterations was chosen by setting $\eta_i = 1.8^i$. A sequence $\{\xi_i\}$ of tolerances for the partial policy evaluation iterations was chosen by setting $\xi_i = 10/2^i$. The partial policy evaluation iterations were stopped as soon as $j > \eta_i$ or $\|x_{i,j} - x_{i,j-1}\|_\infty < \xi_i$.

In all the cases relaxation was used to update the values. Some experimentation was performed to find the relaxation factors ω (to the nearest 0.05) that give the best performance with the value iteration algorithm. Some more experimentation showed that the same relaxation factors give good performance with the other algorithms, and the same relaxation factors were used for all the algorithms. Relaxation factors used varied between 1.65 and 1.80, with the smaller relaxation factors giving better performance with the larger problems.

The computation times on a Sun Sparc 1000 model 1104 (50MHz) for problems with state spaces of different sizes are shown in Figure 11.1, for the four algorithms mentioned above. A stopping tolerance of $\varepsilon = 0.01$ was used in each case, which gives an accuracy at stopping of approximately 3 significant digits. Value iteration took significantly more time than the other algorithms. There does not seem to be a significant difference between the pre-Jacobi and pre-Gauss-Seidel versions of modified policy iteration. The two-decisions algorithm consistently outperformed the other algorithms, although the improvement over modified policy iteration was not by an order of magnitude.

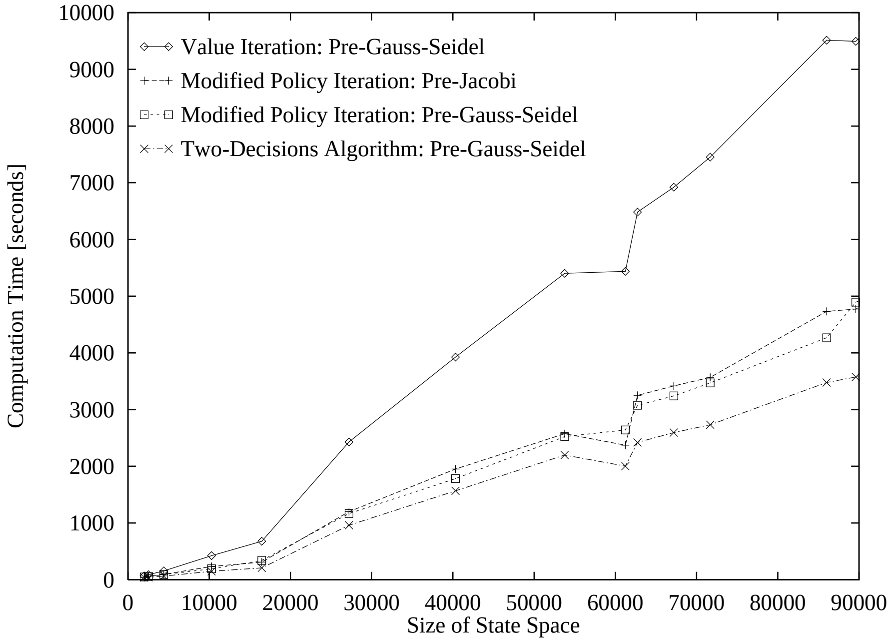


Figure 11.1 Computation times vs. state space size for pre-Gauss-Seidel value iteration, pre-Jacobi modified policy iteration, pre-Gauss-Seidel modified policy iteration, and pre-Gauss-Seidel two-decisions algorithm

The size of the state space grows very quickly with the number of terminals. Therefore, as the number of terminals increases, the "curse of dimensionality" soon causes solving the problems to become unpractical using the techniques described in this paper. Parallel processing or decomposition approaches can be used to extend the sizes of the problems that can be solved in reasonable time.

11.4.2 Example

Results for an example problem are shown in Figures 11.2 and 11.3. The example involves a system with 3 terminals and 3 vehicles. The input data for the problem are shown in Table 11.1.

Figure 11.2 shows the optimal expected value as a function of the number of loads at terminal 3 that have to go to terminal 2, for different numbers of vehicles available at terminal 3. States represented in Figure 11.2 have no loads or vehicles at any other terminal, no loads at terminal 3 destined to terminals other than terminal 2, and vehicles not at terminal 3 are en route to terminal 3. As can be seen in Figure 11.2, the optimal value decreases as the number of waiting loads increases. As shown in Proposition 6, this intuitive property holds if the holding cost c and the transportation cost φ are nondecreasing in the number of loads. Figure 11.3 shows the optimal acceptance threshold for a batch consisting of 1 load destined from terminal 3 to terminal 1 and 2 loads destined from terminal 3 to terminal 2, as a function of the number of loads at terminal 3 that have to go to terminal 2, for different numbers of vehicles available at terminal 3. The states represented in Figure 11.3 are the same as those represented in Figure 11.2. The curves for more than 1 vehicle available at terminal 3 stop because vehicles are dispatched when there are more loads waiting. For example, when there is 1 vehicle available at terminal 3, the vehicle will be held as long as there are 3 or less loads that have to go to terminal 2, and as soon as there are 4 or more loads at terminal 3 that have to go to terminal 2, the vehicle will be dispatched with these loads.

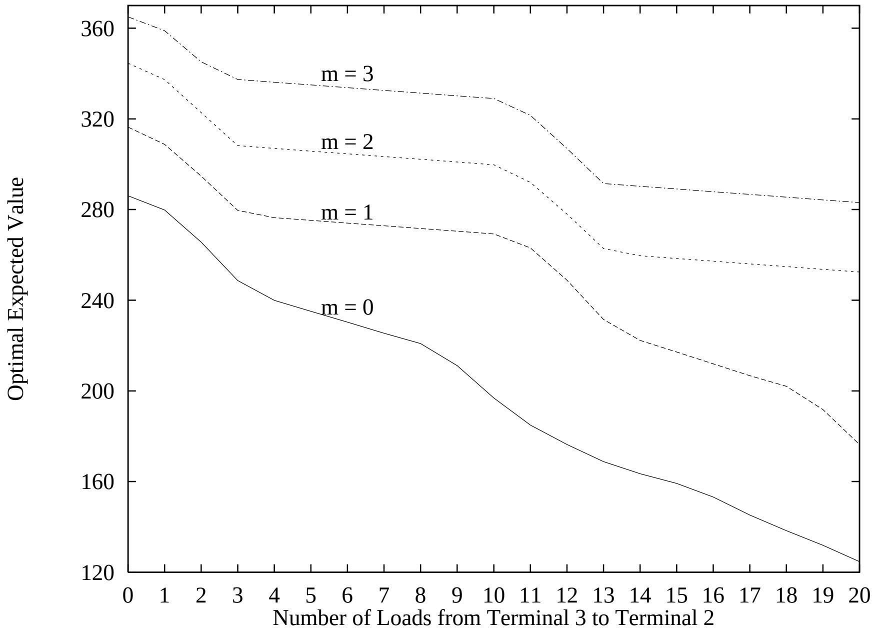


Figure 11.2 Optimal expected value vs. number of loads at terminal 3 destined to terminal 2 for different numbers of vehicles m available at terminal 3

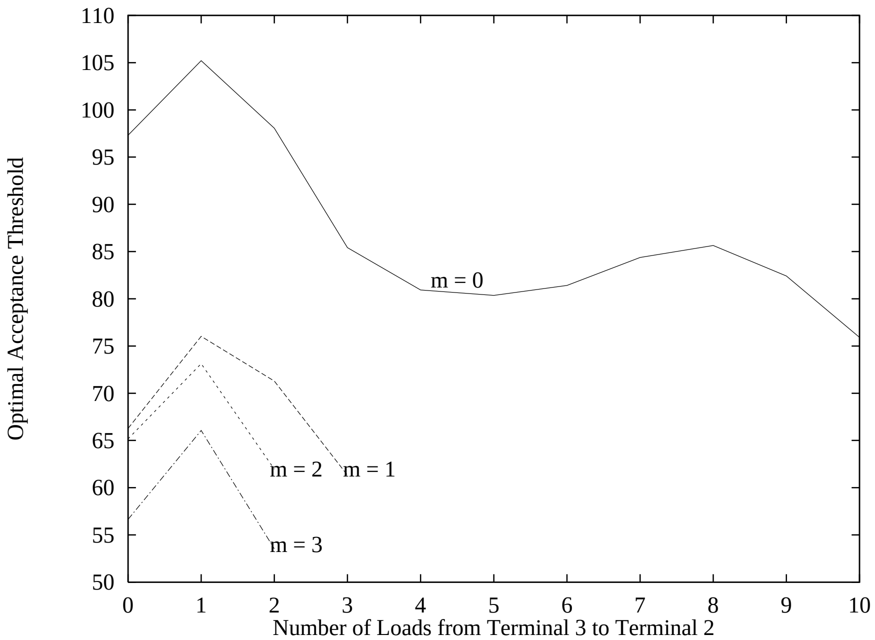


Figure 11.3 Optimal acceptance threshold vs. number of loads at terminal 3 destined to terminal 2 for different numbers of vehicles m available at terminal 3

Table 11.1 Input data of example problem

Number of terminals $K = 3$

Number of vehicles $\overline{M} = 3$

Vehicle capacity $\overline{N}_v = 10$

Terminal	1	2	3
Capacity	20.00	20.00	20.00
Load Arrival Rate	1.00	2.00	1.50
Vehicle Arrival Rate	0.02	0.02	0.02
Vehicle Holding Cost	0.00	0.00	0.00

Load Size:

Terminal	1	2	3
1	—	7.00	5.00
2	7.50	—	5.00
3	8.00	1.00	—

Load Holding Cost Per Unit Time:

Terminal	1	2	3
1	—	1.70	1.50
2	3.00	—	2.00
3	2.00	0.20	—

Fixed Dispatch Cost:

Terminal	1	2	3
1	—	10.00	10.00
2	10.00	—	10.00
3	10.00	10.00	—

Variable Dispatch Cost:

Terminal	1	2	3
1	—	1.60	1.40
2	1.80	—	1.50
3	2.40	1.20	—

Table 1: Input data of example problem (continued)

Rewards are conditionally exponentially distributed

with mean reward proportional to the number of loads

Mean Reward Proportionality Constant:

Terminal	1	2	3
1	—	43.00	32.00
2	55.00	—	40.00
3	45.00	10.00	—

Batch compositions are multinomially distributed

Terminal	1	2	3
Multinomial n	6	4	5

Multinomial p:

Terminal	1	2	3
1	0.30	0.40	0.20
2	0.10	0.50	0.30
3	0.20	0.40	0.20

Penalty cost for refusing load $p = 2.0$

Discount rate $\alpha = 0.1$

Size of state space $|\mathcal{X}| = 218400$

Relaxation factor $\omega = 1.65$

Starting maximum number of partial policy evaluation iterations $\eta_0 = 1$

Multiplier for maximum number of partial policy evaluation iterations $\eta_1 = 1.8$

Starting partial policy evaluation tolerance $\xi_0 = 10.0$

Divider for partial policy evaluation tolerance = 2.0

Stopping tolerance $\varepsilon = 0.01$

12. Research Directions

Many issues relevant to dynamic and stochastic resource allocation problems and freight distribution problems remain unaddressed in the DSKP and DSDP. In this chapter, some directions are given for research to address these issues.

12.1 Dynamic and Stochastic Resource Allocation Problems

1. The optimality result for the DSKP was obtained assuming that the arrival rate λ of resource demands does not depend on time (but may depend on the amount of resource available). In many applications, the arrival process of the demands is time-dependent, for example seasonal with daily, weekly, or annual cycles. Solving dynamic and stochastic resource allocation problems with time-varying parameters over finite and infinite horizons requires additional research.

2. The amount of resources available can itself be a stochastic process. For example, a portfolio manager manages a fund of which the size changes randomly. These problems have attracted much research attention, but much remains to be done.

3. There may be a number of related knapsacks involved in the problem. For example, airline yield management addresses the selling of seats on a number of flights that interact through passengers transfers, aircraft and crew assignments, timetables, and return flights.

4. The process of negotiating the reward for the resource demanded was not modeled. The knapsack problem may be part of a larger competitive environment, in which multiple agents allocate resources from multiple knapsacks to many customers. Examples are problems addressing the selling of assets, and yield management problems. Game theory may throw additional light on the competitive dynamics in these problems.

5. It was assumed that the joint distribution of reward and size is known. In most practical applications, the distributions are not known, but have to be estimated with the available data. As the system develops and more data become available, these estimates can be updated. A Bayesian approach can be followed.

12.2 Dynamic and Stochastic Freight Distribution Problems

1. One of the most important issues in the management of distribution operations concerns the drivers, their needs and their work rules. It was estimated that the average driver turnover in the trucking industry is approximately 80% per year [179]. This very high turnover is attributed to the unfavorable working conditions. Carriers have to take the drivers into account when making dispatching decisions. In the DSDP, no distinction was made between the availability of a driver and the availability of a vehicle. Further research is needed to incorporate driver related issues directly into the model.

2. The simplifying assumption was made that all loads of the same load class are measured in numbers of the same size units. In actual LTL applications, loads come in many different sizes and shapes. Loading a vehicle with these loads is a very complicated three dimensional bin packing problem. As a first step in addressing these problems, dynamic and stochastic one dimensional bin packing problems can be studied.

3. The assumption was made that the fleet of vehicles is homogeneous: all vehicles are the same, and any vehicle can transport any class of loads. In real distribution systems, the fleets consist of many types and sizes of vehicles with different operating characteristics. Also, not all loads can be transported by any type of vehicle; some types of loads, for example perishable products, liquids, and hazardous materials, require specialized equipment. Additional research can be aimed at incorporating heterogeneous fleets and loads with specialized requirements.

4. It was assumed that loads are transported directly from their origin terminals to their destination terminals, with no additional consolidation at intermediate breakbulk

terminals. Daganzo [48, 51] discusses the role of terminals and transhipments in one-to-many and many-to-many freight distribution operations with some intuitive models. Although the DSDP can be extended to include the option of additional consolidation operations at intermediate terminals, the size of the state space will increase a lot, and with that the computational effort required to solve these problems. Additional research can be aimed at alternative approaches to find optimal policies that allow for multiple consolidation operations.

5. Only problems with a very small number of terminals can be solved in reasonable time with the algorithms investigated in this study. The examples given had two or three terminals; beyond that the computation times are unacceptable. There are a large number of carriers with such a small number of terminals, who can benefit from this research. However, the carriers who need sophisticated decision support systems the most for managing their complex operations, are the large regional and national carriers, some of whom have hundreds of terminals. Different approaches are therefore needed for finding optimal or good suboptimal policies for managing complex distribution operations. A number of alternative approaches are the following.

   (a) Decomposition methods. The DSDP has a very natural decomposition. It is intuitive to expect that the most important information for a dispatcher needing to decide whether and to which destinations to dispatch a vehicle, is the information related to the terminal where the vehicle is currently located, i.e., the number of loads of the different load classes at that terminal, the number of vehicles at the terminal, and the number of vehicles en route to the terminal. It may be sufficient to aggregate the information related to all the other terminals for decision making purposes at each terminal. As an example of the motivation for this intuition, consider Figure 12.1. The problem data for this example is given in Table 11.1. The optimal acceptance threshold for a batch consisting of 1 load destined from terminal 3 to terminal 1 and 2 loads destined from terminal 3 to terminal 2, is shown as a function of the number of loads at terminal 3 that have to go to terminal 2, for different numbers of vehicles available at terminal 3. There are no loads or vehicles at terminal 2, and no loads at terminal 3 destined to terminals other than terminal 2, and the vehicles not at terminal 3 are en route to terminal 3. Two sets of curves are shown in Figure 12.1. The one set of curves shows the optimal threshold at terminal 3 under the conditions described above, if there are no loads at terminal 1. The other set of curves shows the optimal threshold at terminal 3 under the same conditions, except that terminal 1 is now loaded to 25% of its capacity with loads destined to terminal 3. Note that the two sets of curves are very close together.

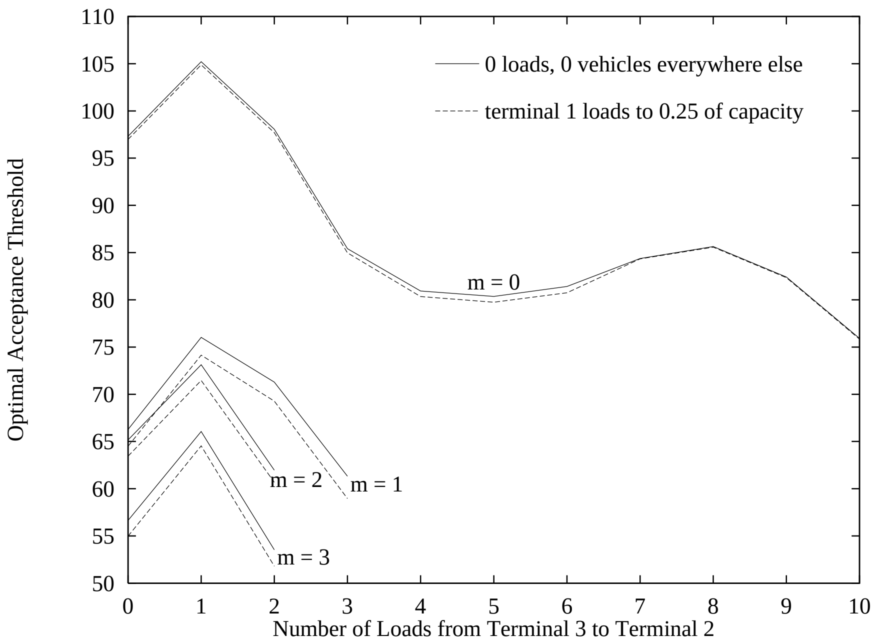


Figure 12.1 Optimal acceptance threshold vs. number of loads at terminal 3 destined to terminal 2 for different numbers of vehicles m available at terminal 3

It therefore seems that the optimal policy for accepting loads and dispatching vehicles at terminal 3 is not very sensitive with respect to the state at other terminals, but quite sensitive with respect to the state at terminal 3. This suggests that a decomposition approach, in which the terminals make decentralized decisions, and where the information regarding the own terminal is taken into account in detail by the terminal, but the information regarding the other terminals is aggregated, should give good results. This decomposition approach has the further advantage that it lends itself to implementation in a parallel computing environment. It is also interesting because it is similar in nature to the way in which decisions are made in many practical systems. Comparing the results of this decomposition approach with optimal solutions obtained for small systems can therefore provide an indication of the quality of decisions, or the degree of suboptimality, that are achieved in practice. Such a decomposition approach was implemented. Initial computational results indicate that the quality of the decomposition solutions are very problem dependent.

(b) Stochastic programming. This includes a very wide variety of techniques, some of which involve analytical approximations to the optimal expected value function, sampling based methods utilizing deterministic optimization techniques for each sample, or stochastic versions of nonlinear programming techniques.

(c) Approximation methods. As an example of these methods, a parametric function depending on the state variables can be fitted to the expected value function. Using this parametric function, estimates of the expected value in a state can be obtained without needing to compute values for each state as in the classical iterative methods. For examples, see Bertsekas [14].

(d) Sampling-based methods. These include Monte-Carlo simulation methods for policy evaluation, the method of temporal differences (TD(λ)), and Q-learning. For more on TD(λ), see Sutton [174], and for a convergence proof for Q-learning, see Tsitsiklis [181].

13. Conclusion

This research was motivated by the need for decision support tools to manage complex freight distribution operations. Due to the vast economic scope of freight transportation, small improvements in efficiency can generate huge savings, making this research both worthwhile and stimulating.

The research was conducted in two parts. In the first part a generic resource allocation problem in a dynamic and stochastic environment was studied. This problem was motivated by the issues facing a carrier who has to decide about the loading and dispatching of a vehicle. Because of the many applications of this problem, it is called the Dynamic and Stochastic Knapsack Problem (DSKP). Both the infinite horizon and finite horizon cases of the DSKP were studied.

It was shown that an optimal acceptance rule for the DSKP is given by a threshold rule, which is easy to compute given the value function. It was also shown that under typical conditions the only times that need to be considered as potential stopping times, are times right after an arrival has been accepted. This gives an optimal policy, consisting of an optimal acceptance rule and an optimal stopping time, with a simple, intuitive structure. Examples were given to demonstrate the structure of the optimal value function and optimal threshold under both typical conditions, when intuitive properties hold, and also for cases when counterintuitive behavior is exhibited.

Efficient algorithms for computing optimal solutions were given for both the infinite horizon and finite horizon case. These algorithms have the advantage that they compute the value for each state only once, in contrast with the iterative algorithms that are generally used. It was shown that the optimal expected value and the optimal acceptance rule have a number of intuitive monotonicity and convexity properties.

In the second part of the research a distribution system with a network of terminals and a fleet of vehicles was modeled and analyzed. This problem is called the Dynamic and Stochastic Distribution Problem (DSDP).

As for the DSKP, it was shown that an optimal acceptance rule is given by a threshold rule. Together with an optimal dispatching rule it gives an optimal policy for the DSDP. Optimality equations were derived for the DSDP, and a number of fundamental structural properties were established. The optimal expected value and an optimal policy can be computed with a number of classical algorithms. The convergence of these algorithms was established, in spite of the fact that the DSDP does not satisfy the assumptions usually made for these types of algorithms. An algorithm that exploits the structure of the DSDP was developed, and its convergence to the optimal solution was shown. The performance of these algorithms were compared in computational experiments.

A variety of issues in freight distribution that require further research, and possible approaches for addressing these issues, were discussed.

List of References

List of References

[1] U.S. industrial outlook 1994. Report s/n 003-009-00635-0, U.S. Department of Commerce/International Trade Administration, Washington, D.C., 1994.

[2] E.A. Adamidou, A.L. Kornhauser, and Y.A. Koskosidis. A game theoretic/network equilibrium solution approach for the railroad freight car management problem. *Transportation Research*, 27B:237–252, 1993.

[3] S.C. Albright. Optimal sequential assignments with random arrival times. *Management Science*, 21:60–67, 1974.

[4] S.C. Albright. A Bayesian approach to a general house selling problem. *Management Science*, 24:432–440, 1977.

[5] S.C. Albright and C. Derman. Asymptotic optimal policies for the stochastic sequential assignment problem. *Management Science*, 19:46–51, 1972.

[6] J. Alstrup, S. Boas, O.B.G. Madsen, and R.V.V. Vidal. Booking policy for flights with two types of passengers. *European Journal of Operational Research*, 27:274–288, 1986.

[7] A. Balakrishnan, T.W. Chien, and R.T. Wong. Selecting aircraft routes for long-haul operations: A formulation and solution method. *Transportation Research*, 24B:57–72, 1990.

[8] M. Ball, L.D. Bodin, and R. Dial. A matching based heuristic for scheduling mass transit crews and vehicles. *Transportation Science*, 17:4–31, 1983.

[9] M. Ball and A. Roberts. A graph partitioning approach to airline crew scheduling. *Transportation Science*, 19:107–126, 1985.

[10] C. Barnhart, E.L. Johnson, G.L. Nemhauser, G. Sigismondi, and P. Vance. Formulating a mixed integer distribution problem to improve solvability. *Operations Research*, 41:1013–1019, 1993.

[11] P.P. Belobaba. Airline yield management. an overview of seat inventory control. *Transportation Science*, 21:63–73, 1987.

[12] P.P. Belobaba. Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, 37:183–197, 1989.

[13] J. Benjamin. An analysis of inventory and transportation costs in a constrained network. *Transportation Science*, 23:177–183, 1989.

[14] D. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, New York, NY, 1995.

[15] D.J. Bertsimas. A vehicle routing problem with stochastic demand. *Operations Research*, 40:574–585, 1992.

[16] D.J. Bertsimas and G. Van Ryzin. A stochastic and dynamic vehicle routing problem in the Euclidean plane. *Operations Research*, 39:601–615, 1991.

[17] D.J. Bertsimas and G. Van Ryzin. Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles. *Operations Research*, 41:60–76, 1993.

[18] R.E. Bixby, J.W. Gregory, I.J. Lustig, R.E. Marsten, and D.F. Shanno. Very large-scale linear programming: A case study in combining interior point and simplex methods. *Operations Research*, 40:885–897, 1992.

[19] D. Blackwell. Discrete dynamic programming. *Annals of Mathematical Statistics*, 33:719–726, 1962.

[20] L. Bodin, B. Golden, A. Assad, and M. Ball. Routing and scheduling of vehicles and crews: The state of the art. *Computers and Operations Research*, 10:63–211, 1983.

[21] K. Boel and P. Varaiya. Optimal control of jump processes. *SIAM Journal on Control and Optimization*, 15:92–119, 1977.

[22] J. Bramel, E.G. Coffman, P.W. Shor, and D. Simchi-Levi. Probabilistic analysis of the capacitated vehicle routing problem with unsplit demands. *Operations Research*, 40:1095–1106, 1992.

[23] P. Brémaud. *Point Processes and Queues, Martingale Dynamics.* Springer-Verlag, New York, NY, 1981.

[24] S.L. Brumelle and J.I. McGill. Airline seat allocation with multiple nested fare classes. *Operations Research*, 41:127–137, 1993.

[25] S.L. Brumelle, J.I. McGill, T.H. Oum, K. Sawaki, and M.W. Tretheway. Allocation of airline seats between stochastically dependent demands. *Transportation Science*, 24:183–192, 1990.

[26] F.T. Bruss. A unified approach to a class of best choice problems with an unknown number of options. *The Annals of Probability*, 12:882–889, 1984.

[27] F.T. Bruss. Advances in best choice modelling. *Operations Research Proceedings*, pages 649–656, 1986.

[28] F.T. Bruss and L.C.G. Rogers. Embedding optimal selection problems in a Poisson process. Working Paper 88-020, Systems and Industrial Engineering Department, The University of Arizona, Tucson, AZ 85721, November 1988.

[29] L.D. Burns, R.W. Hall, D.E. Blumenfeld, and C.F. Daganzo. Distribution strategies that minimize transportation and inventory costs. *Operations Research*, 33:469–490, 1985.

[30] J.F. Campbell. Freight consolidation and routing with transportation economies of scale. *Transportation Research*, 24B:345–361, 1990.

[31] J.F. Campbell. Locating transportation terminals to serve an expanding demand. *Transportation Research*, 24B:173–192, 1990.

[32] J.F. Campbell. One-to-many distribution with transhipments: An analytic model. *Transportation Science*, 27:330–340, 1993.

[33] M. Carey. Extending a train pathing model from one-way to two-way track. *Transportation Research*, 28B:395–400, 1994.

[34] M. Carey. A model and strategy for train pathing with choice of lines, platforms and routes. *Transportation Research*, 28B:333–353, 1994.

[35] R.L. Carraway, R.L. Schmidt, and L.R. Weatherford. An algorithm for maximizing target achievement in the stochastic knapsack problem with normal returns. *Naval Research Logistics Quarterly*, 40:161–173, 1993.

[36] T.W. Chien, A. Balakrishnan, and R.T. Wong. An integrated inventory allocation and vehicle routing problem. *Transportation Science*, 23:67–76, 1989.

[37] G. Choquet. *Topology.* Academic Press, New York, NY, 1966.

[38] G.A. Corea and V.G. Kulkarni. Minimum cost routing on stochastic networks. *Operations Research*, 38:527–536, 1990.

[39] T.G. Crainic, J.A. Ferland, and J.-M. Rousseau. A tactical planning model for rail freight transportation. *Transportation Science*, 18:165–184, 1984.

[40] T.G. Crainic, M. Florian, and J.-E. Léal. A model for the strategic planning of national freight transportation by rail. *Transportation Science*, 24:1–24, 1990.

[41] T.G. Crainic and J.-M. Rousseau. Multicommodity, multimode freight transportation: A general modeling and algorithmic framework for the service network design problem. *Transportation Research*, 20B:225–242, 1986.

[42] T.G. Crainic and J. Roy. Design of regular intercity driver routes for the LTL motor carrier industry. *Transportation Science*, 26:280–295, 1992.

[43] J.R. Current. The design of a hierarchical transportation network with transhipment facilities. *Transportation Science*, 22:270–277, 1988.

[44] J.R. Current, C.S. ReVelle, and J.L. Cohon. The median shortest path problem: A multiobjective approach to analyze cost vs. accessibility in the design of transportation networks. *Transportation Science*, 21:188–197, 1987.

[45] R.E. Curry. Optimum airline seat allocation with fare classes nested by origins and destinations. *Transportation Science*, 24:193–204, 1990.

[46] C.F. Daganzo. An approximate analytic model of many-to-many demand responsive transportation systems. *Transportation Research*, 12:325–333, 1978.

[47] C.F. Daganzo. The distance traveled to visit N points with a maximum of C stops per vehicle: An analytic model and an application. *Transportation Science*, 18:331–350, 1984.

[48] C.F. Daganzo. The break-bulk role of terminals in many-to-many logistic networks. *Operations Research*, 35:543–555, 1987.

[49] C.F. Daganzo. Modeling distribution problems with time windows: Part I. *Transportation Science*, 21:171–179, 1987.

[50] C.F. Daganzo. Modeling distribution problems with time windows: Part II. two customer types. *Transportation Science*, 21:180–187, 1987.

[51] C.F. Daganzo. A comparison of in-vehicle and out-of-vehicle freight consolidation strategies. *Transportation Research*, 22B:173–180, 1988.

[52] C.F. Daganzo. Shipment composition enhancement at a consolidation center. *Transportation Research*, 22B:103–124, 1988.

[53] C.F. Daganzo and R.W. Hall. A routing model for pickups and deliveries: No capacity restrictions on the secondary items. *Transportation Science*, 27:315–329, 1993.

[54] M.S. Daskin and N.D. Panayotopoulos. A Lagrangian relaxation approach to assigning aircraft to routes in hub and spoke networks. *Transportation Science*, 23:91–99, 1989.

[55] R.K. Deb. Optimal control of batch service queues with switching costs. *Advances in Applied Probability*, 8:177–194, 1976.

[56] R.K. Deb. Optimal dispatching of a finite capacity shuttle. *Management Science*, 24:1362–1372, 1978.

[57] R.K. Deb and C.P. Schmidt. Optimal average cost policies for the two-terminal shuttle. *Management Science*, 33:662–669, 1987.

[58] R.K. Deb and R.F. Serfozo. Optimal control of batch service queues. *Advances in Applied Probability*, 5:340–361, 1973.

[59] P. Dejax and T.G. Crainic. A review of empty flows and fleet management models in freight transportation. *Transportation Science*, 21:227–248, 1987.

[60] C. Derman, G.J. Lieberman, and S.M. Ross. A sequential stochastic assignment problem. *Management Science*, 18:349–355, 1972.

[61] C. Derman and R. Strauch. A note on memoryless rules for controlling sequential control processes. *The Annals of Mathematical Statistics*, 37:276–278, 1966.

[62] M. Desrochers, J. Desrosiers, and F. Soumis. A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research*, 40:342–354, 1992.

[63] M. Desrochers and F. Soumis. A generalized permanent labeling algorithm for the shortest path problem with time windows. *INFOR*, 26:191–211, 1988.

[64] M. Desrochers and F. Soumis. A column generation approach to the urban transit crew scheduling problem. *Transportation Science*, 23:1–13, 1989.

[65] J. Desrosiers, F. Soumis, and M. Desrochers. Routing with time windows by column generation. *Networks*, 14:545–565, 1984.

[66] G. Dobson and P.J. Lederer. Airline scheduling and routing in a hub-and-spoke system. *Transportation Science*, 27:281–297, 1993.

[67] O. Drissi-Kaïtouni. Solution approaches for multimode multiproduct assignment problems. *Transportation Research*, 25B:317–327, 1991.

[68] M. Dror, G. Laporte, and P. Trudeau. Vehicle routing with stochastic demands: Properties and solution frameworks. *Transportation Science*, 23:166–176, 1989.

[69] M. Dror, P. Trudeau, and S.P. Ladany. Network models for seat allocation of flights. *Transportation Research*, 22B:239–250, 1988.

[70] Y. Dumas, J. Desrosiers, and F. Soumis. The pickup and delivery problem with time windows. *European Journal of Operational Research*, 54:7–22, 1991.

[71] M.M. Etschmaier and D.F.X. Mathaisel. Airline scheduling: An overview. *Transportation Science*, 19:127–138, 1985.

[72] J.M. Farvolden and W.B. Powell. Subgradient methods for the service network design problem. *Transportation Science*, 28:256–272, 1994.

[73] L.F. Frantzeskakis and W.B. Powell. A successive linear approximation procedure for stochastic, dynamic vehicle allocation problems. *Transportation Science*, 24:40–57, 1990.

[74] P.R. Freeman. The secretary problem and its extensions: A review. *International Statistical Review*, 51:189–206, 1983.

[75] G. Gallego and G. Van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40:999–1020, 1994.

[76] C. Gédéon, M. Florian, and T.G. Crainic. Determining origin-destination matrices and optimal multiproduct flows for freight transportation over multimodal networks. *Transportation Research*, 27B:351–368, 1993.

[77] J. Guélat, M. Florian, and T.G. Crainic. A multimode multiproduct network assignment model for strategic planning of freight flows. *Transportation Science*, 24:25–39, 1990.

[78] A.E. Haghani. Formulation and solution of a combined train routing and makeup, and empty car distribution model. *Transportation Research*, 23B:433–452, 1989.

[79] R.W. Hall. Vehicle scheduling at a transportation terminal with random delay en route. *Transportation Science*, 19:308–320, 1985.

[80] R.W. Hall. The fastest path through a network with random time-dependent travel times. *Transportation Science*, 20:182–188, 1986.

[81] R.W. Hall. Comparison of strategies for routing shipments through transportation terminals. *Transportation Research*, 21A:421–429, 1987.

[82] R.W. Hall. Characteristics of multi-stop/multi-terminal delivery routes, with backhauls and unique items. *Transportation Research*, 25B:391–403, 1991.

[83] R.W. Hall. Route selection on freight networks with weight and volume constraints. *Transportation Research*, 25B:175–189, 1991.

[84] R.W. Hall. Design for local area freight networks. *Transportation Research*, 27B:79–95, 1993.

[85] M.I. Henig. Risk criteria in a stochastic knapsack problem. *Operations Research*, 38:820–825, 1990.

[86] K.L. Hoffman and M. Padberg. Solving airline crew scheduling problems by branch-and-cut. *Management Science*, 39:657–682, 1993.

[87] P. Jaillet and A.R. Odoni. The probabilistic vehicle routing problem. In B.L. Golden and A.A. Assad, editors, *Vehicle Routing: Methods and Studies*, chapter 13, pages 293–318. North-Holland, Amsterdam, Netherlands, 1988.

[88] A.I.Z. Jarrah, G. Yu, N. Krishnamurthy, and A. Rakshit. A decision support framework for airline flight cancellations and delays. *Transportation Science*, 27:266–280, 1993.

[89] J.-J. Jaw, A.R. Odoni, H.N. Psaraftis, and N.H.M. Wilson. A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows. *Transportation Research*, 20B:243–257, 1986.

[90] W.C. Jordan and M.A. Turnquist. A stochastic, dynamic network model for railroad car distribution. *Transportation Science*, 17:123–145, 1983.

[91] D. Jovanović and P.T. Harker. Tactical scheduling of rail operations: The SCAN I system. *Transportation Science*, 25:46–64, 1991.

[92] J.F. Kaufman. Blocking in a shared resource environment. *IEEE Transactions on Communications*, 29:1474–1481, 1981.

[93] M.H. Keaton. Designing optimal railroad operating plans: Lagrangian relaxation and heuristic approaches. *Transportation Research*, 23B:415–431, 1989.

[94] M.H. Keaton. Designing railroad operating plans: A dual adjustment method for implementing Lagrangian relaxation. *Transportation Science*, 26:263–279, 1992.

[95] D.P. Kennedy. Optimal sequential assignment. *Mathematics of Operations Research*, 11:619–626, 1986.

[96] Y. Kinberg and A.G. Rao. Stochastic models of a price promotion. *Management Science*, 21:897–907, 1975.

[97] W.M. Kincaid and D.A. Darling. An inventory pricing problem. *Journal of Mathematical Analysis and Applications*, 7:183–208, 1963.

[98] Y.A. Koskosidis. Clustering algorithms for consolidation of customer orders into vehicle shipments. *Transportation Research*, 26B:365–379, 1992.

[99] Y.A. Koskosidis, W.B. Powell, and M.M. Solomon. An optimization-based heuristic for vehicle routing and scheduling with soft time window constraints. *Transportation Science*, 26:69–85, 1992.

[100] G. Laporte, F. Louveaux, and H. Mercure. The vehicle routing problem with stochastic travel times. *Transportation Science*, 26:161–170, 1992.

[101] G. Laporte, Y. Nobert, and S. Taillefer. Solving a family of multi-depot vehicle routing and location-routing problems. *Transportation Science*, 22:161–172, 1988.

[102] R.C. Larson. Transporting sludge to the 106-mile site: An inventory/routing model for fleet sizing and logistics system design. *Transportation Science*, 22:–, 1988.

[103] T.C. Lee and M. Hersh. A model for dynamic airline seat inventory control with multiple seat bookings. *Transportation Science*, 27:252–265, 1993.

[104] J.M.Y. Leung, T.L. Magnanti, and V. Singhal. Routing in point-to-point delivery systems: Formulations and solution heuristics. *Transportation Science*, 24:245–260, 1990.

[105] C.-L. Li, D. Simchi-Levi, and M. Desrochers. On the distance constrained vehicle routing problem. *Operations Research*, 40:790–799, 1992.

[106] T.L. Magnanti and R.T. Wong. Network design and transportation planning: Models and algorithms. *Transportation Science*, 18:1–55, 1984.

[107] C. Malandraki and M.S. Daskin. Time dependent vehicle routing problems: Formulations, properties and heuristic algorithms. *Transportation Science*, 26:185–200, 1992.

[108] J.W. Mamer. Successive approximations for finite horizon, semi-Markov decision processes with application to asset liquidation. *Operations Research*, 34:638–644, 1986.

[109] S. Martello and P. Toth. *Knapsack Problems. Algorithms and Computer Implementations.* John Wiley & Sons, West Sussex, England, 1990.

[110] H. Mendelson, J.S. Pliskin, and U. Yechiali. A stochastic allocation problem. *Operations Research*, 28:687–693, 1980.

[111] B.L. Miller. Finite state continuous time Markov decision processes with a finite planning horizon. *SIAM Journal on Control*, 6:266–280, 1968.

[112] T. Nakai. An optimal selection problem for a sequence with a random number of applicants per period. *Operations Research*, 34:478–485, 1986.

[113] T. Nakai. A sequential stochastic assignment problem in a partially observable Markov chain. *Mathematics of Operations Research*, 11:230–240, 1986.

[114] T. Nakai. A sequential stochastic assignment problem in a stationary Markov chain. *Mathematica Japonica*, 31:741–757, 1986.

[115] M.F. Neuts. A general class of bulk queues with Poisson input. *The Annals of Mathematical Statistics*, 38:759–770, 1967.

[116] M.F. Neuts. Queues solvable without Rouche's theorem. *Operations Research*, 27:767–781, 1979.

[117] J.D. Papastavrou, S. Rajagopalan, and A.J. Kleywegt. A stochastic model for the knapsack problem with a deadline. Technical report, School of Industrial Engineering, Purdue University, West Lafayette, IN 47907-1287, July 1995.

[118] A.N. Perakis and N.A. Papadakis. Minimal time vessel routing in a time-dependent environment. *Transportation Science*, 23:266–276, 1989.

[119] H. Pirkul, J. Current, and V. Nagarajan. The hierarchical network design problem: A new formulation and solution procedures. *Transportation Science*, 25:175–182, 1991.

[120] S.R. Pliska. Controlled jump processes. *Stochastic Processes and their Applications*, 3:259–282, 1975.

[121] D.A. Popken. An algorithm for the multiattribute, multicommodity flow problem with freight consolidation and inventory costs. *Operations Research*, 42:274–286, 1994.

[122] W.B. Powell. Analysis of vehicle holding and cancellation strategies in bulk arrival, bulk service queues. *Transportation Science*, 19:352–377, 1985.

[123] W.B. Powell. Approximate, closed form moment formulas for bulk arrival, bulk service queues. *Transportation Science*, 20:13–23, 1986.

[124] W.B. Powell. A local improvement heuristic for the design of less-than-truckload motor carrier networks. *Transportation Science*, 20:246–257, 1986.

[125] W.B. Powell. A stochastic model of the dynamic vehicle allocation problem. *Transportation Science*, 20:117–129, 1986.

[126] W.B. Powell. An operational planning model for the dynamic vehicle allocation problem with uncertain demands. *Transportation Research*, 21B:217–232, 1987.

[127] W.B. Powell. A comparative review of alternative algorithms for the dynamic vehicle allocation problem. In B.L. Golden and A.A. Assad, editors, *Vehicle Routing: Methods and Studies*, chapter 12, pages 249–291. North-Holland, Amsterdam, Netherlands, 1988.

[128] W.B. Powell. Optimization models and algorithms: An emerging technology for the motor carrier industry. *IEEE Transactions on Vehicular Technology*, 40:68–80, 1991.

[129] W.B. Powell and P. Humblet. The bulk service queue with a general control strategy: Theoretical analysis and a new computational procedure. *Operations Research*, 34:267–275, 1986.

[130] W.B. Powell and Y.A. Koskosidis. Shipment routing algorithms with tree constraints. *Transportation Science*, 26:230–245, 1992.

[131] W.B. Powell and Y. Sheffi. Design and implementation of an interactive optimization system for network design in the motor carrier industry. *Operations Research*, 37:12–29, 1989.

[132] W.B. Powell and H.P. Simão. Numerical simulation of transient bulk queues with general vehicle dispatching strategies. *Transportation Research*, 20B:477–490, 1986.

[133] G.P. Prastacos. Optimal sequential investment decisions under conditions of uncertainty. *Management Science*, 29:118–134, 1983.

[134] E.L. Presman and I.M. Sonin. The best choice problem for a random number of objects. *Theory of Probability and its Applications*, 17:657–668, 1972.

[135] H.N. Psaraftis. A dynamic programming solution to the single vehicle many-to-many immediate request dial-a-ride problem. *Transportation Science*, 14:130–154, 1980.

[136] H.N. Psaraftis. Analysis of an $O(N^2)$ heuristic for the single vehicle many-to-many Euclidean dial-a-ride problem. *Transportation Research*, 17B:133–145, 1983.

[137] H.N. Psaraftis. An exact algorithm for the single vehicle many-to-many dial-a-ride problem with time windows. *Transportation Science*, 17:351–357, 1983.

[138] H.N. Psaraftis. Dynamic vehicle routing problems. In B.L. Golden and A.A. Assad, editors, *Vehicle Routing: Methods and Studies*, chapter 11, pages 223–248. North-Holland, Amsterdam, Netherlands, 1988.

[139] M.L. Puterman. *Markov Decision Processes.* John Wiley & Sons, Inc., New York, NY, 1994.

[140] K. Rana and R.G. Vickson. A model and solution algorithm for optimal routing of a time-chartered containership. *Transportation Science*, 22:83–95, 1988.

[141] K. Rana and R.G. Vickson. Routing container ships using Lagrangean relaxation and decomposition. *Transportation Science*, 25:201–214, 1991.

[142] C.C. Ribeiro and F. Soumis. A column generation approach to the multiple-depot vehicle scheduling problem. *Operations Research*, 42:41–52, 1994.

[143] R. Righter. A resource allocation problem in a random environment. *Operations Research*, 37:329–338, 1989.

[144] L.W. Robinson. Optimal and approximate control policies for airline booking with sequential nonmonotonic fare classes. *Operations Research*, 43:252–263, 1995.

[145] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[146] D. Ronen. Dispatching petroleum products. *Operations Research*, 43:379–387, 1995.

[147] D.B. Rosenfield, R.D. Shapiro, and D.A. Butler. Optimal strategies for selling an asset. *Management Science*, 29:1051–1061, 1983.

[148] K.W. Ross and D.H.K. Tsang. The stochastic knapsack problem. *IEEE Transactions on Communications*, 37:740–747, 1989.

[149] K.W. Ross and D.D. Yao. Monotonicity properties for the stochastic knapsack. *IEEE Transactions on Information Theory*, 36:1173–1179, 1990.

[150] M. Rothstein. An airline overbooking model. *Transportation Science*, 5:180–192, 1971.

[151] M. Rothstein. Hotel overbooking as a Markovian sequential decision process. *Decision Science*, 5:389–404, 1974.

[152] M. Rothstein. OR and the airline overbooking problem. *Operations Research*, 33:237–248, 1985.

[153] H.L. Royden. *Real Analysis*. Macmillan, New York, NY, 1988.

[154] V.V. Rykov. Markov decision processes with finite state and decision spaces. *Theory of Probability and its Applications*, 11:302–313, 1966.

[155] V. Saario. Limiting properties of the discounted house-selling problem. *European Journal of Operational Research*, 20:206–210, 1985.

[156] M. Sakaguchi. Best choice problems for randomly arriving offers during a random lifetime. *Mathematica Japonica*, 31:107–117, 1984.

[157] M. Sakaguchi. A sequential stochastic assignment problem associated with a non-homogeneous Markov process. *Mathematica Japonica*, 29:13–22, 1984.

[158] M. Sakaguchi. A sequential stochastic assignment problem with an unknown number of jobs. *Mathematica Japonica*, 29:141–152, 1984.

[159] R.F. Serfozo. An equivalence between continuous and discrete time Markov decision processes. *Operations Research*, 27:616–620, 1979.

[160] T.R. Sexton and L.D. Bodin. Optimizing single vehicle many-to-many operations with desired delivery times: I. scheduling. *Transportation Science*, 19:378–410, 1985.

[161] T.R. Sexton and L.D. Bodin. Optimizing single vehicle many-to-many operations with desired delivery times: II. routing. *Transportation Science*, 19:411–435, 1985.

[162] E. Shlifer and Y. Vardi. An airline overbooking policy. *Transportation Science*, 9:101–114, 1975.

[163] E.A. Silver. Operations research in inventory management: A review and critique. *Operations Research*, 29:628–645, 1981.

[164] H.P. Simão and W.B. Powell. Numerical methods for simulating transient, stochastic queueing networks—I: Methodology. *Transportation Science*, 26:296–311, 1992.

[165] H.P. Simão and W.B. Powell. Numerical methods for simulating transient, stochastic queueing networks—II: Experimental design. *Transportation Science*, 26:312–329, 1992.

[166] M.G. Sklar, R.D. Armstrong, and S. Samn. Heuristics for scheduling aircraft and crew during airlift operations. *Transportation Science*, 24:63–76, 1990.

[167] M. Sniedovich. Preference order stochastic knapsack problems: Methodological issues. *Journal of the Operational Research Society*, 31:1025–1032, 1980.

[168] M. Sniedovich. Some comments on preference order dynamic programming models. *Journal of Mathematical Analysis and Applications*, 79:489–501, 1981.

[169] M.M. Solomon and J. Desrosiers. Time window constrained routing and scheduling problems. *Transportation Science*, 22:1–13, 1988.

[170] F. Soumis, M. Sauvé, and L. Le Beau. The simultaneous origin-destination assignment and vehicle routing problem. *Transportation Science*, 25:188–200, 1991.

[171] W. Stadje. A full information pricing problem for the sale of several identical commodities. *Zeitschrift fűr Operations Research*, 34:161–181, 1990.

[172] E. Steinberg and M.S. Parks. A preference order dynamic program for a knapsack problem with stochastic rewards. *Journal of the Operational Research Society*, 30:141–147, 1979.

[173] T.J. Stewart. The secretary problem with an unknown number of options. *Operations Research*, 29:130–145, 1981.

[174] R.S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3:9–44, 1988.

[175] M.R. Swihart. A stochastic and dynamic model for the single-vehicle dial-a-ride problem. Master's thesis, School of Industrial Engineering, Purdue University, August 1994.

[176] M. Tamaki. A full-information best-choice problem with finite memory. *Journal of Applied Probability*, 23:718–735, 1986.

[177] M. Tamaki. A generalized problem of optimal selection and assignment. *Operations Research*, 34:486–493, 1986.

[178] D. Teodorović and E. Krčmar-Nožić. Multicriteria model to determine flight frequencies on an airline network under competitive conditions. *Transportation Science*, 23:14–25, 1989.

[179] L. Totten and N. Tohamy. Home for the holidays. *OR/MS Today*, 22:24–29, 1995.

[180] P. Trudeau and M. Dror. Stochastic inventory routing: Route design with stockouts and route failures. *Transportation Science*, 26:171–184, 1992.

[181] J.N. Tsitsiklis. Asynchronous stochastic approximation and Q-learning. *Machine Learning*, 16:185–202, 1994.

[182] L.J.J. Van der Bruggen, J.K. Lenstra, and P.C. Schuur. Variable-depth search for the single-vehicle pickup and delivery problem with time windows. *Transportation Science*, 27:298–311, 1993.

[183] L.R. Weatherford and S.E. Bodily. A taxonomy and research overview of perishable-asset revenue management: Yield management, overbooking, and pricing. *Operations Research*, 40:831–844, 1992.

[184] L.R. Weatherford, S.E. Bodily, and P.E. Pfeifer. Modeling the customer arrival process and comparing decision rules in perishable asset revenue management situations. *Transportation Science*, 27:239–251, 1993.

[185] R.D. Wollmer. An airline seat management model for a single leg route when lower fare classes book first. *Operations Research*, 40:26–37, 1992.

[186] M. Yasuda. Asymptotic results for the best-choice problem with a random number of objects. *Journal of Applied Probability*, 21:521–536, 1984.

[187] A.A. Yushkevich and E.A. Feinberg. On homogeneous Markov models with continuous time and finite or countable state space. *Theory of Probability and its Applications*, 24:156–161, 1979.

Appendices

Appendix A: Proof of Lemma 1

Lemma 1

$$
\begin{aligned}
V_{DSKP}^{\pi} \;=\; E\Bigg[& \sum_{\{i:A_i \le T\}} e^{-\alpha A_i} \left[R_i D_i^{\pi} - p(N^{\pi}(A_i), A_i)\,(1 - D_i^{\pi}) \right] I^{\pi}(A_i) \\
& + \int_0^T e^{-\alpha\tau} \left[-c(N^{\pi}(\tau), \tau) I^{\pi}(\tau) + \alpha v(N^{\pi}(\tau))\,(1 - I^{\pi}(\tau)) \right] d\tau \\
& + e^{-\alpha T} v(N^{\pi}(T^{+})) \,\Bigg|\, N^{\pi}(0^{+}) = N_0 \Bigg]
\end{aligned}
$$

Proof:

Case 1: $T < \infty$ and $\alpha = 0$:

$$
\begin{aligned}
& E\Bigg[\sum_{\{i:A_i \le T\}} e^{-\alpha A_i} \left[R_i D_i^{\pi} - p(N^{\pi}(A_i), A_i)\,(1 - D_i^{\pi}) \right] I^{\pi}(A_i) \\
& \qquad + \int_0^T e^{-\alpha\tau} \left[-c(N^{\pi}(\tau), \tau) I^{\pi}(\tau) + \alpha v(N^{\pi}(\tau))\,(1 - I^{\pi}(\tau)) \right] d\tau \\
& \qquad + e^{-\alpha T} v(N^{\pi}(T^{+})) \,\Bigg|\, N^{\pi}(0^{+}) = N_0 \Bigg] \\
= \; & E\Bigg[\sum_{\{i:A_i \le \mathcal{T}^{\pi}\}} \left[R_i D_i^{\pi} - p(N^{\pi}(A_i), A_i)\,(1 - D_i^{\pi}) \right] \\
& \qquad - \int_0^{\mathcal{T}^{\pi}} c(N^{\pi}(\tau), \tau)\, d\tau \;+\; v(N^{\pi}(\mathcal{T}^{\pi +})) \,\Bigg|\, N^{\pi}(0^{+}) = N_0 \Bigg] \\
= \; & V_{DSKP}^{\pi}
\end{aligned}
$$

Case 2: $T < \infty$ and $\alpha > 0$:

$$
\begin{aligned}
& E\Bigg[\sum_{\{i:A_i \le T\}} e^{-\alpha A_i} \left[R_i D_i^{\pi} - p(N^{\pi}(A_i), A_i)\,(1 - D_i^{\pi}) \right] I^{\pi}(A_i) \\
& \qquad + \int_0^T e^{-\alpha\tau} \left[-c(N^{\pi}(\tau), \tau) I^{\pi}(\tau) + \alpha v(N^{\pi}(\tau))\,(1 - I^{\pi}(\tau)) \right] d\tau \\
& \qquad + e^{-\alpha T} v(N^{\pi}(T^{+})) \,\Bigg|\, N^{\pi}(0^{+}) = N_0 \Bigg] \\
= \; & E\Bigg[\sum_{\{i:A_i \le \mathcal{T}^{\pi}\}} e^{-\alpha A_i} \left[R_i D_i^{\pi} - p(N^{\pi}(A_i), A_i)\,(1 - D_i^{\pi}) \right] \\
& \qquad - \int_0^{\mathcal{T}^{\pi}} e^{-\alpha\tau} c(N^{\pi}(\tau), \tau)\, d\tau \;+\; \int_{\mathcal{T}^{\pi}}^{T} e^{-\alpha\tau} \alpha v(N^{\pi}(\mathcal{T}^{\pi +}))\, d\tau
\end{aligned}
$$

$$
\begin{aligned}
& \left. + \, e^{-\alpha T} v(N^\pi(\mathcal{T}^{\pi +})) \right| N^\pi(0^+) = N_0 \Bigg] \\
= \;& E \Bigg[\sum_{\{i: A_i \le \mathcal{T}^\pi\}} e^{-\alpha A_i} \left[R_i D_i^\pi - p(N^\pi(A_i), A_i)\,(1 - D_i^\pi) \right] \\
& - \int_0^{\mathcal{T}^\pi} e^{-\alpha\tau} c(N^\pi(\tau), \tau)\, d\tau \; + \; \left[-e^{-\alpha\tau} \right]_{\mathcal{T}^\pi}^{T} v(N^\pi(\mathcal{T}^{\pi +})) \\
& \left. + \, e^{-\alpha T} v(N^\pi(\mathcal{T}^{\pi +})) \right| N^\pi(0^+) = N_0 \Bigg] \\
= \;& V_{DSKP}^\pi
\end{aligned}
$$

Case 3: $T = \infty$ and $\alpha > 0$:

$$
\begin{aligned}
& E \Bigg[\sum_{\{i: A_i \le T\}} e^{-\alpha A_i} \left[R_i D_i^\pi - p(N^\pi(A_i), A_i)\,(1 - D_i^\pi) \right] I^\pi(A_i) \\
& + \int_0^{T} e^{-\alpha\tau} \left[-c(N^\pi(\tau), \tau) I^\pi(\tau) + \alpha v(N^\pi(\tau))\,(1 - I^\pi(\tau)) \right] d\tau \\
& \left. + \, e^{-\alpha T} v(N^\pi(T^+)) \right| N^\pi(0^+) = N_0 \Bigg] \\
= \;& E \Bigg[\sum_{\{i: A_i \le \mathcal{T}^\pi\}} e^{-\alpha A_i} \left[R_i D_i^\pi - p(N^\pi(A_i), A_i)\,(1 - D_i^\pi) \right] \\
& \left. - \int_0^{\mathcal{T}^\pi} e^{-\alpha\tau} c(N^\pi(\tau), \tau)\, d\tau \; + \; \int_{\mathcal{T}^\pi}^{\infty} e^{-\alpha\tau} \alpha v(N^\pi(\mathcal{T}^{\pi +}))\, d\tau \right| N^\pi(0^+) = N_0 \Bigg] \\
= \;& E \Bigg[\sum_{\{i: A_i \le \mathcal{T}^\pi\}} e^{-\alpha A_i} \left[R_i D_i^\pi - p(N^\pi(A_i), A_i)\,(1 - D_i^\pi) \right] \\
& \left. - \int_0^{\mathcal{T}^\pi} e^{-\alpha\tau} c(N^\pi(\tau), \tau)\, d\tau \; + \; \left[-e^{-\alpha\tau} \right]_{\mathcal{T}^\pi}^{\infty} v(N^\pi(\mathcal{T}^{\pi +})) \right| N^\pi(0^+) = N_0 \Bigg] \\
= \;& V_{DSKP}^\pi \qquad \square
\end{aligned}
$$

Appendix B: Proof of Lemma 5

Lemma 5

1. *For all $n \in [0, N_0]$ and $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ satisfies the following Lipschitz condition with respect to y:*

$$|g(n, y_2, \zeta) - g(n, y_1, \zeta)| \quad \leq \quad |y_2 - y_1| \quad \text{for all} \;\; y_1, y_2 \in \Re$$

2. *For all $n \in [0, N_0]$ and $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ is absolutely continuous in y on $\Re$.*

3. *For all $n \in [0, N_0]$ and $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ is nonincreasing in y on $\Re$, and strictly decreasing in y on $\{y \in \Re : g(n, y, \zeta) > 0\}$.*

4. *For any $n \in [0, N_0]$ such that $\int_0^n F_S(ds) > 0$, and $\zeta \in \mathcal{Y}$, $g(n, y, \zeta)$ is unbounded above, i.e., for any $M > 0$, there exists a y_1 such that $g(n, y_1, \zeta) > M$.*

5. *For any $n \in [0, N_0]$, $\zeta \in \mathcal{Y}$, and $\varepsilon > 0$, there exists a y_2 such that $g(n, y_2, \zeta) < \varepsilon$.*

6. *For any Borel set $B \subset \Re \times [0, n]$,*

$$\begin{aligned} & \int\int_B \{r - [y - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \\ \leq \quad & \int_0^n \int_{y-\zeta(n,s)}^{\infty} \{r - [y - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \quad \equiv \quad g(n, y, \zeta) \end{aligned}$$

Proof:

1. Let $y_1 \leq y_2$.

$$\begin{aligned} |g(n, y_1, \zeta) - g(n, y_2, \zeta)| \quad \equiv \quad & \left| \int_0^n \int_{y_1-\zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \right. \\ & \left. - \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y_2 - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \right| \\ = \quad & \left| \int_0^n \int_{y_1-\zeta(n,s)}^{y_2-\zeta(n,s)} \{r - [y_1 - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \right. \\ & + \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \\ & \left. - \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y_2 - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \right| \\ = \quad & \left| \int_0^n \int_{y_1-\zeta(n,s)}^{y_2-\zeta(n,s)} \{r - [y_1 - \zeta(n, s)]\} \; F_{R|S}(dr|s) F_S(ds) \right. \end{aligned}$$

$$
\begin{aligned}
& \quad + \left. \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} (y_2 - y_1)\ F_{R|S}(dr|s)F_S(ds) \right| \\
\leq & \int_0^n \int_{y_1-\zeta(n,s)}^{y_2-\zeta(n,s)} \{[y_2 - \zeta(n,s)] - [y_1 - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) \\
& \quad + \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} (y_2 - y_1)\ F_{R|S}(dr|s)F_S(ds) \\
= & \ (y_2 - y_1) \int_0^n \int_{y_1-\zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds) \\
\leq & \ |y_2 - y_1|
\end{aligned}
$$

2. Absolute continuity of $g(n, y, \zeta)$ in y follows directly from $g(n, y, \zeta)$ satisfying a Lipschitz condition with respect to y.

3. Let $y_1 < y_2$.

$$
\begin{aligned}
g(n, y_1, \zeta) \equiv & \int_0^n \int_{y_1-\zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) \\
= & \int_0^n \int_{y_1-\zeta(n,s)}^{y_2-\zeta(n,s)} \{r - [y_1 - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) \\
& + \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) \\
\geq & \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) \\
\geq & \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y_2 - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) \\
\equiv & \ g(n, y_2, \zeta)
\end{aligned}
$$

For any y, suppose $g(n, y, \zeta) > 0$. Then $\int_0^n \int_{y-\zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds) > 0$, and for any $y_1 < y$, the second inequality above (with $y_2 = y$) is strict, so that $g(n, y_1, \zeta) > g(n, y, \zeta)$. Consider any $y_2 > y$. If $\int_0^n \int_{y-\zeta(n,s)}^{y_2-\zeta(n,s)} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) > 0$, then the first inequality above (with $y_1 = y$) is strict, so that $g(n, y, \zeta) > g(n, y_2, \zeta)$. Otherwise, because $g(n, y, \zeta) \equiv \int_0^n \int_{y-\zeta(n,s)}^{\infty} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) > 0$, it must hold that $\int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s)F_S(ds) > 0$. Then $\int_0^n \int_{y_2-\zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds) > 0$, and the second inequality above (with $y_1 = y$) is strict, so that $g(n, y, \zeta) > g(n, y_2, \zeta)$. Thus, if $g(n, y, \zeta) > 0$, then $g(n, y, \zeta)$ is strictly decreasing at y.

4. For any $M > 0$, and any $r \in \Re$ and $s \in [0, n]$, $I_{\{r-[y-\zeta(n,s)] \geq 2M/\int_0^n F_S(ds)\}} \uparrow 1$ pointwise as $y \downarrow -\infty$. Hence, from the monotone convergence theorem,

$$
\int_0^n \int_{-\infty}^{\infty} I_{\{r-[y-\zeta(n,s)] \geq 2M/\int_0^n F_S(ds)\}} F_{R|S}(dr|s)F_S(ds) \uparrow
$$

$$\int_0^n \int_{-\infty}^{\infty} 1 F_{R|S}(dr|s) F_S(ds) \quad = \quad \int_0^n F_S(ds)$$

as $y \downarrow -\infty$. Thus, there exists a y_1 such that $\int_0^n \int_{-\infty}^{\infty} I_{\{r-[y_1-\zeta(n,s)]\geq 2M/\int_0^n F_S(ds)\}} F_{R|S}(dr|s) F_S(ds) > \int_0^n F_S(ds)/2$. Therefore,

$$\begin{aligned}
g(n, y_1, \zeta) \quad &\equiv \quad \int_0^n \int_{y_1-\zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&\geq \quad \int_0^n \int_{y_1-\zeta(n,s)+2M/\int_0^n F_S(ds)}^{\infty} \{r - [y_1 - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&= \quad \int_0^n \int_{-\infty}^{\infty} \{r - [y_1 - \zeta(n,s)]\} \, I_{\{r\geq y_1-\zeta(n,s)+2M/\int_0^n F_S(ds)\}} \; F_{R|S}(dr|s) F_S(ds) \\
&\geq \quad \frac{2M}{\int_0^n F_S(ds)} \int_0^n \int_{-\infty}^{\infty} I_{\{r-[y_1-\zeta(n,s)]\geq 2M/\int_0^n F_S(ds)\}} \; F_{R|S}(dr|s) F_S(ds) \\
&> \quad \frac{2M}{\int_0^n F_S(ds)} \frac{\int_0^n F_S(ds)}{2} \quad = \quad M
\end{aligned}$$

5. For any $y > 0$

$$\begin{aligned}
0 \quad &\leq \quad \int_0^n \int_{y-\zeta(n,s)}^{\infty} \{r - [y - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&\leq \quad \int_0^n \int_{y-\zeta(n,s)}^{\infty} [r + \zeta(n,s)] \; F_{R|S}(dr|s) F_S(ds) \\
&= \quad \int_0^n \int_{-\infty}^{\infty} [r + \zeta(n,s)] \, I_{\{r\geq y-\zeta(n,s)\}} \; F_{R|S}(dr|s) F_S(ds)
\end{aligned}$$

For every $r \in \Re$ and $s \in [0, n]$, $[r + \zeta(n,s)] \, I_{\{r\geq y-\zeta(n,s)\}} \downarrow 0$ pointwise as $y \uparrow \infty$. If $r < y - \zeta(n,s)$, then $\left|[r + \zeta(n,s)] \, I_{\{r\geq y-\zeta(n,s)\}}\right| = 0 \leq \max\{r, 0\} + \max\{\zeta(n,s), 0\}$. If $r \geq y - \zeta(n,s)$, then $r + \zeta(n,s) \geq y > 0$, and $\left|[r + \zeta(n,s)] \, I_{\{r\geq y-\zeta(n,s)\}}\right| = r + \zeta(n,s) \leq \max\{r, 0\} + \max\{\zeta(n,s), 0\}$. Then

$$\begin{aligned}
&\int_0^n \int_{-\infty}^{\infty} [\max\{r, 0\} + \max\{\zeta(n,s), 0\}] \; F_{R|S}(dr|s) F_S(ds) \\
= \quad &\int_0^n \int_0^{\infty} r \, F_{R|S}(dr|s) F_S(ds) \; + \; \int_0^n \int_{-\infty}^{\infty} \max\{\zeta(n,s), 0\} \; F_{R|S}(dr|s) F_S(ds) \quad < \quad \infty
\end{aligned}$$

from the assumptions that $\int_0^n \int_0^{\infty} r \, F_{R|S}(dr|s) F_S(ds) < \infty$ and $\int_0^n |\zeta(n,s)| \; F_S(ds) < \infty$. Hence, from the Lebesgue dominated convergence theorem,

$$\lim_{y\to\infty} \int_0^n \int_{-\infty}^{\infty} [r + \zeta(n,s)] \, I_{\{r\geq y-\zeta(n,s)\}} \; F_{R|S}(dr|s) F_S(ds) \quad = \quad 0$$

Hence for any $\varepsilon > 0$, there exists a y_2 such that

$$g(n, y_2, \zeta) \quad \equiv \quad \int_0^n \int_{y_2-\zeta(n,s)}^{\infty} \{r - [y_2 - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \quad < \quad \varepsilon$$

6. For any Borel set $B \subset \Re \times [0, n]$,

$$
\begin{aligned}
& \int\int_B \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds) \\
= & \int\int_{\{(r,s)\in B: r < y - \zeta(n,s)\}} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds) \\
& + \int\int_{\{(r,s)\in B: r \geq y - \zeta(n,s)\}} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds) \\
\leq & \int\int_{\{(r,s)\in B: r \geq y - \zeta(n,s)\}} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds) \\
\leq & \int\int_{\{(r,s)\in \Re\times[0,n]: r \geq y - \zeta(n,s)\}} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds) \\
= & \int_0^n \int_{y-\zeta(n,s)}^{\infty} \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds) \\
\equiv & \ g(n, y, \zeta)
\end{aligned}
$$

Therefore, $g(n, y, \zeta) = \sup_{B\in\mathcal{B}_n} \int\int_B \{r - [y - \zeta(n,s)]\}\ F_{R|S}(dr|s) F_S(ds)$, where $\mathcal{B}_n$ denotes the Borel sets of $\Re \times [0, n]$. $\square$

Appendix C: Proof of Lemma 7

Lemma 7

1. *For all* $n \in [0, N_0]$, $y_n \in \Re$, *and* $\zeta \in \mathcal{Y}$, $h(n, y_n, y, \zeta)$ *satisfies the following Lipschitz condition with respect to* y:

$$|h(n, y_n, y_2, \zeta) - h(n, y_n, y_1, \zeta)| \quad \leq \quad |y_2 - y_1| \quad \textit{for all} \;\; y_1, y_2 \in \Re$$

2. *For all* $n \in [0, N_0]$, $y_n \in \Re$, *and* $\zeta \in \mathcal{Y}$, $h(n, y_n, y, \zeta)$ *is absolutely continuous in* y *on* $\Re$.

3. *For all* $n \in [0, N_0]$, $y_n \in \Re$, *and* $\zeta \in \mathcal{Y}$, $h(n, y_n, y, zeta)$ *is nonincreasing in* y *on* $\Re$, *and strictly decreasing in* y *if* $\int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s) F_S(ds) > 0$.

4. *For any* $n \in [0, N_0]$, $y_n \in \Re$, *and* $\zeta \in \mathcal{Y}$ *such that* $\int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s) F_S(ds) > 0$, $h(n, y_n, y, \zeta)$ *is unbounded above and below, i.e., for any* $M > 0$, *there exists a* y_1 *such that* $h(n, y_n, y_1, \zeta) > M$, *and a* y_2 *such that* $h(n, y_n, y_2, \zeta) < -M$.

Proof:

1.

$$\begin{aligned}
|h(n, y_n, y_1, \zeta) - h(n, y_n, y_2, \zeta)| \;\; &\equiv \;\; \left| \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \right. \\
&\quad \left. - \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} \{r - [y_2 - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \right| \\
&= \;\; |y_2 - y_1| \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s) F_S(ds) \\
&\leq \;\; |y_2 - y_1|
\end{aligned}$$

2. Absolute continuity of $h(n, y_n, y, \zeta)$ in y follows directly from $h(n, y_n, y, \zeta)$ satisfying a Lipschitz condition with respect to y.

3. Let $y_1 < y_2$.

$$\begin{aligned}
h(n, y_n, y_1, \zeta) \; - \; h(n, y_n, y_2, \zeta) \;\; &\equiv \;\; \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} \{r - [y_1 - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&\quad - \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} \{r - [y_2 - \zeta(n,s)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&= \;\; (y_2 - y_1) \int_0^n \int_{y_n - \zeta(n,s)}^{\infty} F_{R|S}(dr|s) F_S(ds) \quad \geq \quad 0
\end{aligned}$$

If $\int_0^n \int_{y_n-\zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds) > 0$, then the inequality is strict.

4.

$$h(n, y_n, y, \zeta) \equiv \int_0^n \int_{y_n-\zeta(n,s)}^{\infty} [r + \zeta(n,s)]\ F_{R|S}(dr|s)F_S(ds) - y \int_0^n \int_{y_n-\zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds)$$

Noting that

$$\left| \int_0^n \int_{y_n-\zeta(n,s)}^{\infty} [r + \zeta(n,s)]\ F_{R|S}(dr|s)F_S(ds) \right| < \infty$$

from the assumptions that $\int_0^n \int_{-\infty}^{\infty} |r|\ F_{R|S}(dr|s)F_S(ds) < \infty$ and $\int_0^n |\zeta(n,s)|\ F_S(ds) < \infty$, it follows that, if $\int_0^n \int_{y_n-\zeta(n,s)}^{\infty} F_{R|S}(dr|s)F_S(ds) > 0$, then $h(n, y_n, y, \zeta)$ can be made arbitrarily large (small) by choosing sufficiently small (large) values of y. $\square$

Appendix D: Proof of Lemma 14

Lemma 14 *The function Υ satisfies the following Lipschitz condition with respect to $V_1, V_2 \in \mathcal{V}$:*

$$|\Upsilon(n,t,V_2) - \Upsilon(n,t,V_1)| \quad \leq \quad 2d_t(V_1,V_2) \quad \leq \quad 2d_{\mathcal{V}}(V_1,V_2)$$

Proof: Let $\mathcal{S}_1(n,t) \equiv \{s \in [0,n] : V_1(n,t) - V_1(n-s,t) \leq V_2(n,t) - V_2(n-s,t)\}$, and let $\mathcal{S}_2(n,t) \equiv \{s \in [0,n] : V_2(n,t) - V_2(n-s,t) < V_1(n,t) - V_1(n-s,t)\}$. Then it follows that

$$\begin{aligned}
0 \quad &\leq \quad \int_{\mathcal{S}_1(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{V_2(n,t)-V_2(n-s,t)-p(n,t)} \{r - [V_1(n,t) - V_1(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&\leq \quad \int_{\mathcal{S}_1(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{V_2(n,t)-V_2(n-s,t)-p(n,t)} \{[V_2(n,t) - V_2(n-s,t)] \\
&\qquad\qquad - [V_1(n,t) - V_1(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds)
\end{aligned}$$

and

$$\begin{aligned}
0 \quad &\leq \quad \int_{\mathcal{S}_2(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{V_1(n,t)-V_1(n-s,t)-p(n,t)} \{r - [V_2(n,t) - V_2(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&\leq \quad \int_{\mathcal{S}_2(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{V_1(n,t)-V_1(n-s,t)-p(n,t)} \{[V_1(n,t) - V_1(n-s,t)] \\
&\qquad\qquad - [V_2(n,t) - V_2(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds)
\end{aligned}$$

Hence

$$\begin{aligned}
&|\Upsilon(n,t,V_2) - \Upsilon(n,t,V_1)| \\
= \; &\Bigg| \int_0^n \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{\infty} \{r - [V_2(n,t) - V_2(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&- \int_0^n \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{\infty} \{r - [V_1(n,t) - V_1(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \Bigg| \\
= \; &\Bigg| \int_{\mathcal{S}_1(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{\infty} \{r - [V_2(n,t) - V_2(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&+ \int_{\mathcal{S}_2(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{V_1(n,t)-V_1(n-s,t)-p(n,t)} \{r - [V_2(n,t) - V_2(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&+ \int_{\mathcal{S}_2(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{\infty} \{r - [V_2(n,t) - V_2(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
&- \int_{\mathcal{S}_1(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{V_2(n,t)-V_2(n-s,t)-p(n,t)} \{r - [V_1(n,t) - V_1(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds)
\end{aligned}$$

$$
\begin{aligned}
& - \int_{\mathcal{S}_1(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{\infty} \{r - [V_1(n,t) - V_1(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& - \int_{\mathcal{S}_2(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{\infty} \{r - [V_1(n,t) - V_1(n-s,t) - p(n,t)]\} \; F_{R|S}(dr|s) F_S(ds) \Bigg| \\
\leq \; & \int_{\mathcal{S}_1(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{\infty} \{[V_2(n,t) - V_2(n-s,t)] \\
& \qquad - [V_1(n,t) - V_1(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{V_1(n,t)-V_1(n-s,t)-p(n,t)} \{[V_1(n,t) - V_1(n-s,t)] \\
& \qquad - [V_2(n,t) - V_2(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{\infty} \{[V_1(n,t) - V_1(n-s,t)] \\
& \qquad - [V_2(n,t) - V_2(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_1(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{V_2(n,t)-V_2(n-s,t)-p(n,t)} \{[V_2(n,t) - V_2(n-s,t)] \\
& \qquad - [V_1(n,t) - V_1(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
= \; & \int_{\mathcal{S}_1(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{\infty} \{[V_2(n,t) - V_1(n,t)] \\
& \qquad + [V_1(n-s,t) - V_2(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{\infty} \{[V_1(n,t) - V_2(n,t)] \\
& \qquad + [V_2(n-s,t) - V_1(n-s,t)]\} \; F_{R|S}(dr|s) F_S(ds) \\
\leq \; & \int_{\mathcal{S}_1(n,t)} \int_{V_1(n,t)-V_1(n-s,t)-p(n,t)}^{\infty} 2d_t(V_1, V_2) \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n,t)} \int_{V_2(n,t)-V_2(n-s,t)-p(n,t)}^{\infty} 2d_t(V_1, V_2) \; F_{R|S}(dr|s) F_S(ds) \\
\leq \; & 2d_t(V_1, V_2) \left[\int_{\mathcal{S}_1(n,t)} \int_{-\infty}^{\infty} F_{R|S}(dr|s) F_S(ds) \; + \; \int_{\mathcal{S}_2(n,t)} \int_{-\infty}^{\infty} F_{R|S}(dr|s) F_S(ds) \right] \\
= \; & 2d_t(V_1, V_2) \\
\leq \; & 2d_{\mathcal{V}}(V_1, V_2) \qquad \square
\end{aligned}
$$

Appendix E: Proof of Lemma 15

Lemma 15 *The function f satisfies the following Lipschitz condition with respect to $V_1, V_2 \in \mathcal{V}$:*

$$|f(n,t,V_1) - f(n,t,V_2)| \quad \leq \quad \left(2\overline{\lambda} + \alpha\right) d_t(V_1, V_2) \quad \leq \quad \left(2\overline{\lambda} + \alpha\right) d_{\mathcal{V}}(V_1, V_2)$$

Proof:

Case 1: $-\lambda(n,t)\Upsilon(n,t,V_1) + \alpha V_1(n,t) + \lambda(n,t)p(n,t) + c(n,t) \leq \alpha V_1(n,t) - \alpha v(n)$ and $-\lambda(n,t)\Upsilon(n,t,V_2) + \alpha V_2(n,t) + \lambda(n,t)p(n,t) + c(n,t) \leq \alpha V_2(n,t) - \alpha v(n)$:

Then

$$\begin{aligned}
& |f(n,t,V_1) - f(n,t,V_2)| \\
= \; & |\lambda(n,t)\Upsilon(n,t,V_2) \; - \; \lambda(n,t)\Upsilon(n,t,V_1) \; + \; \alpha V_1(n,t) \; - \; \alpha V_2(n,t)| \\
\leq \; & \lambda(n,t)\,|\Upsilon(n,t,V_2) \; - \; \Upsilon(n,t,V_1)| \; + \; \alpha\,|V_1(n,t) \; - \; V_2(n,t)| \\
\leq \; & \left(2\overline{\lambda} + \alpha\right) d_t(V_1, V_2) \\
\leq \; & \left(2\overline{\lambda} + \alpha\right) d_{\mathcal{V}}(V_1, V_2)
\end{aligned}$$

Case 2: $-\lambda(n,t)\Upsilon(n,t,V_1) + \alpha V_1(n,t) + \lambda(n,t)p(n,t) + c(n,t) > \alpha V_1(n,t) - \alpha v(n)$ and $-\lambda(n,t)\Upsilon(n,t,V_2) + \alpha V_2(n,t) + \lambda(n,t)p(n,t) + c(n,t) > \alpha V_2(n,t) - \alpha v(n)$:

Then

$$\begin{aligned}
|f(n,t,V_1) - f(n,t,V_2)| \quad &= \quad |\alpha V_1(n,t) \; - \; \alpha V_2(n,t)| \\
&\leq \quad \alpha d_t(V_1, V_2) \\
&\leq \quad \alpha d_{\mathcal{V}}(V_1, V_2)
\end{aligned}$$

Case 3: $-\lambda(n,t)\Upsilon(n,t,V_1) + \alpha V_1(n,t) + \lambda(n,t)p(n,t) + c(n,t) \leq \alpha V_1(n,t) - \alpha v(n)$ and $-\lambda(n,t)\Upsilon(n,t,V_2) + \alpha V_2(n,t) + \lambda(n,t)p(n,t) + c(n,t) > \alpha V_2(n,t) - \alpha v(n)$:

Case 3.1: $-\lambda(n,t)\Upsilon(n,t,V_1) + \alpha V_1(n,t) + \lambda(n,t)p(n,t) + c(n,t) \leq \alpha V_2(n,t) - \alpha v(n)$:

Then

$$\begin{aligned}
|f(n,t,V_1) - f(n,t,V_2)| \quad = \quad & [\alpha V_2(n,t) \; - \; \alpha v(n)] \\
& - [-\lambda(n,t)\Upsilon(n,t,V_1) \; + \; \alpha V_1(n,t) \; + \; \lambda(n,t)p(n,t) \; + \; c(n,t)]
\end{aligned}$$

$$\begin{aligned}
&< \left[-\lambda(n,t)\Upsilon(n,t,V_2) \;+\; \alpha V_2(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t)\right] \\
&\quad -\left[-\lambda(n,t)\Upsilon(n,t,V_1) \;+\; \alpha V_1(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t)\right] \\
&\leq \lambda(n,t)\left|\Upsilon(n,t,V_1) \;-\; \Upsilon(n,t,V_2)\right| \;+\; \alpha\left|V_2(n,t) \;-\; V_1(n,t)\right| \\
&\leq \left(2\overline{\lambda}+\alpha\right) d_t(V_1,V_2) \\
&\leq \left(2\overline{\lambda}+\alpha\right) d_{\mathcal{V}}(V_1,V_2)
\end{aligned}$$

Case 3.2: $-\lambda(n,t)\Upsilon(n,t,V_1)+\alpha V_1(n,t)+\lambda(n,t)p(n,t)+c(n,t) > \alpha V_2(n,t)-\alpha v(n)$:
Then

$$\begin{aligned}
|f(n,t,V_1)-f(n,t,V_2)| &= \left[-\lambda(n,t)\Upsilon(n,t,V_1) \;+\; \alpha V_1(n,t) \;+\; \lambda(n,t)p(n,t) \;+\; c(n,t)\right] \\
&\quad - \left[\alpha V_2(n,t) \;-\; \alpha v(n)\right] \\
&\leq \left[\alpha V_1(n,t) \;-\; \alpha v(n)\right] - \left[\alpha V_2(n,t) \;-\; \alpha v(n)\right] \\
&\leq \alpha d_t(V_1,V_2) \\
&\leq \alpha d_{\mathcal{V}}(V_1,V_2)
\end{aligned}$$

Case 4: $-\lambda(n,t)\Upsilon(n,t,V_1)+\alpha V_1(n,t)+\lambda(n,t)p(n,t)+c(n,t) > \alpha V_1(n,t)-\alpha v(n)$ and $-\lambda(n,t)\Upsilon(n,t,V_2)+\alpha V_2(n,t)+\lambda(n,t)p(n,t)+c(n,t) \leq \alpha V_2(n,t)-\alpha v(n)$:
Case 4.1: $\alpha V_1(n,t)-\alpha v(n) > -\lambda(n,t)\Upsilon(n,t,V_2)+\alpha V_2(n,t)+\lambda(n,t)p(n,t)+c(n,t)$
Then it follows in a similar way to Case 3.1 that

$$\begin{aligned}
|f(n,t,V_1)-f(n,t,V_2)| &\leq \left(2\overline{\lambda}+\alpha\right) d_t(V_1,V_2) \\
&\leq \left(2\overline{\lambda}+\alpha\right) d_{\mathcal{V}}(V_1,V_2)
\end{aligned}$$

Case 4.2: $\alpha V_1(n,t)-\alpha v(n) \leq -\lambda(n,t)\Upsilon(n,t,V_2)+\alpha V_2(n,t)+\lambda(n,t)p(n,t)+c(n,t)$
Then it follows in a similar way to Case 3.2 that

$$\begin{aligned}
|f(n,t,V_1)-f(n,t,V_2)| &\leq \alpha d_t(V_1,V_2) \\
&\leq \alpha d_{\mathcal{V}}(V_1,V_2)
\end{aligned}$$

□

Appendix F: Proof of Lemma 17

Lemma 17 *If* $\{p(n,\cdot)\}_{n\in[0,N_0]}$ *and* $\{V(n,\cdot)\}_{n\in[0,N_0]}$ *are equicontinuous families of functions on* $[0,T]$*, then* $\{\Upsilon(n,\cdot,V)\}_{n\in[0,N_0]}$ *is an equicontinuous family of functions on* $[0,T]$*.*

Proof: If $\{V(n,\cdot)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$, then for any $\varepsilon > 0$ there exists a $\delta_1 > 0$ such that $|V(n,t_1) - V(n,t_2)| < \varepsilon$ for all $n \in [0,N_0]$ and for all $t_1, t_2 \in [0,T]$ such that $|t_2 - t_1| < \delta_1$. If $\{p(n,\cdot)\}_{n\in[0,N_0]}$ is an equicontinuous family of functions on $[0,T]$, then for any $\varepsilon > 0$ there exists a $\delta_2 > 0$ such that $|p(n,t_1) - p(n,t_2)| < \varepsilon$ for all $n \in [0,N_0]$ and for all $t_1, t_2 \in [0,T]$ such that $|t_2 - t_1| < \delta_2$. Let $\delta = \min\{\delta_1, \delta_2\}$. Then for any $n \in [0,N_0]$ and any $t_1, t_2 \in [0,T]$ such that $|t_2 - t_1| < \delta$, let $\mathcal{S}_1(n) \equiv \{s \in [0,n] : V(n,t_1) - V(n-s,t_1) - p(n,t_1) \le V(n,t_2) - V(n-s,t_2) - p(n,t_2)\}$, and let $\mathcal{S}_2(n) \equiv \{s \in [0,n] : V(n,t_2) - V(n-s,t_2) - p(n,t_2) < V(n,t_1) - V(n-s,t_1) - p(n,t_1)\}$. Then it follows that

$$
\begin{aligned}
& |\Upsilon(n,t_2,V) - \Upsilon(n,t_1,V)| \\
= & \left| \int_0^n \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{\infty} \{r - [V(n,t_2) \right. \\
& \qquad -V(n-s,t_2) - p(n,t_2)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& - \int_0^n \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{\infty} \{r - [V(n,t_1) \\
& \qquad \left. -V(n-s,t_1) - p(n,t_1)]\} \; F_{R|S}(dr|s) F_S(ds) \right| \\
= & \left| \int_{\mathcal{S}_1(n)} \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{\infty} \{r - [V(n,t_2) \right. \\
& \qquad -V(n-s,t_2) - p(n,t_2)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n)} \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{V(n,t_1)-V(n-s,t_1)-p(n,t_1)} \{r - [V(n,t_2) \\
& \qquad -V(n-s,t_2) - p(n,t_2)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n)} \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{\infty} \{r - [V(n,t_2) \\
& \qquad -V(n-s,t_2) - p(n,t_2)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& - \int_{\mathcal{S}_1(n)} \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{V(n,t_2)-V(n-s,t_2)-p(n,t_2)} \{r - [V(n,t_1) \\
& \qquad -V(n-s,t_1) - p(n,t_1)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& - \int_{\mathcal{S}_1(n)} \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{\infty} \{r - [V(n,t_1) \\
& \qquad -V(n-s,t_1) - p(n,t_1)]\} \; F_{R|S}(dr|s) F_S(ds)
\end{aligned}
$$

$$
\begin{aligned}
& - \int_{\mathcal{S}_2(n)} \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{\infty} \{r - [V(n,t_1) \\
& \qquad\qquad -V(n-s,t_1) - p(n,t_1)]\} \; F_{R|S}(dr|s) F_S(ds) \Big| \\
\leq \; & \int_{\mathcal{S}_1(n)} \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{\infty} \{[V(n,t_2) - V(n-s,t_2) - p(n,t_2)] \\
& \qquad\qquad - [V(n,t_1) - V(n-s,t_1) - p(n,t_1)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n)} \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{V(n,t_1)-V(n-s,t_1)-p(n,t_1)} \{[V(n,t_1) - V(n-s,t_1) - p(n,t_1)] \\
& \qquad\qquad - [V(n,t_2) - V(n-s,t_2) - p(n,t_2)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n)} \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{\infty} \{[V(n,t_1) - V(n-s,t_1) - p(n,t_1)] \\
& \qquad\qquad - [V(n,t_2) - V(n-s,t_2) - p(n,t_2)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_1(n)} \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{V(n,t_2)-V(n-s,t_2)-p(n,t_2)} \{[V(n,t_2) - V(n-s,t_2) - p(n,t_2)] \\
& \qquad\qquad - [V(n,t_1) - V(n-s,t_1) - p(n,t_1)]\} \; F_{R|S}(dr|s) F_S(ds) \\
= \; & \int_{\mathcal{S}_1(n)} \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{\infty} \{[V(n,t_2) - V(n,t_1)] \\
& \qquad + [V(n-s,t_1) - V(n-s,t_2)] + [p(n,t_1) - p(n,t_2)]\} \; F_{R|S}(dr|s) F_S(ds) \\
& + \int_{\mathcal{S}_2(n)} \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{\infty} \{[V(n,t_1) - V(n,t_2)] \\
& \qquad + [V(n-s,t_2) - V(n-s,t_1)] + [p(n,t_2) - p(n,t_1)]\} \; F_{R|S}(dr|s) F_S(ds) \\
< \; & 3\varepsilon \int_{\mathcal{S}_1(n)} \int_{V(n,t_1)-V(n-s,t_1)-p(n,t_1)}^{\infty} F_{R|S}(dr|s) F_S(ds) \\
& + 3\varepsilon \int_{\mathcal{S}_2(n)} \int_{V(n,t_2)-V(n-s,t_2)-p(n,t_2)}^{\infty} F_{R|S}(dr|s) F_S(ds) \\
\leq \; & 3\varepsilon \qquad \square
\end{aligned}
$$

Appendix G: Proof of Lemma 20

Lemma 20 *If F_S is continuous on $[0, N_0]$ and $p \in \mathcal{C}$, then for any $V \in \mathcal{C}$, $\Upsilon(\cdot, \cdot, V) \in \mathcal{C}$.*

Proof: Because $V \in \mathcal{C}$, V is uniformly continuous on the compact set $[0, N_0] \times [0, T]$. Hence, for any $\varepsilon > 0$ there exists a $\delta_1 > 0$ such that $|V(n_1, t_1) - V(n_2, t_2)| < \varepsilon$ for all $(n_1, t_1), (n_2, t_2) \in [0, N_0] \times [0, T]$ with $d_{\Re^2}((n_1, t_1), (n_2, t_2)) < \delta_1$.

Also, V is continuous on the compact set $[0, N_0] \times [0, T]$ implies that V is bounded on $[0, N_0] \times [0, T]$. Hence, $\int_0^{N_0} \int_{-\infty}^{\infty} |V(n, t)| \, F_{R|S}(dr|s) F_S(ds) < \infty$, and therefore, for any $\varepsilon > 0$ there exists an $\eta_1 > 0$ such that $\int_{n_1}^{n_2} \int_{-\infty}^{\infty} |V(n, t)| \, F_{R|S}(dr|s) F_S(ds) < \varepsilon$ for any $0 \leq n_1 \leq n_2 \leq N_0$ with $F_S(n_2) - F_S(n_1) \leq \eta_1$, for all $(n, t) \in [0, N_0] \times [0, T]$. In the same way, $\int_{n_1}^{n_2} \int_{-\infty}^{\infty} |V(n - s, t)| \, F_{R|S}(dr|s) F_S(ds) < \varepsilon$ for any $0 \leq n_1 \leq n_2 \leq N_0$ with $F_S(n_2) - F_S(n_1) \leq \eta_1$, for all $(n, t) \in [0, N_0] \times [0, T]$.

Similarly, it follows from $p \in \mathcal{C}$ that for any $\varepsilon > 0$ there exists a $\delta_2 > 0$ such that $|p(n_1, t_1) - p(n_2, t_2)| < \varepsilon$ for all $(n_1, t_1), (n_2, t_2) \in [0, N_0] \times [0, T]$ with $d_{\Re^2}((n_1, t_1), (n_2, t_2)) < \delta_2$. It also follows that for any $\varepsilon > 0$ there exists an $\eta_2 > 0$ such that

$$\int_{n_1}^{n_2} \int_{-\infty}^{\infty} |p(n, t)| \, F_{R|S}(dr|s) F_S(ds) \quad < \quad \varepsilon$$

for any $0 \leq n_1 \leq n_2 \leq N_0$ with $F_S(n_2) - F_S(n_1) \leq \eta_2$, for all $(n, t) \in [0, N_0] \times [0, T]$.

From the assumption that $\int_0^{N_0} \int_{-\infty}^{\infty} |r| \, F_{R|S}(dr|s) F_S(ds) < \infty$ it follows that for any $\varepsilon > 0$ there exists an $\eta_3 > 0$ such that $\int_{n_1}^{n_2} \int_{-\infty}^{\infty} |r| \, F_{R|S}(dr|s) F_S(ds) < \varepsilon$ for any $0 \leq n_1 \leq n_2 \leq N_0$ with $F_S(n_2) - F_S(n_1) \leq \eta_3$.

Let $\eta = \min\{\eta_1, \eta_2, \eta_3\}$. F_S uniformly continuous on $[0, N_0]$ implies that for any $\eta > 0$ there exists a $\delta_3 > 0$ such that $|F_S(n_2) - F_S(n_1)| < \eta$ for all $n_1, n_2 \in [0, N_0]$ with $|n_2 - n_1| < \delta_3$.

Let $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. For any $(n_1, t_1), (n_2, t_2) \in [0, N_0] \times [0, T]$ such that $d_{\Re^2}((n_1, t_1), (n_2, t_2)) < \delta$, let $n_1 \leq n_2$, let

$$\mathcal{S}_1(n_1) \equiv \{s \in [0, n_1] : V(n_1, t_1) - V(n_1 - s, t_1) - p(n_1, t_1) \leq V(n_2, t_2) - V(n_2 - s, t_2) - p(n_2, t_2)\}$$

and let

$$\mathcal{S}_2(n_1) \equiv \{s \in [0, n_1] : V(n_2, t_2) - V(n_2 - s, t_2) - p(n_2, t_2) < V(n_1, t_1) - V(n_1 - s, t_1) - p(n_1, t_1)\}$$

Then it follows that

$$
\begin{aligned}
& |\Upsilon(n_2,t_2,V) - \Upsilon(n_1,t_1,V)| \\
= & \Bigg| \int_0^{n_2} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{\infty} \{r - [V(n_2,t_2) \\
& \qquad -V(n_2-s,t_2) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& - \int_0^{n_1} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{\infty} \{r - [V(n_1,t_1) \\
& \qquad -V(n_1-s,t_1) - p(n_1,t_1)]\} \; F_{R|S}(dr|s)F_S(ds) \Bigg| \\
= & \Bigg| \int_{\mathcal{S}_1(n_1)} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{\infty} \{r - [V(n_2,t_2) \\
& \qquad -V(n_2-s,t_2) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{\mathcal{S}_2(n_1)} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)} \{r - [V(n_2,t_2) \\
& \qquad -V(n_2-s,t_2) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{\mathcal{S}_2(n_1)} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{\infty} \{r - [V(n_2,t_2) \\
& \qquad -V(n_2-s,t_2) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{n_1}^{n_2} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{\infty} \{r - [V(n_2,t_2) \\
& \qquad -V(n_2-s,t_2) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& - \int_{\mathcal{S}_1(n_1)} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)} \{r - [V(n_1,t_1) \\
& \qquad -V(n_1-s,t_1) - p(n_1,t_1)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& - \int_{\mathcal{S}_1(n_1)} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{\infty} \{r - [V(n_1,t_1) \\
& \qquad -V(n_1-s,t_1) - p(n_1,t_1)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& - \int_{\mathcal{S}_2(n_1)} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{\infty} \{r - [V(n_1,t_1) \\
& \qquad -V(n_1-s,t_1) - p(n_1,t_1)]\} \; F_{R|S}(dr|s)F_S(ds) \Bigg| \\
\le & \int_{n_1}^{n_2} \int_{-\infty}^{\infty} |r| \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{n_1}^{n_2} \int_{-\infty}^{\infty} |V(n_2,t_2)| \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{n_1}^{n_2} \int_{-\infty}^{\infty} |V(n_2-s,t_2)| \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{n_1}^{n_2} \int_{-\infty}^{\infty} |p(n_2,t_2)| \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{\mathcal{S}_1(n_1)} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{\infty} \{[V(n_2,t_2) - V(n_2-s,t_2) - p(n_2,t_2)]
\end{aligned}
$$

$$
\begin{aligned}
& \qquad\qquad - [V(n_1,t_1) - V(n_1 - s, t_1) - p(n_1,t_1)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{\mathcal{S}_2(n_1)} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)} \{[V(n_1,t_1) - V(n_1 - s, t_1) - p(n_1,t_1)] \\
& \qquad\qquad - [V(n_2,t_2) - V(n_2 - s, t_2) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{\mathcal{S}_2(n_1)} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{\infty} \{[V(n_1,t_1) - V(n_1 - s, t_1) - p(n_1,t_1)] \\
& \qquad\qquad - [V(n_2,t_2) - V(n_2 - s, t_2) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{\mathcal{S}_1(n_1)} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)} \{[V(n_2,t_2) - V(n_2 - s, t_2) - p(n_2,t_2)] \\
& \qquad\qquad - [V(n_1,t_1) - V(n_1 - s, t_1) - p(n_1,t_1)]\} \; F_{R|S}(dr|s)F_S(ds) \\
< \; & 4\varepsilon \; + \; \int_{\mathcal{S}_1(n_1)} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{\infty} \{[V(n_2,t_2) - V(n_1,t_1)] \\
& \qquad + [V(n_1 - s, t_1) - V(n_2 - s, t_2)] + [p(n_1,t_1) - p(n_2,t_2)]\} \; F_{R|S}(dr|s)F_S(ds) \\
& + \int_{\mathcal{S}_2(n_1)} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{\infty} \{[V(n_1,t_1) - V(n_2,t_2)] \\
& \qquad + [V(n_2 - s, t_2) - V(n_1 - s, t_1)] + [p(n_2,t_2) - p(n_1,t_1)]\} \; F_{R|S}(dr|s)F_S(ds) \\
< \; & 4\varepsilon \; + \; 3\varepsilon \int_{\mathcal{S}_1(n_1)} \int_{V(n_1,t_1)-V(n_1-s,t_1)-p(n_1,t_1)}^{\infty} F_{R|S}(dr|s)F_S(ds) \\
& + 3\varepsilon \int_{\mathcal{S}_2(n_1)} \int_{V(n_2,t_2)-V(n_2-s,t_2)-p(n_2,t_2)}^{\infty} F_{R|S}(dr|s)F_S(ds) \\
\leq \; & 7\varepsilon \qquad \square
\end{aligned}
$$

Appendix H: Proof of Proposition 2

It is shown that if λ, p, and c are independent of t, then $V^*(n,t)$ is nondecreasing in T and nonincreasing in t. All the results and proofs of this section hold for both the DSKP and the MDP; therefore the subscripts are dropped. Also for the purposes of this section, let $\Pi^{MD}(n,t,T)$ denote the class of policies $\pi : [0,n] \times (t,T] \mapsto \mathcal{I}_{RS} \times \{0,1\}$. Let $V^\pi(n,t,T)$ denote the expected value under policy $\pi \in \Pi^{MD}(n,t,T)$ with deadline T, as defined in Chapter 4.

For any $n \in [0,N_0]$, $T > 0$, and $t \in [0,T]$, let the function $\vartheta : \Pi^{MD}(n,t,T) \mapsto \Pi^{MD}(n,0,T-t)$ be defined by

$$\begin{aligned} D^{\vartheta(\pi)}(n',\tau,r,s) &\equiv D^\pi(n',t+\tau,r,s) \\ I^{\vartheta(\pi)}(n',\tau) &\equiv I^\pi(n',t+\tau) \end{aligned}$$

for all $\pi \in \Pi^{MD}(n,t,T)$, $n' \in [0,n]$, $\tau \in (0,T-t]$, $r \in \Re$, and $s \in [0,\infty)$. Note that if $I^\pi(n',\cdot) \in \mathcal{I}_s$ for all $n' \in [0,n]$, then $I^{\vartheta(\pi)}(n',\cdot) \in \mathcal{I}_s$ for all $n' \in [0,n]$. Thus, if $\pi \in \Pi^{MD}_{DSKP}(n,t,T)$, then $\vartheta(\pi) \in \Pi^{MD}_{DSKP}(n,0,T-t)$.

Lemma 62 *For any $n \in [0,N_0]$, $T > 0$, $t \in [0,T]$, and $\pi \in \Pi^{MD}(n,t,T)$, $V^{\vartheta(\pi)}(n,0,T-t) = V^\pi(n,t,T)$.*

Proof: The result follows from the memoryless property of the Poisson process, and the assumption that $\{(R_i,S_i)\}$ is an i.i.d. sequence independent of $\{A_i\}$. □

Lemma 63 *For any $n \in [0,N_0]$, $T > 0$, and $t \in [0,T]$, $V^*(n,t,T) = V^*(n,0,T-t)$.*

Proof: ϑ is a bijection from $\Pi^{MD}(n,t,T)$ onto $\Pi^{MD}(n,0,T-t)$. From Lemma 62, for any $\pi \in \Pi^{MD}(n,t,T)$, $V^\pi(n,t,T) = V^{\vartheta(\pi)}(n,0,T-t)$. Hence,

$$\begin{aligned} V^*(n,t,T) &\equiv \sup_{\pi \in \Pi^{MD}(n,t,T)} V^\pi(n,t,T) \\ &= \sup_{\pi \in \Pi^{MD}(n,t,T)} V^{\vartheta(\pi)}(n,0,T-t) \\ &= \sup_{\pi \in \Pi^{MD}(n,0,T-t)} V^\pi(n,0,T-t) \\ &\equiv V^*(n,0,T-t) \end{aligned}$$

□

For any $n \in [0, N_0]$, $0 < T_1 \leq T_2 \leq \infty$, and $t \in [0, T_1]$, consider the function $\hat{\vartheta}$: $\Pi^{MD}(n, t, T_1) \mapsto \Pi^{MD}(n, t, T_2)$ defined by

$$D^{\hat{\vartheta}(\pi)}(n', \tau, r, s) \equiv \begin{cases} D^{\pi}(n', \tau, r, s) & \text{if } \tau \in (t, T_1] \\ 0 & \text{if } \tau \in (T_1, T_2] \end{cases}$$

$$I^{\hat{\vartheta}(\pi)}(n', \tau) \equiv \begin{cases} I^{\pi}(n', \tau) & \text{if } \tau \in (t, T_1] \\ 0 & \text{if } \tau \in (T_1, T_2] \end{cases}$$

for all $\pi \in \Pi^{MD}(n, t, T_1)$, $n' \in [0, n]$, $r \in \Re$, and $s \in [0, \infty)$. If $I^{\pi}(n', \cdot) \in \mathcal{I}_s$, then $I^{\hat{\vartheta}(\pi)}(n', \cdot) \in \mathcal{I}_s$. Hence, if $\pi \in \Pi^{MD}_{DSKP}(n, t, T_1)$, then $\hat{\vartheta}(\pi) \in \Pi^{MD}_{DSKP}(n, t, T_2)$.

Lemma 64 *For any $n \in [0, N_0]$, $0 < T_1 \leq T_2 \leq \infty$, $t \in [0, T_1]$, and $\pi \in \Pi^{MD}(n, t, T_1)$, $V^{\hat{\vartheta}(\pi)}(n, t, T_2) = V^{\pi}(n, t, T_1)$.*

Proof: For every sample point $\omega \in \Omega$, the same rewards, penalties and holding cost are accumulated by policies π and $\hat{\vartheta}(\pi)$. That the (discounted) terminal values collected by π and $\hat{\vartheta}(\pi)$ are equal follows in a similar way to the proof of Lemma 1. □

Proposition 8 *If λ, p, and c are independent of t, then for any $n \in [0, N_0]$, and $t \geq 0$, $V^*(n, t, T)$ is nondecreasing in T on $[t, \infty]$.*

Proof: Let $t \leq T_1 \leq T_2$. Then $\hat{\vartheta}(\Pi^{MD}(n, t, T_1)) \subseteq \Pi^{MD}(n, t, T_2)$. From Lemma 64, for any $\pi \in \Pi^{MD}(n, t, T_1)$, $V^{\pi}(n, t, T_1) = V^{\hat{\vartheta}(\pi)}(n, t, T_2)$. Hence,

$$\begin{aligned} V^*(n, t, T_1) &\equiv \sup_{\pi \in \Pi^{MD}(n,t,T_1)} V^{\pi}(n, t, T_1) \\ &= \sup_{\pi \in \Pi^{MD}(n,t,T_1)} V^{\hat{\vartheta}(\pi)}(n, t, T_2) \\ &= \sup_{\pi \in \hat{\vartheta}(\Pi^{MD}(n,t,T_1))} V^{\pi}(n, t, T_2) \\ &\leq \sup_{\pi \in \Pi^{MD}(n,t,T_2)} V^{\pi}(n, t, T_2) \\ &\equiv V^*(n, t, T_2) \end{aligned}$$

□

Proposition 2 *If λ, p, and c are independent of t, then for any $n \in [0, N_0]$, and $T \in (0, \infty]$, $V^*(n, t, T)$ is nonincreasing in t on $[0, T]$.*

Proof: Let $0 \leq t_1 \leq t_2 \leq T$. Then, from Lemma 63 and Proposition 8

$$\begin{aligned} V^*(n, t_2, T) &= V^*(n, 0, T - t_2) \\ &\leq V^*(n, 0, T - t_1) \\ &= V^*(n, t_1, T) \end{aligned}$$

□

Appendix I: Proof of Proposition 3

It is shown that if $c(n)$ is nonincreasing in n and $v(n)$ is nondecreasing in n, then $V^*(n,t)$ is nondecreasing in n. For any $0 \leq n_1 \leq n_2 \leq N_0$, $T \in (0,\infty]$, and $t \in [0,T]$, consider the function $\tilde{\vartheta} : \Pi^{MD}(n_1,t,T) \mapsto \Pi^{MD}(n_2,t,T)$ defined by

$$D^{\tilde{\vartheta}(\pi)}(n',\tau,r,s) \equiv \begin{cases} D^{\pi}(n'+n_1-n_2,\tau,r,s) & \text{if } n' \in [n_2-n_1,n_2] \\ 0 & \text{if } n' \in [0,n_2-n_1) \end{cases}$$

$$I^{\tilde{\vartheta}(\pi)}(n',\tau) \equiv \begin{cases} I^{\pi}(n'+n_1-n_2,\tau) & \text{if } n' \in [n_2-n_1,n_2] \\ 0 & \text{if } n' \in [0,n_2-n_1) \end{cases}$$

for all $\pi \in \Pi^{MD}(n_1,t,T)$, $\tau \in (t,T]$, $r \in \Re$, and $s \in [0,\infty)$. If $I^{\pi}(n',\cdot) \in \mathcal{I}_s$, then $I^{\tilde{\vartheta}(\pi)}(n'+n_2-n_1,\cdot) \in \mathcal{I}_s$. Hence, if $\pi \in \Pi^{MD}_{DSKP}(n_1,t,T)$, then $\tilde{\vartheta}(\pi) \in \Pi^{MD}_{DSKP}(n_2,t,T)$.

Lemma 65 *If $c(n)$ is nonincreasing in n and $v(n)$ is nondecreasing in n, then for any $0 \leq n_1 \leq n_2 \leq N_0$, $T \in (0,\infty]$, $t \in [0,T]$, and $\pi \in \Pi^{MD}(n_1,t,T)$, $V^{\pi}(n_1,t,T) \leq V^{\tilde{\vartheta}(\pi)}(n_2,t,T)$.*

Proof: For every sample point $\omega \in \Omega$, the same rewards and penalties are accumulated by policies π and $\tilde{\vartheta}(\pi)$. The holding cost $c(N^{\pi}(\tau)) \geq c(N^{\tilde{\vartheta}(\pi)}(\tau))$ because $N^{\pi}(\tau) \leq N^{\tilde{\vartheta}(\pi)}(\tau)$ for all $\tau \in (t,T]$. The terminal values $v(N^{\pi}(\mathcal{T}^{\pi+})) \leq v(N^{\tilde{\vartheta}(\pi)}(\mathcal{T}^{\tilde{\vartheta}(\pi)+}))$ because $N^{\pi}(\mathcal{T}^{\pi+}) \leq N^{\tilde{\vartheta}(\pi)}(\mathcal{T}^{\tilde{\vartheta}(\pi)+})$. □

Proposition 3 *If λ is constant, $p(n)$ and $c(n)$ are nonincreasing, and $v(n)$ is nondecreasing, then for any $T \in (0,\infty]$, and $t \in [0,T]$, $V^*(n,t,T)$ is nondecreasing in n on $[0,N_0]$.*

Proof: Let $0 \leq n_1 \leq n_2 \leq N_0$. Then $\tilde{\vartheta}(\Pi^{MD}(n_1,t,T)) \subseteq \Pi^{MD}(n_2,t,T)$. From Lemma 65, for any $\pi \in \Pi^{MD}(n_1,t,T)$, $V^{\pi}(n_1,t,T) \leq V^{\tilde{\vartheta}(\pi)}(n_2,t,T)$. Hence,

$$\begin{aligned} V^*(n_1,t,T) &\equiv \sup_{\pi \in \Pi^{MD}(n_1,t,T)} V^{\pi}(n_1,t,T) \\ &\leq \sup_{\pi \in \Pi^{MD}(n_1,t,T)} V^{\tilde{\vartheta}(\pi)}(n_2,t,T) \\ &= \sup_{\pi \in \tilde{\vartheta}(\Pi^{MD}(n_1,t,T))} V^{\pi}(n_2,t,T) \\ &\leq \sup_{\pi \in \Pi^{MD}(n_2,t,T)} V^{\pi}(n_2,t,T) \\ &\equiv V^*(n_2,t,T) \end{aligned}$$

□

Appendix J: Proof of Lemma 25

Lemma 66 *If f_1 and f_2 are absolutely continuous functions, then $\max\{f_1, f_2\}$ is absolutely continuous.*

Proof: Let $\varepsilon > 0$. Then there exists a $\delta_1 > 0$ such that

$$\sum_{i=1}^{n} |f_1(y_i) - f_1(x_i)| < \varepsilon/2$$

for any finite collection $\{(x_i, y_i)\}_{i=1}^{n}$ of nonoverlapping intervals with

$$\sum_{i=1}^{n} |y_i - x_i| < \delta_1$$

and there exists a $\delta_2 > 0$ such that

$$\sum_{i=1}^{n} |f_2(y_i) - f_2(x_i)| < \varepsilon/2$$

for any finite collection $\{(x_i, y_i)\}_{i=1}^{n}$ of nonoverlapping intervals with

$$\sum_{i=1}^{n} |y_i - x_i| < \delta_2$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Then for any finite collection $\{(x_i, y_i)\}_{i=1}^{n}$ of nonoverlapping intervals with

$$\sum_{i=1}^{n} |y_i - x_i| < \delta$$

we have

$$\begin{aligned} & \sum_{i=1}^{n} |\max\{f_1(y_i), f_2(y_i)\} - \max\{f_1(x_i), f_2(x_i)\}| \\ \leq \; & \sum_{i=1}^{n} |f_1(y_i) - f_1(x_i)| \; + \; \sum_{i=1}^{n} |f_2(y_i) - f_2(x_i)| \quad < \quad \varepsilon \end{aligned}$$

Therefore, $\max\{f_1, f_2\}$ is absolutely continuous. $\square$

Lemma 24 *For all $n > 0$, all $t \in [0, T]$, and all $s \in \{1, \ldots, n\}$,*

$$V^{\psi(n)}(n - s, t) = \max\left\{v(n - s), V^{\psi(n-s)}(n - s, t)\right\}$$

Proof: For $n = 1$:

If $\alpha > 0$:

$$\begin{aligned} V^{\psi(0)}(0,t) &= e^{-\alpha(T-t)}v(0) - \frac{\lambda(0)p(0)+c(0)}{\alpha}\left(1-e^{-\alpha(T-t)}\right) &>& \ v(0) \\ &\Leftrightarrow \ -\left[\lambda(0)p(0)+c(0)\right] &>& \ \alpha v(0) \end{aligned}$$

If $\alpha = 0$:

$$\begin{aligned} V^{\psi(0)}(0,t) &= -\left[\lambda(0)p(0)+c(0)\right](T-t)+v(0) &>& \ v(0) \\ &\Leftrightarrow \ -\left[\lambda(0)p(0)+c(0)\right] &>& \ 0 \end{aligned}$$

for all $t < T$. Thus, if $\lambda(0)p(0)+c(0)+\alpha v(0) < 0$, then $V^{\psi(0)}(0,t) > v(0)$, and $I^{\psi(1)}(0,t) = 1$ for all $t \in (0,T)$. Then $V^{\psi(1)}(0,t) = V^{\psi(0)}(0,t) = \max\left\{v(0), V^{\psi(0)}(0,t)\right\}$ for all $t \in [0,T]$. If $\lambda(0)p(0)+c(0)+\alpha v(0) \geq 0$, then $V^{\psi(0)}(0,t) \leq v(0)$, and $I^{\psi(1)}(0,t) = 0$ for all $t \in (0,T)$. Then $V^{\psi(1)}(0,t) = v(0) = \max\left\{v(0), V^{\psi(0)}(0,t)\right\}$ for all $t \in [0,T]$.

For $n > 1$:

From the boundary conditions $V^{\psi(n)}(n-s,T) = v(n-s) = \max\left\{v(n-s), V^{\psi(n-s)}(n-s,T)\right\}$. It is shown by contradiction that $V^{\psi(n)}(n-s,t) = \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$ for all $t \in [0,T]$.

Case 1: Suppose there exists an $s \in \{1,\ldots,n\}$ and a $t_1 \in [0,T)$ such that $V^{\psi(n)}(n-s,t_1) > \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t_1)\right\}$.

Then, from the continuity of $V^{\psi(n)}(n-s,t)$ and $\max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$ in t, there exists a neighborhood $(t_0,t_2) \subseteq (0,T)$ of t_1 such that

$$V^{\psi(n)}(n-s,t) \quad > \quad \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$$

for all $t \in (t_0,t_2)$. Hence, t_1 can be taken as any point in (t_0,t_2), and at least one of the following three cases must hold.

Case 1.1: $V^{\psi(n)}(n-s,t_1) > v(n-s) > V^{\psi(n-s)}(n-s,t_1)$ for some $t_1 \in (t_0,t_2)$:

Then, from the continuity of $V^{\psi(n)}(n-s,t)$, $v(n-s)$ and $V^{\psi(n-s)}(n-s,t)$ in t, there exists a neighborhood $(t_3,t_4) \subseteq (0,T)$ of t_1 such that $V^{\psi(n)}(n-s,t) > v(n-s) > V^{\psi(n-s)}(n-s,t)$ for all $t \in (t_3,t_4)$. Then $I^{\psi(n)}(n-s,t) = 0$, and

$$\frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) \quad = \quad \alpha V^{\psi(n)}(n-s,t) - \alpha v(n-s) \quad \geq \quad 0$$

for all $t \in (t_3, t_4)$. Also,

$$\begin{aligned} \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) &= \frac{\partial v}{\partial t}(n-s) = 0 \\ \Rightarrow \quad \frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) &\geq \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) \end{aligned}$$

for all $t \in (t_3, t_4)$.

Case 1.2: $V^{\psi(n)}(n-s,t_1) > V^{\psi(n-s)}(n-s,t_1) > v(n-s)$ for some $t_1 \in (t_0, t_2)$:

Then, from the continuity of $V^{\psi(n)}(n-s,t)$, $v(n-s)$ and $V^{\psi(n-s)}(n-s,t)$ in t, there exists a neighborhood $(t_3, t_4) \subseteq (0, T)$ of t_1 such that $V^{\psi(n)}(n-s,t) > V^{\psi(n-s)}(n-s,t) > v(n-s)$ for all $t \in (t_3, t_4)$. Then $I^{\psi(n)}(n-s,t) = 1$, and

$$\begin{aligned} \frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) = & -\lambda(n-s) \int_0^{n-s} \int_{x^{\psi(n-s)}(n-s,t,s')}^{\infty} \left\{ r - \left[V^{\psi(n)}(n-s,t) \right.\right. \\ & \left.\left. - V^{\psi(n-s)}(n-s-s',t) - p(n-s) \right] \right\} F_{R|S}(dr|s') F_S(ds') \\ & + \alpha V^{\psi(n)}(n-s,t) + \lambda(n-s) p(n-s) + c(n-s) \end{aligned}$$

and

$$\begin{aligned} & \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) = \frac{\partial V^{\psi(n-s)}}{\partial t}(n-s,t) \\ = & -\lambda(n-s) \int_0^{n-s} \int_{x^{\psi(n-s)}(n-s,t,s')}^{\infty} \left\{ r - \left[V^{\psi(n-s)}(n-s,t) \right.\right. \\ & \left.\left. - V^{\psi(n-s)}(n-s-s',t) - p(n-s) \right] \right\} F_{R|S}(dr|s') F_S(ds') \\ & + \alpha V^{\psi(n-s)}(n-s,t) + \lambda(n-s) p(n-s) + c(n-s) \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad & \frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) - \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) \\ = & \left[\lambda(n-s) \int_0^{n-s} \int_{x^{\psi(n-s)}(n-s,t,s')}^{\infty} F_{R|S}(dr|s') F_S(ds') + \alpha \right] \times \\ & \left[V^{\psi(n)}(n-s,t) - V^{\psi(n-s)}(n-s,t) \right] \\ \geq & \ 0 \end{aligned}$$

for all $t \in (t_3, t_4)$.

Case 1.3: $V^{\psi(n)}(n-s,t) > v(n-s) = V^{\psi(n-s)}(n-s,t)$ for all $t \in (t_0, t_2)$:

Then $I^{\psi(n)}(n-s,t) = 0$, and

$$\frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) = \alpha V^{\psi(n)}(n-s,t) - \alpha v(n-s) \geq 0$$

for all $t \in (t_0, t_2)$. Also,

$$\begin{aligned}
\frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) &= \frac{\partial v}{\partial t}(n-s) = 0 \\
\Rightarrow \quad \frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) &\geq \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t)
\end{aligned}$$

for all $t \in (t_0, t_2)$.

Thus, in all cases, if $V^{\psi(n)}(n-s,t_1) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t_1)\right\} > 0$ for some $s \in \{1, \ldots, n\}$ and some $t_1 \in [0, T)$, then

$$\frac{\partial\left[V^{\psi(n)} - \max\left\{v, V^{\psi(n-s)}\right\}\right]}{\partial t}(n-s,t) \geq 0$$

for all t in a neighborhood of t_1. By repeating the argument at the right endpoint of any such neighborhood of t_1, it follows that $V^{\psi(n)}(n-s,t) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$ is nondecreasing on $[t_1, T]$, and hence $V^{\psi(n)}(n-s,T) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,T)\right\} \geq V^{\psi(n)}(n-s,t_1) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t_1)\right\} > 0$. But

$$V^{\psi(n)}(n-s,T) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,T)\right\} = v(n-s) - \max\{v(n-s), v(n-s)\} = 0$$

and thus a contradiction has been reached. Therefore,

$$V^{\psi(n)}(n-s,t) \leq \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$$

for all $t \in [0, T]$.

Case 2: Suppose there exists an $s \in \{1, \ldots, n\}$ and a $t_1 \in [0, T)$ such that $V^{\psi(n)}(n-s,t_1) < \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t_1)\right\}$.
Then, from the continuity of $V^{\psi(n)}(n-s,t)$ and $\max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$ in t, there exists a neighborhood $(t_0, t_2) \subseteq (0, T)$ of t_1 such that

$$V^{\psi(n)}(n-s,t) < \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$$

for all $t \in (t_0, t_2)$. Hence, t_1 can be taken as any point in (t_0, t_2), and at least one of the following three cases must hold.

Case 2.1: $V^{\psi(n)}(n-s,t_1) < v(n-s)$ and $V^{\psi(n-s)}(n-s,t_1) < v(n-s)$ for some $t_1 \in (t_0, t_2)$: Then, from the continuity of $V^{\psi(n)}(n-s,t)$, $v(n-s)$ and $V^{\psi(n-s)}(n-s,t)$ in t, there exists a neighborhood $(t_3,t_4) \subseteq (0,T)$ of t_1 such that $V^{\psi(n)}(n-s,t) < v(n-s)$ and $V^{\psi(n-s)}(n-s,t) < v(n-s)$ for all $t \in (t_3,t_4)$. Then $I^{\psi(n)}(n-s,t) = 0$, and

$$\frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) \;=\; \alpha V^{\psi(n)}(n-s,t) - \alpha v(n-s) \;\leq\; 0$$

for all $t \in (t_3,t_4)$. Also,

$$\begin{aligned}\frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) &= \frac{\partial v}{\partial t}(n-s) \;=\; 0\\ \Rightarrow\quad \frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) &\leq \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t)\end{aligned}$$

for all $t \in (t_3,t_4)$.

Case 2.2: $V^{\psi(n)}(n-s,t_1) < V^{\psi(n-s)}(n-s,t_1)$ and $v(n-s) < V^{\psi(n-s)}(n-s,t_1)$ for some $t_1 \in (t_0,t_2)$:

Then, from the continuity of $V^{\psi(n)}(n-s,t)$, $v(n-s)$ and $V^{\psi(n-s)}(n-s,t)$ in t, there exists a neighborhood $(t_3,t_4) \subseteq (0,T)$ of t_1 such that $V^{\psi(n)}(n-s,t) < V^{\psi(n-s)}(n-s,t)$ and $v(n-s) < V^{\psi(n-s)}(n-s,t)$ for all $t \in (t_3,t_4)$. Then $I^{\psi(n)}(n-s,t) = 1$, and

$$\begin{aligned}\frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) &= -\lambda(n-s)\int_0^{n-s}\int_{x^{\psi(n-s)}(n-s,t,s')}^{\infty}\left\{r - \left[V^{\psi(n)}(n-s,t)\right.\right.\\ &\qquad \left.\left.-V^{\psi(n-s)}(n-s-s',t) - p(n-s)\right]\right\} F_{R|S}(dr|s')F_S(ds')\\ &\quad + \alpha V^{\psi(n)}(n-s,t) + \lambda(n-s)p(n-s) + c(n-s)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) &= \frac{\partial V^{\psi(n-s)}}{\partial t}(n-s,t)\\ = -\lambda(n-s)&\int_0^{n-s}\int_{x^{\psi(n-s)}(n-s,t,s')}^{\infty}\left\{r - \left[V^{\psi(n-s)}(n-s,t)\right.\right.\\ &\qquad \left.\left.-V^{\psi(n-s)}(n-s-s',t) - p(n-s)\right]\right\} F_{R|S}(dr|s')F_S(ds')\\ + \alpha V^{\psi(n-s)}&(n-s,t) + \lambda(n-s)p(n-s) + c(n-s)\end{aligned}$$

$$\begin{aligned}\Rightarrow\quad & \frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) - \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t)\\ &= \left[\lambda(n-s)\int_0^{n-s}\int_{x^{\psi(n-s)}(n-s,t,s')}^{\infty} F_{R|S}(dr|s')F_S(ds') + \alpha\right] \times\\ &\quad \left[V^{\psi(n)}(n-s,t) - V^{\psi(n-s)}(n-s,t)\right]\\ &\leq 0\end{aligned}$$

for all $t \in (t_3, t_4)$.

Case 2.3: $V^{\psi(n)}(n-s,t) < v(n-s) = V^{\psi(n-s)}(n-s,t)$ for all $t \in (t_0, t_2)$:
Then $I^{\psi(n)}(n-s,t) = 0$, and

$$\frac{\partial V^{\psi(n)}(n-s,t)}{\partial t} \quad = \quad \alpha V^{\psi(n)}(n-s,t) - \alpha v(n-s) \quad \leq \quad 0$$

for all $t \in (t_0, t_2)$. Also,

$$\begin{aligned} \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) \quad &= \quad \frac{\partial v}{\partial t}(n-s) \quad = \quad 0 \\ \Rightarrow \quad \frac{\partial V^{\psi(n)}}{\partial t}(n-s,t) \quad &\leq \quad \frac{\partial \max\left\{v, V^{\psi(n-s)}\right\}}{\partial t}(n-s,t) \end{aligned}$$

for all $t \in (t_0, t_2)$.

Thus, in all cases, if $V^{\psi(n)}(n-s,t_1) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t_1)\right\} < 0$ for some $s \in \{1, \ldots, n\}$ and some $t_1 \in [0, T)$, then

$$\frac{\partial \left[V^{\psi(n)} - \max\left\{v, V^{\psi(n-s)}\right\}\right]}{\partial t}(n-s,t) \quad \leq \quad 0$$

for all t in a neighborhood of t_1. By repeating the argument at the right endpoint of any such neighborhood of t_1, it follows that $V^{\psi(n)}(n-s,t) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$ is nonincreasing on $[t_1, T]$, and hence $V^{\psi(n)}(n-s,T) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,T)\right\} \leq V^{\psi(n)}(n-s,t_1) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t_1)\right\} < 0$. But

$$V^{\psi(n)}(n-s,T) - \max\left\{v(n-s), V^{\psi(n-s)}(n-s,T)\right\} = v(n-s) - \max\{v(n-s), v(n-s)\} = 0$$

and thus a contradiction has been reached. Therefore,

$$V^{\psi(n)}(n-s,t) = \max\left\{v(n-s), V^{\psi(n-s)}(n-s,t)\right\}$$

for all $t \in [0, T]$. □

Appendix K: Proof of Lemma 30

Lemma 29

1. Υ_1 *satisfies the following Lipschitz condition:*

$$|\Upsilon_1(x_1) - \Upsilon_1(x_2)| \quad \leq \quad |x_2 - x_1| \quad \textit{for all} \;\; x_1, x_2 \in \Re$$

2. Υ_1 *is absolutely continuous on* $\Re$.

3. Υ_1 *is nonincreasing on* $\Re$.

Proof: 1. Let $x_1 \leq x_2$.

$$\begin{aligned}
|\Upsilon_1(x_1) - \Upsilon_1(x_2)| &\equiv \left| \int_{x_1}^{\infty} (r - x_1)\ F_R(dr) \ - \ \int_{x_2}^{\infty} (r - x_2)\ F_R(dr) \right| \\
&= \left| \int_{x_1}^{x_2} (r - x_1)\ F_R(dr) \ + \ \int_{x_2}^{\infty} (r - x_1)\ F_R(dr) \ - \ \int_{x_2}^{\infty} (r - x_2)\ F_R(dr) \right| \\
&\leq \int_{x_1}^{x_2} (x_2 - x_1)\ F_R(dr) \ + \ \int_{x_2}^{\infty} (x_2 - x_1)\ F_R(dr) \\
&\leq |x_2 - x_1|
\end{aligned}$$

2. Absolute continuity of Υ_1 follows directly from Υ_1 satisfying a Lipschitz condition.

3. Let $x_1 \leq x_2$.

$$\begin{aligned}
\Upsilon_1(x_1) &= \int_{x_1}^{\infty} (r - x_1)\ F_R(dr) \\
&\geq \int_{x_2}^{\infty} (r - x_1)\ F_R(dr) \\
&\geq \int_{x_2}^{\infty} (r - x_2)\ F_R(dr) \\
&= \Upsilon_1(x_2) \qquad \square
\end{aligned}$$

Appendix L: Proof of Proposition 6

Proposition 6 *If $c(m,\tilde{m},n)$ is nondecreasing in n for all m and $\tilde{m}$, and $\varphi(k,l,q)$ is nondecreasing in q for all k and l, then $V^*(m,\tilde{m},n)$ is nonincreasing in n for all m and $\tilde{m}$.*

To keep the notation simple, the proposition is proved for the special case with one terminal. In this case, loads have to be transported from the terminal to some unspecified destination. It should then be clear that the result holds for any number of terminals, but the extension of this proof to multiple terminals is omitted because it does not offer any new insight.

Proof: By induction on $\tilde{m}$ and n. Without loss of generality, assume that $s = 1$.
For any $m \in \{0, 1, \ldots, \overline{M}\}$, $\tilde{m} = 0$, and $n = \overline{N}$:
If $L^*(m, 0, \overline{N}) = \emptyset$:
Then, from Equation (10.8), the fact that $\sum_{\{b\,:\,\overline{N}+b\leq\overline{N}\}} f_B(b) = 0$, and $\mu(0) = 0$, it follows that

$$V^*(m,0,\overline{N}) \quad = \quad -\frac{\lambda p + c(m,0,\overline{N})}{\alpha}$$

Then

$$\begin{aligned}
V^*(m,0,\overline{N}-1) \quad &\geq \quad \frac{\lambda}{\alpha} \sum_{\{b\,:\,\overline{N}-1+b\leq\overline{N}\}} f_B(b) \times \\
&\qquad \int_{V^*(m,0,\overline{N}-1)-V^*(m,0,\overline{N}-1+b)-p}^{\infty} \Big\{ r - \Big[V^*(m,0,\overline{N}-1) \\
&\qquad\qquad - V^*(m,0,\overline{N}-1+b) - p \Big] \Big\} \, F_{R|B}(dr|b) \\
&\quad - \frac{\lambda p + c(m,0,\overline{N}-1)}{\alpha} \\
&\geq \quad -\frac{\lambda p + c(m,0,\overline{N})}{\alpha} \\
&= \quad V^*(m,0,\overline{N})
\end{aligned}$$

Note that if $m = 0$, then it has to hold that $L^*(m,0,\overline{N}) = \emptyset$, and the result that $V^*(m,0,\overline{N}-1) \geq V^*(m,0,\overline{N})$ has been established.
If $Q^*(m,0,\overline{N}) > 0$:

$$\begin{aligned}
V^*(m,0,\overline{N}-1) \quad &\geq \quad \max_{q\in\{0,1,\ldots,\min\{\overline{N}-1,\overline{N}_v\}\}} \Big\{ V^*(m-1,0,\overline{N}-1-q) - \varphi(q) \Big\} \\
&= \quad \max_{q\in\{1,2,\ldots,\min\{\overline{N},\overline{N}_v+1\}\}} \Big\{ V^*(m-1,0,\overline{N}-1-(q-1)) - \varphi(q-1) \Big\}
\end{aligned}$$

$$
\begin{aligned}
&\geq \max_{q\in\left\{1,2,\ldots,\min\left\{\overline{N},\overline{N}_v\right\}\right\}} \left\{V^*(m-1,0,\overline{N}-q)-\varphi(q-1)\right\} \\
&\geq \max_{q\in\left\{1,2,\ldots,\min\left\{\overline{N},\overline{N}_v\right\}\right\}} \left\{V^*(m-1,0,\overline{N}-q)-\varphi(q)\right\} \\
&= V^*(m,0,\overline{N}) \qquad \text{(L.1)}
\end{aligned}
$$

If $Q^*(m,0,\overline{N})=0$:

Then $V^*(m,0,\overline{N})=V^*(m-1,0,\overline{N})-\varphi(0)$. It is shown by contradiction that $V^*(m,0,\overline{N}-1)\geq V^*(m,0,\overline{N})$. Suppose $V^*(m,0,\overline{N}-1)<V^*(m,0,\overline{N})$. Then

$$
\begin{aligned}
V^*(m-1,0,\overline{N})-\varphi(0) &= V^*(m,0,\overline{N}) \\
&> V^*(m,0,\overline{N}-1) \\
&\geq V^*(m-1,0,\overline{N}-1)-\varphi(0) \\
\Rightarrow \quad V^*(m-1,0,\overline{N}) &> V^*(m-1,0,\overline{N}-1)
\end{aligned}
$$

As (reverse) induction hypothesis on m, suppose that $Q^*(m',0,\overline{N})=0$ and $V^*(m'-1,0,\overline{N})>V^*(m'-1,0,\overline{N}-1)$ for some $m'\in\{1,\ldots,\overline{M}\}$.

If $L^*(m'-1,0,\overline{N})=\varnothing$:

$$
\begin{aligned}
V^*(m'-1,0,\overline{N}-1) &\geq \frac{\lambda}{\alpha}\sum_{\left\{b\,:\,\overline{N}-1+b\leq\overline{N}\right\}} f_B(b)\times \\
&\quad \int_{V^*(m'-1,0,\overline{N}-1)-V^*(m'-1,0,\overline{N}-1+b)-p}^{\infty}\left\{r-\left[V^*(m'-1,0,\overline{N}-1)\right.\right. \\
&\qquad \left.\left.-V^*(m'-1,0,\overline{N}-1+b)-p\right]\right\} F_{R|B}(dr|b) \\
&\quad -\frac{\lambda p+c(m'-1,0,\overline{N}-1)}{\alpha} \\
&\geq -\frac{\lambda p+c(m'-1,0,\overline{N})}{\alpha} \\
&= V^*(m'-1,0,\overline{N})
\end{aligned}
$$

contradicting the induction hypothesis on m that $V^*(m'-1,0,\overline{N})>V^*(m'-1,0,\overline{N}-1)$. Note that if $m'-1=0$, then it has to hold that $L^*(m'-1,0,\overline{N})=\varnothing$, and the result that $V^*(m,0,\overline{N}-1)\geq V^*(m,0,\overline{N})$ is obtained.

If $Q^*(m'-1,0,\overline{N})>0$, then similar to Equation (L.1) it follows that $V^*(m'-1,0,\overline{N}-1)\geq V^*(m'-1,0,\overline{N})$, contradicting the induction hypothesis on m that $V^*(m'-1,0,\overline{N})>V^*(m'-1,0,\overline{N}-1)$.

Thus the only way in which the induction hypothesis on m that $V^*(m'-1, 0, \overline{N}) > V^*(m'-1, 0, \overline{N}-1)$ can hold, is if $Q^*(m'-1, 0, \overline{N}) = 0$. Then

$$\begin{aligned} V^*(m'-2, 0, \overline{N}) - \varphi(0) &= V^*(m'-1, 0, \overline{N}) \\ &> V^*(m'-1, 0, \overline{N}-1) \\ &\geq V^*(m'-2, 0, \overline{N}-1) - \varphi(0) \\ \Rightarrow \quad V^*(m'-2, 0, \overline{N}) &> V^*(m'-2, 0, \overline{N}-1) \end{aligned}$$

and the induction on m continues. However, since there are only a finite number $\overline{M}$ of vehicles, the induction leads to $m' = 0$, in which case it has to hold that $L^*(m', 0, \overline{N}) = \emptyset$, and a contradiction is obtained as shown above. Therefore, the result that $V^*(m, 0, \overline{N}-1) \geq V^*(m, 0, \overline{N})$ for all $m \in \{0, 1, \ldots, \overline{M}\}$, is obtained.

As induction hypothesis on n, suppose that $V^*(m, 0, n'-1) \geq V^*(m, 0, n')$ for all $m \in \{0, 1, \ldots, \overline{M}\}$ and all $n' \in \{n, \ldots, \overline{N}\}$.

If $L^*(m, 0, n-1) = \emptyset$:

By contradiction. Suppose, for any $m \in \{0, 1, \ldots, \overline{M}\}$, that $V^*(m, 0, n-2) < V^*(m, 0, n-1)$. From the induction hypothesis on n, for any $b > 0$, $V^*(m, 0, n-2+b) \geq V^*(m, 0, n-1+b)$. Thus, for any $b > 0$, $V^*(m, 0, n-2) - V^*(m, 0, n-2+b) < V^*(m, 0, n-1) - V^*(m, 0, n-1+b)$. Hence, from Lemma 5,

$$\begin{aligned} & \int_{V^*(m,0,n-2)-V^*(m,0,n-2+b)-p}^{\infty} \{r - [V^*(m, 0, n-2) - V^*(m, 0, n-2+b) - p]\} \, F_{R|B}(dr|b) \\ \geq & \int_{V^*(m,0,n-1)-V^*(m,0,n-1+b)-p}^{\infty} \{r - [V^*(m, 0, n-1) - V^*(m, 0, n-1+b) - p]\} \, F_{R|B}(dr|b) \end{aligned}$$

for any $b > 0$. Then, from Equation (10.8) and the fact that $\mu(0) = 0$, it follows that

$$\begin{aligned} V^*(m, 0, n-2) \geq \; & \frac{\lambda}{\alpha} \sum_{\{b \,:\, n-2+b \leq \overline{N}\}} f_B(b) \times \\ & \int_{V^*(m,0,n-2)-V^*(m,0,n-2+b)-p}^{\infty} \{r - [V^*(m, 0, n-2) \\ & \qquad - V^*(m, 0, n-2+b) - p]\} \, F_{R|B}(dr|b) \\ & - \frac{\lambda p + c(m, 0, n-2)}{\alpha} \\ = \; & \frac{\lambda}{\alpha} \sum_{\{b \,:\, n-1+b \leq \overline{N}\}} f_B(b) \times \\ & \int_{V^*(m,0,n-2)-V^*(m,0,n-2+b)-p}^{\infty} \{r - [V^*(m, 0, n-2) \end{aligned}$$

$$
\begin{aligned}
& \qquad\qquad\qquad\qquad\qquad\qquad -V^*(m,0,n-2+b)-p]\}\ F_{R|B}(dr|b) \\
& \qquad + \frac{\lambda}{\alpha} f_B(\overline{N}-n+2) \int_{V^*(m,0,n-2)-V^*(m,0,\overline{N})-p}^{\infty} \{r-[V^*(m,0,n-2) \\
& \qquad\qquad\qquad\qquad\qquad\qquad -V^*(m,0,\overline{N})-p\Big]\Big\}\ F_{R|B}(dr|b) \\
& - \frac{\lambda p + c(m,0,n-2)}{\alpha} \\
\geq\ & \frac{\lambda}{\alpha} \sum_{\{b\,:\,n-1+b\leq\overline{N}\}} f_B(b) \times \\
& \quad \int_{V^*(m,0,n-2)-V^*(m,0,n-2+b)-p}^{\infty} \{r-[V^*(m,0,n-2) \\
& \qquad\qquad\qquad\qquad\qquad\qquad -V^*(m,0,n-2+b)-p]\}\ F_{R|B}(dr|b) \\
& - \frac{\lambda p + c(m,0,n-2)}{\alpha} \\
\geq\ & \frac{\lambda}{\alpha} \sum_{\{b\,:\,n-1+b\leq\overline{N}\}} f_B(b) \times \\
& \quad \int_{V^*(m,0,n-1)-V^*(m,0,n-1+b)-p}^{\infty} \{r-[V^*(m,0,n-1) \\
& \qquad\qquad\qquad\qquad\qquad\qquad -V^*(m,0,n-1+b)-p]\}\ F_{R|B}(dr|b) \\
& - \frac{\lambda p + c(m,0,n-1)}{\alpha} \\
=\ & V^*(m,0,n-1)
\end{aligned}
$$

which contradicts the assumption that $V^*(m,0,n-2) < V^*(m,0,n-1)$. Thus $V^*(m,0,n-2) \geq V^*(m,0,n-1)$. If $m=0$, then it has to hold that $L^*(m,0,n-1) = \emptyset$, and the result that $V^*(m,0,n-2) \geq V^*(m,0,n-1)$ has been established.

If $Q^*(m,0,n-1) > 0$, then similar to Equation (L.1) it follows that $V^*(m,0,n-2) \geq V^*(m,0,n-1)$, which contradicts the assumption that $V^*(m,0,n-2) < V^*(m,0,n-1)$, and thus $V^*(m,0,n-2) \geq V^*(m,0,n-1)$.

If $Q^*(m,0,n-1) = 0$:

Then $V^*(m,0,n-1) = V^*(m-1,0,n-1) - \varphi(0)$. It is shown by contradiction that $V^*(m,0,n-2) \geq V^*(m,0,n-1)$. Suppose $V^*(m,0,n-2) < V^*(m,0,n-1)$. Then

$$
\begin{aligned}
V^*(m-1,0,n-1) - \varphi(0) &= V^*(m,0,n-1) \\
&> V^*(m,0,n-2) \\
&\geq V^*(m-1,0,n-2) - \varphi(0) \\
\Rightarrow \quad V^*(m-1,0,n-1) &> V^*(m-1,0,n-2)
\end{aligned}
$$

As (reverse) induction hypothesis on m, suppose that $Q^*(m', 0, n-1) = 0$ and $V^*(m'-1, 0, n-1) > V^*(m'-1, 0, n-2)$ for some $m' \in \{1, \ldots, \overline{M}\}$.

If $L^*(m'-1, 0, n-1) = \emptyset$:

Then, from the induction hypothesis on n, for any $b > 0$, $V^*(m'-1, 0, n-2+b) \geq V^*(m'-1, 0, n-1+b)$. Thus, for any $b > 0$, $V^*(m'-1, 0, n-2) - V^*(m'-1, 0, n-2+b) < V^*(m'-1, 0, n-1) - V^*(m'-1, 0, n-1+b)$. Hence, from Lemma 5, from Equation (10.8), and the fact that $\mu(0) = 0$, it follows that

$$\begin{aligned}
V^*(m'-1, 0, n-2) \quad \geq \quad & \frac{\lambda}{\alpha} \sum_{\{b\,:\,n-2+b \leq \overline{N}\}} f_B(b) \times \\
& \int_{V^*(m'-1,0,n-2)-V^*(m'-1,0,n-2+b)-p}^{\infty} \{r - [V^*(m'-1, 0, n-2) \\
& \qquad -V^*(m'-1, 0, n-2+b) - p]\} \; F_{R|B}(dr|b) \\
& - \frac{\lambda p + c(m'-1, 0, n-2)}{\alpha} \\
\geq \quad & \frac{\lambda}{\alpha} \sum_{\{b\,:\,n-1+b \leq \overline{N}\}} f_B(b) \times \\
& \int_{V^*(m'-1,0,n-1)-V^*(m'-1,0,n-1+b)-p}^{\infty} \{r - [V^*(m'-1, 0, n-1) \\
& \qquad -V^*(m'-1, 0, n-1+b) - p]\} \; F_{R|B}(dr|b) \\
& - \frac{\lambda p + c(m'-1, 0, n-1)}{\alpha} \\
= \quad & V^*(m'-1, 0, n-1)
\end{aligned}$$

contradicting the induction hypothesis on m that $V^*(m'-1, 0, n-1) > V^*(m'-1, 0, n-2)$.

If $m'-1 = 0$, then it has to hold that $L^*(m'-1, 0, n-1) = \emptyset$, and the result that $V^*(m, 0, n-2) \geq V^*(m, 0, n-1)$ is obtained.

If $Q^*(m'-1, 0, n-1) > 0$, then similar to Equation (L.1) it follows that $V^*(m'-1, 0, n-2) \geq V^*(m'-1, 0, n-1)$, contradicting the induction hypothesis on m that $V^*(m'-1, 0, n-1) > V^*(m'-1, 0, n-2)$.

Thus the only way in which the induction hypothesis on m that $V^*(m'-1, 0, n-1) > V^*(m'-1, 0, n-2)$ can hold, is if $Q^*(m'-1, 0, n-1) = 0$. Then

$$\begin{aligned}
V^*(m'-2, 0, n-1) - \varphi(0) \quad &= \quad V^*(m'-1, 0, n-1) \\
&> \quad V^*(m'-1, 0, n-2) \\
&\geq \quad V^*(m'-2, 0, n-2) - \varphi(0)
\end{aligned}$$

$$\Rightarrow \quad V^*(m'-2,0,n-1) \quad > \quad V^*(m'-2,0,n-2)$$

and the induction on m continues. However, since there are only a finite number $\overline{M}$ of vehicles, the induction leads to $m' = 0$, in which case it has to hold that $L^*(m',0,n-1) = \emptyset$, and a contradiction is obtained as shown above. Therefore, the result that $V^*(m,0,n-2) \geq V^*(m,0,n-1)$ for all $m \in \{0,1,\ldots,\overline{M}\}$, is obtained, and the induction on n is established. Hence, $V^*(m,\tilde{m},n-1) \geq V^*(m,\tilde{m},n)$ for all $m \in \{0,1,\ldots,\overline{M}\}$ and all $n \in \{1,2,\ldots,\overline{N}\}$, for $\tilde{m} = 0$.

As induction hypothesis on $\tilde{m}$, suppose that $V^*(m,\tilde{m}',n-1) \geq V^*(m,\tilde{m}',n)$ for all $m \in \{0,1,\ldots,\overline{M}\}$, for all $n \in \{1,2,\ldots,\overline{N}\}$, and for all $\tilde{m}' \in \{0,\ldots,\tilde{m}\}$. For any $m \in \{0,1,\ldots,\overline{M}-\tilde{m}-1\}$, and $n = \overline{N}$:

If $L^*(m,\tilde{m}+1,\overline{N}) = \emptyset$:

Then, from Equation (10.8), the fact that $\sum_{\{b\,:\,\overline{N}+b\leq\overline{N}\}} f_B(b) = 0$, it follows that

$$\begin{aligned}
V^*(m,\tilde{m}+1,\overline{N}-1) \quad &\geq \quad \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b\,:\,\overline{N}-1+b\leq\overline{N}\}} f_B(b) \times \\
&\quad \int_{V^*(m,\tilde{m}+1,\overline{N}-1)-V^*(m,\tilde{m}+1,\overline{N}-1+b)-p}^{\infty} \Big\{ r - \Big[V^*(m,\tilde{m}+1,\overline{N}-1) \\
&\qquad\qquad -V^*(m,\tilde{m}+1,\overline{N}-1+b) - p \Big] \Big\} \, F_{R|B}(dr|b) \\
&\quad + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m+1,\tilde{m},\overline{N}-1\right) \; - \; \frac{\lambda p + c(m,\tilde{m}+1,\overline{N}-1)}{\mu(\tilde{m}+1)+\alpha} \\
&\geq \quad \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m+1,\tilde{m},\overline{N}-1\right) \; - \; \frac{\lambda p + c(m,\tilde{m}+1,\overline{N}-1)}{\mu(\tilde{m}+1)+\alpha} \\
&\geq \quad \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m+1,\tilde{m},\overline{N}\right) \; - \; \frac{\lambda p + c(m,\tilde{m}+1,\overline{N})}{\mu(\tilde{m}+1)+\alpha} \\
&= \quad V^*(m,\tilde{m}+1,\overline{N})
\end{aligned}$$

The last inequality follows from the induction hypothesis on $\tilde{m}$ and the assumption that $c(m,\tilde{m},n)$ is nondecreasing in n for all m and $\tilde{m}$. If $m = 0$, then it has to hold that $L^*(m,\tilde{m}+1,\overline{N}) = \emptyset$, and the result that $V^*(m,\tilde{m}+1,\overline{N}-1) \geq V^*(m,\tilde{m}+1,\overline{N})$ has been established.

If $Q^*(m,\tilde{m}+1,\overline{N}) > 0$, then similar to Equation (L.1) it follows that $V^*(m,\tilde{m}+1,\overline{N}-1) \geq V^*(m,\tilde{m}+1,\overline{N})$.

If $Q^*(m,\tilde{m}+1,\overline{N}) = 0$:

Then $V^*(m,\tilde{m}+1,\overline{N}) = V^*(m-1,\tilde{m}+1,\overline{N}) - \varphi(0)$. It is shown by contradiction that

$V^*(m, \tilde{m}+1, \overline{N}-1) \geq V^*(m, \tilde{m}+1, \overline{N})$. Suppose $V^*(m, \tilde{m}+1, \overline{N}-1) < V^*(m, \tilde{m}+1, \overline{N})$. Then

$$\begin{aligned}
V^*(m-1, \tilde{m}+1, \overline{N}) - \varphi(0) &= V^*(m, \tilde{m}+1, \overline{N}) \\
&> V^*(m, \tilde{m}+1, \overline{N}-1) \\
&\geq V^*(m-1, \tilde{m}+1, \overline{N}-1) - \varphi(0) \\
\Rightarrow \quad V^*(m-1, \tilde{m}+1, \overline{N}) &> V^*(m-1, \tilde{m}+1, \overline{N}-1)
\end{aligned}$$

As (reverse) induction hypothesis on m, suppose that $Q^*(m', \tilde{m}+1, \overline{N}) = 0$ and $V^*(m'-1, \tilde{m}+1, \overline{N}) > V^*(m'-1, \tilde{m}+1, \overline{N}-1)$ for some $m' \in \{1, \ldots, \overline{M}-\tilde{m}-1\}$.
If $L^*(m'-1, \tilde{m}+1, \overline{N}) = \emptyset$:

$$\begin{aligned}
V^*(m'-1, \tilde{m}+1, \overline{N}-1) \quad &\geq \quad \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b\,:\,\overline{N}-1+b \leq \overline{N}\}} f_B(b) \times \\
&\int_{V^*(m'-1,\tilde{m}+1,\overline{N}-1)-V^*(m'-1,\tilde{m}+1,\overline{N}-1+b)-p}^{\infty} \Big\{ r - \Big[V^*(m'-1, \tilde{m}+1, \overline{N}-1) \\
&\qquad - V^*(m'-1, \tilde{m}+1, \overline{N}-1+b) - p \Big] \Big\} F_{R|B}(dr|b) \\
&\quad + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m', \tilde{m}, \overline{N}-1\right) - \frac{\lambda p + c(m'-1, \tilde{m}+1, \overline{N}-1)}{\mu(\tilde{m}+1)+\alpha} \\
&\geq \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m', \tilde{m}, \overline{N}-1\right) - \frac{\lambda p + c(m'-1, \tilde{m}+1, \overline{N}-1)}{\mu(\tilde{m}+1)+\alpha} \\
&\geq \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m', \tilde{m}, \overline{N}\right) - \frac{\lambda p + c(m'-1, \tilde{m}+1, \overline{N})}{\mu(\tilde{m}+1)+\alpha} \\
&= V^*(m'-1, \tilde{m}+1, \overline{N})
\end{aligned}$$

contradicting the induction hypothesis on m that $V^*(m'-1, \tilde{m}+1, \overline{N}) > V^*(m'-1, \tilde{m}+1, \overline{N}-1)$. The last inequality follows from the induction hypothesis on $\tilde{m}$ and the assumption that $c(m, \tilde{m}, n)$ is nondecreasing in n for all m and $\tilde{m}$. If $m'-1 = 0$, then it has to hold that $L^*(m'-1, \tilde{m}+1, \overline{N}) = \emptyset$, and the result that $V^*(m, \tilde{m}+1, \overline{N}-1) \geq V^*(m, \tilde{m}+1, \overline{N})$ is obtained.
If $Q^*(m'-1, \tilde{m}+1, \overline{N}) > 0$, then similar to Equation (L.1) it follows that $V^*(m'-1, \tilde{m}+1, \overline{N}-1) \geq V^*(m'-1, \tilde{m}+1, \overline{N})$, contradicting the induction hypothesis on m that $V^*(m'-1, \tilde{m}+1, \overline{N}) > V^*(m'-1, \tilde{m}+1, \overline{N}-1)$.
Thus the only way in which the induction hypothesis on m that $V^*(m'-1, \tilde{m}+1, \overline{N}) >$

$V^*(m'-1, \tilde{m}+1, \overline{N}-1)$ can hold, is if $Q^*(m'-1, \tilde{m}+1, \overline{N}) = 0$. Then

$$\begin{aligned} V^*(m'-2, \tilde{m}+1, \overline{N}) - \varphi(0) &= V^*(m'-1, \tilde{m}+1, \overline{N}) \\ &> V^*(m'-1, \tilde{m}+1, \overline{N}-1) \\ &\geq V^*(m'-2, \tilde{m}+1, \overline{N}-1) - \varphi(0) \\ \Rightarrow \quad V^*(m'-2, \tilde{m}+1, \overline{N}) &> V^*(m'-2, \tilde{m}+1, \overline{N}-1) \end{aligned}$$

and the induction on m continues. However, since there are only a finite number $\overline{M}$ of vehicles, the induction leads to $m' = 0$, in which case it has to hold that $L^*(m', \tilde{m}+1, \overline{N}) = \emptyset$, and a contradiction is obtained as shown above. Therefore, the result that $V^*(m, \tilde{m}+1, \overline{N}-1) \geq V^*(m, \tilde{m}+1, \overline{N})$ for all $m \in \{0, 1, \ldots, \overline{M} - \tilde{m} - 1\}$, is obtained.

As induction hypothesis on n, suppose that $V^*(m, \tilde{m}+1, n'-1) \geq V^*(m, \tilde{m}+1, n')$ for all $m \in \{0, 1, \ldots, \overline{M} - \tilde{m} - 1\}$ and all $n' \in \{n, \ldots, \overline{N}\}$.

If $L^*(m, \tilde{m}+1, n-1) = \emptyset$:

By contradiction. Suppose, for any $m \in \{0, 1, \ldots, \overline{M} - \tilde{m} - 1\}$, that $V^*(m, \tilde{m}+1, n-2) < V^*(m, \tilde{m}+1, n-1)$. From the induction hypothesis on n, for any $b > 0$, $V^*(m, \tilde{m}+1, n-2+b) \geq V^*(m, \tilde{m}+1, n-1+b)$. Thus, for any $b > 0$, $V^*(m, \tilde{m}+1, n-2) - V^*(m, \tilde{m}+1, n-2+b) < V^*(m, \tilde{m}+1, n-1) - V^*(m, \tilde{m}+1, n-1+b)$. Hence, from Lemma 5,

$$\begin{aligned} & \int_{V^*(m,\tilde{m}+1,n-2)-V^*(m,\tilde{m}+1,n-2+b)-p}^{\infty} \{r - [V^*(m, \tilde{m}+1, n-2) \\ & \qquad -V^*(m, \tilde{m}+1, n-2+b) - p]\} \, F_{R|B}(dr|b) \\ \geq \; & \int_{V^*(m,\tilde{m}+1,n-1)-V^*(m,\tilde{m}+1,n-1+b)-p}^{\infty} \{r - [V^*(m, \tilde{m}+1, n-1) \\ & \qquad -V^*(m, \tilde{m}+1, n-1+b) - p]\} \, F_{R|B}(dr|b) \end{aligned} \tag{L.2}$$

for any $b > 0$. Then, from Equation (10.8) it follows that

$$\begin{aligned} V^*(m, \tilde{m}+1, n-2) \; \geq \; & \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b \,:\, n-2+b \leq \overline{N}\}} f_B(b) \times \\ & \int_{V^*(m,\tilde{m}+1,n-2)-V^*(m,\tilde{m}+1,n-2+b)-p}^{\infty} \{r - [V^*(m, \tilde{m}+1, n-2) \\ & \qquad -V^*(m, \tilde{m}+1, n-2+b) - p]\} \, F_{R|B}(dr|b) \\ & + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*(m+1, \tilde{m}, n-2) - \frac{\lambda p + c(m, \tilde{m}+1, n-2)}{\mu(\tilde{m}+1)+\alpha} \\ = \; & \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b \,:\, n-1+b \leq \overline{N}\}} f_B(b) \times \end{aligned}$$

$$
\begin{aligned}
& \int_{V^*(m,\tilde{m}+1,n-2)-V^*(m,\tilde{m}+1,n-2+b)-p}^{\infty} \{r - [V^*(m,\tilde{m}+1,n-2) \\
& \qquad -V^*(m,\tilde{m}+1,n-2+b) - p]\} \; F_{R|B}(dr|b) \\
+ \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} f_B(\overline{N}-n+2) & \int_{V^*(m,\tilde{m}+1,n-2)-V^*(m,\tilde{m}+1,\overline{N})-p}^{\infty} \{r - [V^*(m,\tilde{m}+1,n-2) \\
& \qquad -V^*(m,\tilde{m}+1,\overline{N}) - p\big]\Big\} \; F_{R|B}(dr|b) \\
& + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*(m+1,\tilde{m},n-2) - \frac{\lambda p + c(m,\tilde{m}+1,n-2)}{\mu(\tilde{m}+1)+\alpha} \\
\geq & \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b\,:\,n-1+b\leq \overline{N}\}} f_B(b) \times \\
& \int_{V^*(m,\tilde{m}+1,n-2)-V^*(m,\tilde{m}+1,n-2+b)-p}^{\infty} \{r - [V^*(m,\tilde{m}+1,n-2) \\
& \qquad -V^*(m,\tilde{m}+1,n-2+b) - p]\} \; F_{R|B}(dr|b) \\
& + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*(m+1,\tilde{m},n-2) - \frac{\lambda p + c(m,\tilde{m}+1,n-2)}{\mu(\tilde{m}+1)+\alpha} \\
\geq & \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b\,:\,n-1+b\leq \overline{N}\}} f_B(b) \times \\
& \int_{V^*(m,\tilde{m}+1,n-1)-V^*(m,\tilde{m}+1,n-1+b)-p}^{\infty} \{r - [V^*(m,\tilde{m}+1,n-1) \\
& \qquad -V^*(m,\tilde{m}+1,n-1+b) - p]\} \; F_{R|B}(dr|b) \\
& + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*(m+1,\tilde{m},n-1) - \frac{\lambda p + c(m,\tilde{m}+1,n-1)}{\mu(\tilde{m}+1)+\alpha} \\
= & V^*(m,\tilde{m}+1,n-1)
\end{aligned}
$$

which contradicts the assumption that $V^*(m,\tilde{m}+1,n-2) < V^*(m,\tilde{m}+1,n-1)$. The last inequality follows from Equation (L.2), the induction hypothesis on $\tilde{m}$, and the assumption that $c(m,\tilde{m},n)$ is nondecreasing in n for all m and $\tilde{m}$. Thus $V^*(m,\tilde{m}+1,n-2) \geq V^*(m,\tilde{m}+1,n-1)$. If $m=0$, then it has to hold that $L^*(m,\tilde{m}+1,n-1) = \emptyset$, and the result that $V^*(m,\tilde{m}+1,n-2) \geq V^*(m,\tilde{m}+1,n-1)$ has been established.

If $Q^*(m,\tilde{m}+1,n-1) > 0$, then similar to Equation (L.1) it follows that $V^*(m,\tilde{m}+1,n-2) \geq V^*(m,\tilde{m}+1,n-1)$.

If $Q^*(m,\tilde{m}+1,n-1) = 0$:

Then $V^*(m,\tilde{m}+1,n-1) = V^*(m-1,\tilde{m}+1,n-1) - \varphi(0)$. It is shown by contradiction that $V^*(m,\tilde{m}+1,n-2) \geq V^*(m,\tilde{m}+1,n-1)$. Suppose $V^*(m,\tilde{m}+1,n-2) < V^*(m,\tilde{m}+1,n-1)$. Then

$$
V^*(m-1,\tilde{m}+1,n-1) - \varphi(0) \quad = \quad V^*(m,\tilde{m}+1,n-1)
$$

$$
\begin{aligned}
& & > \; & V^*(m, \tilde{m}+1, n-2) \\
& & \geq \; & V^*(m-1, \tilde{m}+1, n-2) - \varphi(0) \\
\Rightarrow \; & V^*(m-1, \tilde{m}+1, n-1) & > \; & V^*(m-1, \tilde{m}+1, n-2)
\end{aligned}
$$

As (reverse) induction hypothesis on m, suppose that $Q^*(m', \tilde{m}+1, n-1) = 0$ and $V^*(m'-1, \tilde{m}+1, n-1) > V^*(m'-1, \tilde{m}+1, n-2)$ for some $m' \in \{1, \ldots, \overline{M} - \tilde{m} - 1\}$.

If $L^*(m'-1, \tilde{m}+1, n-1) = \emptyset$:

Then, from the induction hypothesis on n, for any $b > 0$, $V^*(m'-1, \tilde{m}+1, n-2+b) \geq V^*(m'-1, \tilde{m}+1, n-1+b)$. Thus, for any $b > 0$, $V^*(m'-1, \tilde{m}+1, n-2) - V^*(m'-1, \tilde{m}+1, n-2+b) < V^*(m'-1, \tilde{m}+1, n-1) - V^*(m'-1, \tilde{m}+1, n-1+b)$. Hence, from Lemma 5, from Equation (10.8), and the induction hypothesis on $\tilde{m}$, it follows that

$$
\begin{aligned}
V^*(m'-1, \tilde{m}+1, n-2) \;\; & \geq \;\; \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b \,:\, n-2+b \leq \overline{N}\}} f_B(b) \times \\
& \int_{V^*(m'-1,\tilde{m}+1,n-2) - V^*(m'-1,\tilde{m}+1,n-2+b) - p}^{\infty} \{r - [V^*(m'-1, \tilde{m}+1, n-2) \\
& \qquad - V^*(m'-1, \tilde{m}+1, n-2+b) - p]\} \, F_{R|B}(dr|b) \\
& \quad + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m', \tilde{m}, n-2\right) \; - \; \frac{\lambda p + c(m'-1, \tilde{m}+1, n-2)}{\mu(\tilde{m}+1)+\alpha} \\
& \geq \;\; \frac{\lambda}{\mu(\tilde{m}+1)+\alpha} \sum_{\{b \,:\, n-1+b \leq \overline{N}\}} f_B(b) \times \\
& \int_{V^*(m'-1,\tilde{m}+1,n-1) - V^*(m'-1,\tilde{m}+1,n-1+b) - p}^{\infty} \{r - [V^*(m'-1, \tilde{m}+1, n-1) \\
& \qquad - V^*(m'-1, \tilde{m}+1, n-1+b) - p]\} \, F_{R|B}(dr|b) \\
& \quad + \frac{\mu(\tilde{m}+1)}{\mu(\tilde{m}+1)+\alpha} V^*\left(m', \tilde{m}, n-1\right) \; - \; \frac{\lambda p + c(m'-1, \tilde{m}+1, n-1)}{\mu(\tilde{m}+1)+\alpha} \\
& = \;\; V^*(m'-1, \tilde{m}+1, n-1)
\end{aligned}
$$

contradicting the induction hypothesis on m that $V^*(m'-1, \tilde{m}+1, n-1) > V^*(m'-1, \tilde{m}+1, n-2)$. If $m'-1 = 0$, then it has to hold that $L^*(m'-1, \tilde{m}+1, n-1) = \emptyset$, and the result that $V^*(m, \tilde{m}+1, n-2) \geq V^*(m, \tilde{m}+1, n-1)$ is obtained.

If $Q^*(m'-1, \tilde{m}+1, n-1) > 0$, then similar to Equation (L.1) it follows that $V^*(m'-1, \tilde{m}+1, n-2) \geq V^*(m'-1, \tilde{m}+1, n-1)$, contradicting the induction hypothesis on m that $V^*(m'-1, \tilde{m}+1, n-1) > V^*(m'-1, \tilde{m}+1, n-2)$.

Thus the only way in which the induction hypothesis on m that $V^*(m'-1, \tilde{m}+1, n-1) >$

$V^*(m'-1, \tilde{m}+1, n-2)$ can hold, is if $Q^*(m'-1, \tilde{m}+1, n-1) = 0$. Then

$$\begin{aligned} V^*(m'-2, \tilde{m}+1, n-1) - \varphi(0) &= V^*(m'-1, \tilde{m}+1, n-1) \\ &> V^*(m'-1, \tilde{m}+1, n-2) \\ &\geq V^*(m'-2, \tilde{m}+1, n-2) - \varphi(0) \\ \Rightarrow \quad V^*(m'-2, \tilde{m}+1, n-1) &> V^*(m'-2, \tilde{m}+1, n-2) \end{aligned}$$

and the induction on m continues. However, since there are only a finite number $\overline{M}$ of vehicles, the induction leads to $m' = 0$, in which case it has to hold that $L^*(m', \tilde{m} + 1, n-1) = \emptyset$, and a contradiction is obtained as shown above. Therefore, the result that $V^*(m, \tilde{m}+1, n-2) \geq V^*(m, \tilde{m}+1, n-1)$ for all $m \in \{0, 1, \ldots, \overline{M} - \tilde{m} - 1\}$, is obtained, and the induction on n is established. Thus, $V^*(m, \tilde{m}+1, n-1) \geq V^*(m, \tilde{m}+1, n)$ for all $m \in \{0, 1, \ldots, \overline{M} - \tilde{m} - 1\}$ and all $n \in \{1, 2, \ldots, \overline{N}\}$, and the induction on $\tilde{m}$ is established. Hence the result follows. $\square$

Vita

Vita

Anton Kleywegt saw daylight for the first time on July 17, 1964, in Krugerdorp, South Africa. In his early years, he explored the world by frequently running away from home, to the consternation of his parents. After a while he settled down and opted for more formal education; for this he attended the Monument Park Primary School, fortunately just down the road from home. High school took him further afield, to the Afrikaans Boys' High School, Pretoria; being such an excellent institution, it was worth every mile. After tormenting himself with the choice between many alternative directions of study, including physics, medicine, and a variety of engineering fields, he proceeded to study civil engineering at the University of Pretoria. Four years and lots of fun later he graduated with a bachelor's degree, and then the South African Defence Force decided it was time for him to apply his skills to military endeavours. He spent the next two years running around in the dunes near Walvis Bay, Namibia, the mountains around Oudtshoorn in the Cape, and the bushes of southern Angola. Thereafter it was time for the jungle of corporate life, and he joined Van Wyk & Louw Consulting Engineers. He made himself useful doing structural design, transportation engineering, project management, and developing geographic information systems and management systems. At the same time he completed a bachelor's degree in economics and management at the University of South Africa.

For a long time his eyes had been set on pursuing graduate studies in a country beyond the horizon, and on August 2, 1992 he departed for the United States of America. After a year of intense study in the School of Civil Engineering at Purdue University, he obtained a master's degree with transportation as major. His interest was drawn towards operations research, and he joined the Ph.D. program in Purdue University's School of Industrial Engineering. After a fulfilling period of learning and many stimulating experiences, he completed his Ph.D. studies, and now he is ready for new adventures. These will take him to the Georgia Institute of Technology, where he will be an assistant professor in the School of Industrial and Systems Engineering.