DEMRY, Woodrow Steven, 1943–
INVENTORY MANAGEMENT SUBJECT TO FINANCIAL
CONSTRAINTS.

The Ohio State University, Ph.D., 1971
Engineering, industrial

University Microfilms, A XEROX Company, Ann Arbor, Michigan
INVENTORY MANAGEMENT
SUBJECT TO FINANCIAL CONSTRAINTS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Woodrow Steven Demmy, B.I.E., M.S.

The Ohio State University
1971

Approved by

[Signature]
Adviser
Department of Industrial Engineering
PLEASE NOTE:

Some Pages have indistinct print. Filmed as received.

UNIVERSITY MICROFILMS
ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to all those who made this work possible. It is certainly true that the success of one person requires the help and encouragement of many others. A man who is completely self-made is either extremely rare or non-existent.

I am particularly indebted to my advisor, Professor Albert B. Bishop, who guided and encouraged me through five years of graduate studies, including the completion of the Master of Science degree in 1967. His patience and many helpful suggestions and constructive criticisms are greatly appreciated.

The Air Force Logistics Command sponsored the final portion of my graduate studies at the Ohio State University, and I sincerely thank Mr. John Madden and Mr. Raymond Peckham for approving my participation in this program, and for providing many important insights into the problems of military supply management. I am also grateful to Miss Virginia Wyant for her fine typing of the many drafts of this study.

Finally, I am especially grateful to my wife, Barbara Sue, for her patience and understanding through my seemingly endless studies. Without her encouragement and lost weekends, this work would not have been possible.

W. Steven Demmy
Springfield, Ohio
August 2, 1971
## VITA

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 2, 1943</td>
<td>Born - Springfield, Ohio</td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>B.I.E., The Ohio State University</td>
<td>Columbus, Ohio</td>
</tr>
<tr>
<td>1966-1968</td>
<td>Industrial Engineer, Delco Products Division of General Motors Corporation, Dayton, Ohio</td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>M.Sc., The Ohio State University</td>
<td>Columbus, Ohio</td>
</tr>
<tr>
<td>1968-1969</td>
<td>Research Associate, Department of Industrial Engineering, The Ohio State University</td>
<td>Columbus, Ohio</td>
</tr>
</tbody>
</table>

## PUBLICATIONS


FIELDS OF STUDY

Major Field: Industrial Engineering


Studies in Decision Theory. Professor W. T. Morris

Studies in Experimental Design. Professor J. B. Newhart
TABLE OF CONTENTS

ACKNOWLEDGMENTS .......................................................... ii
VITA ...................................................................................... iii
LIST OF TABLES ........................................................................ vii
LIST OF FIGURES ....................................................................... viii

Chapter

I. INTRODUCTION AND LITERATURE REVIEW .................. 1

   Introduction
   Characteristics of Military Supply Systems
   The Federal Budget System
   A Review of the Literature

II. GENERALIZED LAGRANGE MULTIPLIERS ......................... 48

   Introduction
   Theory
   Application to Allocation Problems

III. A MODEL OF A BUDGET-CONSTRAINED INVENTORY SYSTEM 65

   Introduction
   Elements of a Decision Problem
   The Objective Function
   A Model of a Budget Constrained Inventory System
   The General Problem
<table>
<thead>
<tr>
<th>IV. DETERMINISTIC MODELS ..................................</th>
<th>Page 101</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td></td>
</tr>
<tr>
<td>Small-Scale Linear Systems</td>
<td></td>
</tr>
<tr>
<td>Decomposeable Linear Systems</td>
<td></td>
</tr>
<tr>
<td>Generalizations of the Dantzig-Wolfe Algorithm</td>
<td></td>
</tr>
<tr>
<td>Differential Management</td>
<td></td>
</tr>
<tr>
<td>Special Cases</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V. PROBABILISTIC MODELS ..................................</th>
<th>167</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td></td>
</tr>
<tr>
<td>Objectives in Probabilistic Environments</td>
<td></td>
</tr>
<tr>
<td>The Constraint Set</td>
<td></td>
</tr>
<tr>
<td>Solution Procedures</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VI. STEADY-STATE DETERMINISTIC MODELS ....................</th>
<th>187</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td></td>
</tr>
<tr>
<td>A Model with Fill or Kill Demands</td>
<td></td>
</tr>
<tr>
<td>A Rationing Model</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VII. SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH ........</th>
<th>220</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td></td>
</tr>
<tr>
<td>Suggestions for Further Research</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>APPENDIX ..................................................................</th>
<th>230</th>
</tr>
</thead>
</table>

| BIBLIOGRAPHY ................................................................| 246     |
LIST OF TABLES

Table                                      Page

4.1 Dimensions of Problem (3.31)-(3.33)     103

4.2 Computational Experience of Gomory and Dzielinski with the Lot-Size Programming Problem 158
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Information, Material, and Fund Flows in a Typical Inventory System</td>
<td>8</td>
</tr>
<tr>
<td>1.2</td>
<td>A Simplified Military Organization</td>
<td>10</td>
</tr>
<tr>
<td>1.3</td>
<td>Flows of Material and Funds in a Stock Fund Financial Arrangement</td>
<td>13</td>
</tr>
<tr>
<td>2.1</td>
<td>Gaps in Two-Dimensional V-Space</td>
<td>55</td>
</tr>
<tr>
<td>3.1</td>
<td>Sequence of Events within Period i</td>
<td>75</td>
</tr>
<tr>
<td>3.2</td>
<td>Relationship among On-Hand Stock, On-Order Stock, Order Quantity and Shipment Variables</td>
<td>77</td>
</tr>
<tr>
<td>4.1</td>
<td>Network Representation of Subproblem for the Single-Item, Single-Demand Class Case</td>
<td>127</td>
</tr>
<tr>
<td>4.2</td>
<td>The Coefficient Matrix of the Master Problem</td>
<td>143</td>
</tr>
<tr>
<td>5.1</td>
<td>The Probability Densities of g and b for a Feasible Solution to an Absolute Constraint</td>
<td>177</td>
</tr>
<tr>
<td>6.1</td>
<td>Cost per Lost Sale versus the Proportion of Sales Lost</td>
<td>192</td>
</tr>
<tr>
<td>6.2</td>
<td>On-Hand Inventory in the Deterministic Fill-or-Kill Case</td>
<td>193</td>
</tr>
<tr>
<td>6.3</td>
<td>On-Hand Inventory in Rationed Demand Case</td>
<td>205</td>
</tr>
<tr>
<td>6.4</td>
<td>Plot of $F = (f-c)(1-f)^{1/2}$</td>
<td>211</td>
</tr>
</tbody>
</table>
Introduction

Perhaps the most important type of constraint encountered in the management of military supply systems appears in the form of a budgetary limitation on the funds appropriated for the support of different types of activities. A supply system must be operated within whatever budgeted funds are appropriated by Congress and allocated to the system by higher Department of Defense officials. Typically, if a supply system runs short of funds early within a given budget period, no additional funding will be available, and a deterioration of supply support may be anticipated. Conversely, any funds remaining at the end of the period are usually lost to the system, indicating supply support within the period might have been improved through allocation of the lost resources.

Although budget allocations are used to control many real-world inventory systems, relatively little detailed theoretical analysis of decision making in such environments has been performed. In a 1964 review of inventory theory, Hadley and Whitin (1964, p. 152) conclude
that none of the inventory models discussed in the operations research literature up to that time provide an adequate treatment of budgetary constraints. Later in this chapter, we review the results of relevant theoretical investigations that have been conducted since Hadley and Whitin's review. Unfortunately, we conclude that very little is known about the nature of optimum decision-making in the presence of budgetary constraints; in fact, apparently no quantitative studies have been published describing even the operational characteristics of heuristic procedures currently being utilized. As a result, the operations analyst confronted with prescribing policies for the management of an inventory system subject to budgetary constraints is confronted with a number of important questions. Some of these include:

1. What is the nature of inventory decision-making subject to budget constraints? What aspects of this class of problems are susceptible to mathematical analysis?

2. Are computationally feasible methods for determining optimum decisions available? If not, can one compute bounds on the cost of system operation under an optimal policy to use as a check on the near-optimality of an easily computed, but heuristically derived result?

3. Are there any simplified situations or alternate formulations which provide insights into the nature of optimum decision making subject to budget constraints? If so, what are these situations, and what insights can be obtained?

4. What heuristic methods for dealing with budget constraints have been proposed? What are the characteristics of these techniques when employed in an operational environment?
Hadley and Whitin (1963, p. 4) note that there are great differences among existing inventory systems. They differ in size and complexity, in the types of items they carry, in the costs associated with operating the system, in the nature of the stochastic processes associated with the system, and in the nature of the information available to decision-makers at a given point in time. All of these variations may have an important effect upon the type of operating decisions which may be optimal in a given budget-constrained environment. Unfortunately, it is beyond the scope of this work to consider other than a very small subset of the many interesting inventory system structures that may be encountered in practice.

In this paper, we will instead focus our attention on single-location, multi-item military supply systems managed under a revolving fund concept. We are particularly interested in systems in which budgetary constraints on the amount of funds that may be spent on procurement represent the primary form of budgetary limitation, although other forms of budgetary constraints such as cash balance limitations will be considered. Our objective is to provide at least partial answers to all of the questions posed above in the context of such a system.

In pursuing this objective, we will first discuss the qualitative aspects of problems of inventory management subject to budget constraints, and we will review the
approaches to this problem that have been suggested by other researchers. This is the topic of the remainder of this chapter.

As we shall see later in this work, many problems of inventory management subject to budget constraints may be solved by utilizing Everett's (1963) generalized Lagrange multiplier techniques to decompose the original problem into a series of smaller single-item inventory problems. In Chapter II we discuss some of the major characteristics of these techniques. Results presented in this chapter will be used repeatedly in later chapters.

In Chapter III, we formulate the general problem of inventory management subject to budget constraints as a special case of a stochastic constrained dynamic optimization problem. Although computationally feasible methods for obtaining optimal solutions to the general problem do not appear to be available, several important special cases do appear to be economically solvable. In Chapter IV we discuss solution procedures for several important situations in which demand is assumed known, while in Chapter V we discuss solution procedures for situations in which demand is probabilistic. In these chapters we will show that several important special cases of the original problem may be solved by solving a series of smaller single-item problems. A sample problem illustrating the basic computational procedure is summarized in the Appendix.
The results of Chapters IV and V provide techniques for obtaining numerical solutions to the problem of inventory management subject to budget constraints, but they do not provide qualitative insights into the characteristics of an optimal policy. Hence, in Chapter VI we consider two simplified versions of the general problem to obtain insights into the nature of an optimum policy in the presence of a restrictive procurement constraint. First, we consider a multi-item inventory system in which demand is known and constant, and all demands occurring when the system is out of stock are lost. We seek insights into the nature of optimum stocking policies when the procurement constraint is sufficiently tight that at least some lost sales are unavoidable. Second, we consider a multi-item system in which demand is known and constant, but in this case we assume that the procurement constraint is satisfied through a priori rationing decisions. These two models represent extreme cases of procedures used in actual practice for balancing demands and funds in limited funding environments.

Chapter VII summarizes some of the major findings reached in earlier chapters and recommends areas in which further research is needed.

Before we commence our detailed investigation of the quantitative characteristics of problems of military supply management subject to budgetary constraints, it is important
to understand some of the qualitative characteristics of such problems. Hence, in the following sections we will discuss the characteristics of military supply operations managed under a revolving fund concept and the role of budgetary controls in these processes.

Characteristics of Military Supply Systems

Functions and Organization

Problems associated with the maintenance and control of inventories of physical goods are found in almost all types of organizations, including manufacturing firms, retail establishments, agricultural businesses, military forces and even private households. The amount of funds invested in inventories at any one time is immense. For example, investment in inventories by U. S. manufacturing firms exceeded $153 billion in 1969, and the value of inventories managed by Department of Defense revolving funds alone exceeded $12 billion.


Why do organizations carry inventories? Kenneth Arrow (1958, p. 1-14) classifies motives for holding inventories into three major kinds—transaction, precautionary, and speculative motives. These classifications parallel those suggested by Keynes as the motives for holding stocks of cash. The transaction motive results from the fact that even in the case of certainty, it is generally physically impossible or economically unsound to have goods arrive in a given system precisely when demands for them occur. Inventories are therefore held to compensate for this lack of synchronization. The precautionary motive results from an inability to predict demand exactly and the consequent need to maintain some type of safety allowance in order to reduce the costs of shortages. The speculative motive results when changes in prices or costs are anticipated in the future. In certain circumstances, profits may be made by purchasing goods at a low price and holding them until a high price occurs.

Once an organization has decided to maintain stocks of physical goods, a large number of activities may be required to support the flow of goods to and through the inventory system. As illustrated in Figure 1.1, replenishment of system stocks may require the initiation of purchase requests, negotiation of purchase contracts with suppliers, scheduling, production, and delivery activities performed by the supplier, and material-handling activities required to
Figure 1.1. Information, material, and fund flows in a typical inventory system.
receive and store the replenishment shipment. Similarly, filling a customer's order may require checking of stock status records for the availability of parts, packing the order and preparing it for shipment, and final transportation to the customer. In addition to activities required to support the flows of physical goods through the system, many tasks may be required to maintain the information and financial flows associated with the flows of physical goods. Stock status records must be updated and checked for accuracy; suppliers must be paid for their goods; customer accounts must be updated; and financial planning and operating reports must be prepared.

In organizations responsible for the management of hundreds or hundreds of thousands of items, the above tasks may require the coordinated efforts of individuals in several different functional organizations. For example, in the military organization illustrated in Figure 1.2, Supply may be responsible for initiating purchase requests, and for receiving, inspecting, and storing the goods when they arrive. Procurement may be responsible for negotiating all contracts, while the Comptroller may be responsible for paying for the shipment and updating stock status and financial records. Similarly, Supply may acknowledge receipt of a customer requisition, pick the order from warehoused stores, and prepare it for shipment. Transportation may arrange for the shipment of the goods to the customer, while the Comptroller
Figure 1.2. A simplified military organization.
would bill the customer for the shipment and again update stock and financial status records. The Operations, Maintenance, and Civil Engineering organizations might play the role of 'customers', since these groups requisition goods from Supply to support their activities. The names for particular functional groupings may vary from one military service to the next, and the sets of activities performed by a particular functional group may also vary; however, the existence of several specialized organizations to perform particular subsets of the activities associated with the maintenance and control of inventories is a characteristic common to a large number of inventory systems.

**Stock-Funded Supply Systems**

The Secretary of Defense is authorized by Congress under Title 10, United States Code 2208, to "establish working capital funds and to finance the inventories of such stores, supplies, and equipment as he may designate". Such funds are also called 'revolving' or 'rotary' funds since these terms describe the continuing process of expenditure and recovery which characterize the operation of such funds. There are three general types of revolving funds in the Department of Defense -- stock funds, industrial funds, and management funds. Stock funds are concerned with the management of items, while industrial and management funds relate to the repair and/or manufacture of items or to the
management of services, respectively. Although the Department of Defense has employed such funds extensively since 1949, the Navy utilized this financial technique to finance the acquisition and replenishment of goods as early as 1893 (Burkhead, 1956, p. 268).

A stock fund may be created through an initial appropriation, through the capitalization of inventories and associated financial accounts, or through a combination of these activities. Monies from this fund are then used to place orders for goods from private industry or other governmental agencies. Such purchases decrease the cash assets of the fund, and increase the value of its physical inventories. The value of the fund might be decreased by inventory losses, by paying for certain services (e.g., transportation of goods) from the cash account of the fund or by issues of material to customers. When orders for stock are received from user agencies, or 'customers', the physical inventories are decreased, and the cost of the item, plus a 'surcharge', usually computed as a predetermined percentage of the item cost, is then billed to the customer's account. When the bill is paid, the cash in the stock fund is increased, and, assuming the surcharge is set at a level that just compensates for expenses, the total value of the fund is back where it started. Thus, the stock fund has "revolved" by changing its assets from cash to physical inventory to accounts payable and back to cash. Figure 1.3 illustrates
Figure 1.3. Flows of material and funds in a Stock Fund financial arrangement.
this cycle of activity.

As noted above, the customer may pay a surcharge to compensate for certain losses borne by the stock fund. If surcharge revenues are insufficient to compensate for all such losses, the stock fund may be partially supported by an appropriation -- often called a "subsidy" -- if it is desired to maintain the fund at its current level of investment.

Generally, not all expenses incurred in providing supply support are paid for from the stock fund account. For example, in the Air Force Stock Fund, monies paid to suppliers, inventory losses, and certain transportation costs are drawn from the stock fund account, but the costs of such activities as preparation of purchase requests, negotiation of contracts, and receiving, inspection, and storage activities are borne by other accounts.

We have discussed some of the general structural characteristics of military supply systems managed under a stock fund concept. In the next section, we will sketch some of the major features of the Federal budget system and its impact upon stock fund operations.
The Federal Budget System

Phases. The budget cycle of the United States Government consists of four identifiable phases: (1) executive formulation and submission; (2) Congressional authorization and appropriation; (3) budget execution and control; and (4) audit. Each of these phases interrelates and overlaps with the others.

Executive Formulation and Submission. The President's transmission of budget proposals to Congress each January culminates a process of planning and review which generally begins some 14-18 months earlier. For example, formulation of the 1970 budget, which covers the fiscal year beginning 1 July 1969 and ending 30 June 1970, began in the spring of 1968.

Each spring, every federal agency evaluates its programs, identifies policy issues, and makes budgetary projections, giving attention both to important modifications and innovations in its programs, and to alternate long-range program plans. The detailed documents produced during this process are important in two major respects: not

---

3 Much of the material in this section is from The Budget in Brief, Fiscal Year 1970, U.S. Bureau of the Budget, p. 57-60, and from The Budget of the United States Government, Fiscal Year 1970, p. 174-177. For more extensive discussions of the Federal Budgeting System, the reader is referred to Burkhead (1956) and to Wildavsky (1964).
only do they serve as supporting evidence of an agency's need for funds in the forthcoming period, but these documents often become a primary control tool in guiding an agency's actions when the budget execution phase is reached. In the context of stock fund budgetary projections, the formulation process may require careful consideration of major changes in force structures, operating policies, maintenance programs, or other significant policy changes by stock fund customers that may impact on their requirements for supply support. Changes in logistics support systems and policies of the stock fund itself must also be reviewed. Objectives for end-of-year inventory status must be established. Finally, detailed projections of stock fund operations in the coming fiscal year may be prepared. The budgetary documents that result from this process may consist of sales forecasts, projected on-hand and on-order inventories, and statements of estimated financial flows and of the projected end-of-period financial condition of the fund. These documents may be quite similar to the products of the budgetary planning process in commercial inventory operations. Finally, estimates of stock fund budgetary projections are consolidated with budgetary estimates of other agencies of the Department of Defense.

After review in each agency and by the Bureau of the Budget, preliminary plans are presented to the President for his consideration. At about this time, the President also
receives projections of the economic outlook and revenue estimates prepared by the Treasury Department, the Council of Economic Advisers, and the Bureau of the Budget.

After review of both sets of projections, the President establishes general budget and fiscal policy guidelines for the fiscal year which will begin about twelve months later. Tentative policy determinations and planning targets are then given to the agencies as the guidelines for the final preparation of their budgets. In turn, agency heads may provide guidelines for those subordinates who prepare portions of the agency budget.

Throughout the summer the departments and agencies conduct extensive examinations of their budgets in the light of any ceilings or policy directives that have been transmitted to them. Needless to say, a great amount of coordination between the management heirachies of an agency occurs during this phase to bring the composite agency budget into line with the overall policy objectives.

Throughout the fall and early winter, individual agency budgets are finalized and reviewed in detail by the Bureau of the Budget. The documents are then presented to the President for decision. Last minute changes may then be incorporated before the budget is submitted to Congress in January.
Classification of Expenditures. Both the tasks required to formulate the budget and the impact of budget documents as a control tool during the execution phase of the budget cycle are strongly influenced by the particular expenditure categories used to display budgetary requests in the President's budget.

Burkhead (1956, p. 110-132) notes that there are three basic schemes for classifying expenditures arising from governmental activities; these are program, organizational, and objects-oriented schemes. A program classification scheme classifies expenditures by the type of activity to be performed; an organizational scheme arrays expenditures by the administrative unit responsible for the expenditures, while a classification system by objects arrays expenditures to reflect the nature of items or services purchased, regardless of the programs for which they are used. In classification systems involving many levels of detail, any one or a combination of these schemes might be employed at any given level within a particular classification system.

In the Appendix to the 1969 Budget, the first breakout of budgetary requests to support the activities of the Department of Defense (DOD) are based on a combination of the program and organization schemes. For example, the 1969 Budget proposes an appropriation of $6.7 billion under the
heading of Operations and Maintenance, Air Force. The basic functional classifications associated with military appropriation requests in the 1969 Budget are:

1. Military Personnel
2. Reserve Military Personnel
3. Operations and Maintenance
4. Procurement
5. Research, Development, Testing, and Evaluation
6. Military Construction
7. Family Housing
8. Civil Defense
9. Special Foreign Currency Program
10. Revolving and Management Funds

The major DOD organizational categories in the 1969 Budget consist of the Army, Navy, Marine Corps, and Air Force; corresponding National Guard units; and a final category for agencies that operate directly under the control of the Secretary of Defense.

Within a given appropriation, the Appendix to the 1969 Budget presents both program and objects-oriented classifications of requested appropriations. For example, the appropriation request for Operations and Maintenance, Air Force is supported by schedules with proposed expenditures broken out according to the following categories:
Activities Classifications

Aircraft fuel and oil
Logistical Support
Training Support
Operational Support
Medical Support
Servicewide Support
Contingencies

Objects Classifications

Personnel Compensation
Personnel benefits
Benefits for former personnel
Travel and Transportation for persons
Transportation of things
Rent, communications, and utilities
Printing and reproduction
Other services
Services of other agencies
Supplies and materials
Equipment
Grants, subsidies, and contributions
Insurance claims and indemnities
Quarters and subsistence charges
These schedules display obligations by expense category for three different fiscal years. For example, schedules in the 1969 budget displayed enacted or proposed appropriations for fiscal years 1967, 1968, and 1969. These data on what was done in the past and proposed for the future provide the starting point for the next phase of the budget cycle, Congressional authorization and appropriation.

**Congressional Authorization and Appropriation.**

Congressional review begins when the President sends his budget to Congress. The Congress can change programs, delete programs, or add programs not requested by the President. It can increase or decrease appropriations over those requested in the President's budget. It also legislates the means of raising revenue.

Under traditional procedures, two pieces of legislation are required for Congressional approval of budget outlays. Congress first enacts legislation which authorizes an agency to carry out a particular program. Many programs are authorized for a specified number of years, or even indefinitely; other programs, such as foreign aid, atomic energy, and space exploration, require annual authorization legislation. Most stock funds in the Department of Defense have permanent authorizations, i.e., budget authority becomes available from time to time without further action by Congress.
The granting of budget authority -- which permits an agency to enter into obligations requiring either immediate or future payment of money -- usually is a separate, subsequent action. Most budget authority is enacted in the form of appropriations, but smaller amounts of budget authority may be granted in the form of contract authority and authority to spend debt receipts. Contract authority permits an agency to incur obligations, but requires an appropriation "to liquidate" in order to permit payment of obligations.

Requests for appropriation and for changes in revenue laws are considered first in the House of Representatives. The Ways and Means Committee reviews all proposed revenue measures; the Appropriations Committee, through its thirteen subcommittees, studies the proposals for appropriations and examines each agency's performance in detail. Each committee then recommends the action to be taken by the House of Representatives.

As parts of the budget are approved by the House, the bills are forwarded to the Senate, where a similar process is followed. In case of disagreement between the two houses of Congress, a conference committee, consisting of members from both the House and Senate, meets to resolve the issues. The conference report is then submitted to both Houses of Congress for approval, and is finally transmitted
to the President, in the form of an appropriation or tax bill, for his approval or veto.

**Budget Execution and Control.** Once approved, the budget becomes the basis for the program operations of each agency during the fiscal year.

Central control over most of the budget authority made available to the executive branch is maintained through a system of "apportioning" the authority. Under the law, the Director of the Bureau of the Budget must distribute or apportion appropriations and other budget authority to each agency by time periods (usually quarterly), or by activities. Obligations, i.e., arrangements requiring either immediate or future payment of money, may not be incurred in excess of the amounts apportioned. The objective of the apportionment system is to plan the effective and orderly use of available authority and to prevent the need for requesting additional or supplemental authority.

The second step in the process of budget execution is allotment, which extends budgetary authority to administrative units within agencies and departments. Allotment consists in the extension of obligational authority to subdivisions of the agency, usually on a monthly or quarterly basis. Although obligations resulting from the allotment process must be consistent with the apportionment from the Bureau of the Budget the allotment process is internal and
the Bureau has no supervision over its terms or procedures. Customarily, the allotment process is administered by the budget office of the agency, under the authority of the agency head. An allotment may be stated in terms of objects of expenditure or activities within administrative units, and may state conditions for incurring obligations that are far more specific than the wording of enabling legislation. For example, three of the major administrative conditions placed upon obligations by divisions of the Air Force Stock Fund are as follows: (Trapp, 1969, p. 4-55),

1. No over-obligation of any funds limitation is permitted -- under any circumstances.

2. Obligations that would result in expenditure upon delivery of materiel of an amount that would exceed the cash estimated to be available in the cash account of the fund are not permitted. In other words, the Stock Fund must be able to pay its bills when they become due.

3. Actual expenditures (including accounts payable) in excess of available cash are prohibited (no bad checks allowed).

Allotment systems are usually supplemented with detailed manuals which elaborate on policies and procedures that are to guide expenditure decisions. These manuals further define the conditions under which funds may be expended and the flexibility delegated to funds managers for shifting allocated authorizations between expenditure categories.

After appropriation bills have been signed by the President, the Treasurer of the United States establishes
a credit equal to the aggregate amount of the appropriation. These credits are made available to disbursing officers through accounts maintained in Federal Reserve Banks. After such accounts are established, agencies may incur obligations, i.e., enter into contracts for goods and services, providing, of course, that such obligations are consistent with the terms of the appropriation.

As creditors present bills for payment, the appropriate administrative officers in each agency certify the payment and direct disbursing officers to issue checks to liquidate the obligation. The certifying officers who sign vouchers for payment are liable for performance in accordance with applicable statutes and regulations, and are responsible for the adequacy of balances available to meet the payment. Certifying officers are also responsible to adhere to limits and policies established by the apportionment and allotment processes. In many instances, non-compliance with appropriate limits and policies may result in legal penalties against the responsible official.

General responsibility for compliance with the legalities of appropriations is assigned to the agencies to whom appropriations are made. The agencies operate their own review and control systems to assure that the obligations they incur and the resulting outlays are in accordance with the provisions of the authorizing and appropriating legislation. These review and control systems consist of accounting
records to show appropriations, obligations, and disbursements, as well as other types of reports that indicate that each subdivision of the agency is attaining its program objectives. As noted above, the detailed documents developed during the budget formulation phase (with appropriate modifications to reflect Congressional changes and restrictions) often serve as a primary control device during the execution of an agency's program.

In the context of a stock fund operation, reports required by the review and control system may consist of displays of planned versus observed on-hand and on-order inventories, actual sales versus projected sales, and revised projections of end-of-period financial position, as well as summaries of obligations, disbursements, and other financial information. Whenever these reports indicate significant deviations from previously approved programs, actions may be initiated to identify the cause of the deviation and, if necessary, to adjust operations in an appropriate manner.

In addition to the review and control systems within the agencies, the Bureau of the Budget also reviews substantive and financial reports and keeps abreast of an agency's progress in attaining its program objectives. If developments indicate that an agency will not require all authority made available to it, "reserves" may be established by the Bureau to withhold amounts not needed. Such reserves may be released at a later date, if necessary, but only for
purposes of the appropriation. On the other hand, changes in laws or other factors may indicate the need for more authority, and supplemental requests may have to be made of the Congress.

**Audit.** As noted above, general responsibility for assuring that budget execution is conducted within the framework of legal authority is the responsibility of individual agencies. The Congress, however, maintains an independent check through the General Accounting Office (GAO), headed by the Comptroller General.

The General Accounting Office conducts after-the-fact audits of the manner in which Government agencies are discharging their financial responsibilities. The GAO has the responsibility for the settlement of accounts -- for "closing the books" of administrative officers responsible for the custody of public funds. The GAO may also conduct special investigations of the financial affairs and administration of departments and agencies. The Comptroller General summarizes these and other activities of the GAO in an annual report to the Congress, thus completing the complicated chain of events that began several years before.

**Comments on the Federal Budget System**

The above discussion of the Federal Budget Cycle may imply a certain precision regarding the timing and procedures
involved in the budgetary process; in truth, this is not quite the case. Although the general characteristics of the budgeting cycle may possess substantial stability, the budgetary cycle of the U. S. Government is a dynamic, evolutionary process. Procedures change as world events, political moods, and the personnel who prepare, approve, and administer budget programs change. Similarly, major (and minor) modifications in budgetary requests or in appropriations may be instituted at any time during the budget cycle to reflect these changes.

At least two major categories of causes may result in significant difference between the budgetary requests prepared during the formulation phase of the budget cycle and the final budgetary authorizations that limit and sustain the execution of an agency's program.

First, a level of uncertainty may exist in forecasts of future needs. The long time span between the initial preparation of budget estimates and the execution of associated programs makes the accurate estimation of needs that will arise during the forthcoming fiscal period quite difficult, if not impossible. World events such as a Pueblo crisis or the outbreak of Mideastern hostilities may have profound effects upon requirements for both activities and supplies; technological difficulties with hardware systems may arise that require major program modifications; or policy changes by administrators may cause major shifts in resource
requirements.

Another major source of forecasting uncertainty arises from the political nature of the budgeting process. Since budgeting impacts upon the actions and aspirations of many men, the budgeting process is as much a political and psychological exercise as it is an exercise in the rather mechanical operations of estimation and aggregation required to produce budgetary displays. In a study of the politics of the budgeting process, Waldavsky (1964, p. v) notes:

"... Budgeting deals with the purposes of men. How can they be moved to cooperate? How can their conflicts be resolved? How can they find ways of dealing effectively with recalcitrant problems. Serving diverse purposes, a budget can be many things: a political act, a plan of work, a prediction, a source of enlightenment, a means of obfuscation, a mechanism of control, an escape from restrictions, a means to action, a brake on progress, even a prayer that the powers that be will deal gently with the best aspirations of fallible men."

The above comments point out that the operation of a budgeting system presents important psychological problems associated with the inspiration, motivation, persuasion, and coordination of men. Unfortunately, the understanding of the forces that govern human behavior is even less developed that the understanding of the technological characteristics of hardware systems; hence, forecasting problems are made even more difficult by the uncertainties of these political processes.
A second major cause for differences between initial budgetary requests and eventual budgetary authorizations may be attributed to what H. A. Simon (1947, p. 3) refers to as "limited rationality". In the context of the budgetary process the behavior of individual decision makers (Congressmen, Bureau of the Budget, and agency personnel) is severely constrained by the size and complexity of decisions to be made; the limited amount of time, information, and facilities (including both mental and mechanical facilities) available to assist in decision making; and the lack of a clearly-stated, widely-accepted methodology for choosing among alternatives in some optimal sense. You might sympathize with Congressman Laird's plaintive cry, "A lot of things go on in this subcommittee that I cannot understand". In the face of such far-reaching difficulties, decision makers are forced to rely on simple, often routine, methods of choice. For example, decision makers may rely on rule of thumb methods (the House Appropriations Committee may cut items in the President's budget by a fixed percentage) or upon interpersonal relationships in reaching budgetary decisions. Another simplifying technique sometimes employed by decision makers calls for directing one's observations to the responsible administrative officials rather than to the details of the program being managed. For example, a senior

---

Congressman reported that he followed an administrator's testimony looking for "strain in voice or manner", "covert glances", and other such indications, and later followed them up probing for weaknesses (Marvick, 1952, p. 297).

Stromberg (1970, p. 4) observes that another major behavioral characteristic of the budgetary process is the impetus for maintaining stable, predictable relationships among participants in the process. Stromberg points out such tendencies as cutting/padding biases in which figures in the President's budget may be "padded" above what is actually desired because the House is expected to cut the request.

To complete the circle, the House may be expected to cut because the President's budget is expected to be padded. Another stabilizing tendency, Stromberg points out, is the evolution of institutional roles whereby participants in the budgetary process tend to act on the budget only in ways that are well-understood by other participants and that have been legitimized through years of consistent, repeated behavior. For example, the Senate has a reputation for acting as an appeals court on budgetary decisions by the House, generally limiting itself to restoring some portion of the proposed appropriations for those items cut by the House.

In view of the uncertainties and irrationalities of certain portions of the budgetary process, one might conjecture that the ultimate question relating to inventory management subject to budgetary constraints concerns the selection
of strategies for preparing and supporting budgetary requests so that the appropriations that are eventually granted represent an allocation of the national resources that is in some sense 'optimal'. The answer to this question involves much more than considerations of inventory theory; rather, it requires a statement of what activities the government ought to perform at any particular time, since determination of an 'optimal' allocation of resources requires the simultaneous consideration of all alternate uses for those resources. Hence, providing an answer to the above question would be utopian in the fullest sense of the word, since its accomplishment and acceptance would end the conflict over the government's role in society.

A second but simpler question along these same lines might be stated as follows: Suppose an inventory manager has formulated a specific set of objectives, such as the maintenance of a given level of customer service and the establishment of a given end-of-period investment in stocks. What strategy should be followed so that there is a reasonably good chance that sufficient resources will be appropriated so that these objectives might be achieved?

As interesting and important as this second question might be, we leave its consideration to other researchers. Rather, in this work we will focus our attention on the problems of an inventory manager who is presented with a given budget and instructed to abide by it -- to do the best he
can with what he has, consistent with legal and administrative limitations and guidelines established by Congress and other higher authorities. Of course, the term "best" needs further clarification, but we will defer that discussion to later chapters.

At this point, it is appropriate to ask what success other researchers have had in dealing with this latter problem. This is the subject of the next section.

A Review of Literature

Background

Although budget allocations are used to control many real-world inventory systems, relatively little detailed theoretical analysis of decision making in such environments has been performed. In a 1964 review of inventory theory, Hadley and Whitin (1964, p. 152) conclude that none of the inventory models discussed in the operations research literature up to that time provide an adequate treatment of budgetary constraints. First, they note that budgets are often set in terms of actual dollars to be spent, whereas the costs included in inventory control models are often not of the out-of-pocket type. For example, inventory carrying charges include elements of opportunity costs such as the rate of

5 Hadley and Whitin present a detailed discussion of the difficulties of dealing with budget constraints in Reference 50.
return on alternate uses of capital, and costs of placing orders are not typically charged against the procurement budget. Second, they note that although several steady-state models have been developed which deal with constraints on the average annual expenditures over a very long time horizon, attempts to apply such a model by setting the constraint on the average annual expenditures equal to the absolute number of dollars provided by the budget must eventually be unsuccessful. Since demand is a random variable, the number of dollars required to keep the system in operation is also a random variable. In those years in which procurements turn out to be less than the budgeted amount, the money would not be spent and would be lost forever since budget money cannot be carried over from one year to the next. The losses from budget money in certain years might be serious enough to cause the system to run downhill and destroy the steady-state assumption. Another realistic problem is that if the budget money is not completely spent in a given year there is a very good chance that the budget will be cut the following year. Finally, Hadley and Whitin point out that an absolute procurement budget can become a serious problem during the transition from the existing system to a new type of steady-state system operation, particularly if higher levels of stocks are needed in the new system. During such a transition phase, the level of dollar expenditures may be expected to deviate significantly from
an average steady-state value.

**Systems Subject to Absolute Constraints**

Since Hadley and Whitin's review, three authors have presented results of theoretical investigations of inventory systems subject to absolute constraints. These results are discussed below.

Iglehart (1965) has considered a stationary two-product inventory model in which an inventory of capital is maintained to provide production capability for a second product in the inventory system. Capital is allowed to depreciate stochastically from period to period, and random demands are placed on the inventory of the product produced. Unsatisfied demands are backordered. Optimal dynamic policies are obtained for the n-period problem, where the optimal policy in period i is described by three non-increasing functions of two variables. As might be expected, substantial computation is required to obtain numerical results for \( n \geq 1 \).

Evans (1967) considers a two-product model with backorders and a constraint on the amount ordered in the form of a limitation on some resource, e.g., labor, raw materials, capital, etc. Assuming a linear purchase cost, and non-negative partial derivatives of the expected shortage and holding cost function, Evans shows the optimal policy in period i is
characterized by a point, two functions of a single variable, and a third function of a two-state vector. These functions define seven regions on a graph whose ordinate and abscissa are the on-hand inventories at the beginning of period \( i \) for products one and two respectively. Within a given region, a particular set of ordering actions is optimal; hence, determining the optimal policy is straightforward once the regions have been defined. Unfortunately, the computation required to define the regions is appreciable.

Finally, Lalchandani (1965) considers a single-item inventory problem in which the amount ordered in each period is constrained to be less than or equal to a fixed constant.

**Other Approaches**

Although the above studies appear to be the only theoretical investigations that have directly considered absolute budget constraints, several other works have considered related problems of inventory management subject to financial limitations. For ease of exposition, we have roughly categorized the results of these works into three classes: linear models, non-linear models, and heuristic methods.

**Linear Models.** One of the earliest and most active areas of analytical research into problems of inventory management focus on the now classic "warehouse model" and its
various generalizations. This problem was initially formu-
lated by A. S. Cahn(1948), and was considered as "solved" 
when Cahn succeeded as identifying it as a linear program-
ming problem. Subsequently, Dantzig(1951), made a similar 
observation. Later, the problem was reformulated and ex-
tended and its properties subjected to intensive study for 
the purpose of evolving more efficient methods of solution.

Perhaps the simplest kind of warehousing model covers 
the planning of purchases, sales, and storage patterns for 
a single commodity under known prices, where the demands d_j, 
j=1,2,...,N for the commodity over the N periods in the 
planning horizon are assumed known. Let

\[ H = \text{the fixed warehouse capacity} \]
\[ I_0 = \text{the initial stock of inventory in the} \]
\[ \text{warehouse} \]
\[ x_j = \text{the amount to be purchased in period } j \]
\[ y_j = \text{the amount to be sold in period } j \]
\[ s_j = \text{sales price per unit prevailing in} \]
\[ \text{period } j \]
\[ c_j = \text{the purchase price per unit} \]

The problem may be stated as follows: Given H, I_0, 
and s_j, c_j, and d_j, j=1,2,...,N; determine values of x_j and 
y_j so as to maximize the profits of the firm over the plan-
ning horizon without exceeding warehouse capacity.
Symbolically, this problem may be written as:

$$\max_{x_j, y_j} Z = \sum_{j=1}^{N} s_j y_j - \sum_{j=1}^{N} c_j x_j$$  \hspace{1cm} (1.1)$$

subject to the $i=1,2,\ldots,N$ warehouse capacity constraints

$$\sum_{j=1}^{i} x_j - \sum_{j=1}^{i} y_j \leq H - I_0$$  \hspace{1cm} (1.2)$$

and to the $N$ selling constraints

$$-\sum_{j=1}^{i-1} x_j + \sum_{j=1}^{i} y_j \leq I_0$$  \hspace{1cm} (1.3)$$

Also, the variables $x_j$, $y_j$, $j=1,2,\ldots,N$ are constrained to be non-negative, and $y_j$ cannot exceed $d_j$.

The above formulation assumes that the firm can sell only from inventory on-hand at the start of any period and that delivery times for purchases are negligible; hence, cumulative sales through period $i$ must satisfy

$$\sum_{j=1}^{i} y_j \leq I_0 + \sum_{j=1}^{i-1} x_j.$$  Equation 1.3 results from re-arrangement of this expression.

Several generalizations of the problem (1.1)-(1.3) have been developed. First, Charnes and Cooper (1955) extended the original warehousing problem to multi-product and
multi-location situations. In a latter work, Charnes, Cooper, and Miller (1959) showed that the original model could be further extended by adjoining cash and liquidity constraints; thus supplying a new approach for unifying financial planning and the costing of funds.

Although it is relatively easy to formulate quite general warehousing models in a linear programming framework, rather severe difficulties are usually encountered when one attempts to perform the computations required to solve the problem. In particular, the number of constraints prohibits the use of the straightforward simplex method except for very small problems. Hence, a number of authors including Dantzig (1959), Prager (1957), Bellman (1956), Dreyfus (1957), Karlin (1959), Ford and Fulkerson (1962), and Charnes, Cooper, and Miller (1959) have developed algorithms that exploit the structural characteristics of special cases of the general problem.

Perhaps the most comprehensive treatment to date of the warehousing model and its extension is provided by M. R. Rao (1968) in his Ph.D. dissertation. In this work, Rao considers several extensions to the deterministic multi-commodity warehousing model with cash liquidity constraints; some of these extensions include cases in which there are:

1. Time lags in payments and receipts.
2. Time lags in deliveries of purchased and sold goods.
3. Inventory carrying cost proportional to the value of inventory.

4. Varying prices within each time period.

5. Several inputs which are mixed together in fixed proportions to obtain any one of several outputs (but different for each output).

Rao then shows that the above problems may be solved by applying a column-generation scheme based on the Dantzig-Wolfe decomposition principle. This approach requires the solution of several shortest-route subproblems to determine the entering column vector for the standard-simplex pivot step. Rao also discusses computational experience with the procedure; these results indicate the column generation scheme is more computationally efficient than the straightforward linear programming approach. Perhaps even more important, this approach substantially reduces the size of the working basis, thus making it possible to solve problems that were previously too large to fit even modern computing machines using standard simplex formulations.

Rao also considers a multi-commodity warehousing model with cash liquidity constraints in which the available warehouse capacity and/or minimum liquidity requirements are random variables. Deterministic equivalents of this problem are obtained, and solution procedures based on the Dantzig-Wolfe decomposition procedure are provided.

An important deficiency of the warehousing models discussed above is that they ignore the fixed-charge nature
of order processing costs. Such costs are generally recognized as a significant component of the costs of operating an inventory system. Unfortunately, recognition of such a cost destroys the linearity of the criterion function, thus preventing the straightforward application of linear programming algorithms. Fortunately, the difficulty may be resolved by utilizing some results due to Everett (1963) and an appropriate search routine. Proof of this assertion is presented in Chapter IV.

Non-linear Models. Other than the linear programming formulations discussed above, perhaps the earliest treatment of an inventory problem subject to a budget constraint is presented in a 1959 report by Stanford Research Institute (SRI). In the work, SRI presents an approach to handling a budget constraint by minimizing an expression for expected total costs (procurement costs plus inventory carrying charges plus stockout costs), and of varying the stockout costs so as to bring the average holding costs plus average procurement costs to some given level. This is accomplished through the application of a common multiplicative factor, or shortage parameter, applied to the stockout costs. From the theory of Lagrange multipliers, a mathematically equivalent problem is to minimize the ordering costs plus inventory carrying costs subject to a constraint on stockouts. Hadley and Whitin (1961) note, however, that there is little relation
between this procedure and realistic budget constraints, for several reasons: (1) neither carrying charges nor fixed order-processing costs are included in the procurement budget, (2) actual budgets are absolute, not average, and actual costs in any given year resulting from this method may vary over a wide range, and (3) the multiplicative parameter which is required to match expected expenditures to a given budget level may bear little relationship to the actual costs of shortage. A discussion of how a very restrictive budget leads to ridiculously low values for the implied shortage costs in a real-world application of this method is discussed by Hadley and Whitin (1963, p. 402).

Howard (1967), Presutti (1970), and Brown (1967), discuss procedures that are quite similar to the second SRI approach discussed above, although they differ in the form of distribution used to describe demands and in the methods employed to estimate expected expenditures during the budget period. Howard shows how to compute the expected expenditures in the next budget period provided the system was in steady-state operation and was not constrained by the budget in the preceding period. Presutti suggests that the expected costs may be determined by simulation of a sample of items in the inventory. Brown (1967) presents details for a simulation procedure that might be used for this purpose. Unfortunately, none of these authors attempts to evaluate the potential variability associated with their order rules
nor do they discuss what actions are needed if, at some time during the budget period, more or less funds have been expended than were originally planned.

The works of Brown (1967), Sherbrooke (1966), and Lu and Brooks (1968) contain another approach to decision-making subject to a budget constraint. These authors treat the problem of allocating a given safety stock budget as a single-period allocation problem of the following form:

Find a set of decisions \( I = \{I_1, \ldots, I_n\} \) which optimize the performance measure \( P(I) \), where

\[
P(I) = \sum_i P_i(I_i)
\]

subject to the budget constraint

\[
\sum_i R_i(I_i) \leq B
\]

where \( P_i(I_i) \) is a steady-state measure of supply performance (e.g., backorders, fill rate) associated with a given set of decisions related to item \( i \), \( I_i \) is a vector representing such decisions, \( R_i(I_i) \) is the amount of resources expended on item \( i \) as a result, and \( B \) is the budget limitation on safety stock investment. Lu's model considers an additional constraint, the second being a limitation on the average annual orders processed by the supply organization. These problems may be solved by solving an associated Lagrangian problem that is separable by item, and using a procedure to
search over the space of Lagrange multipliers. These methods would be quite satisfactory if decisions related to budget expenditure were to be made at a single point in time, or if demand could be forecast deterministically. However, since demand is stochastic in most real-world systems, it is usually desirable to modify initial planning decisions as demand data is accumulated within the budget period. If such allocations are achieved through repeated application of the above allocation procedure, however, the performance measures used in these methods are invalid since the required steady-state assumptions are violated. Another problem associated with repeated application of these methods is to determine how much of the total budget should be allocated at each review.

Before leaving this section, we would like to point out one serious deficiency of all the models discussed above. All of these models assume that sufficient funding will be available to satisfy (eventually) all requests placed upon the supply system; none assume that customers may have to be "turned away" because of the limited availability of procurement funds. In practice, this assumption is often violated. Since time immemorial, managers of resources have been bedeviled by one persistent problem: namely, needs often exceed the resources available to meet those needs. Hence, it would seem that any procedure proposed for implementation in real-world systems should be capable of handling this
Heuristic Methods. To the author's knowledge, no other theoretical investigations of inventory decision-making in view of an absolute budget constraint have been published. Hence, lacking a theoretical basis to guide decision-making, inventory managers have been forced to develop heuristic methods or rules-of-thumb to achieve acceptable (as opposed to 'optimum') supply performance within a given budget limitation. Of course, heuristic methods may lead to operational results that are quite useful, particularly if such methods result in improved system performance when compared with previously employed methods. As Lee(1966, p. 87) points out:

"An operational solution ... is one that is discriminating enough to lead to a right decision. It may not be based on an awe-inspiring model. It may be derived from some obviously untrue assertions ... The important thing in operational research is not to obtain solutions in the mathematician's sense, but operational solutions -- guides to valid, executive decisions."

Unfortunately, it appears that no studies have been published on the operational characteristics of any of the methods discussed above or of any of the heuristic procedures used in real-world systems.

Summary

Although a number of studies have considered special cases or interpretations of inventory problems in
budget-constrained environments, all contain important deficiencies that limit the applicability of their results to problems that arise in a real-world environment. First, the theoretical results of Iglehart and of Evans in studying two-product systems provide no guidance to inventory managers responsible for thousands of items. Second, linear warehousing models neglect the set-up costs which represent a major segment of operating costs in most inventory systems. Third, studies of inventory systems with constraints on the expected value of expenditures neglect some of the major operational effects of absolute budget limitations. Fourth, authors that treat the problem as an allocation decision to be made at a single point in time ignore the value of information that becomes available within the budget period. Fifth, although some of the methods discussed above may provide satisfactory operational results, it appears that no studies in this area have been published.

In addition to the above comments, most of the observations made by Hadley and Whitin concerning the inadequacies of models of budget-constrained inventory systems available in 1964 are still valid; namely,

1. inventory models generally do not describe the out-of-pocket nature of operating costs.

2. budget constraints are often absolute, and steady-state models with constraints on average annual expenditures do not describe fundamental effects associated with such constraints.
3. adequate methods for dealing with transients have not yet been developed.

In conclusion, it appears that the theoretical understanding of the nature of optimum inventory decision-making in a budget-constrained environment is still at a rather primitive stage. Although much effort has been devoted to such problems, many important questions remain unanswered. In Chapter III, we will explore the characteristics of this problem that have made it difficult to treat. To provide a proper perspective for this discussion, however, we will first discuss several mathematical concepts that will be required repeatedly during later stages of this work. This is the topic of the next chapter.
CHAPTER II  
GENERALIZED LAGRANGE MULTIPLIERS

Introduction

Mathematical models of decision processes require the explicit statement of all aspects of the decision problem: the variables that are involved, the interrelationships among these variables, and the criteria to be employed in selecting a course of action. The criteria for choice must be expressed as a single-valued function, alternately called the objective function, the criterion function, the return function, or the measure of effectiveness, which expresses the value or utility of a particular course of action in the context of the model. In many problems of military supply system management, however, it may be extremely difficult if not impossible to develop such a function. In this case, the best that can be done using analytic approaches is to identify the set of actions that are undominated by any other action, leaving to implicit judgment processes the choice of a particular member of the undominated set.

In this chapter, we will discuss how Everett's generalized Lagrange multiplier techniques might be employed to identify at least a partial set of undominated actions. Applications of these techniques to constrained resource allocation problems will also be discussed. These techniques will be employed repeatedly in later chapters of
this work to develop efficient computational approaches to problems of inventory management in budget-constrained environments.

Theory

Undominated Actions

Suppose a decision maker has a set of objectives or goals that he wishes to achieve. Let $V_n(Q_i)$ denote the measure of accomplishment toward the nth goal associated with the set of actions $Q_i$, where a higher value of $V_n(Q_i)$ denotes a more desirable situation. Then if the actions $Q_i$ and $Q_k$ are such that

$$V_n(Q_i) \geq V_n(Q_k)$$

for all $n$, with strict inequality holding for at least one $n$, then we say $Q_i$ dominates $Q_k$. In words, we say an action $Q_i$ dominates an action $Q_k$ in the multi-goal case if the action $Q_i$ produces performance toward each goal that is greater than or equal to that produced by $Q_k$. Actions that are undominated by any other action are called efficient actions.
Generating Undominated Actions

Everett (1963) has observed that at least a partial set of undominated actions may be identified by solving a series of optimization problems. Specifically, Everett observed that for any arbitrarily selected set of non-negative parameters \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N) \), called "multipliers", the particular action \( Q^*(\lambda) \) that maximizes the Lagrangian function \( L(Q, \lambda) \), i.e., that solves

\[
\text{maximize } L(Q, \lambda) = \sum_{n=1}^{N} \lambda_n V_n(Q)
\]

(2.1)

is undominated, where \( \mathcal{F} \) denotes the set of feasible actions. The proof of this observation is by contradiction. Suppose there is some action \( \overline{Q} \) that dominates \( Q^*(\lambda) \), i.e., \( \overline{Q} \) is such that \( V_n(\overline{Q}) = V_n(Q^*(\lambda)) \) for all \( n \), with strict inequality holding for at least one \( n \). But since the \( \lambda \) parameters are non-negative, this implies that

\[
\sum_n \lambda_n V_n(\overline{Q}) = \sum \lambda_n V_n(Q^*(\lambda)) \quad (2.2)
\]

which is impossible by the definition of \( Q^*(\lambda) \). Thus, \( Q^*(\lambda) \) must be undominated.

Hence, since the optimization (2.1) produces an undominated solution for a given set of multipliers \( \lambda \), one method of generating a partial set of undominated actions is to solve a series of optimization problems of the form (2.1)
for different sets of \( \lambda \). Unfortunately, there may be certain undominated actions that cannot be produced in this manner. The reasons for this will be discussed later in this chapter.

**Cell Problems**

An important subclass of the general problem (2.1) in which undominated actions may be particularly easy to identify occurs when the overall Lagrangian problem (2.1) is separable, i.e., when the problem (2.1) consists of a number, \( M \), of independent areas or cells in which actions may be taken, and for which the overall value measurement for the total problem is simply the sum of the values that occur from each independent area or cell.

In this case, the Lagrangian function (2.1) takes the form:

\[
\text{maximize } \sum_{n=1}^{N} \lambda_n \sum_{m=1}^{M} V_{nm}(Q_{im})
\]

where \( \Gamma_m \) denotes the subset of actions associated with cell \( m \) and \( V_{nm}(Q_{im}) \) denotes the contribution of cell \( m \) toward goal \( n \) if the action \( Q_{im} \) is instituted within that cell, and

\[
\Gamma' = \bigotimes_{m=1}^{M} \Gamma_m \quad \text{(direct product set)}
\]
Note that in this case a feasible action $Q_i$ for the entire problem consists of an ordered $M$-tuple $(Q_{i1}, Q_{i2}, \ldots, Q_{iM})$ of actions $Q_{im}$, one for each set $\mathcal{I}_m$. Further, notice that (2.3) is the same as:

$$
\sum_{m=1}^{M} \max_{Q_{im} \in \mathcal{I}_m} \sum_{n=1}^{N} \lambda_n V_{nm}(Q_{im})
$$

(2.5)

Hence, when the Lagrangian problem (2.1) consists of a set of cell problems, the particular undominated action $Q^*(\lambda)$ associated with a given set of multipliers $\lambda$ may be identified by simply maximizing

$$
\sum_{n=1}^{N} \lambda_n V_{nm}(Q_{im})
$$

(2.6)
in each cell independently of actions taken in other cells.

Suppose $Q^*_{im}$ denotes the particular action that maximizes (2.6) for a particular set of Lagrange multipliers $\lambda^*$. Then the action $Q^*$ that maximizes (2.1) is defined by the $M$-tuple $(Q^*_{i1}, Q^*_{i2}, \ldots, Q^*_{iM})$ and the associated value of the Lagrange function is simply the sum over all $M$ of the values obtained from evaluating (2.6) for $Q^*_{im}$.

In summary, in the case of cell problems, the problem of finding a single action $Q_i$ that maximizes the Lagrangian problem (2.1) decomposes into the inherently
simpler problem of finding independent actions $Q_{im}$ that maximize the Lagrangian associated with each cell. As we shall see later in this work, many important special cases of the problem of inventory management subject to budget constraints may be decomposed into a series of cell subproblems, with one cell for each item in the inventory system.

Gaps

As noted above, there may be certain undominated actions that cannot be produced using the Lagrange multiplier procedures, i.e., for which no set of Lagrange multipliers exist that will produce the action as a solution to (2.1). The reasons for this may be seen by considering the geometric characteristics of the Lagrange function.

The set of goal measurements $V(Q) = \{V_n(Q)\}$ may be interpreted as a point in N-dimensional Euclidean space, which we will term value space, or simply V-space, since each dimension represents progress toward a particular goal. Every action $Q \in \Gamma$ maps into a point in this space. Now consider a particular solution $Q^*$ produced by a set of Lagrange multipliers $\lambda$ through the optimization problem (2.1). Everett calls the associated point $V(Q^*)$ in V-space an accessible point, since it can be obtained from a knowledge of $\lambda$. 
By definition,
\[ \sum_n \lambda_n v_n(Q^*) \geq \sum_n \lambda_n v_n(Q_i) \quad (2.7) \]

This equation may be written as
\[ -\sum_n \lambda_n v_n(Q^*) + \sum_n \lambda_n v_n(Q_i) = 0 \quad (2.8) \]

Now consider the hyperplane in V-space defined by \( H = \alpha + \sum_n \lambda_n v_n(Q) \), where \( \alpha = -\sum_n \lambda_n v_n(Q^*) \). We see that, because of (2.8), none of the points in V-space lies above this hyperplane, and at least one point, namely \( V(Q^*) \), lies on it.

Each solution produced by Lagrange multipliers therefore defines a bounding hyperplane that is tangent to the set of accessible points in V-space at the point corresponding to the solution. This hyperplane thus constitutes an upper bound to the set of accessible points. It is clear that, since no such tangent bounding hyperplane exists in regions where the upper envelope of points in V-space is not concave, the Lagrange multiplier method cannot produce solutions in such a region. Such regions are called gaps. Conversely, for any point on the envelope where a tangent bounding hyperplane does exist (i.e., the envelope is concave at the point), it is obvious that there exists a set of
multipliers (namely the set of coefficients which define the hyperplane) for which the action corresponding to the point in question maximizes the Lagrangian.

In summary, the evaluation of (2.1) for a sufficiently large number of Lagrange multiplier values will succeed in producing all undominated actions that correspond to concave regions of the envelope, and will fail to produce any actions that lie in nonconcave regions of the envelope.

Figure 2.1 illustrates these concepts for a two-dimensional V-space. In the figure, the points A-M lie on the upper envelope of points in V-space. The points E and F, and the points I and J lie in gaps, i.e., non-concavities in the upper envelope. Notice that possibly several tangent hyperplanes may pass through the point D, two of which are shown.

![Figure 2.1. Gaps in two-dimensional V-space.](image)
Fortunately, there are many important problems in which non-concavities in the envelope either do not exist or are of little significance. For example, if the set is convex and the value measures $V_n(Q)$ are concave, then it is easily shown that the upper envelope of all feasible points $\{V_n(Q_i)\}$ is concave (Geoffrion, 1968, p. 11). Thus, in this case, multipliers exist which will produce every point on the upper envelope, except perhaps at the boundary of the set $T$.

Another important situation in which gaps may be of little significance occurs when the overall decision problem consists of a large number of cell subproblems. In this case, even though there may be large convexities in the envelope of the $V$-space for each cell, the result of the overall optimization (2.1) is an envelope in the $V$-space for the total problem in which the convexities are vastly reduced in significance. In fact, since the gap structure of the overall problem simply reflects faithfully the gap structure of individual cells, only degeneracies in which several cells have gaps corresponding to the same multiplier values can cause a larger gap in the overall $V$-space envelope.
Application to Allocation Problems

Consider the allocation problem

$$\max_{Q_i \in \Omega} H(Q_i)$$

subject to

$$C_k(Q_i) \leq c_k \quad \text{all } k$$

where $H(Q_i)$ may be interpreted as the payoff or utility associated with the action $Q_i$, $C_k(Q_i)$ denotes the expenditure of the kth resource which results from employing the action $Q_i$, $c_k$ denotes the amount of the kth resource that is available, and $\Omega$ denotes the set of possible strategies or actions that might be followed. Let us now define the Lagrangian function

$$L(Q, \lambda) = H(Q) - \sum_k \lambda_k C(Q)$$

where $\lambda_k$ denotes a non-negative number. Observe that if we define $V_1(Q) = H(Q)$ and $V_{k+1}(Q) = -C_k(Q)$, $k=1,2,\ldots,K$, then (2.11) has the same form as (2.1). From the previous section, if one picks a particular set of non-negative values $\lambda = \{ \lambda_k, k=1,2,\ldots,K \}$, the solution $Q(\lambda)$ which maximizes the unconstrained Lagrangian problem (2.11) is undominated, and hence is also an optimal solution to the
constrained optimization problem

\[ \text{Maximize } H(Q) \quad \text{subject to } \]

\[ Q \in \Gamma \]

\[ C_k(Q) \leq b_k(\lambda) \quad k=1,2,\ldots,K \quad (2.13) \]

where \( b_k(\lambda) \) denotes the resource expenditure \( C_k(Q(\lambda)) \) associated with the set of decisions \( Q(\lambda) \). Note that if one can identify a particular set of multipliers \( \lambda \) such that \( b_k(\lambda) = c_k \) for all \( k \), an optimum solution to the allocation problem \((2.9)-(2.10)\) will have been obtained.

Unfortunately, for reasons discussed above not all undominated solutions are necessarily obtained using Lagrangian multipliers. Thus it will generally not be possible to identify a set of multipliers such that \( b_k(\lambda) = c_k \) for all \( K \). In most practical applications, however, if one can identify a set of multipliers such that \( \left| b_k(\lambda) - c_k \right| \) is small, a good approximate solution to \((2.9)-(2.10)\) which is sufficient for practical purposes has been obtained. Often, such a solution can be obtained using Lagrange multiplier techniques with far less effort than finding an exact solution would require. Examples of such situations are discussed by Brooks(1967) and by Sherbrooke (1968) and other examples will be presented later in this work.
One generally does not know in advance what levels of resource expenditures \( b_k(\lambda) \) are implied by a given choice of \( \lambda \). Hence, it may be necessary to solve the unconstrained Lagrangian problem several times before a particular set of Lagrange variables are identified which imply resource usage such that \( |b_k(\lambda) - c_k| \) is sufficiently small to justify termination of the calculations. If the points on the upper envelope in payoff-resource space are not sufficiently dense in the region of the constraint, or if the envelope itself is not concave in the region of the constraint, it may not be possible to identify numbers which imply resource usages that satisfy a predetermined criterion for "closeness", i.e., which satisfy \( |b_k(\lambda) - c_k| \leq \epsilon_k \), \( k=1,2,\ldots,K \) where \( \epsilon_k \) is a predetermined small number. The latter condition corresponds to the gaps discussed by Everett. As he points out, in applications to cell problems such gaps correspond to the highly degenerate situation in which many cells have gaps for the same values of the Lagrange multipliers. Since most inventory systems manage items with diverse physical and cost characteristics, however, such degenerate behavior is highly unlikely if the overall problem consists of several cell problems, one for each item in the system. Should gaps in the overall problem appear, Everett and more recently Bellmore, Greenberg and Jarvis (1968) have indicated ways of dealing with them.

Everett suggested that the allocation problem (2.9)
Eq. (2.10) might be solved by performing the following sequence of calculations:

1. Set \( i = 1 \) and arbitrarily select a set of non-negative multipliers \( \lambda^i = \{ \lambda^i_1, \lambda^i_2, \ldots, \lambda^i_K \} \).

2. Find the solution \( Q(\lambda^i) \) that solves the Lagrangian problem

\[
\max_{Q \in \mathcal{F}} H(Q) - \sum_{k} \lambda^i_k C_k(Q)
\]  

(2.14)

3. Compute the resource usage \( b^i_k(\lambda^i) = C_k(Q(\lambda^i)) \), \( k = 1, 2, \ldots, K \), associated with the solution \( Q(\lambda^i) \).

4. If \( b^i_k(\lambda^i) \) is "sufficiently near" \( c^i_k \) for all \( k \), terminate the calculations. Otherwise, go to step 5. (One test for "nearness" is to check if \(|b^i_k(\lambda^i) - c^i_k| \leq \varepsilon_k\), where \( \varepsilon_k \) is a predetermined small number.)

5. Given the sets of Lagrange multipliers \( \lambda^1, \lambda^2, \ldots, \lambda^i \), the set of payoffs \( H^1, H^2, \ldots, H^i \), and the associated measures of resource usage \( b^1, b^2, \ldots, b^i \), where \( H^i = H(Q(\lambda^i)) \) and where \( b^i = \{ b^i_1(\lambda^i), \ldots, b^i_k(\lambda^i) \} \), select a new set of Lagrange multipliers \( \lambda^{i+1} \) for the next cycle of calculations. Replace \( i \) by \( i+1 \) and go to step 2.

In the special case in which the Lagrange function \( (2.14) \) is separable, step 2 may be accomplished by solving a series of subproblems, one for each cell, and step 3 may be performed by computing the resource usage associated with each cell, and aggregating these results over all cells.

Everett has shown that if \( \lambda_k^i \) is fixed,
k=1,2,...,K, k\neq j, then b_j(\lambda^i) is a monotonically nondecreasing function of \lambda_j. Everett suggested that this fact might be exploited by employing extrapolation and interpolation procedures to guide choices of \lambda_j in the above algorithm to yield the value or values of b_j desired, if such values of \lambda_j exist.

As we noted in the previous section, the Lagrange multiplier method will generate all undominated actions that correspond to concave regions of the upper envelope of points in V-space. We also noted that the set of multipliers associated with an accessible point W on the upper envelope defines the slope of a hyperplane that is tangent to the envelope at W. Hence, one way to pick a new trial multiplier in step 5 is to fit a concave surface to the upper envelope of the set of points \((H^1, b^1), (H^2, b^2), \ldots, (H^i, b^i)\) and to use the slope of the curve evaluated where b = \((c_1, c_2, \ldots, c_K)\) as the new value of \lambda.

Brooks and Goeffrion(1966) have suggested that the dual variables associated with a linear approximation to the original problem might be used to guide the choice of \lambda. Later, Geoffrion(1970) further refined these ideas and placed them in perspective with several other approaches to large-scale programming problems. The Brooks and Goeffrion procedure corresponds to approximating the envelope of accessible points in V-space by the supremum of the set of bounding hyperplanes tangent to the upper envelope of the set of
points \( \{(H^1, b^1), \ldots, (H^i, b^i)\} \). In this case, the simplex method of linear programming may be employed in step 5 above to automatically select values of the multipliers for subsequent iterations. The Brooks and Geoffrion procedure may be considered as a special case of the Dantzig-Wolfe algorithm which will be discussed in some detail in Chapter IV.

Although a hyperplane is certainly the easiest surface to fit to the envelope of generated points, Fox and Landi (1966) point out that it may be worthwhile in certain applications to fit some other concave surface.

Fox and Landi (1968) have observed that in the special case of optimization subject to one constraint, the identification of the particular value of \( \lambda \) which satisfies

\[
\sum_k |b_k(\lambda) - c_k| \leq \varepsilon
\]

where \( \varepsilon \) denotes an arbitrarily small number, is quite similar to the problem of finding a zero of a monotone function on a bounded interval (assuming two values \( \lambda' \) and \( \lambda'' \) can be found such that \( b(\lambda') \leq b(\lambda'') \)). They then show that bisection is the unique minimax sequential search procedure for finding a zero of a monotone function known to lie in a given interval; that is, bisection minimizes the maximum length of the interval remaining (at termination) after a fixed number of evaluations, or, equivalently,
minimizes the maximum number of evaluations required to locate the root in an interval of fixed length. When the monotone function $C(Q(\lambda))$ is known to be convex, Fox and Landi point out that a procedure described by Gross and Johnson (1959) is minimax and appears to be superior to bisection. A presentation of their method can also be found in Bellman and Dreyfus (1962).

Everett points out that in most operations research work, one is not simply interested in achieving the optimum payoff for some given resource levels, but rather in exploring the entire range of what can be obtained as a function of the resource commitments. In this case it matters little whether this function is produced by solving a spectrum of problems with constraints stated in advance, or simply by sweeping through $\lambda_k$'s to solve a spectrum of problems whose constraint levels are produced in the course of the solution. The Lagrange multiplier method when applicable is therefore quite efficient if the whole spectrum of constraints is to be investigated.

Everett further points out that when one is primarily interested in exploring the entire range of efficient solutions, the Lagrange multiplier method is particularly well suited to use with computers, where the trial and error variation of the multipliers as well as the maximizations within individual cells of a cell problem can be programmed to be rapidly and automatically executed. An example of
how this technique has been applied to problems of reliability design is presented by Barlow and Proschan (1965).
CHAPTER III

A MODEL OF A BUDGET-CONSTRAINED
INVENTORY SYSTEM

Introduction

As noted in Chapter I, the primary focus of this work is the problem of inventory management during the execution phase of the budget cycle. Specifically, we are interested in identifying appropriate courses of action for the inventory manager who is presented with a given budget and instructed to abide by it—to do the best he can with what he has, consistent with limitations established by Congress and other higher authorities. In this chapter, we will develop an analytical model of this problem. Subsequent chapters will then discuss procedures that are appropriate for obtaining solutions to important special cases of the general problem.

Elements of a Decision Problem

Arrow, Karlin, and Scarf (1958, p. 16) suggest that a decision problem typically has four parts:

1. A model, expressing a set of assumed empirical relations among a set of variables.

2. A subset of decision variables, whose values are to be chosen by the decision maker.
3. An objective function of the variables in the model, a function such that a higher value represents a more desirable state of affairs from the viewpoint of the decision maker.

4. Computing methods for analyzing the effects of alternate values of the decision variables on the objective function.

Ideally, we would like to have computational techniques for reaching optimal solutions—that is, sets of the decision variables that produce higher values of the objective function than any other permissible set of decisions. Often, however, the complexities of the problem or the extent or economics of the computational burden prevent such a solution from being obtained; if so, we are often happy to settle for a descriptive solution, i.e., for an evaluation of the objective function associated with a given set of decision variables. Descriptive solutions are often partial substitutes for optimal solutions, for they aid in the identification of good (rather than "optimal") alternatives. In certain complex situations, however, even descriptive solutions may not be economically obtainable. Arrow, Karlin, and Scarf stress that solutions that are not effectively computable are not properly solutions at all; thus, the acceptability of a solution is often relative to available computing technology. From another point of view, the selection of the method for computing solutions to a decision problem should be considered as one of the
decision variables in the overall decision problem. In this light, the minimum cost mode of operation of the system might involve the use of simplified computing techniques which achieve "good" solutions rather than the use of an expensive computational method which guarantees an "optimal" solution will eventually be found.

The computational efficiency of any solution technique will, of course, depend heavily upon the number of variables and the complexity of the relationships required to describe the system of interest. Hence, let us now consider the specific elements required to model the problems of inventory management during the execution phase of the budget cycle.

The Objective Function

During the Execution phase of the budget cycle, the primary questions that must be resolved by the inventory manager relate to which stocks are to be replenished and to what degree, which stocks are to be purged from the system, and what actions are appropriate to respond to the various demands for stock which may be placed upon the system. These decisions must be reached in such a manner that the aggregate effects of the individual actions do not exceed funding limitations or other administrative constraints imposed by higher authorities, and must be consistent with other resource limitations of the system (e.g., available
warehouse space). If the manager has some flexibility in shifting resources from one category to another, the reallocation decisions might also have important effects.

Our problem then is to identify a set of ordering, disposal, termination, shipment, and reallocation actions which optimize the "effectiveness" of the supply system subject to the funding, administrative, and resource constraints. But just what is supply system effectiveness? Unfortunately, there appears to be no single measure that provides a generally satisfactory answer to this question; in fact, there is no general agreement on what "supply effectiveness" means. One reason for this situation is that it is difficult to separate questions concerning the performance of a military supply organization from questions about the relative desirability or utility of the many military activities that it supports. Although it is generally agreed that some types of activities are more important or of more value than others, widely-accepted quantitative methods for reflecting these differences are yet to be developed. Second, many aspects of military supply operations may have important military consequences in themselves. For example, maintaining lower stock levels to decrease the costs of storage and costs of technical obsolescence involves an increased risk of serious deficiencies in the advent of a military emergency. As another example, the centralization of supply operations to achieve efficiencies of scale and to
pool the uncertainties associated with more dispersed supply arrangements also increases the likelihood that the supply organization itself will be a major enemy target in the event of an attack.

Although there appears to be no single measure that provides a generally acceptable index of the effectiveness of a supply system, there are a number of variables that quantify important operating characteristics of such systems. These variables measure either:

1. The consumption of resources, e.g., funds expended, man-hours used, and floorspace occupied, or

2. The levels of customer service provided, e.g., the percentage of customer demands filled "off-the-shelf", the number of stockouts per year, and the average number of days until a backorder is filled.

In short, these variables quantify the input-output characteristics of the system.

It appears reasonable to assume that once the major structural characteristics of a supply system have been established, e.g., the selection of a centralized stocking concept as opposed to a base self-sufficiency stocking concept, the "effectiveness" of the resulting supply system may be described as a relatively well-behaved function of the measures of its input-output characteristics. Let $X_i$ denote the vector of input-output characteristics associated with the course of action $Q_i$. Hence, a first approximation to
the effectiveness function might take the form:

\[ V(X_i) = \sum_{k=1}^{K} a_k x_i^k + a_0 \]  

(3.1)

where \( x_i^k \), \( k=1,2,\ldots,K \), denotes the kth input-output variable associated with the outcome vector \( X_i \), and the parameters \( a_k \), \( k=0,1,\ldots,K \) denote the coefficients associated with a linear approximation to the surface of the effectiveness function. Notice that (3.1) defines a hyperplane in \( K \)-dimensional Euclidean space.

Suppose that the effectiveness function \( V(X_i) \) is expressed in terms of "utiles", which we define as an arbitrary unit of value or utility. To satisfy the requirement for dimensional consistency in (3.1), the coefficient \( a_k \) must have dimensions of utiles per \( u_k \), where \( u_k \) denotes the dimensional units of the variable \( x_i \). Notice that if we divide each side of (3.1) by \( a_n \), where \( a_n \) denotes a positive element of the set \( \{ a_k \} \), we obtain a function

\[ V'(X_i) = \sum_{k=1}^{K} a'_k x_i^k + a'_0 \]  

(3.2)

where \( a'_k = a_k/a_n \) for \( k \neq n \), and \( a'_n = 1 \). Notice that division by \( a_n \) merely represents a scale transformation of (3.1); hence, the hyperplane described by (3.2) has the same spacial orientation as the hyperplane defined by (3.1), and both produce identical orderings of the outcomes \( X_i \).
Observe that the coefficient $a_k'$ in (3.2) has dimensional units of $u_k'^n / u_n$. For example, if $x^k_i$ denotes system backorders and $x^n_i$ denotes operating costs in dollars associated with the outcome $X_i$, then $a_k'$ has dimensional units of dollars per backorder. Hence, rather than requiring estimates of the absolute utility associated with a single dimensional unit of a given variable, we may settle for estimates of relative value using a dimensional unit of one of the variables $x^k_i$ as a standard of measurement.

Unfortunately, it appears to be very difficult to assign numerical values to the coefficients $a_k$ through measurements of real world operations, with the possible exception of variables measuring direct operating costs of the system. Determination of the value of stockouts or other measures of customer service appear to be particularly difficult to quantify. Hence, if direct measurements for the coefficients $a_k$ are not available, either subjective estimates of these coefficients must be developed, or alternate procedures must be employed.

Hadley and Whitin (1963, p. 217) have suggested a procedure which avoids the requirement for assigning explicit values to measures of customer service. Rather, they suggest that such values might be determined implicitly, through the specification of constraints on desired levels of service. Using this procedure, one criterion for selecting a course of action might be to minimize the expected
costs of operating the supply system subject to constraints on the minimum acceptable levels of service. This is mathematically equivalent to identifying the minimum cost action from the subset of undominated actions which satisfy the specified constraints.

Brown (1967, p. 359) and Feeney (1955) look at the problem in a slightly different manner. They suggest that the appropriate role for analytic approaches is the identification of the set of efficient actions, i.e., those actions that are undominated by any other action. This set of efficient actions defines exchange curves between costs of system operations and measures of customer support. They then suggest that the appropriate mechanism for the final selection of a course of action is the mind of the decision maker himself. As noted by Feeney (1955, p. 397).

"We conclude, therefore, that in situations in which it is not possible by direct methods to obtain meaningful measures of the cost parameters contained in the decision rules, the strategic decision might best be made by confronting the executive with a picture of the efficient surface. In essence, we present him with the set of all possible alternatives that have the property of being optimal under some set of costs. If, as is generally the case, the efficient surface is convex, we may make a very strong statement with respect to the properties of these alternatives, namely, that so long as total variable cost is non-decreasing over the range of each of the outcome variables, the efficient surface derived under a linear cost model must contain the optimal outcome regardless of the nature of the cost function that is actually employed by the decision maker. That is, even if the cost function is non-linear and the outcome variables themselves are subject to complex restrictions, some point on the efficient surface must provide minimum total variable costs."
In summary, there appear to be three basic methods for combining analytic and judgmental processes to identify appropriate courses of action. All require the application of perceptual and judgmental processes to formulate both formal and mental models of the decision problem, to select appropriate principles of choice, and to perhaps modify the analytic results to account for factors not represented in the formal model. The three approaches then involve the following steps:

1. Apply analytic methods to identify the set of actions that are optimal under a variety of cost or utility structures. Then apply judgmental processes to select a course of action from the efficient set.

2. By judgment, specify numerical values for the coefficients of the criterion function. Apply analytic methods to find the action which optimizes this function.

3. By judgment, specify a set of constraints on all but one of the value measures. Apply analytic methods to identify the maximum-valued measure subject to these constraints.

The usefulness of all of these approaches will, of course, be greatly influenced by the magnitude of the computational burden required to "solve" the analytic model. To gain insights into the theoretical and computational difficulties which might be involved in this aspect of the problem, let us investigate the characteristics of an analytic model of an inventory system subject to budget constraints.
A Model of a Budget Constrained Inventory System

The Variables

Consider an inventory system in which stocks are received periodically, and let \( t_1, t_2, \ldots, t_I \) denote the \( I \) times within the planning horizon that the reviews take place. For convenience, we assume \( t_1 = 0 \), and we refer to the interval between the times \( t_i \) and \( t_{i+1} \) as period \( i \).

We assume that events within a given period, say period \( i \), are sequenced as illustrated in Figure 3.1. At the beginning of the period, all stocks are reviewed and, if necessary, replenishment orders are placed with outside suppliers. Within the period, customer requisitions are received and recorded. Receipts of replenishment orders placed in previous periods for delivery during period \( i \) may also be received at various times during the period, and moved to their appropriate storage locations. Finally, at the end of the period, both newly received and backordered requisitions are reviewed and compared with available stocks, and appropriate shipments to fill customer demands are initiated.
Let $q_{ij}^k$ denote the quantity of item $j$ ordered at time $t_i$ for delivery within the $k$th future period, i.e., period $i+k$. We assume that $k$ is restricted to some set $\mathcal{K}_j$ of real non-negative integers. Negative values of $q_{ij}^k$ for $k \leq 0$ signify the termination (i.e., cancellation) of an order that had been placed in a previous period, but has not yet been delivered. Negative values of $q_{ij}^0$ denote the disposal of on-hand stocks. Such disposal and termination actions may be desirable if there are drastic downward revisions in estimates of stocks required to meet future needs relative to the estimates used for planning purposes in previous periods.

Figure 3.1. Sequence of events within period $i$. 
We assume that \( q_{ij}^k \) must be chosen so that

\[
q_{ij}^k \in S_{ij}
\]

for all \( ij \), where \( S_{ij} \) is a specified set of real numbers which identify feasible ordering, disposal, or termination actions for item \( j \) during period \( i \). For example, if \( S_{ij} = \{0, A, 2A, 3A, \ldots\} \) orders for item \( j \) placed at time \( t_i \) must be placed in multiples of some fixed positive batch size, say a box, a carload, or a dozen. Further, since for this example there are no negative numbers in the set, disposal of excess stock of item \( j \) during this period would not be permitted.

Let \( w_{ij} \) denote the stock on-hand of item \( j \) immediately before ordering at the beginning of period \( i \) and before receipt of any replenishment stocks due to arrive during the period. Negative values of \( w_{ij} \) denote backorders for item \( j \). Let \( z_{ij}^k, k \geq 0 \), denote the amount of stock of item \( j \) on order immediately before ordering at time \( t_i \) and due to be received during that \( k \)th future period, i.e., during period \( i+k \). Finally, we define \( s_{ij} \) as the amount of item \( j \) removed from on-hand stocks at the end of period \( i \) as a result of shipments to fill customer orders or as a result of pilferage, damage, or other losses beyond the direct control of the inventory manager.

The relationships among the variables \( w_{ij}, s_{ij}, z_{ij}^k \) and \( q_{ij}^k \) are illustrated in Figure 3.2.
Figure 3.2. Relationship among on-hand stock, on-order stock, order quantity and shipment variables.

In works on inventory theory the sum of on-hand and on-order stocks is often referred to as the inventory position of the item. In this work, we denote the inventory position of item $j$ before ordering at the beginning of period $i$ as $y_{ij}$; hence, in terms of our other variables

$$y_{ij} = w_{ij} + \sum_{k \in J} z_{ij}^k$$

for all $ij$.

We have now introduced notation to represent important variables describing or affecting the status and
replenishment of system stocks. We will now present notation for variables affecting the depletion of system stocks, i.e., variables representing customer demands placed upon the system and resulting shipping decisions.

Let $d_{im}$, $m=1,2,...,M$, denote the magnitude of the $m$th type of random influence to occur during period $i$. For example, $d_{im}$ might represent the number of demands of a particular demand class $m$, the cost per unit of replenishment stocks for item $m$ on the open market, or the number of units of on-hand stock of item $m$ that are lost during period $i$ due to pilferage or material handling damage. As a final example, suppose that each requisition for a given item is assigned a priority code ranging from 1 to 20, where a code of one denotes the highest priority. In this case, $d_{im}$ might denote the number of demands for a particular item in a given priority class.

Without loss of generality, we may assume that the random variables $d_{im}$ are numbered so that the first $M^*$ variables, $d_{i1}, d_{i2},...,d_{iM^*}$, represent classes of demands for system stocks, and that higher numbered variables denote other forms of random influences.

For notational convenience, we define the vectors

$$D_i = \{d_{i1}, d_{i2},...,d_{iM}\}$$
$$D^i = \{D_l, D_2,..., D_i\}$$
$$D = D^i$$

(3.5)
Note that the vector $D_i$ is $M$-dimensional, whereas the vectors $Q_i$, $W_i$, $S_i$, and $Z_i$ are $N$-dimensional.

Let $\mathbb{F}(D)$ denote the probability distribution function of the random vector $D$, and let $\mathbb{F}_i(D_i)$ denote the marginal distribution function of the random vector $D_i$.

We define $x_{im}^k$ as the number of demands of class $m$ received during the $k$th past period (i.e., during period $i-k$) that remained unsatisfied before initiating shipments at the end of period $i$. The sum of the $x_{im}^k$ over all $k$ thus represents the total backlog of unsatisfied class $m$ demands before period $i$ shipments are initiated. We define $p_{im}^{kj}$ as the quantity of item $j$ shipped at the end of period $i$ to satisfy class $m$ demands that had been backlogged for $k$ periods. Hence, from our above discussion,

$$s_{ij} = \sum_{m=1}^{M^*} \sum_{k} p_{im}^{kj} + d_{ij}$$

where $d_{ij}$ denotes the net effect of random losses from the system during period $i$. Equation (3.6) notes that total losses of on-hand stocks during a given period are the sum of shipments to meet customer demands and of other losses such as those resulting from the damage of stored goods. (The reader will recall that disposal of stock is accounted for by negative values of $q_{ij}^o$.)

To simplify our notation, we define the vectors
The family of vectors $V_{ij}$, $V_i$, $V^i$, and $V$ describe various aspects of the stock status of the inventory system. $V_{ij}$ denotes the on-hand and on-order status of item $j$ at the beginning of period $i$, and $V_i$ denotes the collection of such status information for all items in the system. $V^i$ simply denotes the history of stock status records at the beginning of each period up to and including $i$.

In a similar manner, we denote the status of unfilled requisitions by the following vectors:

$$X_{im} = \{ x^0_{im}, x^1_{im}, \ldots, x^I_{im} \}$$
$$X_i = \{ X_{i1}, X_{i2}, \ldots, X_{iM^*} \}$$
$$X^i = \{ X_1, X_2, \ldots, X_i \}$$
$$X = X^I$$

Notice that $V_i$ is $N$-dimensional, whereas $X_i$ is $M^*$-dimensional.

To further simplify our notation, we define the
decision vectors

\[
Q_{ij} = \{q_{ij}^1, q_{ij}^2, \ldots, q_{ij}^K\}
\]
\[
Q_i = \{Q_{i1}, Q_{i2}, \ldots, Q_{iN}\}
\]
\[
Q^i = \{Q_1, Q_2, \ldots, Q_i\}
\]
\[
Q = Q^i
\]

and

\[
P_{im} = \{p_{im}^1, p_{im}^2, \ldots, p_{im}^N\}
\]
\[
P_{im} = \{P_{im}^1, P_{im}^2, \ldots, P_{im}^K\}
\]
\[
P_i = \{P_{i1}, P_{i2}, \ldots, P_{iM^*}\}
\]
\[
p^i = \{P_1, P_2, \ldots, P_i\}
\]
\[
p = p^i
\]

Hence, \(Q_i\) and \(P_i\) denote the set of ordering and shipping decisions, respectively, made during period \(i\).

We have now introduced notation to represent the status of on-hand, on-order, and backordered stocks in the system, and to represent the ordering and shipping decisions and random influences which modify stock status. These variables may be found in most models reported in the rather extensive literature of inventory theory, though perhaps in a slightly different form than defined here.

In addition to the above, we must also introduce variables to reflect the status of resources available to
each organizational component of the inventory system, and
decision variables that describe the flexibility of the in-
ventory manager or managers to shift resources either between
organizational components or between resource categories
within an organizational component. For example, it may be
possible to shift part of the funds that were originally
budgeted to support procurement activities to cover unfore-
seen workloads in the shipping, receiving and warehousing
organization. Within the shipping, receiving and warehousing
organization, on the other hand, it may be possible to shift
manpower and/or funds from the support of shipping dock
activities to the support of receiving dock activities.

The exact nature of the variables required to de-
scribe this latter aspect of budget-constrained inventory
systems depends heavily upon the organizational and financial
structure of the system, the position of the decision-maker
within this structure, and the interchangeability of re-
sources managed by the decision-maker. Later in this section
we will introduce specific examples of the forms which these
relationships may take. For the moment, however, we will
simply introduce notation required in our later development.

As noted in Chapter I, most supply organizations
operate under administrative and legal limitations.
Generally, these limitations are known in advance, although
they may be modified if future events dictate such adjust-
ments to original plans are required. We will term the set
of documents which depict the current set of limitations as the "budget program" of the system. Consider the budget program at the beginning of period $i$. We define $b_{ief}$ as the quantity of resource $f$ authorized or planned to be available to organizational unit $e$ during period $i$, and we define $r_{ief}$ as the status of this resource at the beginning of the $i$th period. For example, $b_{ief}$ may denote the total amount of funds available to the procurement organization for salaries of civilian personnel during the current fiscal year, while $r_{ief}$ might denote the total amount of this funding expended by the $i$th month of this budgeting interval.

To simplify our notation, we define the vectors

$$R_{ie} = \{r_{ie1}, r_{ie2}, \ldots, r_{ieE}\}$$

$$R_i = \{R_{i1}, R_{i2}, \ldots, R_{iE}\}$$

$$R^i = \{R_{1}, R_{2}, \ldots, R_{i}\}$$

and the vectors

$$B_{ief} = \{b_{ie1}, b_{ie2}, \ldots, b_{ieE}\}$$

$$B_i = \{B_{i1}, \ldots, B_{iE}\}$$

$$B^i = \{B_{1}, B_{2}, \ldots, B_{i}\}$$

$$B = \{B_{1}, B_{2}, \ldots, B_{i}\}$$
where $E$ denotes the number of distinct organization units, and $F_e$ denotes the number of resources utilized within a given organizational unit $e$.

Hence, we may interpret the vector $B_{ie}$ as the quantities of resources programmed to be available to or authorized to organizational unit $e$ during period $i$ while the vector $R_{ie}$ denotes the status of these resources at the beginning of period $i$. The vectors $R_i$ and $B_i$ denote, respectively, the collection of status and budgetary information associated with period $i$, while $R^i$ and $B^i$, respectively, denote the collection of this information for periods up to and including $i$.

As a final notational simplification, we define the vector $U_i = \{V_i, X_i, R_i\}$. The vector $U_i$ thus represents the status of on-hand, on-order, and backordered stocks and the status of other system resources at the beginning of period $i$. We assume that this vector completely defines the state of the system at this point in time.\(^1\)

The Model

Classes of Relationships. Let us now consider the manner in which the variables defined in the previous section

\(^1\)If additional status variables are required, they might be treated as "dummy" resources.
are interrelated. One major class of relationships is the balance equations which relate on-hand and on-order stocks, backorders, and unexpended resources at the beginning of period (i+1) to the status of these quantities at the beginning of period i and to the ordering, disposal, termination, shipment, and reallocation actions taken during the period. The collection of these relationships define the transition function of period i. A second major category of relationships define bounds or limits upon the decision or status variables of the system. It is the particular nature of these constraints which distinguish problems of inventory management subject to budget constraints from other forms of inventory problems.

The Balance Equations. In this work, we assume that on-hand and on-order stocks satisfy the following material balance equations.

\[ w_{i+1,j} = w_{ij} + q^o_{ij} + z^o_{ij} - s_{ij} \quad \text{all } ij \quad (3.13) \]

\[ z^k_{i+1,j} = z^k_{ij} + q^{k+1}_{ij} \quad \text{all } ijk \quad (3.14) \]

These relationships follow immediately from the definitions and from Figure 3.2.

We assume that backorders obey the following balance
equations
\[ x_{im}^0 = d_{im} \]
\[ x_{i+1,m}^k = \begin{cases} 
  x_{im}^{k-1} - \sum_j p_{im}^{j,k-1} & 1 \leq k \leq K_p \\
  0 & \text{otherwise}
\end{cases} \quad (3.15) \]

where (3.15) holds for all \( im \). The summation in (3.15) is over all items \( j \) in the system. The system of equations (3.15) denotes an inventory system in which a given demand may be backordered for up to \( K_p \) periods, where \( K_p \) is an appropriate non-negative integer. After \( K_p \) periods have elapsed and a requisition remains unfilled, we assume the requisition is dropped from the system and the demand is lost. Hence, \( x_{i+1,m}^{k_p} \) denotes the number of demands of class \( m \) lost during period \( i \) due to exceeding the maximum backorder delay time. Notice that in the special case in which \( K_p = 1 \), the system (3.15) denotes a situation in which all demands that are not filled immediately are lost, while in the case that \( K_p = 1 \), the number of periods in the planning horizon, the system (3.15) indicates all demands are backordered until filled. If desired, \( K_p \) could vary by customer class; however, since this would unnecessarily complicate our notation, we will not follow such a convention here.

Let us now consider the relationships among the resource status variables. In general, the following system
of resource balance equations will hold:

\[ r_{i+1,ef} = a_{ief} \left[ r_{ief} + \Delta_{ief} + b_{ief} \right] \] (3.16)

where \( r_{ief} \), the beginning level of resource \( f \) for organizational unit \( e \), is assumed known, and where

\[ b_{ief} = \text{increases in } r_{ief} \text{ due to budget allocations at the beginning of period } i, \text{ or if negative, decreases in } r_{ief} \text{ due to budget reallocations or resource transfers directed by higher authorities.} \]

\[ \Delta_{ief} = \text{increase in resource } r_{ief} \text{ during period } i \text{ due to replenishment, disposal, termination, or shipping actions taken during period } i \text{ or preceding periods.} \]

and where \( a_{ief} \) is a coefficient which indicates the degree of transferability of the quantity of resource \( f \) remaining at the end of period \( i \) to the next period. For example, suppose \( r_{ief} \) denotes the amount of funds invested in interest-bearing bonds at the beginning of period \( i \). In this case, (3.16) might take the form

\[ r_{i+1,ef} = (1 + \alpha) \left[ r_{ief} + b_{ief} \right] \] (3.17)

where \( \alpha \) denotes the rate of interest during period \( i \), and where the investment is assumed to be unaffected by ordering and shipping actions taken during the period. As another example, suppose \( r_{ief} \) denotes unexpended resources of a "use-or-lose" variety, such as unobligated funds covered by a fixed-year appropriation. In this case, (3.16) might take
The form

\[
  r_{i+1,ef} = \begin{cases} 
  \Delta_{ief} + b_{ief} & \text{if } i=i^* \\
  r_{ief} + \Delta_{ief} + b_{ief} & i^* < i < i^{**}
\end{cases}
\]  

(3.18)

where \(i^*\) and \(i^{**}\) denote, respectively, the first and last periods covered by the appropriation. In this example, any funding authority that remains unobligated at the end of period \(i^{**}\) is lost, but funding authority unobligated at the end of period \(i\), \(i^* \leq i < i^{**}\), may be carried over into the succeeding period.

The mathematical relationships required to define the increase \(\Delta_{ief}\) in resource \(r_{ief}\) during period \(i\) as a function of the replenishment and shipping actions in previous periods is dependent upon the peculiar physical and procedural characteristics of the system being studied. In many systems, the following relationship may hold:

\[
  \Delta_{ief} = \sum_{jkt} g_{i-t,jke}^t(q_{ij}^k) + \sum_{jmt} h_{i-t,jme} f\left(\sum_k p_{im}\right)
\]  

(3.19)

where \(g_{i-kef}^{t}(q_{ij}^{k})\) denotes a function which defines the increase in resource \(f\) of organizational unit \(e\) during period \((i+t)\) resulting from the initiation of a replenishment, disposal, or termination action in period \(k\) for \(q_{ij}^{k}\) units of item \(j\). For example, although we assume that replenishment orders for \(q_{ij}^{k}\) units \((q_{ij}^{k} > 0)\) will be delivered in
period \( i+k \), processing the replenishment action might require the use of limited resources such as manpower, equipment, storage facilities, or funds both before and/or after the delivery of goods. Hence, the set of values \( g_{ijkef}^t(q_{ij}^k) \), \( t = i, i+1, \ldots, I \) defines the requirements for resource \( f \) of organizational unit \( e \) in future periods that result from the replenishment action \( q_{ij}^k \). Similar comments apply to termination and disposal actions. Similarly, \( h_{imef}^{jt}(z) \) denotes a function which defines the effect of shipping \( z \) units of item \( j \) at the end of period \( i \) upon resource \( f \) of organizational unit \( e \) during period \((i+t)\).

In later chapters of this work, we will study the case in which the functions \( g_{ijkef}^t(q_{ij}^k) \) and \( h_{imef}^{jt}(\sum_k p_{im}^j) \) take the form

\[
f(z) = \begin{cases} 
\bar{\alpha} + \bar{b} \mid z \mid & \text{if } z < 0 \\
0 & \text{if } z = 0 \\
\bar{\alpha} + \bar{b}z & \text{if } z > 0
\end{cases}
\]

where \( \bar{\alpha} \) denotes setup costs associated with disposal or termination actions, \( \bar{\alpha} \) denotes setup costs associated with replenishment or shipping actions, and \( \bar{b} \) and \( \bar{b} \) correspondingly denote costs that are proportional to the number of units affected. The constants may all be appropriately subscripted to agree with the particular function under consideration.

Although one might easily identify situations in
which the functions $g(.)$ and $h(.)$ take more complicated forms, in many real-world systems (3.20) will provide an adequate approximation to the actual relationship. In addition, although it might be theoretically interesting to consider more complex forms than (3.20), in practice the problems of data collection and analysis required to develop and periodically revise estimates for the coefficients in the relatively simple formula presented above may be substantial and little benefit may accrue from more accurate representations.

In the special case in which $r_{ief}$ denotes unobligated funding authority of a stock fund account, equation (3.19) might take the form:

$$\Delta_{ief} = - \sum_{jk} c_{jk}(q_{ij}^k) + \sum_{jkm} s_j p_{im}^{jk} \quad (3.21)$$

where the function $c_{jk}(q_{ij}^k)$ denotes the loss in obligation authority associated with the action $q_{ij}^k$ and the scaler $s_j$ denotes the sales price per unit of item $j$ to a stock fund customer. (Although stock funds might offer price breaks for quantity purchases, this does not appear to be a common practice). If $q_{ij}^k > 0$, indicating a replenishment action, $c_{jk}(q_{ij}^k)$ denotes the costs of the replenishment action borne by the stock fund account. On the other hand, if $q_{ij}^k < 0$, indicating a termination or disposal action, $-c_{jk}(q_{ij}^k)$ denotes the obligation authority recovered through the disposal or termination action.
In the special case in which \( r_{ief} \) denotes the balance in a stock fund cash account, \( \Delta_{ief} \) will be defined in a manner quite similar to (3.21); the elements included in the summations must be modified, however, to account for time delays between the actions \( q_{ij}^k \) and \( p_{im}^{jk} \) and associated cash flows.

**Constraints.** Let us now consider the types of constraints which limit an inventory manager's flexibility in providing supply support. There are, of course, the rather obvious constraints that disposals in excess of on-hand stocks or terminations in excess of on-order stocks are not possible, while shipments can neither exceed orders or stocks on-hand nor can shipments be negative. If the planning horizon is short, there might also be a lower bound on the desired on-hand inventory at the end of the planning horizon. We may express these constraints symbolically as:

\[
\begin{align*}
  w_{ij} + q_{ij}^o & \geq 0 \quad \text{all } ij \\
  z_{ij}^k + q_{ij}^k & \geq 0 \quad \text{all } ij, \text{ and all } k \geq 0. \\
  w_{ij} + q_{ij}^o + z_{ij}^o - \sum_j p_{im}^{jk} & \geq 0 \quad \text{all } ij \\
  x_{im}^k - \sum_k p_{im}^{jk} & \geq 0 \quad \text{all } ikm
\end{align*}
\]
\[ p_{im}^{jk} \geq 0 \quad \text{all } ijkm \quad (3.26) \]

\[ w_{I+1,j} \geq \gamma_j \quad \text{all } j \quad (3.27) \]

where \( \gamma_j \) denotes the minimum end-of-period inventory for item \( j \).

In addition to the above, there may be possibly several restrictions upon either the levels or rates of change of resources utilized or managed by the inventory system. These constraints may be classified into the following categories:

1. Stock fund financial constraints.
2. Operating budget constraints.
3. Supply system capacity constraints.

**Stock Fund Financial Constraints.** As discussed in Chapter I, there may be possibly several legal and administrative limitations on the levels or rates of change of funds or materials managed under a stock fund concept. For example, R. E. Trapp(1969, p. 4-55) points out that divisions of the Air Force Stock Fund satisfy the following restrictions:

1. No over-obligation of any funds limitation is permitted.
2. The balance in the stock fund cash account is not permitted to fall below a prescribed minimum level.
3. The aggregate value of on-hand and on-order stocks cannot exceed prescribed levels.

Each of the above constraints might be written symbolically as

\[ r_{\text{ief}} \leq b_{\text{ief}} \quad (3.28) \]

where, say, \( f=1 \) denotes the unobligated stock fund appropriation authority, \( f=2 \) denotes the stock fund cash account, \( f=3 \) denotes the measure of aggregate inventory value, and where \( e \) denotes the organizational unit responsible for management of the stock fund resources. As noted above, the balance equations for unobligated funding authority are defined by (3.18) and (3.21), and the balance equations for the cash account are defined similarly with appropriate modifications for the delays in cash flows associated with different ordering and shipping actions. The constraint on the aggregate value of on-hand and on-order stocks may be written as

\[ \sum_j c_j w_{ij} + \sum_k z_{ij}^k \leq b_{\text{ief}} \quad (3.29) \]

where \( c_j \) denotes the book value of a unit of item \( j \). The balance equations associated with this resource measure follow easily from (3.13) and (3.14).

Operating Budget Constraints. In addition to the financial constraints on stock fund operations discussed above, the manager may also be constrained by the funding
limitations of the operating budget for his organization. These constraints may establish limits on the amount of funds the manager may obligate for goods and services to support his organization and may specify the maximum number of civilian and military positions within the organization to be manned during the fiscal period. Similar budgetary limitations restrict the capabilities of each organizational component of the military supply system. These constraints also take the form of (3.23), with the resource balance equations defined by (3.18) and (3.19).

**Capacity Constraints.** Another important category of operational limitations is the capacity characteristics of the military supply system. A warehouse can hold only so much material; a shipping-receiving facility can process only a certain number of trucks at a time. Although such constraints may be subject to change in the long run, in the short term such characteristics of the military supply system may be essentially fixed. Symbolically, the capacity constraints of the supply system may be written in the same form as (3.28), with the capacity balance equation defined by the general relationships (3.18) and (3.19). An important special case occurs when it is not possible to carry unused capacity over into succeeding periods. In this case, \( i^* = i^{**} \) in (3.18). For example, if available man-hours or machine hours are not utilized within a given period, the unused capability is lost.
Another important form of capacity constraint that might be encountered is a limitation on the amount that a given resource can change during a given period. Symbolically, such a constraint takes the form

$$j_{ief} \leq r_{i+1,ef} - r_{ief} \leq j_{ief}$$ (3.30)

where \(j_{ief}\) and \(j_{ief}\) are specified constants. For example, there may be policy limitations on the number of personnel that can be hired or fired in a given period.

**Resource Reallocation Constraints.** Although an inventory manager often possesses some flexibility in shifting or reallocating his resources to compensate for unanticipated shifts in workloads or other requirements, such flexibility is seldom absolute. For example, there may be legal or administrative restrictions on the amount of funds that may be shifted from one expenditure category to another. As another example, there may be personnel policies or union contracts that limit the amount of overtime that may be worked or the number of personnel that may be shifted from one organization or work group to another. In this case the constraints again take the general form of (3.28) with resource balance equations defined by the general relationships (3.18) and (3.19).
The General Problem

As noted above, our problem is to identify the set of ordering, disposal, termination, shipment and reallocation actions that optimize the "effectiveness" of the supply system subject to funding, administrative, and resource constraints. Hence, if all demands D are known, we may state our problem symbolically as:

\[
\text{Optimize } c(U, Q, P, B, D) \quad (3.31)
\]

subject to the system laws

\[
U_{i+1} = T_i(U_i, Q_i, P_i, B_i, D_i) \quad i=1, 2, \ldots, I-1 \quad (3.32)
\]

as well as the constraints

\[
G(U, Q, P, B, D) \leq 0 \quad (3.33)
\]

where the initial conditions \( U_1 \) are assumed known, \( T_i(\cdot) \) denotes a vector function consisting of the balance equations (3.13)-(3.16), and \( G(\cdot) \) represents a vector function of the constraints (3.22)-(3.30). The scalar function \( c(U, Q, P, B, D) \) represents an appropriate measure of effectiveness to be optimized.

Observe that in the special case in which the equations defining \( \Delta_{\text{eff}} \) consist entirely of linear functionals and in which \( c(U, Q, P, B, D) \) is also a linear function,
(3.31)-(3.33) defines a standard linear programming problem and hence may be solved (at least theoretically) using the simplex algorithm of G. B. Dantzig. Because of the special structure of this problem, however, it is possible to develop alternate solution techniques that are more efficient than the general methods. In certain important situations, it is possible to develop similar procedures for the case in which D is known and c(.) is linear, but in which $\Delta_{\text{dof}}$ is defined by piece-wise linear relationships of the form (3.20). The development of such procedures for both the linear and piece-wise linear cases is the topic of Chapter IV.

Now suppose that some or all of the elements of the vector D are random variables rather than constants. This necessitates a reformulation of the objective function, since $c(U,Q,P,B,D)$ is a random variable if any of its arguments are random variables. Hence, $c(U,Q,P,B,D)$ must be replaced by an appropriate deterministic function. There are many possible choices for this function, each of which may be appropriate under certain circumstances. Perhaps the most natural choice is the expected value of $c(U,P,Q,B,D)$, i.e.,

$$
E \left[ c(U,P,Q,B,D) \right] = \int_D c(U,P,Q,B,D)d\mathcal{F}(D) \quad (3.34)
$$

where $E$ denotes the expectation operator.

Similarly, if D is a random vector the constraints with random components defined by (3.33) must also be
reinterpreted. One interpretation is that a solution is considered feasible only if it satisfies all the constraints for all possible combinations of the parameter values. This interpretation is equivalent to requiring that the constraints \((3.33)\) be satisfied with probability one. In this case, it appears that no practical solution procedure has been developed for obtaining optimal solutions to the general problem presented here. In this case, it appears that the analyst who seeks appropriate courses of action which are feasible under this interpretation is limited to the use of descriptive solution procedures such as simulation techniques to evaluate heuristic approaches to such problems. This approach permits the analyst to compare the relative performance of procedures which produce solutions that are feasible but not necessarily optimal responses to the decision problem.

An alternate approach to the above problem when \(D\) is a random vector may be employed if it is highly desirable, but not absolutely necessary, that the constraints \((3.33)\) hold. This approach is based on the ideas of chance-constrained programming first advanced by Charnes and Cooper (1959) and Charnes, Cooper, and Symonds (1958). In contrast to the interpretation of the above paragraph which required that all the constraints must hold for all possible observations of the random vector \(D\), the chance-constrained programming interpretation requires only that each constraint must hold for most of the possible values of \(D\). In this
case, the original set of constraints (3.33) are replaced by the requirement that

$$\text{Prob} \left[ G_i(U,Q,P,B,D) \leq 0 \right] \geq \alpha_i \quad i=1,2,\ldots$$

(3.35)

where $G_i(.)$ denotes the $i$th element of the vector function $G(.)$, and where the $\alpha_i$ are specified constants between zero and one. In this interpretation, a set of actions $\{Q,P,B\}$ are considered feasible if and only if (3.35) is satisfied. Hence, under this formulation a given set of actions are considered feasible if the constraints (3.33) will "probably" be satisfied once the random vector $D$ is observed.

Although no procedure is now available for obtaining optimum solutions to all chance-constrained programming problems, certain important special cases are solvable. Some of these situations will be discussed in Chapter V.

A third interpretation of the constraint set (3.33) when $D$ is a random vector is that the constraints need be satisfied only over the long term, but that any given constraint might be violated in any given period. In this interpretation, the constraints might be more appropriately termed "targets". In this case, the $i$th inequality of constraint set (3.33) is replaced by its expectation, i.e., (3.33) is replaced by

$$\mathbb{E} \left[ G_i(U,Q,D,B,D) \right] \leq 0 \quad \text{all } i.$$  

(3.36)

This latter interpretation is the basis of most of the
probabilistic models of inventory management subject to budget constraints that have appeared in the operations research literature to date. Solution procedures for situations in which constraints of the form (3.36) are appropriate are discussed in Chapter V.
CHAPTER IV
DETERMINISTIC MODELS

Introduction

In the general case, problem (3.31)-(3.33) represents a non-linear, discontinuous, dynamic, stochastic optimization problem of the following dimensions:

decision variables: $IJ(K_q + MK_p) + IEF$
status variables: $IJ(K_q + MK_p + 1) + IEF$
balance equations: $IJ(K_q + MK_p + 1) + IEF$
inequality constraints: $IJ(K_q + MK_p + 1) + L + J$

where

$E = \text{number of organizational units in the supply system}$

$F = \text{number of limited financial and physical resources in each organizational component. (In practical situations, F will vary from one organization to the next; for notational convenience, however, we will assume all organizations have the same number of limited resources.)}$

$I = \text{number of periods in the planning horizon}$

$J = \text{number of items in the inventory system}$

$K_q = \text{maximum number of periods in the delivery lead time for any item in the system}$

$K_p = \text{maximum number of periods that a requisition may be backordered}$

$M = \text{number of demand classes}$

$L = \text{number of joint constraints, i.e., funding and capacity constraints}$
Table 4.1 presents detailed breakouts of the number of status and decision variables and the number of constraints associated with components of the system of equations (3.31)-(3.33). These counts indicate the very large number of variables and constraints required to describe real-world inventory systems subject to budget limitations. For example, consider a system that reviews stocks on a monthly basis and that operates under obligation, cash balance, and inventory aggregate constraints applied on a quarterly basis. Hence, if a one-year planning horizon is employed, \( I = 12 \), \( E = 1 \), \( F = 3 \), and \( L = 3 \cdot 4 = 12 \) (one obligation, one for cash balance, and one inventory aggregate constraint per quarter, times four quarters per year), assuming all other funding and capacity constraints are negligible. If the maximum delivery lead time is six months \( (K_q = 6) \), if requisitions may be backordered for up to three months \( (K_p = 3) \), if there are 100 items in the system \( (J = 100) \), then (3.31)-(3.33) involves \( IJ(K_q + MK_p) + IEF = 12(100)(6+2 \cdot 100 \cdot 3) + 12 \cdot 3 = 727,236 \) decision variables and 728,512 inequality constraints, neglecting non-negativity requirements and limitations on maximum order sizes. If there are 10,000 items managed under this system, the number of constraints and decision variables each grow by factors of about 100.
<table>
<thead>
<tr>
<th>Decision Variables*</th>
<th>Number of Variables</th>
<th>CONSTRANTS</th>
<th>Number of Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>orders</td>
<td>IJK_q</td>
<td>on-hand stocks</td>
<td>IJ</td>
</tr>
<tr>
<td>terminations</td>
<td>IJKM_p</td>
<td>on-order stocks</td>
<td>IJK_q</td>
</tr>
<tr>
<td>disposals</td>
<td>IJ</td>
<td>backorders</td>
<td>IJKM_p</td>
</tr>
<tr>
<td>shipping</td>
<td>IEF</td>
<td>resources</td>
<td>IEF</td>
</tr>
<tr>
<td>reallocation</td>
<td>Total: IJ(K_q+MK_p)+IEF</td>
<td>Total: IJ(K_q+MK_p+1)+IEF</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Status Variables</th>
<th>Number of Variables</th>
<th>Balance Equations</th>
<th>Number of Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>on-hand stocks</td>
<td>IJ</td>
<td>on-hand stocks</td>
<td>IJ</td>
</tr>
<tr>
<td>on-order stocks</td>
<td>IJK_q</td>
<td>on-order stocks</td>
<td>IJK_q</td>
</tr>
<tr>
<td>backorders</td>
<td>IJKM_p</td>
<td>backorders</td>
<td>IJKM_p</td>
</tr>
<tr>
<td>resources</td>
<td>IEF</td>
<td>resources</td>
<td>IEF</td>
</tr>
<tr>
<td>Total</td>
<td>IJ(K_q+MK_p+1)+IEF</td>
<td>Total: IJ(K_q+MK_p+1)+IEF</td>
<td></td>
</tr>
</tbody>
</table>

*In linear programming applications, it is necessary for all variables to be non-negative. Hence, in an LP application, it is necessary to introduce additional variables for termination and disposal actions, thus increasing the total number of decision variables to $IJ(2K_q + MK_p) + IEF$. 
General problems with the level of detail included in the examples presented above appear to be insoluble within the capabilities of current analytic techniques and computing machinery, for there appears to be no technique that can adequately deal with the very large number of variables contained in the examples discussed above. Hence, if practical solutions to problems of inventory management in budget-constrained environments are to be obtained at all, methods for reducing the original problem to a size that is analytically and computationally tractable must be identified.

Two basic categories of simplification processes might be employed to reduce the dimensionality and complexity of decision problems. The first category involves conceptual simplifications associated with the definition of the problem. This category includes such actions as the treatment of some of the decision variables as parameters of the analytical model or the elimination of interaction equations which do not appear to be significant. The second category, which we term analytic simplification processes, involve the identification of fundamental relationships that compactly characterize the structure of the decision problem, and the development of interactive procedures or closed-form expressions which reduce the computational burden required to evaluate a given course of action or, better yet, to identify the particular course of action associated with a given principle of choice. Simplification processes included in this second
category are often based upon the exploitation of important special cases of the general problem.

Perhaps the most obvious form of conceptual simplification is to suppress one's necessary uncertainty about future events from the analytical model, and to use the model to develop plans which are optimal if a given set of future events come to pass. By exercising the model several times, one might develop several plans which are optimal if given assumptions about future events come to pass. Analysis of these plans might then indicate the sensitivity of one's plans to future events. Implicit judgment processes might then be employed to modify the model results to account for the uncertainties of future events, and to insure that sufficient flexibility is inherent in the final version of the plan to permit adequate responses to possible contingencies. When this approach is taken, the computational effort required to "solve" the analytic model is often substantially reduced.

When deterministic models are employed to prepare plans for use in uncertain environments, seldom are the plans used in their entirety. Usually, only the plan for the first week or month of the planning horizon is employed. After that time, new information is usually available. If so, the model may be re-exercised to develop updated plans which account for this additional information.

In this chapter, we will consider solution procedures for the problem (3.31)-(3.33) for situations in which the
future is assumed certain, i.e., for situations in which all elements of the external vector D are assumed known. These procedures are important both for use in inventory management systems in which the uncertainties of future events are accounted for by judgment processes and for solving certain probabilistic problems that possess deterministic equivalents. This latter class of problems will be discussed in Chapter V.

Small-Scale Linear Systems

In certain situations, it may be possible to solve (3.31)-(3.33) directly using the powerful computational procedures of the simplex method of linear programming. Specifically, if

1. the criterion function \( c(U, Q, P, B, D) \) may be expressed as a piecewise-linear function of the decision and status variables,

2. the system of equations (3.32),(3.33) consists entirely of linear functionals, and

3. demand may be treated as known,

then currently existing linear programming computer codes might be employed to solve the problem provided, of course, that the system of equations is sufficiently small to be handled by available computing equipment. Examples of the size of problems that can be handled by current large scale computing equipment will be presented below. The steps required to put our problem in a linear programming format will
also be presented later in this section.

Most major computer manufacturers have developed sophisticated software for solving linear programming problems which fully exploit the computational capabilities of their machines. These manufacturers often provide these software packages free of charge or at a nominal cost to the users of their equipment. The availability of such a package may prove a significant advantage if the analyst must obtain useful answers in a short period of (calendar) time.

Although solution techniques for solving linear programming problems have been available since the early 1950's, the usefulness of these techniques has always been limited by the speed, precision, and storage capabilities of available computing equipment. This is still true today, although the explosive growth in computer technology during the last two decades has significantly reduced these limitations. For example, the LP/600 linear programming package for the General Electric 625/635 line of computing systems can solve up to 4,095-row and 262,000-column standard problems, and 1000-order matrices can be solved in core. Although problems of this size are rarely solved (costs of data collection, editing, and computer run time can be quite high for problems of these dimensions), it does appear that standard linear programming problems with up to 1,000 rows and 30,000 columns can be solved routinely in less than two hours of computing time on current large scale computing equipment. For
problems with special structural characteristics, even larger problems can be solved.

To place our problem into a standard linear programming format we must express it in terms of non-negative variables. To do this, we replace the decision variables \( q_{ij}^k \) and \( b_{ij}^e \) by

\[
q_{ij}^k = q_{ij}^{k+} + q_{ij}^{k-} \quad (4.1)
\]

\[
b_{ij} = b_{ij}^+ + b_{ij}^- \quad (4.2)
\]

where \( q_{ij}^{k+} \) denotes the number of units of item \( j \) ordered in period \( i \) for delivery in period \( i+k \), \( q_{ij}^{k-} \) \((k \geq 0)\) denotes the number of on-order units of item \( j \) due to be delivered in period \( i+k \) that are to be cancelled during period \( i \), \( q_{ij}^{o-} \) denotes the number of units of item \( j \) that are to be disposed of during period \( i \) and where \( b_{ij}^+ \) and \( b_{ij}^- \) denote the amount by which reallocation decisions increase or decrease, respectively, the authorized or available quantity of resource \( r_{ij} \). Slight adjustments to the resource usage equations (3.19) and to the corresponding objective function coefficients may also be required to properly account for these substitutions.

If there is an upper limit \( 0_{ij}^k \) on the maximum order size \( q_{ij}^k \), an alternative to (4.1) is to replace the variable \( q_{ij}^k \) by

\[
q_{ij}^k = -q_{ij}^k + 0_{ij}^k \quad (4.3)
\]
where $q_{i,j}^k$ is the new non-negative decision variable. The substitution (4.3) is preferred to (4.1) since (4.3) does not increase the number of decision variables. In many of the more sophisticated LP computer codes, the above transformations may be performed automatically.

We may reduce the number of variables and constraint equations required to define our problem by expressing the status variables $w_{i,j}^k$, $z_{i,j}^k$, $x_{i,m}^k$, and $r_{i,e,f}$ as functions of the decision variables. Once this is done, the balance equations are accounted for implicitly in (3.31) and (3.33), and hence the explicit statement of these relationships by (3.32) is redundant and may be dropped. To do this, consider the system of balance equations

$$h_i = a_{i-1} \left[ h_{i-1} + \delta_{i-1} \right] \quad i=2,3,... \quad (4.4)$$

where $h_i$ denotes the quantity of a given resource available at the beginning of period $i$, $\delta_i$ denotes changes to the resource during period $i$, and $a_i$ denotes a constant which reflects the degree of transferability of the resource remaining at the end of the period $i$ to the next period. Note that

$$h_i = a_{i-1} \left[ a_{i-2}(h_{i-2} + \delta_{i-2}) + \delta_{i-1} \right] \quad (4.5)$$

By continuing the substitution process initiated by (4.5) and
\[ h_i = \prod_{k=1}^{i-1} a_k \cdot h_1 + \sum_{t=1}^{i-1} \left[ \prod_{k=1}^{t} a_{i-k} \right] S_{i-t} \] (4.6)

where \( h_1 \) denotes the beginning level of the resource.

Hence, the system of balance equations defined by (3.16) may be written as

\[ r_{ief} = a_{i-1,ef} r_{i-1,ef} + \sum_{t=1}^{i-1} a_{t,ef} (\Delta_{i-t,ef} + b_{i-t,ef}) \] (4.7)

where

\[ a_{t,ef} = \prod_{k=1}^{t} a_{i-k,ef} \] (4.8)

Using similar arguments, it is straightforward to show that

\[ x_{i+k,m}^k = d_{im}^0 - \sum_{t=0}^{k-1} \sum_{j} p_{i+t,m}^j \quad k=1,2,\ldots,K_p \] (4.9)

where \( x_{im}^0 = d_{im} \) and all other \( x_{im}^k \) are zero. Similarly,

\[ z_{ij}^k = z_{ij}^{i+k-1} + \sum_{t=1}^{i-1} q_{i-t,j}^{k+t} \quad \text{all } ijk \] (4.10)

where \( z_{ij}^{i+k-1} \) denotes the stock of item \( j \) on order at the beginning of period one and due to be delivered during period \( i+k \).

Finally, on-hand stock at the beginning of period \( i \) simply equals stock on-hand at the beginning of period one plus deliveries in the periods prior to period \( i \) less shipments made during those periods.
Hence,

\[ w_{ij} = w_{lj} + \sum_{t=0}^{i-2} z_{lj} + \sum_{t=1}^{i-1} \left[ \sum_{k=0}^{i-t} q_{tj} - \sum_{km} p_{tm} \right] \]  

(4.11)

for all \( i \neq 1 \) and all \( j \). Values of \( w_{ij} \) and \( z_{lj} \) are assumed known.

Equations (4.7)-(4.11) may now be used to eliminate the status variables from the constraint equations of Chapter III. Hence, constraints (3.22)-(3.28) may be written as follows, where equation (3.22a) below corresponds to equation (3.22), and so on.

\[ w_{lj} + \sum_{t=0}^{i-2} z_{lj} + \sum_{t=1}^{i-1} \left[ \sum_{k=0}^{i-t} q_{tj} - \sum_{km} p_{tm} \right] \]  

(3.22a)

\[ + q_{ij}^0 \geq 0 \text{ all } ij \]

\[ z_{ij} + \sum_{t=1}^{i-1} q_{i-t,j} \geq 0 \text{ all } ijk, k \geq 0 \]  

(3.23a)
\[
\begin{align*}
    w_{lj} + \sum_{t=0}^{i-2} z_{lj} + \sum_{t=1}^{i-1} \left[ \sum_{k=0}^{i-t} q_{tj} + \sum_{km} p_{tm} \right] \\
    + z_{lj} + \sum_{t=1}^{i-l} q_{i-t,j} - \sum_{j} p_{im} \geq 0 \\
    \text{all } ij \\
    d_{im} - \sum_{t=0}^{k} \sum_{j} p_{i+t,m} \geq 0 \text{ all } ikm \\
    p_{im} \geq 0 \text{ (no change) all } ijk \text{m} \\
    w_{lj} + \sum_{t=0}^{I-1} z_{lj} + \sum_{t=1}^{I} \left[ \sum_{t=0}^{I-t} q_{tj} + \sum_{km} p_{tm} \right] \geq \varsigma_j \\
    a_{l-1,ef} r_{lef} + \sum_{t=1}^{i-l} a_{t}^{*} \left[ \Delta_{i-t,ef} + b_{i-t,ef} \right] \leq b_{ief} \text{ all } ief
\end{align*}
\]

where \( a_{ief}^{*} \) is defined by \((4.8)\), \( \Delta_{ief} \) and \( \varsigma_j \) are defined as in Chapter III, and where \( r_{lef} \) is assumed known.

Observe that by regrouping terms, \((3.24a)\) may be
written as

\[ w_{i+1,j} + \sum_{t=0}^{i-1} z_{ij}^t + \sum_{t=1}^{i} \left[ \sum_{k=0}^{i-t} q_{tk}^j - \sum_{km} p_{km}^{jk} \right] \geq 0 \]

which, from (4.11), is simply the requirement that on-hand stock of item j at the beginning of period i+1 cannot be negative, i.e., \( w_{i+1,j} \geq 0 \).

**Decomposable Linear Systems**

**Structure**

Let us again assume that the system of equations (3.31)-(3.32) consists entirely of linear functionals and that demands may be treated as deterministic. In this case, we assume that all items in the system may be classified into a limited number of "families" of items, and that a given item belongs to one and only one family. We assume that a demand for a member of a given family might possibly be satisfied by some other item in the family, but that the demand cannot be satisfied by a shipment of an item from any other family. In other words, items within a given family may be substitutes for one another, but items in different families are not substitutes.

In this case, we may classify the constraint equations of (3.31)-(3.32) into two types: (I) constraints involving only one family of commodities and (II) constraints involving
two or more commodity families. Type I constraints include the material balance equations, the non-negativity requirements, the equations linking demands for items in a given family to the stocks of items within the family, limitations on individual ordering and shipping decisions, and capacity and funding limitations (if any) that apply to a single family of items. Type II constraints consist of the remaining funding and capacity limitations.

The Type I constraints associated with a given family of items \( S \) include

(A) the stock balance equations,

\[
\begin{align*}
\frac{d_i}{d_i} &= w_{ij} + q_{ij} + z_{ij} - \sum_{km} p_{jm}^{jk} \\
\end{align*}
\]  

(4.12)

(B) limitations on individual ordering and shipping decisions:

\[
\begin{align*}
{x_i}^0_{im} &= d_{im} \\
{x_i}^k_{im} &= x_{im}^{k-1} - \sum_{j \in S_{im}} p_{jm}^{j,k-1} & l \leq k \leq K_p \\
{x_i}^k_{im} &= 0 & \text{otherwise} \\
\end{align*}
\]  

(4.13)

\[
\begin{align*}
0 &= q_{ij}^0 = w_{ij} \\
0 &= q_{ij}^k = z_{ij}^k & l \leq k \leq K_q
\end{align*}
\]  

(4.14)

(4.15)

(4.16)
\[ 0 \leq q_{i}^{k+} \leq q_{\text{max}}^{k} \quad (4.17) \]
\[ 0 \leq \lambda_{j}^{k} \quad (4.18) \]

(C) minimum end-of-period stocking constraints

\[ w_{i+1,j} \equiv \gamma_{j} \quad (4.19) \]

(D) limitations that shipments can neither exceed orders nor on-hand stocks

\[ w_{ij} + q_{ij}^{0} + z_{ij}^{0} - \sum_{km} p_{im}^{jk} \geq 0 \quad (4.20) \]
\[ x_{im}^{k} - \sum_{j \in S} \lambda_{j}^{k} \geq 0 \quad (4.21) \]

where the above constraints include only those equations with item subscripts \( j \) included in the set of items \( S \) that are members of the family. The Type I constraints also include capacity and funding limitations and associated resource balance equations (if any) that apply to a single family.

Type II constraints consist of the remaining balance equations and funding and capacity constraints.

In the preceding section, we noted that the number of variables and constraint equations required to define the problem might be reduced by expressing the status variables as functions of the decision variables, thus making the
explicit statement of the balance equations redundant. Similar comments are appropriate here. In our remaining discussion, we assume such actions have been taken, although we will utilize the status variable notation when it is convenient to do so.

Letting $A_f$ denote the coefficient matrix of the set of Type I constraints that apply only to family $f$, the set of all Type I constraints may be written as:

$$A_f Y_f = B_f \quad f = 1, 2, \ldots, F$$

(4.22)

where $Y_f$ and $B_f$ are column vectors whose elements correspond to the decision variables and to the constraint limitations, respectively, associated with commodity family $f$, and where $F$ denotes the total number of such families. We assume that appropriate slack and artificial variables have been added to the system (4.22) to convert any inequality constraints associated with commodity family $f$ into equivalent equality constraints.

The Type II constraints will include all equations involving more than one family of items. In matrix form, they will appear as:

$$\bar{A}_0 \bar{Y}_0 + \bar{A}_1 \bar{Y}_1 + \bar{A}_2 \bar{Y}_2 + \ldots + \bar{A}_F \bar{Y}_F + IS - IR = \bar{B}$$

(4.23)

where $\bar{A}_0$ and $\bar{Y}_0$ will be explained below, $\bar{A}_f$, $f = 1, 2, \ldots, F$ corresponds to the matrix of coefficients of Type II constraints associated with commodity family $f$, $\bar{Y}_f$, $f = 1, 2, \ldots, F$
is the corresponding column vector of decision and status variables, and \( \mathbf{B} \) is the column vector corresponding to the right-hand side of the Type II constraint equations. \( \mathbf{I} \) is identity matrix and \( \mathbf{S} \) and \( \mathbf{R} \) are column vectors of slack and artificial variables, respectively, required to insure equality in the system of equations.

The symbol \( \mathbf{Y}_0 \) in (4.23) denotes the vector of reallocation variables (i.e., the \( b^+_{\text{ief}} \) and \( b^-_{\text{ief}} \) variables defined by equation (4.2) and \( \mathbf{A}_0 \) denotes the corresponding matrix of coefficients. The product \( \mathbf{A}_0 \mathbf{Y}_0 \) thus defines the role of the reallocation variables in the Type II constraints.

Let \( \mathbf{C}_f \) denote the column vector of utility coefficients associated with the decision vector \( \mathbf{Y}_f, f=1,2,...,F \), and let \( \mathbf{C}_0 \) denote the column vector of utility coefficients corresponding to the reallocation variables \( \mathbf{Y}_0 \). Then the problem (3.31)-(3.33) may be written in matrix form as:

\[
\text{minimize } Z \\
Z = \mathbf{C}_0' \mathbf{Y}_0 + \mathbf{C}_1' \mathbf{Y}_1 + \mathbf{C}_2' \mathbf{Y}_2 + ... + \mathbf{C}_F' \mathbf{Y}_F \quad (4.24)
\]
Subject to the constraints

\[
\begin{align*}
A_0 X_0 + A_1 Y_1 + A_2 Y_2 + \ldots + A_F Y_F + IS - IR &= B_0 \quad (4.25) \\
A_1 Y_1 &= B_1 \\
A_2 Y_2 &= B_2 \\
&\vdots \\
A_F Y_F &= B_F \quad (4.26) \\
S, R, Y_f, \bar{Y}_f &\geq 0 \quad (4.27)
\end{align*}
\]

Assuming (4.25) represents L Type II constraints, and that the coefficient matrix \( A_f \) associated with family \( f \) consists of \( m_f \) rows and \( n_f \) columns, then equations (4.22)-(4.27) represent a linear program of size \((L + \sum_{f=1}^{F} m_f) \times (2L + n_0 + \sum_{f=1}^{F} n_f)\), where \( n_0 \) denotes the number of columns in the coefficient matrix \( A_0 \).

Although this problem could be solved using Dantzig's simplex procedure, the great number of computations required for problems of this size and the limited memory capabilities of present day computers are two of the practical problems that are encountered. Hence, we seek alternate methods that will overcome these difficulties.

The Parametric Decomposition Algorithm

The parametric decomposition algorithms developed by Orchard-Hays (1968, p. 271-295) and his colleagues presents one approach for solving linear programming problems with
the special block angular structure of (4.24)-(4.27). This algorithm is also known as the block-product algorithm because of its resemblance to the product form of the inverse method. The block-product algorithm is based on two main techniques, one well known, and the other rather new. The overall strategy of this approach is to first solve a modified problem in which \( \bar{B} \) has been altered and then to use the well known parametric right-hand-side algorithm to return to the original problem. This strategy is sometimes employed even when decomposition is not a consideration. The new technique is a means to maintain the inverse of the decomposition matrix in compact form, thus reducing the amount of computer memory required to store and update the basis during iterations of the parametric right-hand-side algorithm.

Since details of the computational procedure for the block-product algorithm are rather involved, they will not be presented here. The interested reader is referred to Orchard-Hays.

Perhaps the most important feature of the parametric decomposition algorithm (at least to the practicing operations analyst who seeks numerical solutions) is that the procedure has been incorporated into the linear programming packages of at least two major computer manufacturers. The algorithm has been implemented in the LP/600 system for General Electric's 625/635 line of computers and in the OPTIMA Mathematical Programming System for Control Data's
6400/6600 line. If either of these systems is available to the analyst, the problem of solving (4.24)-(4.27) is reduced to that of preparing input in proper form for the mathematical programming system being employed and in specifying the formats for displaying problem solutions, since the difficult questions of data manipulation and computer coding required to implement the algorithm have already been resolved. In practical terms, this generally means a significant reduction in the amount of effort required to obtain a solution. Another advantage of the block-product algorithm is that parametric postoptimality procedures are quite easy to implement since they are in essence merely a continuation of the original solution procedure. Such procedures are often quite useful for investigating the sensitivity of the model results to changes in the coefficients of the model. Parametric postoptimality procedures might also provide a computationally efficient method for obtaining the solution to a new problem by starting with a previously-obtained solution to a similar problem.

The Dantzig-Wolfe Decomposition Algorithm

Method. The Dantzig-Wolfe decomposition algorithm (1961) presents an alternate approach for reducing the dimensionality of linear programming problems with the special block angular structure of (4.24)-(4.27). Rather than reducing computer memory requirements through compact methods
for storing and updating the inverse basis of a standard linear program as is done in the Orchard-Hays procedure, the Dantzig-Wolfe procedure reduces dimensionally by decomposing the original problem into a series of smaller optimization problems. Basically, the Dantzig-Wolfe principle decomposes the general problem above into:

(a) A series of subproblems associated with each family of commodities which ignore joint constraints, and

(b) A master program which determines the optimum combination of subproblem results subject to all constraints.

Assuming the solution space is bounded, the set of feasible solutions to the system of linear equations

$$A_f Y_f = B_f \tag{4.28}$$

form a convex set for all $f$. Thus any $Y_f \succeq 0$ solving $A_f Y_f = B_f$ may be represented as a convex combination of the set of extreme points of the set of feasible solutions to (4.28). Since the set of feasible solution vectors $W_f$ of different basic solutions, $W_f = \{W_{f1}, W_{f2}, \ldots, W_{fH_F}\}$, defines the finite set of extreme points, we can represent any solution $Y_f$ by

$$Y_f = \sum_{h=1}^{H_f} \lambda_{fh} W_{fh} \tag{4.29}$$

where $H_f$ represents the number of different extreme point solutions to (4.28), and $\lambda_{fh}$ represents a scalar quantity
such that

\[ \sum_{h=1}^{H_f} \lambda_{fh} = 1 \quad \lambda_{fh} \geq 0 \quad \text{all } fh \quad (4.30) \]

Conversely, any solution \( Y_f \) represented by (4.29) is feasible for (4.28).

Hence, from (4.29) the contribution to the criterion function associated with the solution \( Y_f \) is

\[ C'_f Y_f = C'_f \begin{bmatrix} H_f \\ \sum_{h=1}^{H_f} \lambda_{fh} W_{fh} \end{bmatrix} \quad (4.31) \]

Rearranging terms,

\[ C'_f Y_f = \sum_{h=1}^{H_f} \lambda_{fh} C'_f W_{fh} = \sum_{h=1}^{H_f} \lambda_{fh} \mathbf{c}_{fh} \quad (4.32) \]

where \( \mathbf{c}_{fh} \) denotes the cost or utility associated with the extreme point \( W_{fh} \). Similarly, the contribution of the solution \( Y_f \) to the Type II constraint equations may be written as

\[ A'_f Y_f = \sum_{h=1}^{H_f} \lambda_{fh} A'_f W_{fh} = \sum_{h=1}^{H_f} \lambda_{fh} D_{fh} \quad (4.33) \]

where the kth element of \( W_{fh} \) corresponds to the kth element of \( Y_{fh} \). The vector \( D_{fh} \) thus represents the contribution of the extreme point solution \( W_{fh} \) to the Type II constraints. For example, if the rth joint constraint is a restriction on
available floorspace then the rth element of $D_{fh}$ denotes that floorspace requirement of the hth extreme point solution associated with family f.

From the above discussion, the problem (4.24)-(4.27) may be expressed in terms of $\lambda_{fh}$ as

$$Z = \sum_{h=1}^{H_o} \lambda_{oh} \phi_{oh} + \sum_{h=1}^{H_1} \lambda_{1h} \phi_{1h} + \ldots + \sum_{h=1}^{H_f} \lambda_{fh} \phi_{fh} \quad (4.34)$$

subject to the Type II constraints

$$\sum_{h=1}^{H_o} \lambda_{oh} D_{oh} + \sum_{h=1}^{H_1} \lambda_{1h} D_{1h} + \ldots + \sum_{h=1}^{H_f} \lambda_{fh} D_{fh} = B \quad (4.35)$$

and the convexity requirements

$$\sum_{h=1}^{H_1} \lambda_{1h} = 1$$

$$\sum_{h=1}^{H_2} \lambda_{2h} = 1 \quad (4.36)$$

$$\sum_{h=1}^{H_F} \lambda_{fh} = 1$$

where we define $\lambda_{oh}$ and $\phi_{oh}$, respectively, as the hth elements of the vectors $Y_o$ and $C_o$, and where $D_{oh}$ denotes the
hth column of $\overline{A}_o$. Note that the variables $\lambda_{oh}$ are not subject to a convexity constraint in (4.36). For notational convenience, we assume that the last $L$ elements of the set $\{D_{oh}\}$ are the $L$ unit vectors associated with the vector of slack variables $S$, and that the last $L$ elements of the set $\{\lambda_{oh}\}$ denote the corresponding slack variables.

The formulation (4.34)-(4.36) generated from the extreme point solutions of $A_{f}Y_{f} = B_{f}$ is called the full master program to (4.24)-(4.27).

Assuming there are $L$ Type II constraints, the extremal problem (4.34)-(4.36) is a linear programming problem with $(L + F)$ rows and a very large number of columns. The number of columns is probably far too numerous to be explicitly expressed; however, using the revised simplex method only those vectors which the simplex method brings into the respective bases are needed at any one time.

Assume we have a basic feasible solution to (4.34)-(4.36) and assign dual variables $\theta = (\theta_1, \theta_2, \ldots, \theta_L)$ to the $L$ rows of (4.35) and $\alpha = (\alpha_1, \ldots, \alpha_F)$ to the $F$ rows of (4.36). The according to the revised simplex column-selection criterion (Dantzig(1963), p. 210), the criterion function might be improved if the relative cost $\Pi_{fh}$ associated with some extreme-point solution $W_{fh}$ is negative; i.e., the criterion function might be improved if

$$\Pi_{fh} = \varepsilon_{fh} - \theta D_{fh} - \alpha_f < 0 \quad (4.37)$$
for some $W_{fh}$, where $\alpha_0 = 0$.

From (4.32) and (4.33), note that (4.37) may be written as

$$\Pi_{fh} = (C_f' - \theta'\bar{A}_f)W_{fh} - \alpha_f$$  \hspace{1cm} (4.38)

for $f=1,2,...,F$, while for the $f=0$ case,

$$\Pi_{oh} = (C_o' - \theta'D_{oh})$$  \hspace{1cm} (4.39)

From (4.38), the extreme point solution associated with commodity family $f$ which produces the greatest rate of improvement to the criterion function may be found by solving the subproblem:

$$\text{minimize} \ (C_f - \theta\bar{A}_f)'W_{fh}$$  \hspace{1cm} (4.40)

subject to the constraints

$$A_fW_{fh} = B_f$$  \hspace{1cm} (4.41)

$$W_{fh} \geq 0$$

Note that any appropriate procedure may be employed to solve this subproblem.

The problem (4.40)-(4.41) may be interpreted as the problem of identifying the set of procurement and shipping actions associated with commodity family $f$ which results in the minimum cost or disutility to the supply system, where the direct and implied costs associated with the $t$th element of $W_{fh}$ is specified by the $t$th element of the vector
\((C_f - \Theta \bar{A}_f)\). Hence, we might term the extreme point \(W_{fh}\) as the hth supply plan for commodity family f.

The subproblem (4.40) (4.41) is a linear programming problem with \(m_f\) rows and \(n_f\) columns, and may thus be solved using Dantzig's simplex algorithm. In many cases, however, the subproblem may possess special structural characteristics which may be exploited to develop solution procedures more efficient than the general methods.

In particular, if (4.41) contains no budget or capacity limitations and if initial stocks are sufficiently low that disposal and termination actions may be ruled out a priori, then the subproblem may be interpreted as the problem of finding a set of minimum cost (or maximum profit) flows of commodities through time, and may thus be treated as an uncapacitated network flow problem.

For example, Figure 4.1 illustrates the network representation of the subproblem for the case in which there is only one item in the commodity family, one demand class for the item, and where deliveries may be made either in period i or in period i+1 if the order is placed at the beginning of period i. In Figure 4.1, it is assumed demands may be backordered for up to one period; if not filled by that time, the demands are lost. The nodes labeled "L" denote dummy shipments which compensate for the lost demand. For subproblems in which there are several items and several demand classes, the associated network representation is no longer
planar, and is difficult to display pictorially.

A number of efficient procedures for solving network flow problems have been developed, including the redefinition of the problem into the standard transshipment or transportation problem format. A number of these techniques are discussed by Dantzig (1963). When (4.41) contains budget or capacity constraints, the subproblem may often be interpreted as a capacitated multi-commodity flow problem. Procedures for solving such problems are presented by Tomlin (1966), Glassy (1967), and by Demmy (1969). The generalized upper-bounding technique of Dantzig and Van Slyke (1967), also known as the group algorithm, is also appropriate for solving the latter class of problems.

source:
deliveries:
on-hand:
shipments:
demand:

Figure 4.1. Network representation of subproblem for the single-item, single-demand class case.
In the special case in which (4.41) contains a relatively small number of extreme points, the subproblem might be solved by enumerating all the extreme points and evaluating each of these solutions using the modified cost vector \((C_f - \Theta A_f)\). This approach has been employed by Dzielinski, Baker and Manne (1963) to solve deterministic multi-item production planning problems. Since their method is applicable to certain cases of budget-constrained inventory problems, details of their procedure will be discussed later in this chapter.

Thus, a number of procedures for solving the subproblems associated with a given commodity family are available. In a given application one might utilize a solution procedure that is most efficient for the particular structural characteristics of the commodity subproblem being considered.

Suppose we have solved the subproblem (4.40)(4.41) for \(f = 1, \ldots, F\), and let us denote the extreme-point solution to the \(f\)th subproblem by \(W_{fh}^*\). Then the relative cost \(\prod_{fh}^*\) associated with the extreme point solution \(W_{fh}^*\) may be evaluated by substitution into (4.38). Similarly, suppose we have evaluated the relative costs \(\prod_{oh}\) directly using (4.39), and let \(\prod_{oh}^*\) denote the lowest value of \(\prod_{oh}\).

If the relative cost factors \(\prod_{fh}^*\) for all \(f\) are non-negative, then increasing any \(\prod_{fh}^*\) in the extremal program will cause no improvement to the criterion function (4.34). In this case, an optimal solution to the extremal problem
has been reached and the solution \( Y = (Y_0, Y_1, \ldots, Y_F) \), where, for \( f = 1 \), \( Y_f \) is given by (4.29), is the optimal solution to (4.24)-(4.27).

On the other hand, if \( \prod f^* h^* \) is negative for some \( f \), say \( f^* \), the current solution might be improved by introducing the variable \( \lambda_f^* h^* \) into the basis. This might most readily be done using a standard simplex pivot step. Introducing a new variable into the basis results in a new set of dual variables \( \theta, \alpha \), which in turn lead to a new set of subproblems which are identical to the subproblems of the preceding iteration except for the modified cost vector \((C_f - \theta A_f)\). The new subproblems may then be solved to determine new candidates for entry into the basis of the extremal problem, and the process continues until an optimal solution is reached.

Equations (4.34)-(4.36) are written as if all extreme-point solutions to the subproblems are known throughout the solution process. At any given time in the process, however, all that is required to continue the solution procedure is the set of extreme points associated with the inverse of the basis vectors and the extreme point \( W_{f^*}^* h^* \) about to enter the basis. Thus, we may obtain an equivalent to the extremal problem (4.34)-(4.36) by dropping all columns except those in the inverse of the basis and the column about to be introduced. The reduced linear programming tableau is termed the restricted master program.
Obtaining an Initial Basic Feasible Solution. To obtain an initial basic feasible solution to (4.34)-(4.36), one may apply the same technique used by Tomlin (1966) for solving multi-commodity flow problems. Thus, we shall add artificial variables $\lambda_i, i = 1, \ldots, F$, to the $F$ rows of (4.36), and minimize $V$,

$$V = \lambda_1 + \lambda_2 + \ldots + \lambda_F$$

subject to

$$\begin{align*}
\sum_{h=1}^{H_0} \lambda_{oh} D_{oh} + \sum_{h=1}^{H_1} \lambda_{1h} D_{1h} + \ldots + \sum_{h=1}^{H_F} \lambda_{Fh} D_{Fh} &= B \\
\sum_{h=1}^{H_1} \lambda_{1h} &= 1 \\
\sum_{h=1}^{H_2} \lambda_{2h} + \lambda_2 &= 1 \\
\ldots
\sum_{h=1}^{H_F} \lambda_{Fh} + \lambda_F &= 1
\end{align*}$$

$$\lambda_{fh}, \lambda_f \geq 0 \hspace{1em} f = 0, \ldots, F$$
Note that if the solution to (4.42)-(4.45) yields $V = 0$, a basic feasible solution to (4.32)-(4.36) will have been obtained.

The solution to (4.42)-(4.45) may be obtained using a decomposition procedure similar to the procedure described above. As an initial basis to the linear programming problem (4.42)-(4.45) we may use the L + F slack and artificial vectors. Assigning new simplex multipliers $\sigma = (\sigma_1, \ldots, \sigma_L)$ and $\varphi = (\varphi_1, \ldots, \varphi_F)$, column $fh$ will decrease the value of $V$ if

$$-1 -\sigma' \overline{w}_{fh} - \varphi_f < 0$$

The subproblem to be solved is thus:

minimize $-\sigma' \overline{w}_{fh}$

subject to

$$A_f w_{fh} = B_f$$

$$w_{fh} \geq 0$$

Again, any appropriate solution technique may be used to solve this problem.

The General Procedure. In general, the procedure for solving linear programs using the Dantzig-Wolfe decomposition algorithm involves the following steps. First, a feasible solution which satisfies all constraints is obtained, and from this solution modified cost coefficients are generated
for each of the commodity family subproblems. Each of the subproblems is then solved independently using the generated cost coefficients to obtain a minimum-cost solution which ignores joint (Type II) constraints. The master program is then solved to determine the minimum-cost convex combination of these extreme point solutions which satisfies both sectional and joint constraints. The dual variables associated with the optimal solution to the master program may then be used to define new cost coefficients for each of the subprograms. These new subprograms are then solved, and an optimality test is applied.

If the relative cost associated with any of the subprogram solutions is negative, the criterion function may be decreased by incorporating the corresponding extreme point solutions into the master program. The process continues until no further improvements to the criterion function are possible.
Generalizations of the Dantzig-Wolfe Algorithm

Orchard-Hays (1968, p. 249) points out that the Dantzig-Wolfe decomposition algorithm described above may be generalized in at least two major senses. First, the decomposition may be compounded. For example, the sub-problems may themselves be decomposable; in particular, when the subproblems may be interpreted as network flow problems, they may be solved by decomposing the problem into a series of shortest route problems (see, for example, Tomlin (1966) and Rao (1968)). Alternately, the master problem \((4.34)-(4.36)\) might also possess a decomposable structure. Dzielinski and Gomory (1965) consider a large scale production planning problem in which this is the case. Second, the various subproblems may be of different forms and, in particular, they may be non-linear. Orchard-Hays notes that the Dantzig-Wolfe decomposition principle may be applied if the subproblems possess the following characteristics:

A.1. The \(\theta\) vector can be interpreted as a set of Lagrange multipliers, or at least optimality in the subproblem can be determined in terms of \(\theta\).

A.2. If the subproblem is not optimal, an improved candidate can be generated for the master problem in the form of a column of coefficients in linear equations.
A.3. A convex combination of candidate columns from a given subproblem can be reinterpreted to provide meaningful answers to the subproblem. To see this, consider the problem

\[
\max c(x) \quad (4.46)
\]

subject to

\[
g_i(x) \leq b_i \quad i=1,2,\ldots,L \quad (4.47)
\]

\[
x \in S \quad (4.48)
\]

where \( S \) is an arbitrary set in \( \mathbb{R}^n \) and \( c \) and \( g_i \) are real-valued functions. Let \( g = (g_1,\ldots,g_L) \) and \( b = (b_1,\ldots,b_L) \) and assume the constraint set is not empty. We shall refer to \((4.46)-(4.48)\) as problem P1.

Suppose we are given a set of grid points \( x^j \in S, j=1,2,\ldots,p \), and define \( c(x^j) = c^j \) and \( g_i(x^j) = g_i^j \). Using this information, a linear approximation to P1 on the grid of points \( (x^1,\ldots,x^p) \) is

\[
\max \sum_{j=1}^{p} \lambda_j \quad (4.49)
\]
\[ \sum_{j=1}^{\rho} g^j_i \lambda_j \leq b_i, \quad i=1,2,\ldots,L \quad (4.50) \]

\[ \sum_{j=1}^{\rho} \lambda_j = 1 \quad \lambda_j \geq 0 \quad (4.51) \]

We call (4.49)-(4.51) the "approximating problem," which we denote as problem P2.

Let \( \Theta^{p+1} = (\Theta^p_1, \ldots, \Theta^p_L) \) and \( \alpha^{p+1} \) denote, respectively, the optimal dual variables associated with the constraints (4.50) and (4.51). Now suppose that we can identify an additional grid point \( x^{p+1} \in S \) such that

\[ c(x^{p+1}) - \Theta^{p+1} g(x^{p+1}) - \alpha^{p+1} > 0 \quad (4.52) \]

Then the criterion function to P2 may be improved by introducing the column \((c^{p+1}, g^{p+1}_1, \ldots, g^{p+1}_L, 1)\) into the basis. Using the standard simplex column selection procedure, if there are several points which satisfy (4.52), the variable associated with the particular point which solves

\[ \text{Maximize } c(x) - \Theta^{p+1} g(x) - \alpha^{p+1} \quad x \in S \quad (4.53) \]

would be selected for entry into the basis, since that variable produces the greatest rate of improvement to the objective function. This process might then be repeated
until no additional $x \in S$ can be found that satisfies (4.52). At that time, the computation terminates.

Brooks and Geoffrion (1966) have suggested that P2 might be employed to determine Everett's (1963) generalized Lagrange multipliers for problem P1. Recall that Everett's procedure was discussed at some length in Chapter II. In the context of that discussion, Brooks and Geoffrion's suggestion is that the optimum dual variables associated with a given grid $(x^1, \ldots, x^p)$ provide an appropriate choice of Lagrange multipliers in step 5 of Everett's procedure (see Chapter II) for use in the next iteration. Also, the identification and construction of a particular column satisfying (4.53) to enter the basis of P2 is identical to steps 2 and 3 of Everett's procedure. Brooks and Geoffrion point out that if the set S contains N points, where N is finite, the use of P2 to find Everett's generalized Lagrange multipliers is finitely convergent to the optimum solution $\lambda^* = (\lambda^*_1, \ldots, \lambda^*_N)$ of P2.

In the case where the set S is not a finite discrete set, the analysis above is complicated by the fact that there are an infinite number of variables in P2. However, Dantzig (1963, Chapter 24) has shown that if S is a bounded convex set, c is concave, g is convex and P2 has a non-degenerate basic feasible solution, then P2 converges to an optimal solution to P1. If the constraint set (4.47) contains no strict equalities, then the non-degeneracy requirement is
not necessary to prove convergence. Unfortunately, the convergence may be asymptotic.

To see how the optimal solution to P2 relates to the original problem P1, recall the discussion of "payoff-resource space" in Chapter II. Observe that (4.53) is equivalent to finding the maximum of a Lagrangian function with multipliers \( \theta \) and \( \alpha \). Hence, the maximizing point lies on the upper envelope in the \((L+1)\)-dimensional payoff-resources space of the set of points \( \{c(x^j), g_1(x^j), \ldots, g_L(x^j)\} \), \( j=1, \ldots, N \). If these points are sufficiently dense near the boundary of their convex hull then some of the policies \( x^j \) corresponding to \( x^* \neq 0 \) (and there will be no more than \( L+1 \) of these) will be good approximate solutions to P1.

Now consider the problem

\[
\max \ c_0(x_0) + c_1(x_1) + \ldots + c_F(x_F) \quad (4.54)
\]

subject to

\[
\bar{g}_{oi}(x_0) + \bar{g}_{li}(x_1) + \ldots + \bar{g}_{Fi}(x_F) = \bar{b}_i \quad \text{all } i \quad (4.55)
\]

\[
g_{lk}(x_1) = b_{lk} \quad \text{all } k \quad (4.56)
\]

\[
x_F \in \mathbf{F} \quad (4.57)
\]
where $c_f$, $g_{fi}$ and $g_{fk}$ are continuous real-valued functions, and $\mathcal{F}$ is an arbitrary set. We denote this problem as P3. Note that P3 may be interpreted as a cell problem with additional constraints for each cell specified by (4.56).

Let $S_f$ denote the set of points $x_f \in \mathcal{F}$ that satisfy

$$g_{fk}(x_f) \leq b_{fk} \quad k=1,2,\ldots,K \quad (4.58)$$

Observe that P3 is equivalent to P1 if we define $x = (x_0, x_1, \ldots, x_F)$, $c(x) = \sum_{f=0}^{F} c_f(x_f)$ and $g_i(x) = \sum_{f=0}^{F} g_{fi}(x_f)$

where $S = S_0 \times S_1 \times \ldots \times S_F$, i.e., $S$ is the direct product set of the $S_f$. Hence, from our preceding discussion, if we have a set of points $x^j = (x^j_0, x^j_1, \ldots, x^j_F)$, $j=1,2,\ldots,p$, where $x^j \in S$, a linear approximation to P3 takes the form

$$\text{maximize} \quad \sum_{j=1}^{p} \lambda_j \left[ \sum_{f=0}^{F} c_f(x^j_f) \right] \quad (4.59)$$

subject to

$$\sum_{j=1}^{p} \lambda_j \left[ \sum_{f=0}^{F} g_{fi}(x^j_f) \right] \leq b_i \quad i=1,\ldots,L \quad (4.60)$$

$$\sum_{j=1}^{p} \lambda_j = 1 \quad (4.61)$$

$$\lambda_j \geq 0$$
We denote the approximating problem (4.59)-(4.61) as problem P4.

Since P4 is a special case of P1, the solution procedure discussed above also applies here. In this case, however, since the sets $S_f$ are disjoint the column-selection sub-problem (4.53) separates into a series of $F+1$ subproblems of the form:

$$\text{Maximize} \quad c_f(x_f) - \sum_{i=1}^{p+1} \theta_i \bar{g}_{fi}(x_f) \quad (4.62)$$

By the definition of $S_f$, however, (4.62) is equivalent to the constrained optimization problem

$$\text{Maximize} \quad c_f(x_f) - \sum_{i=1}^{p+1} \theta_i \bar{g}_{fi}(x_f) \quad (4.63)$$

subject to

$$g_{fk}(x_f) \leq b_{fk} \quad \text{all } k \quad (4.64)$$

Hence, in solving the approximating problem to P3, an improving column (if it exists) may be found by solving a series of cell problems, one for each $f$. Since any appropriate technique might be employed to solve the subproblem (4.63)-(4.64) for a given $f$, computational efficiency might be improved by employing a solution procedure that exploits the peculiar structural characteristics of the subproblem under consideration. Examples of special structures and solution procedures that may be present when P3 is a
budget-constrained inventory problem will be presented later in this chapter.

Suppose we have solved (4.63)-(4.64) for all F, and let $x_f^{p+1}$ denote the solution of the subproblem $f$. Then from (4.52), if

$$
\sum_f c_f(x_f^{p+1}) - \sum_1 \Theta_i \left[ \sum_f \bar{g}_{fi}(x_f^{p+1}) \right] - \alpha > 0 \quad (4.65)
$$

the criterion function might be improved by entering the column

$$
\begin{bmatrix}
\sum_f c_f(x_f^{p+1}) \\
\sum_f \bar{g}_{f1}(x_f^{p+1}) \\
\vdots \\
\sum_f \bar{g}_{fl}(x_f^{p+1}) \\
1
\end{bmatrix}
$$

into the basis of $P_4$. If (4.65) is not satisfied, the computational process terminates.
In the Appendix, we present a sample problem illustrating the computational procedure.

Nemhauser and Widhelm (1970), Orchard-Hays (1968, p. 251), and Wagner (1969, p. 616) note that on large problems the procedure described above may converge very slowly. In particular, Nemhauser and Widhelm (1970, p. 4) note that when the Dantzig-Wolfe procedure is applied to large problems, near optimal solutions may be achieved rather quickly, but thereafter progress towards achieving or verifying an optimal solution is often extremely slow. Orchard-Hays notes that he has seen many computer runs where scores or even hundreds of iterations occurred with the objective function changing in the fourth or fifth decimal place.

On the other hand, despite the possibility of slow convergence, very large problems have been solved using the above techniques. As an example, Dzielsinski and Gomory (1965) report that a three-period production planning problem of the above structure with 26 joint constraints linking 963 items (i.e., \( F = 963 \)) was solved in 24 minutes on an IBM 7090/7094. Further, Orchard-Hays suggests that problems involving up to 50 joint constraints and 10,000 subproblems, with each \( S_f \) containing 5 or 6 points, may be solved rather easily with a good computer code; however, he does not present data on the computation times that might be experienced for problems of this size.
Wagner (1969, p. 613) suggests that for separable problems such as P3, convergence might be improved by employing a linear approximation of the payoff-resource space for each cell. In this case, the linear approximating problem takes the form

\[
\text{maximize} \quad \sum_{f} \sum_{j} \lambda_{fj} C_{f}(x^{j}_{f}) \tag{4.66}
\]

subject to

\[
\sum_{f} \sum_{j} \lambda_{fj} \bar{E}_{fi}(x^{j}_{f}) \leq \bar{b}_{i} \quad i=1, \ldots, L \tag{4.67}
\]

\[
\sum_{j} \lambda_{fj} = 1 \quad f=0, \ldots, F \tag{4.68}
\]

\[
\lambda_{fj} \geq 0 \quad \text{all } f, j
\]

where \( \{ \lambda_{fj} \} \) denotes the set of weights associated with cell \( f \). Note that (4.66)-(4.68) exactly parallels the master program for the linear decomposable problem considered earlier in this chapter. Also observe that the above formulation is a linear programming problem with \((L+F)\)-row constraints, whereas problem P4 has \((L+1)\)-row constraints. If \( F \) is large, the basis to (4.66)-(4.68) will be significantly larger than the basis to P4; however, as we will discuss below, if
Generalized Upper Bounding techniques (GUB) are employed in solving (4.66)-(4.68), the differences in computer memory requirements and computation times between the two formulations may be negligible.

In addition to the possibility of improved convergence characteristics, the formulation (4.66)-(4.68) has another desirable feature; namely, when F is large relative to L, as it will be in many budget-constrained inventory problems, the effect of characteristic A.3 above may be negligible. That this is so may be seen by considering the nature of the basis to the master program (4.66)-(4.68).

\[
\begin{bmatrix}
G_{01} & \ldots & G_{0H_0} & G_{11} & \ldots & G_{1H_1} & \ldots & G_{F1} & \ldots & G_{FH_F} \\
1 & \ldots & 1 & 1 & \ldots & 1 & 1 & \ldots & 1
\end{bmatrix}
\]

Figure 4.2. The coefficient matrix of the master problem.

Figure 4.2 illustrates the coefficient matrix of the linear program (4.66)-(4.68). Pictorially, the matrix appears as a set of F rows of 1's with an upper border of L-dimensional column vectors \( G_{fj} \) that describe the contribution of the jth extreme point solution associated with family
f to the Type II constraints. Symbolically,

\[
G_f = \begin{bmatrix}
\bar{g}_{f1}(x_f^1) \\
\bar{g}_{f2}(x_f^2) \\
\vdots \\
\bar{g}_{fL}(x_f^L)
\end{bmatrix}
\]

Linear programming models with the structure illustrated in Figure 4.2 are known as "group" problems. The term "semi-transportation" has also been used to describe linear programs of this structure because the lower portion of the coefficient matrix of the classical transportation problem consists of similar rows of 1's.

Observe that for each commodity family \( f \) (i.e., \( f \neq 0 \)), the \( \lambda_{fj} \) are non-negative and sum to one; hence they are all effectively bounded to the range \((0,1)\). Now notice that at least one \( \lambda_{fj} \) from each commodity group must be in any valid solution, and hence, if it is the only basic variable from the group, its value is one. Not counting the variable \( Z \), which is always in the basis, there are \( L \) more basis positions. These must be filled from the \( H_o \) columns associated with the \( \lambda_{0j} \) variables or with "extra" columns from each group. If \( F \) is greater than \( L \), at most \( L \) groups can have more than one column in the basis. If for example, there are 50 Type II constraints \((L=50)\) and there are 1,000 commodity
families (F=1000), then at least 950 groups will have only one column in the basis. For these groups, the basis variable will have a value of one, and all other variables \( \lambda_{fj} \) associated with the group will be zero. For the remaining groups, i.e., those with more than one basis member, a simple rounding rule or other heuristic technique might be employed to obtain meaningful subproblem solutions if a convex combination of the basis members is not meaningful. Dzielinski, Baker, and Manne (1963) describe one way in which this may be done. Often, a solution procedure which provides exact solutions for a majority of groups and easily computed heuristic solutions for a small number of groups will provide a near-optimal solution that is good enough for practical purposes.

For the interested reader, Orchard-Hays (1968, p. 224) discusses some methods of resolving the tactical problems encountered in coding the GUB technique for computing equipment.

Differential Management

One of the most important aspects of the Dantzig-Wolfe decomposition procedure in the context of budget-constrained inventory problems is that the subproblems may take many different forms and that different procedures may be used to obtain subproblem solutions. Most real-world inventory systems manage large numbers of items with diverse characteristics, and seldom can a simple model be developed that adequately describes all items in the system and that
can also be solved quickly and easily. The model developed in Chapter III may be sufficiently general to describe most items of a given system adequately, but the associated sub-problems may be quite costly to solve. The Dantzig-Wolfe procedure provides an appropriate mechanism to obtain computation efficiency by using different models and solution procedures for different classes of items, while still assuring that the overall problem is solved.

Brown(1963), Hadley and Whitin(1963, p.424) and others have observed that in most military, retail, and industrial inventory systems a very small fraction of the total number of items stocked account for a very large fraction of the total dollar value of business involved. Frequently, something like ten percent of the items will account for eighty to ninety percent of the dollar volume. For military supply systems with a large number of technical items with high obsolescence probabilities, even more extreme situations may be observed. If the Dantzig-Wolfe decomposition procedure were used to compute ordering and shipping decisions in such systems, a very detailed model might be used to describe item families with high dollar impact, and rather simple, easy-to-solve models (such as those to be discussed in Chapter VI) might be employed for item families with low dollar impact. Implementing a sophisticated model to describe high impact items will often provide more than enough savings to compensate for the increased computing and data handling costs
associated with such a model. On the other hand, if there are
a large number of low impact items, such items might best be
managed by employing simpler, but easier to solve models,
perhaps even at the cost of a little extra inventory. Reduc­
ing the computational costs associated with a large number
of items may produce significant savings in computation costs.
If each item treated in this manner has a very low dollar re­
quirement, the cost in extra inventory required to produce
equal service will be less than these savings, resulting in a
net improvement to the system. Of course, the idea of apply­
ing varying degrees of management attention to items with
varying characteristics is not new; differential management
techniques have been applied in inventory systems for many
years. What is new is the capability to automatically coor­
dinate these activities through use of the Dantzig-Wolfe
procedure.

General Considerations

Although the Dantzig-Wolfe decomposition algorithm
may be generalized to handle non-linear subproblems, the use­
fulness of the approach depends upon the ease with which the
subproblems may be solved. Generally, non-linear problems
are much more difficult to solve than their linear counter­
parts. In certain situations, however, it may be possible to
develop efficient computational algorithms or even closed­
form expressions which provide the required subproblem
solutions with a small amount of computational effort. For example, in some instances the solution to a subproblem may be obtained in a given iteration quite efficiently if one uses the solution from the preceding iteration as a starting point. Because of the generality of the Dantzig-Wolfe technique, it is beyond the scope of this work to review all the forms that subproblems might take on; however, in this chapter and the next we will discuss some important special cases which may appear in many budget-constrained inventory problems.

In many inventory systems, it is an adequate approximation to the real situation to assume that:

1. There is only one item per commodity class and one demand class per item.

2. Only one order may be placed per period and there is only one possible delivery lead time per item.

3. Termination and disposal decisions may be ruled out a priori.

4. All demand may be backlogged until filled, and the cost or disutility of a backorder is independent of its age.

When the above conditions hold, the system of Type I constraints defined by (4.12)-(4.21) reduce to the system

\[ w_{i+1,j} = w_{ij} + q_{i-k,j} - d_{ij} \quad (4.69) \]

\[ w_{i+1,j} = \delta_j \quad (4.70) \]

\[ 0 \leq q_{ij} \leq q_{\text{max}} \quad (4.71) \]
where negative values of \( w_{ij} \) are interpreted as backorders, \( y_j \) denotes inventory of item \( j \) required to be on-hand at the end of the planning horizon, \( k_j \) denotes the number of periods in the lead time for item \( j \), and where \( d_{ij} \) denotes the demand for item \( j \) during period \( i \).

**The Lot Size Programming Problem**

The Problem. Consider a production-inventory system in which the above conditions hold. Further, assume the Type II constraints take the following form:

\[
W_{ki}^1 + W_{ki}^2 \leq M_k \\
W_{ki}^3 + W_{ki}^4 \leq M_k \quad \text{all } k, i \quad (4.72)
\]

\[
W_{ki}^5 + W_{ki}^6 \leq M_k
\]

\[
\sum_r w_{ki}^r - W_{ki}^+ + W_{ki}^- - \sum_r W_{ki}^{r,i-1} = 0 \quad \text{all } k, i \quad (4.73)
\]

\[
\sum_j y_{jk} (q_{ij}) - \sum_r H_{kr_k} W_{ki} \leq 0 \quad \text{all } k, i \quad (4.74)
\]

where \( W_{ki}^r \) denotes the number of workers of skill Type \( k \) and payment class \( r \) working in period \( i \). The superscript \( r=1 \) means straight time for the first shift and \( r=2 \) means straight time and overtime on the first shift. The superscripts \( 3,4,5,6 \) refer to the same payment classes on the second and third shifts. \( H_{kr_k} \) is the number of hours this type worker will work in a period and \( M_k \) is the maximum number of workers of Type \( k \) that can be employed during the planning period. \( W_{ki}^r \) is the number of workers hired into
labor type $k$ during period $i$, and $W^{-}_{ki}$ is the number laid off.
The variable $q_{ij}$ denotes the quantity of item $j$ ordered in
period $i$, and $g_{jk}(q_{ij})$ denotes the man-hours of skill type $k$
required to process the order.

Equations (4.72) limit the number of workers that can
be used on any one shift, and equations (4.73) provide that
work force changes can occur only through hiring and lay-off.
Equations (4.74) link the labor required to the labor avail­
able. Observe that equations (4.72), (4.73), and (4.74) are
simply special cases of the general resource capacity and
balance equations presented in Chapter III. In this case,
we interpret the resource as workers.

Consider an inventory system in which the constraints
(4.69)-(4.74) hold. Suppose the inventory manager wishes to
identify the set of ordering, hiring, and lay-off decisions
that satisfy all forecasted requirements at a minimum variable
operating cost, i.e., backorders are not permitted. Symboli­
cally, we wish to find a set of $q_{ij}$, $W^{+}_{ki}$, $W^{-}_{ki}$, and $W^{r}_{ki}$ which
minimize $Z$,

$$Z = \sum_{ij} c_{ij}(q_{ij}) + \sum_{ikr} R^{r}_{ki} W^{r}_{ki} + \sum_{kr} n^{+}(W^{+}_{ki}) + n^{-}(W^{-}_{ki}) \quad (4.75)$$
subject to (4.69)-(4.74) and to the backorder constraint:

\[ w_{ij} \geq 0 \quad \text{all } ij \quad (4.76) \]

where \( c_{ij}(q_{ij}) \) denotes the cost of procuring \( q_{ij} \) units of item \( j \) in period \( i \), \( R_{ki}^{r} \) is the wage rate of workers of type \( k \) and payment class \( r \) and \( \tilde{\gamma}_{ki}^{+}(W_{ki}^{+}) \) and \( \tilde{\gamma}_{ki}^{-}(W_{ki}^{-}) \) are, respectively, functions defining the hiring and firing costs in period \( i \) associated with labor class \( k \).

Observe that when backlogging is not permitted, the end-of-period inventory requirement (4.70) will be satisfied if we simply replace the demand \( d_{ij} \) for item \( j \) during period 1 with the requirement for \( d_{ij}' = (d_{ij} + s_{j}) \) units during period 1. Hence, in our following discussion, we assume \( d_{ij}' \) has been substituted in (4.69) for \( d_{ij} \) and that (4.70), which is then redundant, has been dropped from the problem.

The above problem is known as the economic lot size programming problem. This problem was initially studied by A. S. Manne (1958). Additional studies have been reported by Dzielski, Baker, and Manne (1963) and Dzielski and Gomory (1965). More recently, Gornstein (1970) has shown how to augment the above problem to account for interactions among assemblies and their component parts in a manufacturing process. All of these authors have restricted their attention to the case where delivery lead times are negligible (i.e., \( k_{j}=0 \) for all \( j \)) and \( q_{\max} = \infty \). They also assume that the
cost function $Z$ is linear, and hence that $C_{ij}(q_{ij})$, 
$\bar{F}_{ki}(w_{ki}^+), \bar{F}_{ki}(w_{ki}^-)$ are linear, and that $
\bar{g}_{jk}(q_{ij})$ takes the form:

$$
\bar{g}_{jk}(q_{ij}) = \begin{cases} 
0 & \text{if } q_{ij} = 0 \\
 a_{jk} + b_{jk} q_{ij} & \text{if } q_{ij} > 0 
\end{cases} \quad (4.77)
$$

where $a_{jk}$ and $b_{jk}$ denote, respectively, the setup and production time per unit required of skill type $k$ to procure or produce a lot of item $j$. Until noted otherwise, we will assume these conditions hold for the remainder of this section.

Dzielinski and Gomory reformulated the above problem into the form of problem P4. This requires multiplication of (4.75) by $-1$ to obtain an equivalent maximization problem and the introduction of auxiliary variables $\lambda_h$ corresponding to the decision vectors $Q_h = (Q_{i1}^h, Q_{i2}^h, ..., Q_{ij}^h)$, where $Q_{ij}^h = (q_{ij}^h, ..., q_{ij}^h)$ denotes the production schedule for item $j$ associated with the $h$th grid point included in the approximating problem P4. Dzielinski and Gomory term $Q_{ij}^h$ the $h$th "production plan", since their work is primarily concerned with a production system which manufactures rather than procures its inventories. If an item is procured from outside suppliers, we might term $Q_{ij}^h$ as the $h$th procurement schedule for the item.
Dzielinski and Gomory observed that when the lot size problem is stated in the form of P4, the column selection problem associated with the \( \{q\} \) variables decomposes into a series of \( J \) single-item problems of the form

\[
\text{Minimize} \quad \sum_i \left[ A_{ij} \delta(q_{ij}) + B_{ij} q_{ij} \right] \quad (4.78)
\]

subject to (4.69)-(4.71), where

\[
A_{ij} = - \sum_k \theta_{ij} a_{jk} \; \quad (4.79)
\]

\[
B_{ij} = c_{ij} - \sum_k \theta_{ij} b_{jk} \quad (4.80)
\]

where \( \theta_{ik} \) denotes the dual (Lagrangian) variable associated with the \((i,k)\)th row of (4.74), \( \delta(\cdot) \) denotes the Dirac delta function, and \( c_{ij} \) denotes the linear material cost associated with \( q_{ij} \). This result follows if one substitutes (4.77) into (4.63) and rearranges terms. The problem of minimizing a functional of the form of (4.78) subject to the requirement that the demands \( d_{ij} \) in each period are satisfied is precisely the dynamic economic lot size problem considered by several authors, including Bhatia and Garg(1960), Kernar(1961), Wagner and Shetty(1962), and Wagner and Whitin(1958). Since Dzielinski and Gomory follow the approach of Wagner and Whitin, we will also.
Characteristics of an Optimal Schedule. Manne (1958) observed that an optimal solution to the dynamic lot size problem (4.78) is characterized by the fact that ordering (production) may take place in a period only if no inventory enters that period. Thus, if ordering occurs in period $r$ and next in period $s > r$, and if these ordering decisions are components of the optimal solution to the dynamic lot size problem, then

$$q_r = \sum_{t=r}^{s-1} d_t$$

where $d_t$ denotes the requirement in period $t$ and $q_r$ is the order quantity in period $r$. That is, under an optimal policy the order placed in period $r$ is just sufficient to satisfy requirements in periods $r, r+1, \ldots, s-1$. If $d_t > 0$ for all $t$, then an optimal solution to the dynamic lot size problem is one of the $2^{I-1}$ schedules of this form.

One method to determine a particular schedule that minimizes (4.78) is to simply enumerate all $2^{I-1}$ schedules of the above form, to evaluate all of these schedules using the cost parameters specified by (4.79)(4.80), and to select the particular schedule which solves (4.78). This approach was employed by Dzielsinski, Baker, and Manne (1963) in earlier work on the multi-item lot size problem. When $I$ is large, however, say $I=12$, the enumeration approach requires a substantial amount of computation.
The Dynamic Lot Size Algorithm. Dzielinski and Gomory observed that the dynamic lot size algorithm devised by Wagner and Whitin (1958) provides an efficient method for determining an optimal schedule from the above set. Specifically, Wagner and Whitin devised an algorithm to determine the optimum solution to the problem of minimizing $C(Q)$,

$$C(Q) = \sum_{i=1}^{I} \left[ c_{i}(q_{i}) + h_{i}(w_{i+1}) \right]$$  \hspace{1cm} (4.81)

where $Q = (q_{1},...,q_{I})$ and where $c_{i}(.)$ and $h_{i}(.)$ are concave continuous, real-valued functions defined on $(0, +\infty)$ and satisfy

$$c_{i}(q) = K_{i} \delta(q) + c_{i}q \text{ and } h_{i}(w) = h_{i}w$$  \hspace{1cm} (4.82)

for non-negative $q$ and $w$. In the above, $\delta(q)$ denotes the Dirac delta function, $c_{i}(q)$ denotes the cost of ordering $q$ units in period $i$ and $h_{i}(w)$ denotes the cost of holding $w$ units of stock in period $i$. Observe that in (4.78), $h_{i}(w)=0$. Actually, as noted by Veinott (1966, p. 749), (4.82) is not necessary to the proof of Wagner and Whitin's procedure and so will not be imposed in discussing their algorithm.

Let $g_{rs}$ denote the cost of ordering in periods $r$, $r+1,\ldots,s-1$ when the order(production) in period $r$ satisfies the requirements for those periods.
Hence,

\[ g_{rs} = c_r(D_r, s-1) + \sum_{t=r+1}^{s-1} c_t(0) + \sum_{t=r}^{s-1} h_t(D_{t+1}, s-1) \]  \hspace{1cm} (4.83)

where

\[ D_{rs} = \sum_{t=r}^{s} d_t \]  \hspace{1cm} (4.84)

where \( d_t \) denotes the requirement for period \( t \). Let \( f_i \) be the minimum cost over periods \( i, \ldots, I \) when an order is placed in period \( i \). Then, defining \( f_{I+1} = 0 \), the recursive solution of

\[ f_i = \min_{r < s < I+1} \left[ g_{rs} + f_s \right] \]  \hspace{1cm} (4.85)

for \( i=1,2,\ldots,I \) yields the optimal schedule. Wagner and Whitin(1958) also establish planning horizon theorems under the additional hypothesis of (4.82).

Dzielinski and Gomory note that the recursion (4.85) requires approximately \( I(I+1)/2 \) computations and comparisons, and that backtracking to determine the \( q_{ij} \) that give the optimal cost is only a very small additional calculation. This is the amount of calculation that is substituted for evaluating the \( 2^{I-1} \) scalar products required by the enumerative approach.
Forming the Production Plan. After the subproblems (4.78) are solved for each item, the results may be accumulated to construct the composite vector that is a candidate for entry into the basis. If the relative cost of this vector given by (4.62) is positive, the objective function might be improved by entering the vector into the basis. If the relative cost is not positive, and if none of the variables $W_{ki}^+$, $W_{ki}^-$, or $W_{ki}$ have positive relative cost, the process terminates. Otherwise, the vector with the greatest relative cost may be entered into the basis, thus defining new values for the dual variables, and the cycle repeats.

Computational Results. Dzielsinski and Gomory developed an experimental computer code for performing the grand cycle of calculations described above. The program is a combination of several FORTRAN and FAP programming language routines, and was designed for use on an IBM 7090/94 Data Processing System. Details of the program are given by Dzielsinski and Gomory (1965).

Table 4.2 summarizes computational experience reported by Dzielsinski and Gomory. As indicated in the table, it appears that problems involving around 400 items, 20-30 joint constraints and 6 planning periods may be solved in less than 10 minutes of IBM 7090 time. For many inventory systems, the capability to solve problems of this size might be sufficient to manage the requirements of a large fraction of the total dollar activity of the system.
Table 4.2. Gomory and Dzielsinski's computational experience with the lot-size programming problem.

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: Parameters that Determine Size of Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) I: Number of Items</td>
<td>35</td>
<td>963</td>
<td>428</td>
<td>428</td>
<td>428</td>
<td>428</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(2) K: Number of Facilities</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(3) I: Number of Periods</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>(4) S: Number of Shifts</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>(5) (2<em>K</em>I)*(S+1): Work Force Columns</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>60</td>
<td>84</td>
<td>96</td>
<td>72</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>(6) (K*I)^3: Slack Columns</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>30</td>
<td>42</td>
<td>48</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>(7) (S+2)-(K-I)+2 Number Interacting Constraint Rows</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>42</td>
<td>58</td>
<td>66</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>(8) (2^I-1)*I: Item Production Scheduled</td>
<td>140</td>
<td>3847</td>
<td>1712</td>
<td>6784</td>
<td>27392</td>
<td>54434</td>
<td>204800</td>
<td>204800</td>
<td>204800</td>
</tr>
</tbody>
</table>

II: Size of Problems as Ordinary Linear Programming Problems

| (9) (1)+(7): Number of Rows | 61 | 989 | 454 | 466 | 486 | 494 | 150 | 150 | 150 | 150 |
| (10)(5)+(6)+(8) Number of Columns | 194 | 3901 | 1766 | 6874 | 27518 | 54928 | 204908 | 204908 | 204908 | 204908 |

continued
Table 4.2 - Continued

<table>
<thead>
<tr>
<th>III: Size of Problems with Decomposition and Dynamic Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td>(11) (7): Number of Rows........  27 27 27 43 59 66 51 51 51 51</td>
</tr>
<tr>
<td>(12) (5)+(6): Number of Columns  54 54 54 90 126 144 108 108 108 108</td>
</tr>
<tr>
<td>(13) (1) X (I(I+1)/2): Number of Elementary Steps in DP ALG.............. 210 .5778 .2568 .6420 .11984 .15408 .7800 .7800 .7800 .7800</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IV: Computations of Problems with Decomposition and Dynamic Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td>(14) Total Time to Compute Optimal Solution (Minutes of 7090 time)........... 3.36*# 24.5*# 5.66* 9.0* 10.70* 17.92* 5.90* 8.87* 8.07* 9.00*</td>
</tr>
<tr>
<td>(15) Total Phase II Iterations...... 48 34 22 35 52 65 45 49 46 51</td>
</tr>
<tr>
<td>(16) Average time to create a Production Plan (Minutes)........... 0.07# 0.72# 0.264 0.345 0.483 0.510 0.183 0.225 0.223 0.215</td>
</tr>
<tr>
<td>(17) Number of Production Plans Created........ 16 12 19 24 20 29 26 34 31 36</td>
</tr>
</tbody>
</table>

*This does not include 7090 time to setup the problem and output the results.

#This problem was run using a version of the dynamic programming subroutine that was later improved. The improved routine, used in the remaining problems, requires about \( \frac{1}{2} \) less time than the original routine.
Other Special Cases

General Considerations. As noted previously, one of the most important aspects of the Dantzig-Wolfe decomposition procedure in the context of budget-constrained inventory problems is that the sub-problems may take different forms and that different procedures may be used to obtain subproblem solutions. When the system of Type I constraints take the form of (4.69)-(4.71), the subproblems take the form of deterministic, dynamic, single-item inventory problems. When the subproblem objective function is concave and backlogging is not permitted, the Wagner-Whitin algorithm provides an efficient mechanism for obtaining subproblem results, as was discussed above. When the subproblem objective function is not concave and/or backlogging is permitted, however, alternate solution procedures must be employed.

The extensive literature of inventory theory abounds with computational algorithms for solving important special cases of single item inventory problems. Any of these methods may be appropriate in certain situations to determine subproblem solutions. In this section, we will briefly mention procedures which are quite likely to appear in budget-constrained inventory problems.
Concave Holding Costs. Consider the lot-size programming problem discussed above. Suppose that in addition to the constraints discussed above the Type II constraints include limitations on the total dollar value of inventory held in each of the $I$ periods of the planning horizon. Symbolically, these constraints might be written as

$$\sum_j c_j w_{ij} \leq M_i \quad \text{all } i$$

(4.86)

where $c_j$ denotes the book value of item $j$, $w_{ij}$ is the amount of stock of item $j$ on hand before ordering at the beginning of period $i$, and $M_i$ denotes the investment limit. Let $\sigma_i$ denote the dual variable corresponding to the $i$th constraint of (4.86). In this case, the column selection problem leads to a series of single-item subproblems of the form

$$\min \sum_i \left[ A_{ij} f(q_{ij}) + B_{ij}q_{ij} + \sigma_i c_j w_{ij} \right]$$

(4.87)

subject to (4.69)-(4.71), where the symbols of (4.87) are defined above. Observe that this problem is of the form (4.81), with $h_i(w_i) = \sigma_i c_j w_{ij}$. Hence, the dual variable $\sigma_i$ may be interpreted as the implied inventory holding cost associated with period $i$. Since (4.87) is of the form (4.81), the Wagner-Whitin algorithm is appropriate for solving subproblems of this form.
Dynamic Lot Size Algorithm with Backlogging. In the above models, we sought policies which satisfied all forecasted requirements at the lowest variable operating costs. Under this policy, backorders are not permitted.

Wagner (1969, p. 303) presents a generalization of the dynamic lot size algorithm which may be used to determine optimal single-item procurement schedules when backlogging is permitted (or is unavoidable due to Type II constraints). Assume that $h_i(w_i)$, the cost of holding $w_i$ units during period $i$, is concave for both $w_i=0,1,2,...$ and for $w_i=0,-1,-2,...$, where negative values of $w_i$ denote backorders; however, $h_i(w_i)$ need not be concave for $..., -2, -1, 0, 1, 2,...$. What continues to hold true is that $q_i$ is either zero or the sum of demands in consecutive periods containing $q_i$; but here the sum can include demands prior to, as well as beyond, period $i$. The key observation is that an optimum schedule is still characterized by a sequence of regeneration points at which entering inventory is zero. Ordering (production) occurs only once between these periods. The definition of $grs$ is generalized to

$$grs = \text{minimum total cost of ordering once during the periods } r, r+1, ..., s-1 \text{ to meet all demand in this interval.}$$

Thus an intermediate optimization is required to find the
costs:

\[ g_{rs} = \begin{cases} 
  c_r(d_r) \quad & \text{for } s=r+1 \\
  \min_{r \leq t < s-1} \left[ c_t(D_{t,s-1}) + \sum_{u=r+1}^{s-1} c_u(o) \right] + h_{r+1}(w_{r+1}) + \ldots + h_{s-1}(w_{s-1}) \quad & \text{otherwise}
\end{cases} \]  

(4.88)

In (4.88), the beginning inventory \( w_r \) is zero, and for the other periods on-hand inventory is determined by the recursion \( w_{i+1} = w_i + q_i - d_i \).

If one is solving an inventory problem subject to a Type II constraint on the number of backorders in the system, this constraint will appear in the subproblem as a component of the \( h_i(w) \) function for negative values of \( w \).

**Subproblems with Convex Costs.** Now consider a situation in which the subproblem objective function takes the form of (4.81), but now assume that the ordering and holding cost functions \( c_i(q_i) \) and \( h_i(w_i) \), respectively, are convex. Convex ordering (production) costs arise when there are several sources of limited production at different unit costs during a period. If one uses these sources up to capacity in order of ascending unit costs, the resulting production schedule is convex in the total amount produced (Veinott, 1966, p. 746). A similar remark applies to convex storage
costs. In the convex cost situation it is quite easy to impose upper bound limitations on on-order quantities \( q_i \) or on-hand inventory \( w_i \). Specifically, one may simply define \( c_i(q_i) \) and \( h_i(w_i) \) as infinitely large for values of \( q_i \) or \( w_i \), respectively, that exceed the upper limits. This convention preserves the convexity of these functions, while guaranteeing that the associated values of the variables will not appear in an optimum solution. Our problem may then be stated as: Given \( c_i(q) \) and \( h_i(w) \) are, respectively, convex functions for all non-negative real values of \( q_i \) and \( w_i \). Find \( Q=(q_1,q_2,\ldots,q_I) \), where \( q_i \geq 0 \) for all \( i \), such that \( C(Q) \) is minimized,

\[
C(Q) = \sum_{i=1}^{I} \left[ c_i(q_i) + h_i(w_{i+1}) \right]
\]  

subject to

\[
w_i \geq 0 \quad i=1,\ldots,I
\]  

where

\[
w_{i+1} = w_i + q_i - d_i \quad i=1,\ldots,I
\]  

and where \( w_1 \) is known, as is the demand \( d_i \) in period \( i \), \( i=1,2,\ldots,I \). The constraints (4.91) are the usual material balance equations while (4.90) reflects the requirement that all demands must be satisfied, i.e., backorders are not
permitted. Later in this section we will discuss how the algorithm might be extended to treat the backlog case.

We shall present an algorithm of Wagner (1969, p. 295) for solving the above problem. An extensive set of references for solving generalized versions of (4.89)-(4.91) for convex $C(Q)$ is presented by Veinott (1966, p. 746). Wagner's algorithm for solving (4.89)-(4.91) is quite simple. In brief, the algorithm starting in period 1 proceeds period-by-period to fill each unit of demand required as cheaply as possible, given the production already scheduled and the resultant inventory. The details of the algorithm for the above problem are as follows:

Step 1. Let $p$ be the earliest period in which the current demand requirement value is $d_p \geq 0$. For each of the periods 1, 2, ..., $p$ consider increasing production by one unit in the current trial schedule so that one unit of $d_p$ is filled.

Step 2. For each of the possible $p$ revisions, calculate the entire incremental cost from the increased production and inventory holding. Select an alternative with minimum incremental cost and revise the trial schedule accordingly. If there is more than one such alternative, schedule production in as late a period as possible.

Step 3. Reduce the current value of $d_p$ by one unit. Examine whether the current values of all $d_p$ have now been reduced to zero. If so, stop; otherwise return to Step 1.

An illustration of the above algorithm is presented
by Wagner (1967, p. 295-299). Wagner also notes that only a few simple changes in the wording of the above algorithm are necessary to treat the backlog case, although doing so increases the computational burden considerably. Let a backlog be denoted by a negative value for $w_i$. Now assume that $h_i(w_i)$ is convex for all values of $i$, that is, for $-2, -1, 0, 1, 2, \ldots$. Alter Step 1 so that one considers increasing production by one unit for each of the periods $1, 2, \ldots, I$. Alter Step 2 to include the incremental cost from backlogging. The bulk of added computation arises in Step 2 in calculating the entire incremental cost for each possible revision. Given the production amounts in a revised schedule, one must allocate them to the demands already considered so that the corresponding inventory and backlogging costs are minimal. Step 3 remains unchanged.

Additional Techniques. In the preceding paragraphs, we have reviewed several special cases which may arise in problems of inventory management and for which particularly efficient solution procedures are available. The literature of inventory theory abounds, however, with efficient solution procedures for other special situations that may arise in real-world environments. For an excellent review of many of the techniques that are available, the reader is referred to Veinott (1966).
CHAPTER V
PROBABILISTIC MODELS

Introduction

In the preceding chapter, we discussed solution procedures for situations in which the future is assumed certain. In several important circumstances, however, essentially the same procedures developed in Chapter IV may be employed for solving inventory management problems in which the uncertainties associated with future events are explicitly recognized. The purpose of this chapter is to review some of these special cases.

In the preceding chapter, we discussed solution procedures for maximizing $Z$,

\[
Z = c_0(x_0) + c_1(x_1) + \ldots + c_F(x_F) \quad (5.1)
\]

subject to

\[
\begin{align*}
\bar{g}_{1i}(x_1) + \bar{g}_{2i}(x_2) + \ldots + \bar{g}_{Fi}(x_F) &= \bar{b}_i & i=1, \ldots, I \\
g_{1k}(x_1) + \ldots & = b_{lk} & k=1, \ldots, K \\
\vdots \quad \cdot \quad \cdot \quad \cdot & \quad = b_{FK} \\
\bar{g}_{FK}(x_F) & = b_{FK} \\
x_F & \in \Gamma_f 
\end{align*}
\]

(5.3)
where \(c_f, \bar{c_i}, \) and \(g_{fk}\) are real-valued functions and \(F\) is an arbitrary set. In our previous discussion, we assumed that all aspects of the decision problem were known and we then sought efficient procedures to identify optimal policies in the given deterministic situation.

In this chapter, we will consider solution procedures for situations in which at least some elements of (5.1)-(5.3) are random variables. In general, this necessitates a reformulation of the objective function \(Z\), since \(Z\) is a random variable if any of its components are random variables and "minimization" of a random variable is meaningless. Hence, the functional (5.1) must be replaced by an appropriate deterministic function. Similarly, if any of the elements of (5.2)(5.3) are random variables, the corresponding constraints must be reinterpreted. In the following sections, we will discuss alternate objective functions and constraint interpretations that may be appropriate in special cases of inventory management problems. In these discussion, we are particularly interested in identifying special cases of stochastic versions of (5.1)-(5.3) which are separable, and thus susceptible to the same type of optimization procedures discussed in the previous chapter.
Types of Objectives

As noted above, if any of the components (5.1)-(5.3) are random variables, we must in general replace $Z$ by an appropriate deterministic function, since (5.1) is a random variable if any of its elements or elements of the constraints (5.2)(5.3) are random variables. There are many choices for this function, each of which may be appropriate under certain circumstances. Some of the most commonly employed types of objective functions are:

1. maximize the expected value of the objective function ($E$ model)

2. minimize the expected squared error about some preferred value ($V$ model)

3. maximize the probability of obtaining at least a certain specified value of the objective function ($P$ model)

The expected-value criterion is particularly appropriate for situations in which a given decision must be repeated many times, or when many semi-independent decisions must be made and no one decision affects a significant portion of the objective function value or requires a significant portion of the limited resources which link the otherwise independent decision areas. The minimum-variance criterion is appropriate for situations in which an inventory manager is assigned a set of "targets", rather than a set of
absolute constraints, and small deviations about the target values are permissible but large deviations from the target values are highly undesirable. The third objective function corresponds to the "satisficing" or aspiration level criterion suggested by H. A. Simon (1957, Chapters 14 and 15). In this approach, the objective function is based upon a set of values which, if achieved, would be regarded as satisfactory. In this approach, an optimal policy is one which is most likely to equal or exceed the specified set of aspiration levels.

Charnes and Cooper (1963) have established deterministic equivalents in the form of convex programming problems for the above objectives for the special cases in which $Z$ is linear, i.e., in which

$$ Z = DX $$

(5.5)

where $D$ denotes a random vector with known distribution function and $X$ denotes the vector of decision variables. Rao (1968) discusses the above objectives for the case in which $Z$ is linear in the context of the classical warehousing problem. In this case, the vector $D$ is interpreted as the set of random purchase and selling prices prevailing in future time frames and the vector $X$ denotes corresponding quantities of goods to be purchased or sold in those periods. $Z$ is then the profit associated with the warehouse operation.

Because of the large number of variables involved in
most problems of inventory management subject to budget con-
straints, it is important that the objective function (and 
constraints) be separable to reduce the computational effort 
and computer memory required to find an optimal solution.
As we shall see below, the above objectives lead to separable 
criterion functions in several commonly-encountered situa-
tions.

The E-Model

Suppose our objective is to maximize the expected 
value of $Z$ defined by (5.1). In this case, since expectation 
is a linear operator,

$$E[Z] = E\left[\sum_{f} c_f(x_f)\right] = \sum_{f} E\left[c_f(x_f)\right]$$  \hspace{1cm} (5.6)

that is, the expected value of a sum equals the sum of the 
expected values. If for a given $f$, $c_f(x_f)$ is statistically 
independent of the values $x_h$ for all $h \neq f$, then (5.6) is 
separable by family $f$.

The V-Model

Now suppose that our objective is to minimize the 
expected squared deviation of the value measure $Z$ about some 
target value $W$. Symbolically, we wish to minimize

$$E \left[(Z-W)^2\right] = E \left[\left(\sum_{f} c_f(x_f)-W\right)^2\right]$$  \hspace{1cm} (5.7)
For notational convenience, let $u^f = E[c_f(x_f)]$ for a given value of $x_f$. Then (5.7) may be written as

$$E[(Z-W)^2] = E\left[\left(\sum_f c_f(x_f) - \sum_f u^f\right)^2\right]$$

$$+ 2E\left[\left(\sum_f c_f(x_f) - \sum_f u^f\right)\left(\sum_f u_f - W\right)\right]$$

$$+ \left(\sum_f u_f - W\right)^2$$

(5.8)

since the second term on the right of (5.8) vanishes by the definition of $u^f$. In the special case in which $c_f(x_f)$ is statistically independent of the values of $x_h$ for $h \neq f$, (5.9) may be written as

$$E[(Z-W)^2] = E\left[\left(\sum_f c_f(x_f) - \sum_f u^f\right)^2\right]$$

$$+ \left(\sum_f u_f - W\right)^2$$

(5.9)

$$= \sum_f \sigma^2_f + \left(\sum_f u_f - W\right)^2$$

(5.10)

where $\sigma^2_f = E\left[(c_f(x_f) - u_f)^2\right]$, i.e., $\sigma^2_f$ denotes the variance of $c_f(x_f)$. The term $(\sum_f u_f - W)^2$ is called the bias.

Notice that in general (5.10) is not separable by family $f$ because of cross-product terms associated with the bias, even when the families are statistically independent. However, observe that if the decision variables $x_f$ are selected so that the bias equals some specified constant $B$,
i.e., so that

\[ E \left[ \sum_f c_f(x_f) - W \right] = \sum_f u_f - W = B \quad (5.10a) \]

then the resulting function is separable by item family. Hence, if the item families are statistically independent, we may replace the objective function (5.7) by the modified objective

\[
\text{minimize} \quad \sum_f \sigma_f^2 + B^2 \quad (5.11)
\]

subject to

\[
\sum_f u_f - B = W \quad (5.12)
\]

where the minimization is over B as well as over the decision vectors \( x_f \). Note that the functionals in (5.11) and (5.12) are each separable.

The P-Model

Now suppose our objective is to maximize the probability of attaining at least a given value \( W \) for the objective function \( Z \). Symbolically, we may write this objective as

\[
\text{Max Prob} \left[ Z = \sum_f c_f(x_f) \geq W \right] \quad (5.13)
\]

Deterministic equivalents for (5.13) when \( Z \) is a linear functional (i.e., when \( Z \) takes the form of (5.5)) in the form of a linear functional subject to a convex constraint
set have been developed by Charnes and Cooper (1963) and by Rao (1968, p. 132). Unfortunately, the equivalent formulations discussed by these authors are not separable and manipulations of their results to induce separability are not readily apparent. Analysis of other functional forms of $Z$ leads to similar difficulties.

Suppose, on the other hand, that aspiration levels $W_f$ are established for each item family. In this case, a reasonable objective might be to maximize the sum of the probabilities that each of these aspiration levels are attained. Symbolically, this objective may be written as

$$\text{maximize } \sum_f \text{Prob} \left[ c_f(x_f) \geq W_f \right]$$

(5.14)

This criterion function is separable provided the utility measure $c_f(x_f)$ associated with a given family $f$ is independent of actions $x_h$ taken in family $h$ for all $h \neq f$. This is often the case. Suppose, for example, that an inventory manager aspires to fill all demands "off-the-shelf", that is, he aspires to fill all demands at the time of occurrence from on-hand stocks. Such an objective may be appropriate when demand cannot be backlogged, i.e., when all demand during a stockout condition is lost. In this case, $\text{Prob} \left[ c_f(x_f) \geq W_f \right]$ might be interpreted as the probability that a randomly-selected demand for a member of item family $f$ occurring during the planning horizon will be filled "off-the-shelf".
In steady-state systems, this probability is often called the "fill rate", since in this case the probability is equivalent to the proportion of demands per unit time that are expected to be filled from on-hand stocks. Several authors, including Brooks, Gillen, and Lu (1969) and Brooks (1967), have employed this objective in studying stocking policies for several steady-state inventory systems.

The Constraint Set

Interpretations

When the system of relationships (5.2)(5.3) contains random elements, the corresponding constraints must be reinterpreted. As in the previous section, several interpretations are possible, depending upon the particular physical and procedural characteristics of the system of interest. Some of the commonly encountered interpretations are:

. the absolute interpretation
. the chance-constrained interpretation
. the expected-value interpretation

In the following sections, we will consider the analytical characteristics of constraints under each of the interpretations.
The Absolute Interpretation

Consider the constraint

\[ g(X,D) \leq b \]  \hspace{1cm} (5.15)

where \( g \) is a real, single-valued function, \( X \) is a vector of decision variables to be determined, \( D \) is a vector of random variables, and \( b \) is a scalar random variable. Since \( D \) is a random vector, \( g(X,D) \) is also a random variable. Let \( f_g(g) \) and \( f_b(b) \), respectively, denote the probability density functions for \( g(X,D) \) and \( b \). One interpretation of the constraint (5.15) is that the decision vector \( X \) must be selected so that the constraint (5.15) cannot be violated, regardless of the particular values of \( D \) and \( b \) that are eventually observed. This is equivalent to requiring that (5.15) hold with probability one. This interpretation requires that \( X \) be selected so that the tails of the probability density functions for \( g \) and \( b \) do not overlap, as illustrated in Figure 5.1. Technically, for certain forms of probability density functions it may be impossible to satisfy a constraint under an absolute interpretation for any value of \( X \). For example, if \( f_g(g) \) and \( f_b(b) \) are normal probability density functions, there is some finite area under the tail of each pdf for all values of \( x \). Hence, the curves must overlap. Hence, if \( f_g(g) \) and \( f_b(b) \) are normal, the chance-constrained interpretation of the next section may be more appropriate.
A deterministic equivalent of (5.15) under an absolute interpretation may be developed as follows. Let \((D_k, b_k)\) denote the \(k\)th possible combination of the random elements \(D\) and \(b\). Then (5.15) will be satisfied regardless of the particular values of \(D\) and \(b\) that are eventually observed if, and only if

\[
g(X, D_k) \leq b_k
\]  

(5.16)

for all \(k\). Illustrations of this approach for the case where (5.1)-(5.3) are systems of linear equations are presented by Cocks (1968) and by Hillier and Liberman (1967, p. 531-6). Observe that if this approach is employed for

\[ f_g(g) \]
\[ f_b(b) \]

\(g\) and \(b\)

Figure 5.1. The probability densities of \(g\) and \(b\) for a feasible solution to an absolute constraint.
each constraint of the system (5.2)-(5.3), the separable structure of the system is not changed, but the number of constraints may be substantially increased. In general, the total number of possible combinations \((D_k, b_k)\) that might be observed will be very large, if not infinite. In this event, finding a particular \(X\) that satisfies (5.16) for all \(k\) may be extremely difficult, if not impossible. In many situations, however, only a small subset of the constraints (5.16) need be considered explicitly. For example, consider the constraint

\[
dx \leq b
\]  

(5.17)

where \(b\) is uniformly distributed on the interval \((100, 400)\), \(d\) is uniformly distributed on the interval \((5, 10)\), and \(x\) is a scalar decision variable. In this case, since the probability densities of \(b\) and \(d\) are continuous, an infinite number of combinations of \(b\) and \(d\) are possible. Observe, however if the constraint

\[
10 \ x \leq 100 \  \text{or} \  x \leq 10
\]  

(5.18)

is satisfied, the constraint (5.17) will be satisfied for any combination of \((d, b)\) that are eventually observed. Hence, in this problem, only the constraint (5.18) need be explicitly considered to insure that (5.17) is satisfied.

An absolute interpretation is particularly appropriate for situations in which (5.15) denotes limits on the
physical capacity of the supply system. For example, a transportation system may have a fixed limit on the tonnage that may be handled in a given amount of time and a shipping facility may have capacity limitations which bound the maximum number of orders that may be processed in a given time frame. Because of possibilities of absenteeism or of machine breakdowns, however, the capacity of the transportation system or the shipping facility at a given point in time may be a random variable.

The Chance-Constrained Interpretation

The General Approach. In the previous section, we assumed that (5.15) must hold absolutely, i.e., the constraint must be satisfied for all possible outcomes for the random elements D and b. Let us now consider situations in which it is highly desirable, but not absolutely necessary, that the constraint hold. This interpretation may be particularly appropriate when (5.15) represents constraints imposed by the higher management. Often, such constraints may be violated if extreme situations develop and violations of the original constraints are justified. For example, Congress may grant supplemental budgetary authority during a given fiscal year to cover expenditures for unanticipated but justified requirements.

When it is highly desirable, but not absolutely necessary that a constraint involving random elements be
satisfied, deterministic equivalents may be obtained by applying the ideas of chance-constrained programming first advanced by Charnes and Cooper (1959) and Charnes, Cooper, and Symonds (1958). Extensions of this early work have since been presented by Charnes, Cooper, and Thompson (1964), Agnew, Agnew, Rasmussen, and Smith (1969), Resh (1970), Eisner, Kaplan, and Soden (1971), Rao (1968), and others.

In contrast to the interpretation of the previous section which required that the constraint (5.15) must hold for all possible observations of D and b, the chance-constrained programming interpretation requires only that the constraint must hold for most of the possible combinations of (D,b). Specifically, the chance-constrained programming approach considers a given decision vector X as feasible relative to a given set of constraints if and only if

$$\text{Prob} \left[ g_t(X,D) \leq b_t \right] \geq \alpha_t, \quad t=1,2,\ldots,T \quad (5.19)$$

where $\alpha_t$ denotes a specified constant between zero and one, and the subscript t denotes the tth constraint. Hence, under this formulation a given action vector X is considered feasible relative to the constraint if the constraint will "probably" be satisfied once the random variables D and b are known, where $\alpha_t$ denotes the required level of confidence associated with constraint t. In applications of chance-constraints to inventory management problems, the required level of confidence $\alpha_t$ is generally determined by implicit
judgmental processes. Note that in the special case for which \( \alpha_t = 1 \), (5.19) is equivalent to an "absolute" constraint.

In general, it is quite difficult to specify deterministic equivalents of (5.19). However, in the special case in which

(a) all elements of the vector \( D \) are known, either because \( D \) is not a random variable or because \( D \) is observed prior to the selection of the decision vector \( X \),

(b) the right-hand side elements \( b_t \) are independent random variables, and

(c) all elements of the decision vector \( X \) are selected before any of the random variables \( b_t \) become known,

deterministic equivalents of (5.19) may be obtained quite easily, as we shall see below.

**Random Right-Hand Side Case.** Suppose the \( b_t \) are the only random components of (5.19), and let \( F_t(\cdot) \) denote the distribution function of \( b_t \). Further, suppose all elements of the vector \( X \) must be determined before the value of any of the \( b_t \) are known. In this case, a deterministic equivalent of (5.19) is

\[
g_t(X,D) \leq F_t^{-1}(\alpha_t) \quad t=1,2,\ldots,T
\]

(5.20)

where \( F_t^{-1}(\cdot) \) denotes the inverse of \( F_t(\cdot) \) and where \( D \) is assumed known. \( F_t^{-1}(\alpha_t) \) is often called the \( \alpha_t \)th fractile of \( F_t(\cdot) \). Observe that if \( F_t(\cdot) \) is not known, Chebycheff's inequality or the Camp-Mendell inequality might be employed to obtain an (approximate) deterministic equivalent similar
to (5.20). For an example, see Rao (1968, p. 115).

Now consider the system of relationships (5.2)(5.3). Suppose the chance-constrained interpretation (5.19) is applied to each constraint relationship in this system. Further, assume that $b_1, b_2, \ldots, b_I; b_{11}, \ldots, b_{Fk}$ are independent random variables, that all other aspects of (5.2)(5.3) are known, and that all the decision variables $x_f, f=0, \ldots, F$ must be established before any of the $b_i$ or $b_{fk}$ values are known. In this case, a deterministic equivalent for (5.2) (5.3) is of the form (5.2)(5.3) with the right-hand sides $b_i$ and $b_{fk}$ replaced by the required fractiles of their respective distribution functions. Hence, in this case the deterministic equivalent also possess a separable structure.

**Random Right-Hand Side Joint Constraint Case.** In some applications it may be appropriate to specify a single joint constraint rather than a set of individual constraints of the form (5.19). For example, it may be desirable to specify a probability with which all cash liquidity and obligation constraints are to be met. If a joint constraint is to be specified, the system of constraints (5.19) is replaced by the single constraint

$$
\text{Prob} \left[ g_t(X, D) = b_t, t=1, 2, \ldots, T \right] \geq \alpha \quad (5.21)
$$

where $\alpha$ denotes the probability with which all constraints are to be satisfied. If the distributions of the random
elements $b_t$ are independent of each other and if the $b_t$ represent the only random components of (5.21), then a method suggested by Miller and Wagner (1955) can be used to obtain a deterministic equivalent of (5.21). A brief sketch of this method may also be found in Wagner (1969, p. 670).

Following Miller and Wagner's approach, let $H_t(y) = P\left[b_t \leq y\right]$. Then a deterministic equivalent of (5.21) under the above assumptions is

$$g_t(X,D) - y_t \leq 0 \quad t = 1, \ldots, T \quad (5.22)$$

$$\prod_{t=1}^{T} H_t(y_t) \geq \alpha \quad (5.23)$$

where the $y_t$ are unrestricted in sign. Relationship (5.22) defines the fractiles of $F_t(.)$ that the random variable $b_t$ must exceed if the constraint is not to be violated, and (5.23) denotes the deterministic equivalent of (5.21). Observe that (5.22) (5.23) possesses a separable structure.

Wagner and Miller note that (5.23) is non-linear and that rarely will it be concave, as is required by most non-linear algorithms. However, they note that a transformed version,

$$\sum_{t=1}^{T} \ln H_t(y_t) = \ln \alpha \quad (5.24)$$

is frequently concave -- for example, when the density for
$b_t$ is given by a Normal, Gamma, or uniform distribution.

The Expected-Value Interpretation

A third interpretation of a set of constraints involving probabilistic elements is that the constraints need be satisfied only over the long term, but that any given constraint might be violated in any given period. In this case, deterministic equivalents for the set of constraints

$$g_t(X,D) = b_t \quad t=1,\ldots,T \quad (5.25)$$

under this interpretation are given by the corresponding expectations, i.e., by

$$E \left[ g_t(X,D) \right] = E \left[ b_t \right] t=1,\ldots,T \quad (5.26)$$

assuming $b_t$ is statistically independent of $X$ for all $t$.

The expected-value interpretation is the basis of most of the probabilistic models of inventory management subject to budget constraints that have appeared in the operations research literature to date.

One reason for the popularity of the expected-value interpretation is that many real-world constraints are accurately described in this manner. Another major reason is that if one couples an expected-value constraint interpretation with the assumption of stationary demand processes, substantial mathematical simplifications may be possible. Quite often closed-form expressions may be obtained for
optimum stocking policies in these situations. The works of Gerson and Brown (1971) and Presutti and Trepp (1970) are examples of this approach.

It should be pointed out, however, that the operational characteristics of inventory systems subject to expected-value constraints may be significantly different from the operational characteristics of the same system under absolute operational constraints, even when demand processes are stationary. Systems subject to absolute constraints may be characterized by peaks and valleys of buying behavior induced by the cyclical nature of the budgeting system, even when the rate of demands for system stocks remain constant. In systems subject to expected-value constraints, such cyclical patterns should not be present. In short, a model based on an expected-value interpretation of budgetary constraints may be a very poor approximation to a system in which budget constraints are actually absolute.

Solution Procedures

In the previous section, we have discussed several objective functions which may be appropriate mechanisms for evaluating alternate courses of action in probabilistic environments. We noted that in several important situations, these objective functions are either separable or may be replaced by a modified objective function that is separable subject to a separable constraint (the V-Model). Similarly,
we have discussed several interpretations of constraints involving probabilistic elements, and we have shown that in several important situations, one may obtain deterministic equivalents of these constraints that are separable by item family $f$.

In any given problem of inventory management in probabilistic environments, one must select an appropriate deterministic objective function that describes the goals of the decision-maker as a single-valued measure and the relationship of these goals to alternate courses of action. One must also specify a deterministic set of constraints that limit the decision-maker's flexibility in choosing a course of action. If the objective function and each of the constraints are separable by item family $f$, the problem will be of the same form as the problem (5.1)-(5.4), and thus the same techniques discussed in Chapter IV may be applied to identify optimal policies. It should be noted that there is no requirement that all constraints be interpreted in the same manner, e.g., not all constraints need be interpreted as, say, chance constraints. What is required is that all members of the constraint set possess deterministic equivalents that are separable.
CHAPTER VI
STEADY-STATE DETERMINISTIC MODELS

Introduction

As noted in Chapter I, all models that have been proposed to date for use in budget-constrained systems assume that sufficient funds will be available to satisfy all requests placed upon the supply system, and that all demands for stock may be backordered indefinitely if necessary until stock becomes available to fill the request. In practice, these assumptions are often violated. First, because of the difficulties in forecasting future requirements and because of certain irrationalities of the Federal budgeting process, demands within a given budget period often exceed the resources available to meet those demands. Even supply systems managed under a stock fund concept, which theoretically should be little bothered by funding constraints, may in practice find authorized obligation authority consistently less than the funding levels required to meet all demands placed upon the system. This latter situation may occur if the price structure which determines reimbursements to the fund do not fully compensate for all expenses borne by the fund, and if subsidies to the fund are insufficient to compensate for the net losses. Second, many of the requirements for goods from a military supply system are transient in nature; in such cases if delays in filling a customer
request are excessive, the requirement which created the request may no longer exist when stock eventually becomes available. Although delays of hours or perhaps days may sometimes be acceptable, delays of weeks or of months in filling a customer request may be too long to satisfy a transient need.

In situations in which demands for goods exceed available funding for long periods of time, at least some lost sales are unavoidable. Such lost sales may result from one of two major mechanisms:

1. The supply system and/or the customer may determine explicitly what proportion of total potential demands will be filled, say through the review and reduction of the authorized usage for each line item or through the use of administrative controls that restrict the situations in which goods may be obtained from the supply system, and

2. The proportion of demands for each item to be filled may be determined implicitly, through the usual workings of the supply system. For example, all requests might be backordered until they are either filled or cancelled. Since almost all requirements are transitory when considered over a sufficiently long time period, a point will eventually be reached where cancellations are just sufficient to balance the funding rate and the demand rate of uncancelled requests. There may, of course, be much more complex mechanisms which achieve this same end, i.e., the balancing of funding and uncancelled requirements.

Regardless of whether an explicit or implicit mechanism for creating lost sales is employed, the inflow of funds must balance the outflow of goods over the long run if the system is to operate on an even keel. Hence, in
restricted funding situations only a portion of the total needs or requirements for at least some items managed by the supply system can be satisfied.

In this chapter, we seek insights into appropriate operating policies when some lost sales are unavoidable due to long term imbalances of authorized funding and total potential demands for goods. To this end, we will study two steady-state models of multi-item inventory systems subject to budget constraints. These models represent two extremes to the spectrum of mechanisms employed in real world supply systems which produce balance between the inflows of funds and the outflows of goods. First, we will study a system in which demand is known and constant and all demands are of a "fill-or-kill" variety, that is, we assume that all demands that occur when the system is out of stock are lost. Hence, in this system lost sales are generated by stockout conditions during a portion of each inventory cycle. In our second model, we assume that the required reductions in demands are achieved by an a priori rationing procedure. We again assume that demands are known and constant, but since reductions in demands are known in advance, we assume that the system is managed so that stockouts for the remaining demands never occur. Taken together, these models provide important insights into appropriate operating policies in systems which have long-term funding limitations.
A Model with Fill-or-Kill Demands

The Problem

Consider an N item inventory system subject to a constraint on average annual procurements over an infinite planning horizon. We assume that this constraint is sufficiently tight that it is impossible to meet all demands placed on the system, i.e., we assume the limitation B on average annual expenditure is less than the total dollar demand rate for all items in the system. Hence, some demand must go unsatisfied.

In all of the inventory models known to the author, the cost of a lost sale is assumed constant. This assumption is usually made both because of mathematical convenience and because in many commercial systems the loss due to lost sales are essentially constant and equal to the profit which would have been realized had the sale been completed. Hence, in these models, the annual cost of lost sales is directly proportional to the number of sales lost, i.e., annual cost of lost sales is generally stated as

\[ \alpha \cdot f \cdot D \]

where \( \alpha \) is the cost of a lost sale, \( D \) is the annual demand rate, and \( f \) is the percentage of annual sales lost, where \( \alpha \) and \( D \) are assumed known.
In a large number of real-world situations, however, costs of lost sales are not directly proportional to the number of sales lost. For example, in a military supply system with a large number of priority demand classes, initial lost sales may be levied upon the lowest priority demand customers. In such cases, a small proportion of lost sales may have little effect on performance of the military mission. As the fraction of lost sales approaches a large portion of total potential demands, however, rather severe degradation of the military mission may occur. Such a situation is illustrated by the cost curve A in Figure 6.1. Curve C illustrates an opposite situation, that is, a situation in which the loss of a relatively small proportion of the total potential demands may cause a serious degradation of the military mission. Curve D illustrates the limiting case in which the cost per lost sale is constant, regardless of the fraction of sales lost. Curve E denotes the opposite limit, in which almost all sales can be lost without cost to the supply system.

The family of curves illustrated in Figure 6.1 may be described mathematically by a family of functions of the form \( df^B \), where \( f \) denotes the fraction of sales lost and \( a \) and \( B \) are constants. For example, curves A, B, C, and D of Figure 6.1 correspond to the parameter values for \( B \) of 2,1, \( \frac{1}{2} \), and 0, respectively, and \( a=1 \). Curve E corresponds to the parameter values \( B = \infty \) and \( a=1 \). Hence, \( B \) is a shape
parameter that determines the rate at which the cost of a lost sales increases, while $\alpha$ is a scale parameter that describes the cost of losing the very last and highest priority unit of demand.

Hence, we may describe the cost of a single lost sale of item $j$ by the function

$$a_j f_j^{B_j}$$

(6.1)

where $f_j$ denotes the fraction of total potential sales for item $j$ that are lost and $a_j$ and $B_j$ are known constants which characterize the costs of lost sales for item $j$. Obviously,

Figure 6.1. Cost per lost sale versus the proportion of sales lost.
the fraction of sales lost cannot be less than zero nor more than one, i.e.,

$$0 \leq f_j \leq 1$$  \hspace{1cm} (6.2)

Hence, from (6.1) the average annual cost of lost sales for item $j$ is

$$\alpha_j f_j B_j \cdot f_j D_j = \alpha_j D_j f_j B_j + 1.$$  \hspace{1cm} (6.3)

since on the average there are $f_j D_j$ lost sales per year.

In a supply system in which all demands are known, constant and of a fill-or-kill variety and in which some lost sales are unavoidable, the pattern of on-hand inventory

\begin{figure}
\centering
\includegraphics{on_hand_inventory}
\caption{On-hand inventory in the deterministic "fill-or-kill" case.}
\end{figure}
through time for a given item \( j \) will appear as in Figure 6.2. Immediately after an order for an item is received, on-hand stocks will be raised to a level of \( Q_j \). After a time period \( Q_j/D_j \) years, these stocks will be exhausted and all demands occurring from this point in time until the next order is received are lost. Finally, another order is received and the cycle repeats.

From Figure 6.2, the time in years between arrivals of orders, \( R_j \), for item \( j \) is

\[
R_j = T_j + \frac{Q_j}{D_j} \quad (6.4)
\]

where \( T_j \) denotes the length of time in years each cycle that item \( j \) is out of stock. Since the average number of reorder cycles for item \( j \) per year is just \( 1/R_j \), the average annual total cost of ordering all items is just

\[
\sum_{j=1}^{N} \frac{A_j}{R_j} = \sum_{j=1}^{N} \frac{D_j}{Q_j + D_j T_j} A_j \quad (6.5)
\]

where \( A_j \) denotes the cost of placing an order for item \( j \).

From (6.3), the average annual cost of lost sales is

\[
\sum_{j=1}^{N} \phi_j D_j \left( \frac{D_j T_j}{Q_j + D_j T_j} \right)^{B_j+1} \quad (6.6)
\]

since the average fraction of demands lost per year equals the fraction of demands lost each cycle, i.e., \( f_j = T_j/R_j \).

The average annual holding costs for all items is
where \( I \) denotes the inventory carrying charge and \( C_j \) denotes the cost per unit of item \( j \). This expression may be derived by noting that the average number of units of item \( j \) on hand during that portion of the reorder cycle with positive stock levels is just \( Q_j/2 \), and the length of time the stock level is positive is \( Q_j/D_j \). Hence, the average holding cost per cycle of item \( j \) is just

\[
\sum_{j=1}^{N} \frac{IC_j Q_j^2}{2(Q_j+D_j T_j)}
\]

(6.7)

Multiplication of the cost per cycle times the average number of cycles per year, \( 1/R_j \), gives the average annual cost for item \( j \), and summing these results over all items yields the above expression.

Finally, the procurement cost of an order for item \( j \) is \( C_j Q_j \); hence, the average annual procurement expenditure for all items is

\[
\sum_{j=1}^{N} \frac{D_j}{Q_j + D_j T_j} \cdot C_j Q_j
\]

(6.9)
We are interested in minimizing the average annual costs of placing orders, holding inventory, and shortages subject to a constraint on average annual procurement expenditures. This problem may be written formally as

Minimize $\sum_{j=1}^{N} \left[ \frac{D_j}{Q_j + D_j T_j} \cdot A_j + \frac{IC_j Q_j^2}{2(Q_j + D_j T_j)} \right]$

subject to the budget constraint

$$\sum_{j=1}^{N} \frac{D_j}{Q_j + D_j T_j} \cdot C_j Q_j \leq b$$

provided that

$$T_j \geq 0 \quad Q_j \geq 0$$

The equivalent Lagrangian problem may be written as follows:

Minimize $L(Q, T, \delta)$
where

$$L(Q, T, \delta) = \sum_{j=1}^{N} \frac{D_j}{Q_j + D_j T_j} \left[ A_j + \frac{IC_j Q_j^2}{2D_j} + \alpha_j D_j T_j \cdot \left( \frac{D_j T_j}{Q_j + D_j T_j} \right) B_j \right]$$

$$+ \delta \left[ \sum_{j=1}^{N} \frac{D_j}{Q_j + D_j T_j} \cdot C_j Q_j - b \right]$$

(6.14)

Notice that (6.14) is separable over \(j\). Hence, if the Lagrange multiplier \(\delta\) is fixed we may solve the Lagrangian problem (6.13) by solving a series of \(N\) single-item subproblems of the form:

$$\text{Minimize} \quad L_j(Q_j, T_j, \delta)$$

(6.15)

where

$$L_j(Q_j, T_j, \delta) = \frac{D_j}{Q_j + D_j T_j} \left[ A_j + \frac{IC_j Q_j^2}{2D_j} \right]$$

$$+ \alpha_j D_j T_j \left( \frac{D_j T_j}{Q_j + D_j T_j} \right) B_j + \delta C_j Q_j$$

(6.16)

subject to the constraints that \(Q_j\) and \(T_j\) be non-negative.
Solving the Subproblem. We will now investigate how the solution to the subproblem (6.15) might be obtained. Since all variables in the subproblem have the same item subscript \( j \), we will drop the subscript throughout this section.

Two necessary conditions that \( Q^*, T^* \) be optimal in (6.15) is that they satisfy

\[
\frac{\partial L}{\partial T} = \frac{\partial L}{\partial Q} = 0 \tag{6.17}
\]

provided that \( T \) and \( Q \) are each positive.

As noted in Chapter II, one method to solve our original problem (6.10)-(6.12) is the develop efficient methods to solve the Lagrangian problems (6.15) for fixed values of the multiplier \( \lambda \), and to utilize an appropriate search procedure to identify the particular value of \( \lambda \) associated in a given procurement constraint. In Chapter II, we discussed several search procedures that might be employed to guide choices of \( \lambda \); hence, in this chapter, we will focus our attention on finding efficient methods to solve the subproblems (6.15) and to interpret how the solutions change with changing values of the procurement constraint.

It is tedious but straightforward to show that
\[
\frac{\partial L}{\partial Q} = D(Q+DT)^{-2} \left[ -A + \frac{ICQ^2}{2D} + ICQT - \alpha(B+1)DT \left( \frac{DT}{Q+DT} \right)^B \right]
\]

and that
\[
\frac{\partial L}{\partial T} = D(Q+DT)^{-2} \left[ -AD - \frac{ICQ^2}{2} - \delta CDQ + \alpha(B+1)DQ \left( \frac{DT}{Q+DT} \right)^B \right]
\]

Notice that the partial derivations (6.18) and (6.19) vanish if \( T \) is infinite. In this extreme case, all demands are lost, producing an average annual shortage cost of \( \alpha D \). In this case, \( \alpha D \) also represents the total average annual cost associated with the item since holding costs and shortage costs would be zero. This solution could produce the minimum value of the Lagrangian if costs of ordering and holding inventories and the implied costs of procurement are significantly greater than the costs of shortage. Otherwise, the solution with \( T = \infty \) represents either a maximum or saddle point of the Lagrangian (6.15).

(At first glance, the reader might also expect the derivatives to vanish as \( Q \) becomes very large. Note, however, that as \( Q \) increases (6.18) and (6.19) each approach \( ICD/2 \) due to the \( Q^2 \)-term associated with the holding cost. Hence, the derivatives (6.18) and (6.19) cannot vanish with large values of \( Q \).)
Let us momentarily assume that $Q$ and $T$ are each finite and positive. In this case, (6.18) and (6.19) imply:

\[-AD + \frac{ICQ^2}{2} + ICDQT - \alpha(B+1)D^2T \left(\frac{DT}{Q+DT}\right)^B + \delta CD^2T = 0 \quad (6.20)\]

and

\[-AD - \frac{ICQ^2}{2} - \delta CDQ + \alpha(B+1)DQ \left(\frac{DT}{Q+DT}\right)^B = 0 \quad (6.21)\]

The latter equation implies

\[\alpha(B+1) \left(\frac{DT}{Q+DT}\right)^B = \frac{A}{Q} + \frac{ICQ}{2D} + \delta C \quad (6.22)\]

Substituting (6.22) into (6.20), we obtain

\[-AD + \frac{ICQ^2}{2} + ICDQT - D^2T \left(\frac{A}{Q} + \frac{ICQ}{2D} + \delta C\right) + \delta CD^2T = 0 \quad (6.23)\]

Multiplying each side of (5.23) by $Q$ and rearranging terms, we obtain

\[(Q+DT) \left(\frac{ICQ^2}{2} - AD\right) = 0 \quad (6.24)\]

From our above assumptions, $(Q+DT)$ must be positive. Hence, (6.24) implies

\[Q = \sqrt{\frac{2AD}{IC}} \quad (6.25)\]
which is the well-known Wilson lot size formula.

Substituting (6.25) into (6.22), we obtain, after some rearranging,

\[ \alpha(B+1) \left( \frac{DT}{Q_w + DT} \right)^B = \sqrt{\frac{2AIC}{D}} + \delta C \quad (6.26) \]

where \( Q_w \) denotes the Wilson lot size defined by (6.25).

Equation (6.26) may now be solved for \( T \). The result is

\[ T = \frac{1}{D} \sqrt{\frac{2AD}{IC}} \left[ \left( \frac{\alpha D(B+1)}{\sqrt{2AICD + \delta CD}} \right)^{1/B} - 1 \right]^{-1} \quad (6.27) \]

From (6.25) and (6.27), the fraction \( f \) of sales lost each cycle is

\[ f = \frac{Q_w g}{Q_w + Q_w g} = \frac{g}{1+g} \quad (6.28) \]

where

\[ g = \left[ \left( \frac{\alpha D(B+1)}{\sqrt{2AICD + \delta CD}} \right)^{1/B} - 1 \right]^{-1} \quad (6.29) \]

\[ = \frac{\left( \sqrt{2AICD + \delta CD} \right)^{1/B}}{\left( \alpha D(B+1) \right)^{1/B} - \left( \sqrt{2AICD + \delta CD} \right)^{1/B}} \quad (6.30) \]

Let us momentarily assume that the value of \( f \) given by (6.31) is such that 0 \( \leq f \leq 1 \). If we now substitute (6.31) into (6.27), we obtain the following relationship between the
fraction of sales lost each cycle and the length of stockout period under an optimal policy:

\[
T = \frac{Q_w}{D} \left( \frac{f}{1-f} \right) \quad (6.32)
\]

Hence, the total length of a cycle under an optimal policy will be

\[
\frac{Q_w}{D} + T = \frac{Q_w}{D} + \frac{Q_w}{D} \frac{f}{1-f} = \frac{Q_w}{D} \left( \frac{1}{1-f} \right) \quad (6.33)
\]

It is interesting to note that \( \frac{Q_w}{D} \) is simply the length of an inventory cycle when no lost sales are incurred.

Suppose we were to solve a series of subproblems using different values of the Lagrange multiplier \( \gamma \). From (6.31), even if \( \gamma = 0 \), implying the procurement constraint is not active, some lost sales will be incurred, although the number of sales lost will be quite low if the costs of ordering and holding inventories are quite small relative to the costs of stockouts. From (6.31), as \( \gamma \) increases, indicating tighter procurement limitations, the fraction of lost sales should increase. Finally, when

\[
\gamma = \frac{aD(B+1) - \sqrt{2AICD}}{CD} \quad (6.34)
\]

equation (6.31) indicates that all sales for the item should be lost, and the item should be dropped from the inventory
system. The value of the Lagrange function (6.16) associated with this action is $\alpha D$, the cost of incurring lost sales all the time. For even higher values of $\alpha$, the minimum value of the Lagrangian function is still $\alpha D$, the cost incurred when all sales are lost.

Discussion

The reader might observe that our results include the special case in which the cost of a lost sale is a constant, regardless of the fraction of sales lost. In our model, this corresponds to $B=0$ in the cost expression (6.1). Notice from (6.31) that as $B$ approaches zero, the term $1/B$ approaches infinity. Hence, in a system in which the cost of a lost sale is a constant, if an item is to be carried at all under a given procurement limitation, then it is never optimal to incur stockouts on that item. Further as the procurement constraint is tightened in such a system, the number of items that satisfy (6.34) will increase. Hence, under a restrictive procurement limitation there will be fewer items carried in the inventory system than there will be in a less restrictive situation. Hence, the above equations quantify what we would intuitively expect to find in a long term restricted-funding situation.

Let us now assume that $B$ is positive, i.e., that the cost of a lost sale increases as the fraction of sales lost increases. Suppose that $\alpha$ and $B$ are identical for all items
in the system. Then (6.31) indicates that items with high unit costs should have a higher proportion of lost sales than items which are less expensive. Similarly, if all else is equal, items with high demand rates should have a smaller proportion of lost sales than low activity items. Finally, items with high ordering or storage costs relative to other items should have higher proportions of lost sales, while items with high stockout costs relative to other items should have lower proportions of lost sales.

A Rationing Model

The Problem

Consider an N item inventory system in which an a priori rationing mechanism is employed to balance the outflows of goods and the inflows of funds over the long term. Specifically, we assume that an explicit decision is made which reduced the demands for item j placed upon the supply system from some theoretical maximum level \( D_j \) which would occur if there were no procurement limitation to a level \( D_j(l-f_j) \), where \( f_j \) denotes the fraction of demands that are reduced, cancelled, or "lost" as a result of the rationing procedure. The reduction in demands placed upon the supply system might be accomplished by instituting administrative controls which restrict the situations in which requisitions for goods may be initiated or by modifying maintenance policies so that failure rates are reduced or so that items are repaired which
would otherwise be discarded and replaced with new items. (These latter actions require more maintenance resources, and hence larger maintenance budgets if they are to be implemented, but they do reduce the pressure for procurement of new items.) If even more severe actions are required to meet the procurement constraint, entire aircraft squadrons might be deactivated, thus decreasing demands for supply support.

In this section, we assume that the demands for item \( j \) placed upon the supply system after the application of the rationing process are known and constant with a rate of \( D_j(1-f_j) \); hence, it is possible to manage the system so that stockouts for nonrationed demands never occur. Figure 6.3 illustrates the pattern of on-hand stocks through time for a typical item in this system.

Figure 6.3. On-Hand Inventory in Rationed Demand Case.
We are interested in policies for ordering and rationing stocks which minimize the average annual costs of placing orders, holding inventory, and lost sales subject to a constraint on average annual procurement expenditures. Using arguments and notation similar to that used in the previous section, this problem may be written formally as

\[ \text{Minimize} \sum_{j=1}^{N} \left[ \frac{A_j D_j (1-f_j)}{Q_j} + \frac{IC_j Q_j}{2} + \alpha_j D_j f_j^{B_j+1} \right] \quad (6.35) \]

subject to the procurement limitation

\[ \sum_{j=1}^{N} C_j D_j (1-f_j) \leq b \quad (6.36) \]

The associated Lagrangian problem may be written as

\[ \text{Minimize} \sum_{j=1}^{N} \left[ \frac{A_j D_j (1-f_j)}{Q_j} + \frac{IC_j Q_j}{2} + \alpha_j D_j f_j^{B_j+1} \right] + \gamma \left[ \sum_{j=1}^{N} C_j D_j (1-f_j) - b \right] \quad (6.37) \]

If the Lagrange multiplier \( \gamma \) is fixed, the problem (6.37) may be solved by solving the series of N cell problems

\[ \text{Minimize} \quad L_j(Q_j, f_j, \gamma) \quad j=1, 2, \ldots, N \quad (6.38) \]
where

\[ L_j(Q_j, f_j, \gamma) = \frac{A_j D_j (1 - f_j)}{Q_j} + \frac{IC_j Q_j}{2} + \alpha_j D_j f_j^{B_j + 1} \]

subject to the constraints that \( Q_j \) be positive and that \( f_j \) lie in the \((0,1)\) interval.

As noted above, one method to solve our original problem is to develop efficient methods to solve the Lagrangian problems (6.38) for fixed values of the multiplier \( \gamma \), and to utilize an appropriate search procedure to identify the particular value of \( \gamma \) associated with a given procurement constraint. As in the previous sections, we will focus our attention on finding efficient methods to solve the subproblems (6.38) and to interpret how the solutions change with changing values of \( \gamma \).

Solving the Subproblem

Necessary Conditions. Since all variables in the subproblem have the same item subscript \( j \), we will drop the subscript throughout this section to simplify our notation.

Two necessary conditions that \( Q^*, f^* \) be optimal in (6.38) is that they satisfy

\[ \frac{\partial L}{\partial Q} = \frac{\partial L}{\partial f} = 0 \]  

(6.40)
provided $Q^* > 0$ and $0 < f < 1$. Hence, it is necessary that

$$\frac{\partial L}{\partial Q} = \frac{-AD(1-f)}{Q^2} + \frac{IC}{2} = 0$$

(6.41)

and

$$\frac{\partial L}{\partial f} = \frac{-AD}{Q} + \alpha D(B+1)f^B - \gamma CD = 0$$

(6.42)

which is simply the Wilson lot size for an annual demand rate of $D(1-f)$ units per year.

The Constant Cost Case: $B=0$. Before solving for $f$ in the general case, let us first consider the special situation in which $B=0$, i.e., the case in which the cost of a lost sale is a constant, regardless of the fraction of sales lost. In this special situation, (6.39) is a linear function of $f$, and hence the Lagrangian will be minimized when $f$ lies on a boundary, i.e., when $f = 0$ or $f = 1$. If we set $f = 0$, indicating no lost sales are incurred, the optimum $Q$ is the Wilson lot size, and the resulting value of the Lagrangian (6.39) is $\sqrt{2ADIC + \gamma CD}$. On the other hand, if $f = 1$, indicating all sales are lost, the value of the Lagrangian is $\alpha D$. By comparing these values, one can easily determine which solution minimizes (6.39).
From the above, if one is solving a series of Lagrangian problems and the cost of a lost sale is a constant, and if

\[ \delta < \frac{\alpha D - \sqrt{2AICD}}{CD} \]  \hspace{1cm} (6.44)

then the optimum policy is to order the item using the Wilson lot size formula and to never incur lost sales on the item. On the other hand, if (6.44) is not satisfied, the item should be dropped from the system and lost sales incurred all the time.

Note that if the cost of a lost sale is constant, i.e., if \( B = 0 \), the set of optimum ordering decisions for an item in an inventory system in which an a priorirationing policy is employed are identical to the set of optimum decisions implied by (6.34) for the same item if a lost sales mechanism were active. As we shall see below, however, slightly different ordering patterns will be associated with these two types of lost sales mechanisms if \( B = 0 \).

**The Variable Cost Case: \( B \neq 0 \).** Let us now consider the case in which \( B \neq 0 \). Substituting (6.43) into (6.42) we obtain, after some manipulation,

\[
\left[ f^B - \left( \frac{\delta C}{\alpha (B+1)} \right) (1-f)^{\alpha} - \frac{\sqrt{2AICD}}{2\alpha D (B+1)} \right] = 0 \hspace{1cm} (6.45)
\]

Unfortunately, it does not appear possible to manipulate (6.45) to obtain a closed-form solution for \( f \). Since \( f \)
is restricted to the (0,1) range, however, the solution to (6.45) (if it exists) may be found without too much effort by employing an appropriate search procedure in this interval.

Notice that the function on the left-hand side of (6.45) is of the form:

$$F(f) = (f^B - h)(1-f)^{\frac{1}{2}} - g$$

where h and g are appropriate constants. Figure 6.4 illustrates the behavior of this function for $B=1$, $g=0$ and several values of h. Observe that for $0 \leq f \leq 1$, $f^B - h$ is continuously increasing and that $(1-f)^{\frac{1}{2}}$ is continuously decreasing for increasing values of f. Hence, $F(f)$ is unimodal for $0 \leq f \leq 1$. Further, observe that the equation $F(f) = 0$ may have two, one, or zero roots in the (0,1) interval, depending upon the particular values of the parameters h and g.

Notice that in the special case in which $h=g=0$, $F(f)$ has roots of $f=0$ and $f=1$. In the context of (6.45), $h=0$ indicates the procurement constraint is not binding, and $g=0$ indicates that costs of ordering and holding inventories are negligible relative to the costs of stockouts. Hence, in this special case, the solution which minimizes the Lagrange function (6.38) is $f=0$, i.e., no lost sales should be incurred. If $\gamma$ is increased in (6.45), indicating a more restrictive procurement limitation, while $g=0$, then the real positive root
Figure 6.4. Plot of $F = (F-c)(1-f)^{1/3}$
to (6.45) which minimizes (6.38) is

$$f = h^{1/B} = \left( \frac{\delta C}{Q(B+1)} \right)^{1/B} \quad (6.46)$$

while the alternate root $f=1$ maximizes the Lagrange function. (We have ignored negative or imaginary roots associated with (6.46) since these have no physical significance in our problem.)

Now observe that if the parameter $g$ is increased slightly above zero, while $h<1$, the real roots to $F(f) = (f^B-h)(1-f)^{1/2} - g$ will be $f' + \varepsilon'$ and $f'' - \varepsilon''$, where $f'$ and $f''$ denote the roots for the $g=0$ case and $\varepsilon'$ and $\varepsilon''$ denote appropriate positive constants. Since $f'$ minimizes the Lagrange function and $f''$ produces the maximum Lagrange value, and since $L(Q,f,\delta)$ is well-behaved and unimodal in the $(0,1)$ interval, the root $f' + \varepsilon'$ will produce the minimum to the new Lagrange problem with $g>0$. This argument may be repeated for increasing values of $g$ until the roots are identical. Since $F(f)$ is unimodal, this will occur when $F(f)$ is a maximum. For even larger values of $g$, there will be no real root which satisfies (6.45). Hence, in this case the solution will lie at an extreme point.

If no lost sales are incurred, the optimum $Q$ is the Wilson lot size, and the resulting value of the Lagrangian function is $\sqrt{2AICD} + \delta CD$. On the other hand, if $f=1$, indicating all sales are lost, the value of the Lagrangian is $\delta D$. By comparing these values, one can easily determine which
solution minimizes $L(Q,f,y)$.

Observe that a solution to (6.45) will be $f^* = 0$ whenever

$$\alpha = -\frac{\sqrt{2AI\alpha D}}{2\alpha D}$$

From the theory of Lagrange multipliers, however, the value of $\alpha$ which minimizes the Lagrange function (6.37) will never be negative. Hence, if $B > 0$, every item in the system will have at least some lost sales, although the proportion of sales a given item should lose might be small if the costs of lost sales are high relative to the costs of ordering and holding inventories.

Now observe that since $f$ is restricted to the $(0,1)$ interval, $(1-f)$ cannot exceed 1. Hence, the function

$$f^B - \frac{\gamma C}{\alpha (B+1)} - \frac{\sqrt{2AI\alpha D}}{2\alpha D (B+1)}$$

must always be greater than or equal to the function on the left-hand side of (6.45). The function (6.48) vanishes when $f=f^*$, where

$$f^* = \left(\frac{\sqrt{2AI\alpha D} + 2\alpha D}{2\alpha D (B+1)}\right)^{1/B}$$

Since (6.48) is an upper bound to the function on the left-hand side of (6.45) and since the function defined by (6.48) increases monotonically in the $(0,1)$ interval, the real root which satisfies (6.45) (if it exists) will be greater than $f^*$, i.e., $f^*$ is a lower bound on the minimizing root to (6.45).
Hence, one does not need to search the entire \((0,1)\) interval to locate the root which minimizes \((6.45)\), but merely the interval between \(f^*\) and the maximum value of the function on the left-hand side of \((6.45)\). Further, if

\[
\delta \geq \left( \frac{2aD(B+1) - \sqrt{2AICD}}{2CD} \right) 
\]  

(6.50)

there is no value of \(f\), \(0 \leq f \leq 1\), which satisfies \((6.45)\). In this case, the minimizing root to the subproblem is \(f=1\), i.e., lost sales should be incurred on the item all the time. Since the condition \((6.50)\) is based upon an upper bound to the function on the left-hand side of \((6.45)\), we would expect to find the root \(f=1\) in \((6.45)\) for values of \(\delta\) slightly smaller than indicated in \((6.50)\). However, evaluation of \((6.50)\) should provide a useful check to see if search for the root is justified.

Discussion

Although it does not appear possible to obtain a closed-form solution for the optimum value of \(f\) when \(f\) does not lie on a bound, we can obtain insights into the behavior of the optimum \(f\) as the parameters \(A\), \(I\), \(D\), \(D\), \(\delta\) and \(a\) change. Notice that if we substitute \((6.43)\) into \((6.39)\), the Lagrangian takes the form

\[
L(Q^*f, \delta) = \sqrt{2AICD(1-f)} + aDf^{B+1} + \delta CD(1-f) 
\]  

(6.51)
where $Q^*$ denotes the order quantity given by (6.43). The first term on the right of (6.51) denotes the costs of ordering and holding inventory. The second term denotes costs of shortages, and the third term denotes the implied cost of the procurement constraint. Observe that the terms $\sqrt{2AICD(1-f)}$ and $\gamma CD(1-f)$ are strictly decreasing functions of $f$, while $\alpha Df^{B+1}$ is a strictly increasing function by $f$.

Let $f^*$ denote the value of $f$ that minimizes (6.51) for a given set of parameters, and assume $f^*$ does not lie on a boundary, i.e., $0 \leq f^* \leq 1$. Now suppose the ordering cost $A$ is increased. From (6.51), increasing $A$ increases the cost of ordering and holding inventory which in turn causes the minimum point of the Lagrange function to shift to the right. Hence, items with high ordering costs should have a higher proportion of lost sales than items that are less expensive to order, all else being equal. Similarly, items with high unit costs or with high holding costs should incur more lost sales under an optimal policy than items that are less expensive or items that are cheaper to store. On the other hand, items with high stockout costs relative to other items should have lower proportions of lost sales, all else being equal.

**A Comparison**

In the "fill-or-kill" model considered earlier, the
average fraction of sales lost per year equals the fraction of each cycle the system is out of stock. Hence, in the fill-or-kill model,

\[ f = \frac{T}{Q/D + T} = \frac{DT}{Q + DT} \]  

(6.52)

Substituting (6.52) and (6.25) into (6.15), the Lagrangian function for the fill-or-kill model may be written as

\[ L(Q, f, \lambda) = (1-f)\sqrt{2AICD} + \lambda D^{B+1} + \lambda CD(1-f) \]  

(6.53)

Now compare (6.53) and (6.51). Observe that these two functions are identical except for the order and holding cost term. Since \( f \) is bounded to the \((0,1)\) interval, \((1-f)^{\frac{1}{2}} \geq (1-f)\). Hence, for given values of \( Q \) and \( f \) the order and holding cost contribution to the Lagrange function is greater in the rationing case than in the "fill-or-kill" case.

Hence, for a given value of the Lagrange multiplier the value of \( f \) that minimizes (6.51) will be greater than the value which minimizes (6.53), all else being equal. In other words, for a given value of the Lagrange multiplier, the optimum fraction of lost sales in the rationing case for a given item will be greater than the optimum fraction of lost sales in the fill-or-kill case. Hence, as the value of \( \lambda \) is increased, the point at which it becomes optimal to completely discontinue carrying a given item and incur lost sales all the time is reached earlier (i.e., for a smaller \( f \)) than in the fill-or-kill case. Hence, for a given value of the Lagrange
multiplier, there will be more items that incur lost sales all the time under an optimal policy in the rationing case than in the fill-or-kill case. Hence, the average annual procurement budget associated with a given value of the Lagrange multiplier will be lower in the rationing case than in the fill-or-kill situation. Conversely, the Lagrange multiplier $\gamma_R$ associated with a given procurement budget in the rationing case will be smaller than the Lagrange multiplier $\gamma_F$ associated with the same level of annual procurements in the fill-or-kill case. Hence, the implied cost of a given procurement limitation is higher in the case in which demands are rationed.

Summary

In this chapter, we have sought insights into appropriate operating policies when some lost sales are unavoidable due to long-term imbalances of authorized funding and total potential demands for goods. To this end, optimal policies were derived for two models of multi-item inventory systems subject to procurement constraints. In both models it was assumed that demand was known, constant and independent of actions taken by the supply system. In the fill-or-kill model, it was assumed that inflows of funds and outflows of goods where balanced by lost sales generated by stockout conditions during a portion of each inventory cycle. In the rationing model, it was assumed that balance between flows
of funds and goods was achieved through an a priori rationing decision.

Optimal operating policies were derived for both models. When the cost per lost sale is constant, the same operating policy is optimal for both models. In the constant cost case, if an item is to be carried at all under a given procurement limitation, then no stockouts should be incurred by that item. The optimal order quantity in this case is the well-known Wilson lot size. Conversely, in the constant case there will always be some items that should be completely dropped from the supply system when procurement funds are limited. For these items, all sales are lost. As the procurement restriction becomes tighter, there should be fewer items carried by the inventory system than in a less restrictive situation.

When the cost of a lost sale increases as the fraction of sales lost increases, the optimal operating policies for the fill-or-kill and rationing models are similar but not identical. In both situations, all items should incur at least some lost sales when procurement funds are limited and the fraction of sales lost for each item should increase as the procurement limitation is tightened. In both cases, items with high unit costs should incur a higher proportion of lost sales than items that are less expensive. Similarly, if all else is equal, items with high ordering or storage costs relative to other items should have higher proportions
of lost sales, while items with high stockout costs relative to other items should have lower proportions of lost sales.

In both the fill-or-kill and rationing models, the optimal order quantity is given by the classical Wilson lot size formula,

\[
Q = \sqrt{\frac{2AD^*}{IC}}
\]

where \( D^* \) denotes the annual rate at which units are withdrawn from the supply system and \( A, I, \) and \( C \) are as defined above. For the fill-or-kill situation, \( D^* = D \), the potential annual demand rate for the item. For the rationing model, \( D^* = D(1-f) \), where \( f \) denotes the fraction of demands that are reduced, cancelled, or "lost" as a result of the rationing procedure.

The above observations provide interesting insights into the nature of optimum policies in restricted funding environments. Unfortunately, these observations are only a start toward understanding the nature of optimum policies for real-world inventory systems. Much work remains to be done. In the next chapter, we will discuss some areas in which additional research is required.
CHAPTER VII

SUMMARY AND AREAS FOR FURTHER RESEARCH

Summary

This work focuses attention on single-location, multi-item military supply systems managed under a revolving fund concept. In Chapter I, the qualitative aspects of the problem of inventory management in such systems are discussed, and the approaches to this problem that have been suggested by other researchers are reviewed. Unfortunately, it appears that the theoretical understanding of optimum decision making in such environments is still at a very primitive stage. Although much effort has been devoted to this problem, many important questions remain unanswered.

Perhaps the most important aspect of inventory management in budget-constrained environments is the determination of the categories and levels of expenditures that are authorized for operation of the supply system. Ultimately, one would like to develop a system of procedures for preparing and evaluating budgetary proposals and for controlling and auditing the expenditures of funds so that the resulting allocation of national resources is in some sense "optimal". The development of such a system involves much more than questions of inventory theory, since determination
of an "optimal" allocation of national resources requires the simultaneous consideration of all alternate uses for those resources.

It appears that at least three major difficulties must be overcome if such a Utopian system is ever to be developed. First, methods must be developed which recognize the levels of uncertainty associated with the forecasts of future needs, and that are sufficiently flexible to adjust for inaccurate forecasts as new information becomes available. There is also a continuing need to improve the accuracy and consistency of such forecasts. Second, methods must be developed which eliminate or reduce difficulties attributed to "limited rationality", i.e., to the limited information processing capabilities of both men and machines. Such methods might require the development of management procedures or computing techniques which reduce the information processing requirements placed upon individual decision makers or computing machines, or which improve the information processing capabilities of the decision makers or of their machines. Finally, there is a need to develop a clearly-stated, widely-accepted methodology for choosing among alternatives in some optimal sense.

As interesting and important as the above questions might be, we have left their consideration to other researchers. Rather, in this work we have focused our attention on the problems of an inventory manager who is
presented with a given budget and instructed to abide by it --to do the best he can with what he has, consistent with legal and administrative limitations established by Congress and other higher authorities. In Chapter III, an analytical model of this latter problem was developed, and Chapters IV and V discussed solution procedures for important special cases of the general problem. Chapter IV discussed solution procedures for situations in which demand is assumed known, while Chapter V discussed solution procedures for probabilistic situations. In these chapters, it was shown that good approximate solutions may be obtained for several important special cases of the general problem by using either the Dantzig-Wolfe decomposition algorithm of linear programming or a generalized version of that procedure. Basically, the generalized procedure requires the solution of a series of single-item subproblems and a "master" linear programming problem which finds the optimum linear combination of subproblem results. The master problem results may be used to (1) test for the optimality of the current solution or (2) to identify a new set of subproblems whose solution might be used to improve the objective function of the original problem.

An important feature of the Dantzig-Wolfe decomposition procedure is that any appropriate solution procedure may be employed to obtain subproblem results. Hence, in a given application one might utilize a procedure that is most
efficient for the particular structural characteristics of
the commodity subproblem being considered. In Chapter IV,
solution procedures were discussed for several important
cases of subproblems that may appear in budget-constrained
inventory problems.

Computational experience reported by Dzielinski and
Gomory(1965) suggests that deterministic problems involving
around 400 items, 20 to 30 joint budgetary and capacity con­
straints and 6 planning periods may be solved using this
technique in less than 10 minutes of IBM 7090 computing time.
For many inventory systems, the capability to solve problems
of this size might be sufficient to manage the requirements
of a large fractions of the total dollar activity of the
system. If large-scale "third-generation" computing equip­
ment is available (e.g., IBM 360/65, CDC 6600, or GE 635
computing systems) even larger problems might be solved at
a reasonable computational cost.

As noted in the introduction, two major inadequacies
of models of budget-constrained inventory systems that have
been proposed to date are that available models generally do
not describe the out-of-pocket nature of operating costs nor
do they present adequate methods for dealing with transients
which might arise during the changing of funding environments.
The models presented in Chapters IV and V accurately repre­
sent both of these characteristics of real-world systems.
The price paid for this increased resolution, however, is
the need to obtain and maintain much more information about the characteristics of and interrelationships among items in the inventory system than is required using less detailed models. Also, the cost of computing optimum stocking policies is undoubtedly much higher than the computational costs associated with less detailed models. Whether the improvements in operating results justify these increased information processing costs is a question that has not yet been addressed.

Unfortunately, although the results of Chapters IV and V provide techniques for obtaining numerical solutions to many problems of inventory management subject to budget constraints, they do not provide qualitative insights into the characteristics of an optimal policy. Hence, in Chapter VI we considered two simplified versions of the general problem to obtain insights into the nature of an optimum policy in the presence of a restrictive procurement constraint. First, we considered a multi-item system in which demand was known and constant and where all demands when the system is out of stock are lost. We referred to this as the "fill-or-kill" case. Closed-form expression were obtained for optimal stocking policies when the procurement constraint is sufficiently tight that at least some lost sales are unavoidable. Next, we considered a multi-item system in which demand is known and constant, but in this case it was assumed that the procurement constraint is satisfied through an a priori
rationing decision. Although it does not appear possible to obtain closed-form expressions for all elements of an optimum stocking policy in this latter system, important quantitative insights were obtained. In both the fill-or-kill and rationing cases, the optimum order quantity takes the form of the classical Wilson lot size formula, computed using the rate at which stock is withdrawn from the system. Other similarities among optimal policies for these systems are discussed at the conclusion of Chapter VI.

Suggestions for Further Research

As noted above, very little is known about the nature of optimum decision-making in the presence of budgetary constraints, and there are several major areas in which further research is required. Some of the general problems were sketched above; other, more specific problem areas in which research is needed will be sketched in the following paragraphs.

First, there is obviously a need for continuing theoretical investigations of the nature of optimal decisions in budget-constrained inventory systems. The work reported in Chapter VI might be considered as a small step in this direction. Ultimately, one would like to identify the characteristics of optimal policies for multi-item budget-constrained systems in which the underlying demand processes are stochastic and non-stationary, and to develop efficient
computing methods for determining numerical values for the policy variables. Perhaps the investigation of deterministic systems will lead to valuable insights into this problem. Since a direct attack on multi-item stochastic problems appears to be extremely complicated (see Hadley and Whitin (1961, p. 215)), another approach might be to obtain upper and lower bounds on the optimal decisions. If such bounds can be obtained, and if they prove to be sufficiently close, one might not need to identify exactly optimal solutions to obtain satisfactory operational results.

Second, there is a need to investigate the usefulness of alternate mathematical programming techniques for obtaining solutions to the models discussed in Chapters IV and V. The generalized decomposition procedure discussed in those chapters is but one of several approaches to problems of large-scale optimization that are currently available. Although the Dantzig-Wolfe technique appears to be the most general method currently available, it may not be the most efficient method for solving all types of problems. Other techniques such as the partitioning procedures developed by Rosen and by Bender are applicable to certain special cases of budget-constrained inventory problems. (For an excellent review of these and other techniques, see Geoffrion (1970)). One or more of these procedures may have computational characteristics superior to the generalized Dantzig-Wolfe procedure. The situations in which this is true need to be
identified.

Third, there is a need for continued investigations of the applicability of the Dantzig-Wolfe (D-W) decomposition principle to problems of inventory management. Studies in this area may concentrate on one of a number of important questions, including:

1. How much computational effort is required when the D-W procedure is applied to problems of the size encountered in real-world systems?

2. Are further computational refinements possible? For example, can the sub-problems for many items be solved simultaneously on computing systems with multi-programming capability? Can the solutions to previously-solved problems be used to reduce the computational effort required to solve a similar problem?

3. Do the benefits achieved in applying the Dantzig-Wolfe procedure justify the increased computational costs? What are the cost/benefits of the D-W procedure relative to other (perhaps heuristic) approaches to this problem?

Answers to these questions are required if the D-W procedure is to be applied routinely and economically to solve day-to-day problems of inventory management.

Fourth, there is a need to investigate the variability of actual resource usage associated with approaches that utilize constraints on the expected usage, such as models proposed by Presutti and Trepp(1970) and by Howard(1967). Under a given ordering rule, total resource usage for all
items in the inventory is the sum of individual item resource usage. Hence, the moments of the probability distribution of total resource usage within a budget period might be obtained from the probability generating functions of individual item resource usage (assuming the items are independent). If the variance-to-mean ratio of this distribution is sufficiently small, actual expenditures will almost always be very close to the expected value, and special actions required to align actual and budgeted expenditures might be negligible. Identification of which situations, if any, in which these conditions hold would provide a major advance to the understanding of inventory system performance under budget constraints.

Another area of needed research related to expected-value methods is the development of procedures for modifying the decision rules when actual expenditures are more or less than was originally anticipated.

Finally, and perhaps most important in terms of immediate usefulness to managers of real-world systems, is the need to conduct empirical investigations of the operational characteristics of the various approaches that have been suggested for dealing with budget-constrained inventory problems. Brown (1963) has observed that certain properties of inventory systems, such as a log-normal distribution of annual dollar activity, are common to entire industries. Initial research might then be directed toward analysis of systems with such aggregate features. Certain characteristics of rules that
have been proposed, such as the variability of actual resource usage, might be studied analytically for certain simple systems. For more complex problems, such as the evaluation of a system's dynamic response in the face of major changes in funding or demand levels, simulation methods will probably be required.
APPENDIX

ILLUSTRATION OF GENERALIZED DECOMPOSITION

Introduction

In Chapters IV and V it was observed that several special cases of inventory management problems might be solved using a generalization of the Dantzig-Wolfe decomposition algorithm of linear programming. In this Appendix, a sample problem is presented to illustrate the steps involved in applying this algorithm.

The Problem

Consider a multi-item inventory system subject to procurement constraints. Further, assume that

1. initial inventories are zero.
2. delivery lead times are negligible.
3. backorders are not permitted, and
4. sufficient funding is available so that assumption 3 may be satisfied.
5. costs of holding inventories during a given period are proportional to the dollar value of stock on hand at the end of the period.
We define:

\[ c_j = \text{unit cost} \]
\[ w_{ij} = \text{inventory on hand at the beginning of period } i \text{ before ordering.} \]
\[ q_{ij} = \text{quantity ordered at the beginning of period } i. \]
\[ d_{ij} = \text{demand during period } i. \]

where \( j \) denotes an item subscript and the \( d_{ij} \) are assumed known. Further, let

\[ A = \text{cost of placing an order.} \]
\[ h = \text{holding cost factor arising from assumption 5. above.} \]

We wish to identify procurement schedules \( Q_j = (q_{1j}, q_{2j}, \ldots, q_{IJ}) \) for each item \( j \) in the system so as to minimize the total variable costs of ordering stocks and holding inventories subject to the procurement constraints. Symbolically, we wish to minimize \( Z \),

\[
Z = \sum_{j=1}^{J} \sum_{i=1}^{I} A (q_{ij}) + hc_{j}w_{ij} \tag{A.1}
\]

subject to the obligation constraints

\[
\sum_{j=1}^{J} \sum_{i \in S(k)} c_{j}q_{ij} = b_{k} \quad k=1, \ldots, K \tag{A.2}
\]
where $S(k)$ denotes the set of periods associated with the kth procurement constraint, $b_k$ denotes the associated procurement limit, and $K$ denotes the total number of procurement limitations. We assume the sets $S(j)$ and $S(k)$ are mutually exclusive for $j \neq k$. The symbol $\delta(q_{ij})$ denotes the delta function, i.e.,

$$
\delta(q_{ij}) = \begin{cases} 
1 & \text{if } q_{ij} \geq 0 \\
0 & \text{otherwise}
\end{cases}
$$

Further, the solution must satisfy the material balance equations,

$$
\begin{align*}
  w_{i+1,j} &= w_{ij} + q_{ij} - d_{ij} \quad \text{all } ij \\
\end{align*}
$$

and the initial conditions

$$
\begin{align*}
  w_{1,j} &= 0 \quad \text{all } j \\
\end{align*}
$$

Suppose we have identified a set of "procurement plans" $Q^p = (Q^p_1, Q^p_2, \ldots, Q^p_j)$, $p=1,2,\ldots,P$, where $Q^p_j$ denotes the procurement schedule for item $j$ associated with the $p$th procurement plan. We assume that each item procurement schedule $Q^p_j$ satisfies (A.3) and (A.4), and that at least one of the procurement plans $Q^P$ is feasible in (A.2). Then a linear approximation to the problem (A.1)-(A.4) is

$$
\begin{align*}
\text{Min} \quad & \sum_{p=1}^{P} \lambda_p c^p \\
\end{align*}
$$

(A.5)
subject to

$$\sum_{p=1}^{P} \lambda_p g^p_k \leq b_k \quad k = 1, 2, \ldots, K$$  \hspace{1cm} (A.6)

$$\sum_{p=1}^{P} \lambda_p = 1$$  \hspace{1cm} (A.7)

where

$$c^p = \sum_{j=1}^{J} \sum_{i=1}^{I} \left[ a^p_{ij} q^p_{ij} + h c^p_{ij} w^p_{ij} \right]$$  \hspace{1cm} (A.8)

$$g^p_k = \sum_{j=1}^{J} \sum_{i \in S(k)} c^p_{ij} q^p_{ij}$$  \hspace{1cm} (A.9)

In the above, $q^p_{ij}$ denotes the quantity of item $j$ purchased during period $i$ under procurement plan $p$, $w^p_{ij}$ denotes the units of stock of item $j$ on hand at the beginning of period $i$ under plan $p$, and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_P)$, where $\lambda_p$ denotes the weighting factor associated with plan $p$. Note that (A.5)-(A.7) is of the same form as the approximating problem $P_2$ discussed in Chapter IV.

The problem (A.5)-(A.7) is a linear programming problem in the variables $\lambda_p$. Suppose we have an optimal basic feasible solution to this problem, and let $\theta = (\theta_1, \ldots, \theta_k)$ and $\sigma$ denote, respectively, the dual variables associated
with the constraints (A.6) and (A.7). Now suppose we can identify a procurement plan $P_{p+1} = (Q_{P+1}, \ldots, Q_{j})$ satisfying (A.3) and (A.4) such that

$$c_{P+1} - \sum_k \theta_k e_{k}^{P+1} - \sigma < 0 \quad (A.10)$$

Then the criterion function (A.5) may be improved by introducing the variable $\lambda_{p+1}$ associated with $Q_{P+1}$ into the basis. Using the standard simplex column selection criterion, if there are several procurement plans which satisfy (A.10), the procurement plan which produces the minimum value for the left-hand side of (A.10) would be selected for entry into the basis. The required procurement plan may be identified by solving the problem

$$\text{Min} \sum_{j=1}^{J} \sum_{i=1}^{I} (A_{j}(q_{ij}) + h_{j}w_{ij})$$

$$+ \sum_{k} \theta_k \left[ - \sum_{j=1}^{J} \sum_{i \in S(k)} c_{j}q_{ij} \right]$$

subject to (A.3) and (A.4). In turn, the solution to (A.11) may be obtained by solving $J$ subproblems of the form:

$$\text{Min} \sum_{i=1}^{I} \left[ A_{j}(q_{ij}) + h_{j}w_{ij} + \delta_{i}c_{j}q_{ij} \right] \quad (A.12)$$
subject to

\[ w_{i+1,j} = w_{ij} + q_{ij} - d_{ij} \]  \hspace{1cm} (A.13)

\[ w_{lj} = 0 \]  \hspace{1cm} (A.14)

where

\[ \xi_i = \begin{cases} -\theta_k & \text{if } i \in S(k) \\ 0 & \text{otherwise} \end{cases} \]  \hspace{1cm} (A.15)

The problem (A.12)-(A.14) is a single-item dynamic lot size problem, and hence may be solved using any one of possibly several techniques. As noted in Chapter IV, the Wagner-Whitin lot size algorithm is one efficient technique for solving subproblems of this form.

After each of the \( J \) subproblems are solved, the variable cost \( c_{P+1} \) and the total procurement expenditures \( g_{P+1,k} \) associated with each budget constraint may be computed by substitution into (A.8) and (A.9), and the test (A.10) may then be applied. If the newly-created procurement schedule satisfies (A.10), the criterion (A.5) may be improved by entering the variable \( \lambda_{P+1} \) into the basis, replacing \( P \) by \( P+1 \) and repeating the above process. If the procurement plan \( Q_{P+1} \) does not satisfy (A.10), the process terminates.

As noted in Chapter IV, solving the subproblem (A.9) is equivalent to minimizing the Lagrangian function with multipliers \( \theta \) and \( \sigma \). Hence, the minimizing procurement plan
is efficient and may be represented as a point on the convex hull of the \((K+1)\)-dimensional payoff resources space associated with the problem \((A.1)-(A.4)\). If these points are sufficiently dense, then some of the procurement plans associated with the basis variables at termination will be good approximate solutions to the original problem.

**Computational Experience**

To investigate the computational characteristics of the above procedure, a computer program was written to perform the required calculations. The program is written in FORTRAN IV and compiles and executes on a General Electric 635 computing system. The basic steps performed by the program are as follows:

1. Read problem data from cards and print this information for later reference.

2. Construct an initial basis to the problem \((A.5)-(A.7)\) using the \(K\) slack vectors associated with \((A.6)\) and an artificial vector. The artificial vector is of the form \((CC,0,0,...,0,1)\), where \(CC\) denotes a very large number. (In the sample problems to be discussed below, \(CC\) was arbitrarily set to \(10^6\)). That is, the artificial vector has a very large element in the cost position \((A.5)\), a "1" in the convexity constraint position, and zeros elsewhere. Symbolically, this step may be performed by setting \(P=K+1\) and applying the following formulas for \(p=1,2,...,P\):

\[
c^p = \begin{cases} 
CC & \text{if } p=P \\
0 & \text{if } 1 \leq p \leq K 
\end{cases}
\]
and

\[ g^p_k = \begin{cases} 
1 & \text{if } p=k \\
0 & \text{otherwise}
\end{cases} \]

3. Solve the problem (A.5)-(A.7) using the P vectors that have been identified to date. Let \( \theta_1, \theta_2, \ldots, \theta_K \) and \( \sigma \) denote the dual variables associated with the optimal solution.

4. Solve the subproblem (A.12)-(A.15) for each of the J items in the system using the Wagner-Whitin algorithm.

5. Compute the composite vector elements by substitution into (A.8)-(A.9).

6. Use (A.10) to test for optimality. If the value on the left-hand side of (A.10) is negative, append the newly created procurement plan vector to the linear programming problem (A.5)-(A.7), replace P by P+1, and go to step 3. Otherwise, terminate the computational process.

Note that in performing step 3, the optimal solution from cycle i provides a basic feasible solution for cycle i+1. By employing this fact, the optimum solution in step 3 may often be obtained after a single simplex pivot operation.

To illustrate the computational procedure, suppose that

\[ I = 9 \]
\[ J = 13 \]
\[ A = \$300.00 \]
\[ K = 4 \]
\[ h = 0.01 \]

and suppose that the unit cost and demand requirements for each item are as displayed in Table A.1. Further, assume the
TABLE A.1.
Item Data for 13-Item Problem.

<table>
<thead>
<tr>
<th>Item</th>
<th>Unit Cost</th>
<th>d_{i1}</th>
<th>d_{i2}</th>
<th>d_{i3}</th>
<th>d_{i4}</th>
<th>d_{i5}</th>
<th>d_{i6}</th>
<th>d_{i7}</th>
<th>d_{i8}</th>
<th>d_{i9}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.14</td>
<td>22</td>
<td>148</td>
<td>20</td>
<td>10</td>
<td>17</td>
<td>171</td>
<td>86</td>
<td>274</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>1.38</td>
<td>1000</td>
<td>15881</td>
<td>3955</td>
<td>9001</td>
<td>3075</td>
<td>4727</td>
<td>4110</td>
<td>8973</td>
<td>8101</td>
</tr>
<tr>
<td>3</td>
<td>1.09</td>
<td>367</td>
<td>45</td>
<td>515</td>
<td>56</td>
<td>860</td>
<td>174</td>
<td>481</td>
<td>161</td>
<td>308</td>
</tr>
<tr>
<td>4</td>
<td>0.81</td>
<td>139</td>
<td>50</td>
<td>235</td>
<td>191</td>
<td>239</td>
<td>120</td>
<td>157</td>
<td>314</td>
<td>432</td>
</tr>
<tr>
<td>5</td>
<td>1.54</td>
<td>33</td>
<td>41</td>
<td>238</td>
<td>150</td>
<td>62</td>
<td>307</td>
<td>132</td>
<td>218</td>
<td>123</td>
</tr>
<tr>
<td>6</td>
<td>2.50</td>
<td>36</td>
<td>74</td>
<td>68</td>
<td>83</td>
<td>93</td>
<td>90</td>
<td>100</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>0.09</td>
<td>117</td>
<td>65</td>
<td>62</td>
<td>30</td>
<td>128</td>
<td>42</td>
<td>66</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>0.28</td>
<td>85</td>
<td>390</td>
<td>633</td>
<td>713</td>
<td>480</td>
<td>412</td>
<td>712</td>
<td>319</td>
<td>177</td>
</tr>
<tr>
<td>9</td>
<td>0.84</td>
<td>-0-</td>
<td>-0-</td>
<td>10</td>
<td>12</td>
<td>-0-</td>
<td>2</td>
<td>-0-</td>
<td>-0-</td>
<td>-0-</td>
</tr>
<tr>
<td>10</td>
<td>0.41</td>
<td>85</td>
<td>100</td>
<td>180</td>
<td>403</td>
<td>233</td>
<td>1329</td>
<td>1179</td>
<td>639</td>
<td>1270</td>
</tr>
<tr>
<td>11</td>
<td>0.45</td>
<td>112</td>
<td>14</td>
<td>74</td>
<td>220</td>
<td>699</td>
<td>686</td>
<td>1357</td>
<td>2719</td>
<td>2104</td>
</tr>
<tr>
<td>12</td>
<td>3.39</td>
<td>-0-</td>
<td>33</td>
<td>2</td>
<td>8</td>
<td>28</td>
<td>32</td>
<td>8</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>13</td>
<td>53.00</td>
<td>-0-</td>
<td>-0-</td>
<td>-0-</td>
<td>15</td>
<td>12</td>
<td>-0-</td>
<td>-0-</td>
<td>-0-</td>
<td>151</td>
</tr>
</tbody>
</table>
obligation constraints take the form

\[
\sum_{j=1}^{J} \sum_{i=n-1}^{n} c_{ij} q_{ij} \leq 1.10 D_n \tag{A.16}
\]

where \( n = 2k \) for \( k = 1,2,3,4 \), and where

\[
D_n = \sum_{j=1}^{J} \sum_{i=n-1}^{n} d_{ij} \tag{A.17}
\]

The constraints (A.16) state that the \( k \)th budget interval consists of periods \( 2k-1 \) and \( 2k \), and that total obligations during a given budget interval cannot exceed 110% of the total dollar value of demands against the system during the budget interval.

Some of the major events in the solution of the above problem are as follows:

**Cycle 1**

**Step 1.** Applying (A.17), the vector of elements in the right-hand side of (A.6) and (A.7) are:

\[
\begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
b_4 \\
l
\end{bmatrix} = \begin{bmatrix}
27,174 \\
23,572 \\
17,167 \\
25,417 \\
1
\end{bmatrix} \tag{A.18}
\]
Step 2. A starting basis may be formed from both slack variables and one artificial variable. Hence, set 
\( P = 4 + 1 = 5 \).

Step 3. In the initial cycle, the optimal solution to step 3 is obtained immediately, since all \( P \) vectors are in the basis. The optimal dual variables are \( \theta_k = 0 \), \( k = 1, 2, 3, 4 \) and \( \sigma = CC = 10^6 \).

Step 4. Each of the thirteen subproblems are now solved using the Wagner-Whitin algorithm. The procurement schedule for items 1, 2, and 13 produced by this process and the ordering and holding costs of these schedules over the 9-period planning horizon are:

<table>
<thead>
<tr>
<th>Period</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>706</td>
<td>20,836</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>20,913</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>17,074</td>
<td>--</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>--</td>
<td>151</td>
</tr>
</tbody>
</table>

Cost $305  $1,683  $606
The procurement schedules for items 3-12 produced during this cycle are of the same form as the schedule for item 1, i.e., because of the very high ordering cost it is optimal to order these items but once during the planning horizon. This order is placed during the first period in which there is a positive level of demand.

Step 5. Substituting the solutions from step 4 into (A.8) and (A.9), we obtain

$$
\begin{bmatrix}
    c^6 \\
    g^6_1 \\
    g^6_2 \\
    g^6_3 \\
    g^6_4
\end{bmatrix}
= 
\begin{bmatrix}
    6,382 \\
    44,738 \\
    30,311 \\
    -0- \\
    23,562
\end{bmatrix}
$$

(A.19)

Thus, the production schedules generated during this cycle require $44,738 worth of orders to be placed during the first budget interval, and obligations of $30,311, $0, and $23,562, respectively, during the second, third, and fourth budget intervals. The total ordering and holding costs over the planning horizon for all items is $6,382. Items 2, 9, and 13 are the only items with obligations in the last three budget intervals under procurement schedules produced during this cycle.
Step 6. Substituting into (A.10), we obtain

\[ c^6 - \sum_k \theta_k g_k^6 - \sigma = 6,382 - 0 - 0 - 0 \]

\[-0 - 10^6 = -993,617\]

Hence, the criterion (A.5) may be improved by entering a vector corresponding to this procurement plan into the basis of the problem (A.5)-(A.7). The required vector is of the form of (A.16), with a "1" appended below the \( g_4^6 \) element. The "1" corresponds to the convexity constraint (A.6). The procedure now returns to step 3 for the second cycle of calculations.

**Cycle 2**

Step 3. The vector corresponding to (A.16) enters the basis, replacing the slack variable associated with the first budget interval. The new dual variables are \( \theta_1 = -22.2, \theta_2 = \theta_3 = \theta_4 = 0, \) and \( \sigma = 10^6. \)

Step 4. In the subproblems, \( \xi_1 = -\theta_1 = 22.2, \) and \( \xi_2 = \xi_3 = \xi_4 = 0. \) The value of \( \xi_1 \) represents a rather severe penalty for purchases during the first budget interval. The subproblem solutions reflect this fact. The procurement schedules for items 1, 2, and 13 are:
<table>
<thead>
<tr>
<th>Period</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>170</td>
<td>16,881</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>590</td>
<td>16,031</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>8,837</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>17,074</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>151</td>
</tr>
</tbody>
</table>

Cost $603 $1,796 $606

Step 5. Substituting the subproblem solutions into (A.8) and (A.9), we obtain

\[
\begin{bmatrix}
  c_7 \\
g_1^7 \\
g_2^7 \\
g_3^7 \\
g_4^7 
\end{bmatrix}
= \begin{bmatrix}
  9,208 \\
  24,704 \\
  38,150 \\
  12,195 \\
  23,562 
\end{bmatrix}
\]

Step 6. Using the above values in (A.10), we obtain

\[
9,208 - (-22.2)(24,704) - 0 - 0 - 0 - 10^6 = -442,120.
\]

Hence, the computations return to step 3.
Cycles 3 through 34

The calculations for these cycles follow the pattern of cycles 1 and 2. During cycle 7, the artificial vector leaves the basis.

The calculations converge asymptotically to the optimum value of (A.5). Initial improvement in the criterion is quite rapid. In cycle 10, the criterion (A.5) has a value of $9,070; in cycle 15, the value is $8,454; while the optimal value, obtained in cycle 34, is $8,235. In the last 10 cycles, the objective improves by less than $50. The production plans produced each cycle, however, appear to be significantly different until the final cycles of the computation.

Optimality is reached during cycle 34, after .047 hour of computing time. At termination, the following vectors are in the basis of the approximating problem:

<table>
<thead>
<tr>
<th>$\lambda_9$</th>
<th>$\lambda_{32}$</th>
<th>$\lambda_{30}$</th>
<th>$\lambda_{33}$</th>
<th>$\lambda_{34}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8,406</td>
<td>8,032</td>
<td>8,063</td>
<td>8,019</td>
<td>8,313</td>
</tr>
<tr>
<td>26,541</td>
<td>27,784</td>
<td>27,161</td>
<td>28,129</td>
<td>26,819</td>
</tr>
<tr>
<td>23,574</td>
<td>23,976</td>
<td>23,976</td>
<td>22,609</td>
<td>24,941</td>
</tr>
<tr>
<td>19,262</td>
<td>14,835</td>
<td>18,240</td>
<td>15,858</td>
<td>14,835</td>
</tr>
<tr>
<td>18,054</td>
<td>32,014</td>
<td>29,333</td>
<td>32,014</td>
<td>32,014</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Comparing the above solutions with the constraint vector (A.18), we see that each of the associated procurement plans provide an approximate solution to the original problem (A.1)-(A.4). If the constraints are not absolute, one of these plans, say plan 9, might be selected for implementation. Otherwise, these plans provide a good starting point for applying techniques suggested by Everett (1963) for exploring the gaps in the payoff-resource space envelope associated with this problem.
BIBLIOGRAPHY


46. ________, GLM and Integer Programming, Computer Sciences Center, Southern Methodist University, Dallas, Texas, October 1969, 11 pp.


