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

SOLUTION METHODS FOR STOCHASTIC DYNAMIC LINEAR PROGRAMS

by

John R. Birge

TECHNICAL REPORT SOL 80-29

December 1980

Research and reproduction of this report were supported by the Department of Energy Contract DE-AC03-76SF00326, PA# DE-AT03-76ER72018; Office of Naval Research Contract N00014-75-C-0267; National Science Foundation Grants MCS76-81259, MCS-7926009 and ECS-8012974 (formerly ENG77-06761).

PREFACE

Many practical problems in industrial and social planning require optimal decisions to be made periodically through time. Linear programs, called *dynamic linear programs*, can often be formulated to model the requirements of these decision processes. These programs are generally quite large and difficult to solve. The search for efficient methods in finding their optimal solutions has been a major topic in operations research for the past 25 years.

Often the problem modeled by a dynamic linear program involves uncertainties which can complicate and exponentially increase the program's size. There may be several possible outcomes in the future, but deterministic linear programs usually only consider the most likely outcome. In this dissertation, we present methods for solving the *stochastic dynamic linear program*, the dynamic linear program with uncertainties explicitly included.

Our methods take advantage of the program structure. Dynamic linear programs are characterized by a *staircase* structured coefficient matrix, in which non–zero elements appear only in blocks along the diagonal or adjacent to the diagonal. This structure makes many efficient techniques possible. We will show that the stochastic model's specific structure can lead to additional procedures, and that these procedures may improve upon complicated "brute force" solution methods.

We begin in Chapter I by presenting sufficient conditions for a deterministic problem's optimal solution to solve a stochastic problem. The second chapter discusses the difficulties involved in using deterministic solutions in general. We also explore the possibilities for combining separate deterministic solutions and give examples of problems that require the stochastic dynamic linear program to be solved.

Chapters III, IV, and V present methods for solving the full stochastic program. The first method follows from the decompositions approach to large–scale programming. The next two methods employ different large–scale structured programming techniques, in which, the basis is partitioned but not completely decomposed.

Chapter VI demonstrates that the methods we present are actually dynamic programming approaches. They only differ in their strategies for approximating the optimal state space solution at each stage.

In the final chapter, we present some computational results for our algorithms and discuss potential areas of applications. We also state our conclusions on the use of stochastic dynamic linear programs and suggest areas of future research.

In this dissertation, we use standard mathematical notation. More specific notational conventions are defined in the text. Within each chapter, we refer to equations and propositions by their order of presentation in that chapter (eg.,equation (12)). In reference to equations in other chapters, we prefix the equation by the chapter's roman numeral (eg., (II.12)).

TABLE OF CONTENTS

CHAPTER I

Deterministic Solutions for Stochastic Dynamic Linear Programs

1. Introduction

Dynamic linear program models have been formulated for many different practical situations. When these models involve uncertain quantities, the solution of the resulting stochastic dynamic linear program can be very difficult. Under some circumstances, however, an associated deterministic problem can be stated that is easier to solve and can be used as the "solution" to the stochastic program.

In this chapter, we will state conditions that imply that this *deterministic solution* is in fact optimal for the stochastic dynamic linear program. We call a solution "deterministic" it solves a program that does not allow for any uncertainty in the program parameters. We also will use *myopic solution* to refer to a solution of a program in period t that does not make use of information from periods after t. A solution that considers uncertainty in the future, is called *stochastic*.

In our development of an optimal deterministic solution, we first present the basic multi-stage model and the various approaches taken for its solution. Also, in Section 2, we introduce some terms and notation that appear in the following chapters and discuss the value of having information about the random variables. To do this, we present inequalities that measure the superiority of the stochastic solution over a deterministic solution.

The chapter concludes in Section 3 with the description of conditions for an optimal deterministic solution. We also give an example of a problem in which a deterministic solution is optimal and illustrate how added complications can necessitate a stochastic solution.

2. The Multi–Stage Problem

The dynamic, or multi–stage, linear program, to which we shall refer, has the following form:

$$
\begin{array}{llllll}
\min & z_1 = c_1x_1 + c_2x_2 & \cdots & + c_Tx_T & & \\
\text{subject to} & A_1x_1 & & & = b_1, & \\
 & -B_1x_1 + A_2x_2 & & & = \xi_2 & (DLP) \\
 & & \ddots & & \vdots & \\
 & & & -B_{T-1}x_{T-1} + A_Tx_T & = \xi_T, & \\
 & & & x_t & \geq 0 \text{ for all } t, &
\end{array}
$$

where $x_t \in \Re^{n_t}$ (n_t-dimensional Euclidean space), $b_1 \in \Re^{m_1}$, $\xi_t \in \Re^{m_t}$, and the vectors, c_t, and matrices, A_t and B_t, are dimensioned to conform.

For a DLP problem, the right–hand sides, ξ_t, are given. For a situation in which ξ_t is random, the DLP becomes one possible program out of the possible outcomes for ξ_t. In our analysis here, we will not let any of the other quantities be random, so c_t, A_t, B_t, and b_1, will be assumed known. Moreover it is assumed ξ_t and $\xi_{t'}$ are independent $t \neq t'$.

We can also view DLP as an optimal control problem by assuming more structure for the matrices. In this case, we would have $x_t = (y_t, u_t)$, where y_t is a state variable and u_t is a control variable, and we would partition the matrices and vectors as

$$A_t = \begin{pmatrix} I & 0 \\ G_t & D_t \end{pmatrix},$$

$$\xi_t = \begin{pmatrix} 0 \\ \xi_t^2 \end{pmatrix} \tag{1}$$

and

$$B_t = \begin{pmatrix} A_t & B_t \\ 0 & 0 \end{pmatrix}. \tag{2}$$

This formulation includes transitions from state to state according to

$$-A_t y_t - B_t u_t = y_{t+1}$$

and interperiod requirements of

$$G_{t+1} x_{t+1} + D_{t+1} u_{t+1} = \xi_{t+1}.$$

The random vector y_{t+1} represents unknown state to state transitions and future requirements. Although some computational efficiency may be afforded by the special structure of this model, we will restrict our discussion to the general case of DLP.

The first approach we consider for DLP is to solve it for all possible ξ_t and then take the expected value of z_1 as the measure of cost. This approach requires *perfect information* of the outcomes of all future events and is known in the literature as the "wait–and–see" solution (see Madansky [38]). It would be the optimal solution, if one could somehow wait and see until the end of the planning horizon and could beforehand make decisions based on what will occur. Obviously such a perfect information solution is not implementable. When averaged over every possible ξ_t, it provides a measure of the best expected value one could achieve given advanced information about the random variables. We write the expected value of this solution as

$\bar{z}_1$, where,

$$\bar{z}_1 = \mathrm{E}_{\xi_T,\ldots,\xi_2}[z_1(\xi_T,\ldots,\xi_2)], \tag{3}$$

the expectation, "E", is with respect to the random variables, $\xi_2,\ldots,\xi_T$, and z_1 is defined as in DLP with parameters $\xi_2,\ldots,\xi_T$.

An alternative and more realistic approach to the stochastic problem is to consider that during each period t the value ξ_t is known and a decision on x_t must be made without knowledge of the realizations of the future periods' uncertainties. This is known as the "here–and–now" solution because it reflects the need for current decisions. The program can be written as

$$\begin{array}{llllllll}
\min & z_2 = & c_1x_1 & + \mathrm{E}_{\xi_2}[c_2x_2 & + \mathrm{E}_{\xi_3}[c_3x_3 & ++ \mathrm{E}_{\xi_T}[c_Tx_T]] & & \\
\text{subject to} & & A_1x_1 & & & & & = b_1, \\
 & & -B_1x_1 & +A_2x_2 & & & & = \xi_2, \\
 & & & & & & & = \vdots \\
 & & & & -B_{T-1}x_{T-1} & & +A_Tx_T & = \xi_T, \\
 & & & & & & x_t & \geq 0, \\
 & & & & & & \xi_t & \in \Xi_t,
\end{array}$$

for $t = 2,\ldots,T$.

(4)

Problem (4) states that for $t = T$, $\hat{x}_{T-1}$ is given , $\xi_T = \hat{\xi}_T$ is observed and $\hat{x}_T \geq 0$ chosen so that c_Tx_T is minimum. Assuming that $\hat{x}_T$ will be so chosen, $\hat{x}_{T-1}$ is chosen so that, $c_{T-1}\hat{x}_{T-1} + \mathrm{E}_{\xi_T}[c_Tx_T]$ is minimum given $\hat{x}_{T-2}$ and ξ_{T-1} observed, etc.

In general, the decision process for an actual implementation proceeds as follows:

A. A decision, $\hat{x}_1$, is made and implemented.

B. A realization, $\hat{\xi}_2$, of ξ_2 is observed, and a decision, $\hat{x}_2$, is made and implemented.

C. The process is repeated to find each $\hat{x}_t$ in period t, given the past decision $\hat{x}_{T-1}$ and the outcome of the random vector ξ_{T-1}.

We are going to investigate the effects of using different methods for determining the decisions, $\hat{x}_t$. Not all of these methods exactly solve (4), so we must evaluate the expected value of the objective function for solutions by each method. For a given method, μ, of choosing $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_T$, we first define a function of the random variables, ξ_t, which gives the cost of using those decisions. We write this as

$$
\begin{aligned}
z(\hat{x}, \xi \mid \mu) = {} & c_1 \cdot \hat{x}_1(\mu) \\
& + c_2 \cdot \hat{x}_2(\hat{x}_1, \hat{\xi}_2 \mid \mu) + \cdots + c_T \cdot \hat{x}_T(\hat{x}_1, \ldots, \hat{x}_{T-1}, \hat{\xi}_2, \ldots, \hat{\xi}_{T-1} \mid \mu),
\end{aligned} \tag{5}
$$

where $\hat{x}_t(\hat{x}_1, \ldots, \hat{x}_{t-1}, \hat{\xi}_2, \ldots, \hat{\xi}_t \mid \mu)$ is the decision chosen by μ given the previous decisions, $\hat{x}_1, \ldots, \hat{x}_{t-1}$, and observations, $\hat{\xi}_2, \ldots, \hat{\xi}_t$.

We can then take expectations over the random vectors to determine the expected cost of the solution found by method μ. We write this as

$$\overline{z}(\mu) = \mathrm{E}_{\xi}[z(\hat{x}, \xi \mid \mu)]. \tag{6}$$

An exact (but expensive) method for finding an optimal solution to (4) is to proceed by a dynamic programming scheme with backward iteration. We will call this "Method 2". We begin by setting the terminal valuation function:

$$
\begin{aligned}
z_2^T(x_{T-1}, \xi_T) = {} & \min c_T x_T \\
\text{subject to} \quad & A_T x_T = \xi_T + B_{T-1} x_{T-1}, \\
& x_T \geq 0
\end{aligned} \tag{7}
$$

and let $z_2^T(x_{T-1}) = \mathrm{E}_{\xi_T}[z_2^T(x_{T-1}, \xi_T)]$. We further define the recursion on the valuation function as:

$$\begin{aligned} z_2^t(x_{t-1}, \xi_t) = \ & \min\ c_t x_t + z_2^{t+1}(x_t) \\ \text{subject to} \quad & A_t x_t = \xi_t + B_{t-1} x_{t-1}, \\ & x_t \geq 0 \end{aligned} \tag{8}$$

and let $z_2^t(x_{t-1}) = \mathrm{E}_{\xi_t}[z_2^t(x_{t-1}, \xi_t)]$. We finally arrive at

$$\begin{aligned} z_2^1(b_1) = \ & \min\ c_1 x_1 + z(x_1) \\ \text{subject to} \quad & A_1 x_1 = b_1, \\ & x_1 \geq 0. \end{aligned} \tag{9}$$

Method $\mu = 2$ leads to a sequence of decisions, $\hat{x}_t(\mu = 2)$. We can take expectations of the outcome of these decisions as in (6). This yields

$$\overline{z}_2 = \mathrm{E}_\xi[z_2(\hat{x}, \xi \mid \mu = 2)]. \tag{10}$$

We observe that for some outcome $\hat{\xi} = (\hat{\xi}_2, \ldots, \hat{\xi}_T)$,

$$z_2(\hat{x}, \hat{\xi} \mid \mu = 2) = c_1 \hat{x}_1(z_2^1) + c_2 \hat{x}_2(z_2^2(\hat{x}_1, \hat{\xi}_2)) + \cdots + c_T \hat{x}_T(z_2^T(\hat{x}_{T-1}, \hat{\xi}_T))$$

where, for each t, $\hat{x}_t(z_2^t(\hat{x}_{t-1}, \hat{\xi}_t))$ is the optimal solution of (8). Hence, by integrating, we obtain

$$\overline{z}_2 = z_2^1(b_1).$$

So, for Method 2, the result in (9) is the minimum expected cost found by the method.

The expected value, $\overline{z}_2$, is the best possible solution for situations in which decisions must be implemented over time. (For details on this result, see Chapter VI on dynamic programming.) Enumerating the states, x_{t-1}, in general, can be very difficult, especially when ξ_t can have a continuous distribution. For this reason, we assume that either the distribution of ξ_t is discrete for all t or that we can approximate the continuous distribution by a discrete sample of size $\overline{k}_t$ where $\overline{k}_t$ is not too large. We assume, therefore, that for all t and some $\hat{\xi}_t \in \Re^{m(t)}$,

$$P(\xi_t = \hat{\xi}_t) = \begin{cases} p_t^1, & \text{if} \quad \xi_t = \xi_t^1; \\ p_t^2, & \text{if} \quad \xi_t = \xi_t^2; \\ ., & \quad . \\ p_t^{\overline{k}_t}, & \text{if} \quad \xi_t = \xi_t^{\overline{k}_t}; \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

where by our independence assumptions the probabilities p_t^i do not depend on earlier outcomes.

We have then $\overline{k}_t$ possible outcomes for the random right-hand sides in period t. The outcomes form a tree of possible values (see Figure 1.).

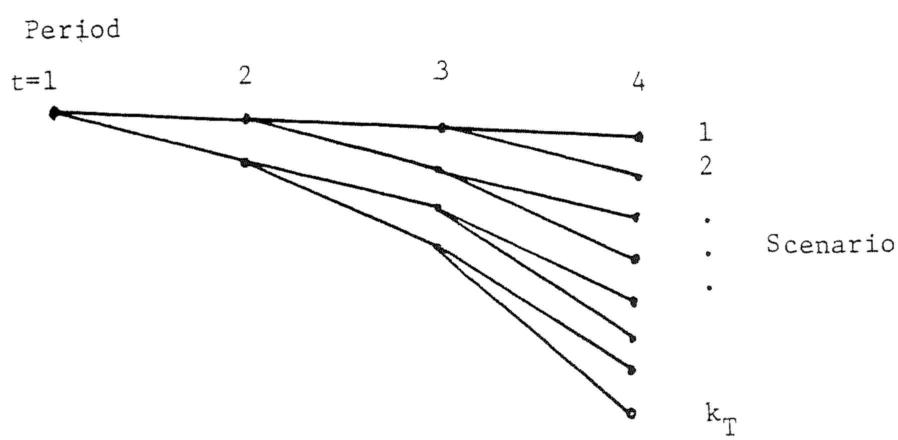

Figure 1. The outcome tree.

We also define k_t as the total number of possible outcomes from period 1 to period t, thus

$$k_t = \Pi_{\tau=1}^{t} \overline{k}_\tau. \tag{12}$$

The tree of outcomes includes k_t nodes in each period t. We call each node a *scenario*. In period t then, a scenario corresponds to a realization of outcomes, $\xi_1 = \hat{\xi}_1, \ldots, \xi_t = \hat{\xi}_t$.

Using this framework, a *descendant* of a scenario j in period t, is defined as any node in periods $t+1$ to T on the branch connected to node j. We adopt the notation $\overline{j}$ for descendants of j. An *ancestor* of j is then a node on the same branch as j in periods 1 to $t-1$. We denote an ancestor of j as $\hat{j}$. These definitions will apply in this and all subsequent chapters.

Given the discrete distribution, we can write (4) as an explicit linear program. (4) becomes

$$\begin{array}{llllll}
z_3 = \min & c_1 x_1 + \sum_{j=1}^{k_2} p_2^j c_2 x_2^j & + \cdots + & & \sum_{j=1}^{k_T} p_T^j c_T x_T^j & \\
\text{subject to} & & & & & \\
& A_1 x_1 & & & & = b_1, \\
& -B_1 x_1 & +A_2 x_2^j & & & = \xi_2^j, \text{ for all } j, \\
& & & \ddots & & \vdots \\
& & & -B_{T-1} x_{T-1}^{\hat{j}} & +A_T x_T^j & = \xi_T^j, \text{ for all } j, \\
& & & & x_t & \geq 0 \text{ for all } t.
\end{array} \tag{SDLP}$$

This program, SDLP, is the primary focus of our presentation. It represents a formulation of the general stochastic dynamic linear program given a discrete distribution. By defining p_t^j as the probability of a node in the tree of outcomes, we can also incorporate interdependence of the time periods into the model.

SDLP will be analyzed as a structured linear program (presented in Figure 2). Its structure resembles the staircase structure of deterministic dynamic linear programs, but the repetition of the $-B_t$ blocks for each descendant scenario forms spikes below the diagonal. This property makes the strict application of staircase approaches difficult.

Another complication of the stochastic model is that the number of *blocks*, non-zero partitions of the coefficient matrix, grows exponentially with the number of periods, as we see in (12). We will present methods for solving SDLP that reduce the effects of this complication.

The decision process of solving SDLP will be called Method $\mu = 3$. The expected cost found by this method is

$$\bar{z} = \mathrm{E}[z_3(\hat{x}, \xi \mid \mu = 3)], \tag{13}$$

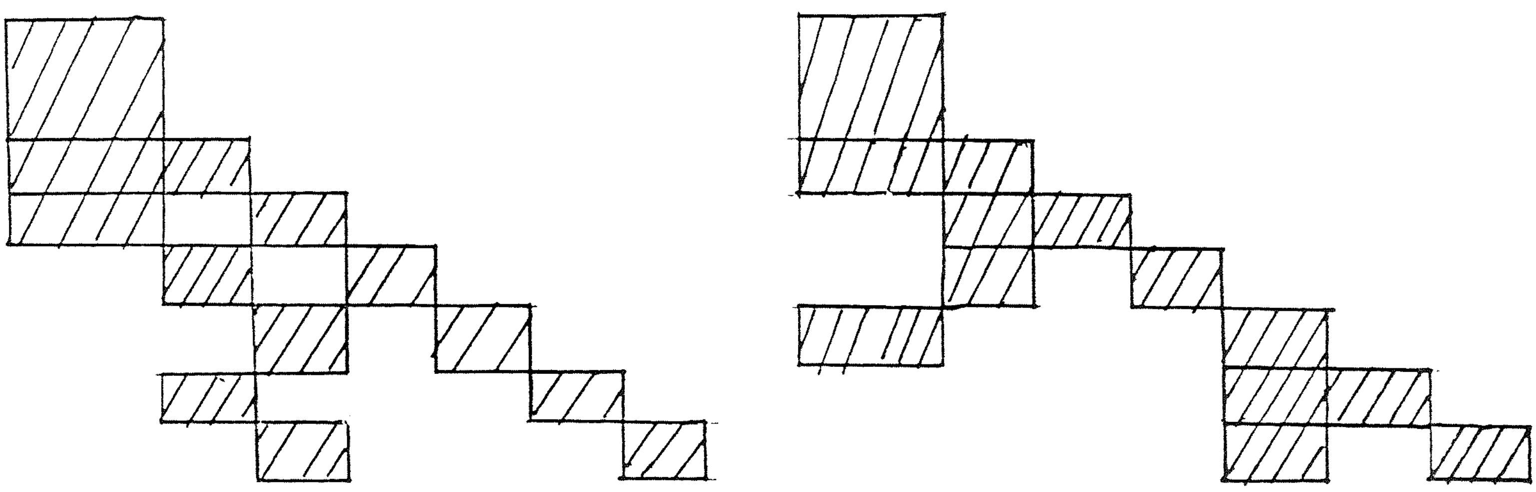

Figure 2. Alternative structures.

An extreme simplification in SDLP would be to consider only one possible outcome for each random variable. A reasonable choice would be the expectation, $\overline{\xi}_t$. The corresponding program has the same form as DLP in which the means, $\overline{\xi}_2, \ldots, \overline{\xi}_T$, appear on the right–hand side. The solution of this problem eliminates the increasing size problem of SDLP and is, therefore, the form usually chosen in practice. In some instances, as we show below, it can even yield the optimal solution to the stochastic program.

We call this deterministic approach of substituting the expectations for the random variables, Method $\mu = 4$. We compute the expected cost of Method 4 then as

$$\overline{z}_4 = \mathrm{E}_\xi [z_4(\hat{x}, \xi \mid \mu = 4)]. \tag{14}$$

The above four approaches to the stochastic dynamic linear program are all related to one another as the following lemma states.

Lemma 1. *The expected cost of the approaches presented above are ordered by*

$$\overline{z}_1 \leq \overline{z}_2 \leq \overline{z}_3, \quad \overline{z}_1 \leq \overline{z}_2 \leq \overline{z}_4, \tag{15}$$

and, for (discrete) distributions without approximations $\overline{z}_1 \leq \overline{z}_2 \equiv \overline{z}_3 \leq \overline{z}_4$.

Proof. Each method can be shown to improve upon the solution by the previous method. For $\overline{z}_1 \leq \overline{z}_2$, we observe that $z_1(x, \xi \mid \mu = 1) \leq z_2(x, \xi \mid \mu = 2)$ since Method 1 chose the optimum solution for each ξ whereas $\mu = 2$ does not in general. Integration preserves the inequality, hence, $\overline{z}_1 \leq \overline{z}_2$.

The next inequality involves either $\mu = 3$, where an incorrect discrete distribution can be used to approximate the correct one, or $\mu = 4$,where the distribution is replaced by one calculated at the expected value. Since

Method 2 optimizes (8) for all ξ_t, $\mathrm{E}_{\xi_t}[z_2^t(x_{t-1}, \xi_t)] \leq \mathrm{E}_{\xi_{t-1}}[z_3^t(x_{t-1}, \xi_{t-1})]$ for all t. Hence, $\overline{z}_2 \leq \overline{z}_3$ or $\overline{z}_2 \leq \overline{z}_4$ follows. ∎

The lemma includes the obvious result that, for the expected value of perfect information, EVPI, where EVPI$= \overline{z}_1 - \overline{z}_2$ if the correct distribution is used or $\overline{z}_1 - \overline{z}_3$ if not,

$$\text{EVPI} \geq 0.$$

This difference represents the maximum amount that one would pay for information about the future. When the EVPI is low there may be little necessity in refining forecasts, but, when it is high, incomplete information about the uncertainties may be costly. In the following chapter, we present an example of this possibility.

A second quantity that we want to examine here is the difference $\overline{z}_2 - \overline{z}_4$ or $\overline{z}_3 - \overline{z}_4$, which we call the *value of the stochastic solution* , VSS. VSS measures the benefit from solving a stochastic program over solving its deterministic approximation. A low VSS indicates that the more complicated SDLP might not be worth the extra effort. VSS can be bound without solving SDLP, as we show in the next chapter.

We note that in Lemma 1 to guarantee that $\overline{z}_3 \leq \overline{z}_4$, we had to assume that the discrete distribution was correct. If the distribution used was only an approximation, then it is possible to make an estimate of the distribution that would lead to $\overline{z}_3 > \overline{z}_4$. This anomaly can occur because some scenarios could lie under sections of the piecewise linear curve, $z_3^t(x_{t-1}, \xi_t)$, that lead to high penalties. We discuss these scenario results more closely in Chapter 2. For our purposes, we assume that the discrete distribution used suffices because more information about the distribution is not available. The consideration of individual scenarios is the only alternative.

Given these assumptions about the distribution, we would still like to know when SDLP should be solved. To show when SDLP need not be solved, in the next section, we present some conditions that imply VSS=0. In the following chapter, we will give bounds on VSS that also aid in evaluating whether SDLP is worth solving.

3. Optimal Deterministic Solutions

When the numerical costs of solving a stochastic problem are high, a deterministic solution technique is attractive. Since decisions often cannot wait for the detailed analysis of all future possibilities, the method based on assuming some "best guess" of the future environment, is most often the one implemented. In fact, the even simpler policy of using a myopic solution may provide a good basis for decisions. By finding conditions for VSS=0, we are trying to avoid the effort of correctly solving a general stochastic program. With the conditions below, we can check whether the stochastic program need be tried at all. The following lemma, a well-known result from sensitivity analysis, is fundamental in our development.

Lemma 2. *Let B be an optimal basis for DLP with $\xi = (\overline{\xi}_2, \overline{\xi}_3, \ldots, \overline{\xi}_T)$. If B remains feasible for all $\xi \in \Xi$, then B is an optimal basis for DLP for all ξ.*

Proof. Partition the coefficient matrix and cost row according to basic variables, x_B, and non-basic variables, x_N. DLP becomes

$$
\begin{array}{lrl}
\min & c_B x_B + c_N x_N & \\
\text{subject to} & B x_B + N x_N & = (b_1, \bar{\xi})', \\
& x_B & \geq 0, \\
& x_N & \geq 0,
\end{array}
\tag{16}
$$

where we use the notation v' to indicate the transpose of v, so $(b_1, \bar{\xi})'$ is the column vector of right–hand sides in DLP.

For B optimal, there exist prices, π, such that

$$
\begin{aligned}
\pi B &= c_B, \\
\pi N &\leq c_N,
\end{aligned}
\tag{17}
$$

and

$$
\begin{aligned}
x_B = B^{-1}(b_1, \bar{\xi})^T &\geq 0, \\
x_N &\geq 0.
\end{aligned}
\tag{18}
$$

If x_B remains feasible for all ξ in (18), (17) still holds, guaranteeing dual feasibility and complementarity. Hence, B is still an optimal basis. ∎

The next problem we might encounter is that of testing whether B is indeed feasible for all values of ξ. An enumeration of all possible ξ is not necessary. Garstka and Rutenberg [25] showed that simple computations for many practical problems, could be performed quickly to find the probability that a given basis is optimal. Their process involves fixing some components in the lattice of discrete values of ξ and then finding the feasible range for the remaining components. This method also proves valuable in the subproblem solutions we investigate in Chapter 3.

To use a basis which satisfies the conditions in Lemma 2 in SDLP, still other conditions must be met. The next lemma helps us find these conditions. For this lemma, we will use a solution from DLP in SDLP. We do this by letting the set of basic activities in DLP, $\{x_t^B : t = 1, \ldots, T\}$, be repeated to form a basic set in SDLP. This basic set in SDLP is $\{x_t^{B_j} : j = 1, \ldots, k_t; t = 1, \ldots, T\}$, where $x_t^{B_j} = x_t^B$ for all j.

Lemma 3. *Let the set of activities for a feasible basis, B, in DLP be $\{x_1^B, \ldots, x_T^B\}$ where each x_t^B represents activities from period t. Also, let SDLP have at least two distinct new scenarios at each period (i.e., $\overline{k}_t \geq 2$). The activities, $\{x_t^{B_j}\}$, where $x_t^{B_j} = x_t^B$ for all j, form a feasible basis in SDLP if and only if x_t^B consists of $m(t)$ activities for all t.*

Proof. (See Figure 3.) Essentially we shall show that if the count on the number of basic x_t is not $m(t)$ for all t that the corresponding SDLP candidate for basis will be singular. For the necessity of the condition, first let $\mu(t)$ be the number of elements in x_t^B. Assume $x(t)$ does not have $m(t)$ elements for all t. Thus, there exists some $\mu(t) > m(t)$. (If not, since $\sum_{t=1}^T m(t) = m$, $\mu(t) = m(t)$ for all t.)

We note that, for $\mu(t) > m(t)$, $t < T$. This is true because, if $\mu(T) \geq m(T)$, then $\sum_{t=1}^{T-1} \mu(t) < \sum_{t=1}^{T-1} m(t)$, which contradicts the fact that $\{x_t^{B_j}\}$ corresponds to a basis. Set $t' = \min\{t : \mu(t) > m(t)\}$. We note that there exists no $t'' < t'$ such that $\mu(t'') < m(t'')$. Again, this would mean the basis was not of full rank. Now, we have $\sum_{t=1}^{t'} \mu(t) - \sum_{t=1}^{t'} m(t) = \overline{k}_{t'} \cdot \delta_t$, where $\delta_t = \mu(t') - m(t')$. But, for $t > t'$, $\sum_{t=t'+1}^{T} \mu(t) - \sum_{t=t'+1}^{T} m(t) = -\overline{k}_{t'} \cdot \overline{k}_{t'+1} \cdot \delta_t$, since each deficiency is repeated $\overline{k}_{t'+1}$ times. Hence, $\sum_{t=1}^{T} \mu(t) = \overline{k}_{t'}(1 - \overline{k}_{t'+1}) \cdot \delta_t + \sum_{t=1}^{T} m(t)$, which, by our assumption, implies the columns of the activities, $\{x_t^B\}$, do not form a basis in SDLP. To show sufficiency, first note that if x_t^B has $m(t)$ elements in each period t, then, for all t, there exists a square non-singular partition of the basis, B_t, with columns and rows only in t. (If not, B does not span the row space in period t.) Therefore, in SDLP, the set of columns, $\{B_t(j)\}$, is linearly independent. By construction, the x_t^B

correspond to $m(1) + \sum_{t=2}^{T} k_t m(t)$ columns, so the activities form a basis. ∎

From this lemma, we obtain our result as stated in the following theorem.

Theorem. *If the optimal basis, B, for the program DLP, with $\xi = \bar{\xi}$, is feasible for all $\xi \in \Xi$, and if B has as many columns as there are rows in each period, then the set of activities in B forms an optimal basis in SDLP, and $VSS = 0$.*

Proof. First, to show primal feasibility, let $A_t^{B_j}$ be the square non-singular submatrix associated with the activities x_t^B in scenario j. For all t and j, we have

$$\hat{x}_t^{B_j} = (A_t^{B_j})^{-1} (\hat{\xi}_t^j + B_{t-1} \hat{x}_{t-1}^{B_j}), \tag{19}$$

which is the same value as $\hat{x}_t^B$ in DLP for $\xi_t = \hat{\xi}_t$. Hence, $\hat{x}_T^{B_j} \geq 0$ by Lemma 1 for all j and t.

Let π_t^B be the dual variables in DLP for the basis, B. Next, define $\pi_t^{B_j} = p^j \pi_t^B$. In period T, we obtain

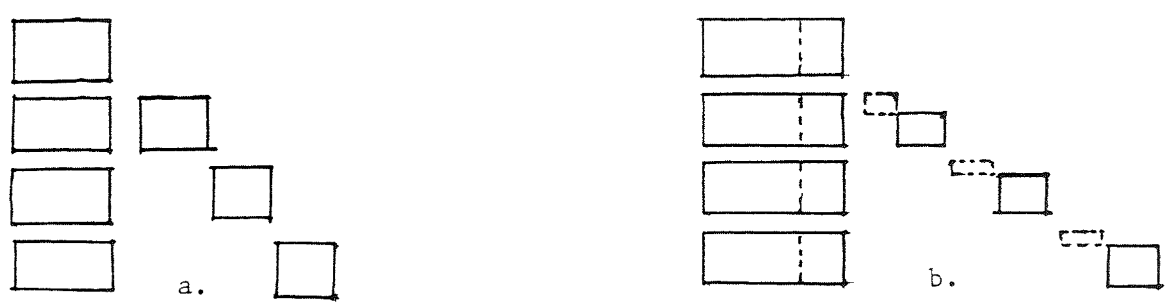

Figure 3. Example of (a) a basis and (b) non-spanning columns for a 2 period exampl

$$\pi_t^{B_j} = p^j \pi_t^B \tag{20}$$

and

$$\pi_T^{B_j} A_T^{N_j} \leq p^j c_T^N, \tag{21}$$

where the c_T^N are the non-basic costs in scenario j for $j = 1, \ldots, k_T$. In general, for $\pi_t^{B_j}$ in SDLP, we have

$$-\sum_{\bar{j}=1}^{\bar{k}_{t+1}} \pi_{t+1}^{B_{\bar{j}}} B_t^{B_j} + \pi_t^{B_j} A_t^{B_j} = -p^j \pi_{t+1}^B B_t^B + p^j \pi_t^B A_t^B = p^j c_t^B \tag{22}$$

and

$$-\sum_{\bar{j}=1}^{\bar{k}_{t+1}} \pi_{t+1}^{B_{\bar{j}}} B_t^{N_j} + \pi_t^{B_j} A_t^{N_j} \leq c_t^N. \tag{23}$$

By (22) and (23), for $x_t^{B_j}$ feasible, B is an optimal basis in SDLP. (19) implies that the solution in SDLP given scenario j is the same as in DLP. Hence, $\bar{z}_3 = \bar{z}_4$, and VSS $= 0$. We further note that this also implies the optimal solution is myopic. ∎

This result gives us a method to check for deterministic optimality, but it may be difficult to satisfy these conditions in practical examples. Even when not satisfied, they could be useful, however, in finding an optimal solution with confidence α, where α is the probability of being optimal. The Garstka and Rutenberg procedure mentioned above would be useful for these computations. (This could also be applied to a problem of the form of a *chance-constrained program* as in Charnes and Cooper [14], but we wish to restrict our development to the *recourse problem*, SDLP.)

The next result presents an alternative set of conditions that may prove useful when the conditions in our theorem are not satisfied. We state them as a corollary.

Corollary. *Let $\{B(i)\}$ be a family of bases for DLP, where $B(i)$ is optimal for $\xi_i \in \Xi$. Assume also that, for all periods t and nodes j of scenario i, the set of basic activities, $\{x_t^{B(i(j))}\}$, is the same for all $i(j)$ that include the nodes, $\{\bar{j}_l : l = 1, \ldots, \bar{k}\}$, the descendants of j. If each $B(i)$ additionally has square blocks in each period, then the set of basic activities chosen from $\{B(i)\}$ is optimal in SDLP, and VSS $= 0$.*

Proof. Since $x_t^{B(i(j))}$ is the same for all $\bar{j}$, we can define a set of activities for SDLP as $x_t^{B_j} = x_t^{B(i(j))}$ for all t and j. Now, for primal feasibility, we again have (19) for all t and j, so $\hat{x}_t^{B_j} \geq 0$.

For the dual, define

$$\pi_t^{B_j} = p^j \pi_t^{B(i(j))}. \tag{24}$$

Hence, at period T,

$$\pi_T^{B_j} A_T^{B_j} = p^j c_T^B, \tag{25}$$

and

$$\pi_T^{B_j} A_T^{N_j} \leq p^j c_T^N. \tag{26}$$

For general t, we have

$$\begin{aligned} -\sum_{\bar{j}=1}^{\bar{k}_{t=1}} \pi_{t+1}^{B_{\bar{j}}} B_t^B + \pi_t^{B_j} A_t^{B_j} &= \sum_{\bar{j}=1}^{\bar{k}_{t+1}} p^{\bar{j}}(-\pi_{t+1}^{B(i(\bar{j}))} B_t^B + \pi_t^{B(i(j))} A_t^{B_j}), \\ &= p^j c_t^B, \end{aligned} \tag{27}$$

and

$$\sum_{j=1}^{\overline{k}_{t+1}} \pi_{t+1}^{B_j} B_t^N + \pi_t^{B_j} A_t^{N_j} \leq p^j c_t^N. \tag{28}$$

(25), (26), (27), and (29) give us dual feasibility and complementarity, proving that the set of variables $\{x_t^{B_j}\}$ is optimal. Again, from (19), the values are the same as in solving any deterministic form DLP, so VSS $= 0$.
∎

The corollary gives us more conditions for finding the optimal solution to SDLP without actually solving it. An example of a model which meets these requirements is the Hotelling–Nordhaus model of exhaustible resources (see [34] and [46]) and its extension by Chao [12].

In Chao's model, a dynamic production schedule is chosen to minimize the cost of satisfying an increasing sequence of demand requirements over time. The demands may be satisfied by any of $m-1$ technologies, each using one distinct resource, with finite availability and one "backstop" technology with no resource limits. The program is

$$\min \sum_{i=1}^{m} \sum_{t=0}^{\infty} \beta^t c_i y_{it} + \sum_{i=1}^{m} \sum_{t=0}^{T} \beta^t k_i x_{it} \tag{29}$$

subject to

$$\sum_{t=0}^{\infty} y_{it} \leq R_i, i = 1, 2, \ldots, m,$$
$$\sum_{i=1}^{m} y_{it} = D_t, t = 1, 2, \ldots, T,$$
$$y_{i,t+1} = y_{it} + \sum_{s=0}^{\infty} (\delta_s - \delta_{s-1}) x_{i,t-s},$$
$$y_{it} \geq 0, t = 0, 1, \ldots,$$
$$x_{it} \geq 0, i = 1, 2, \ldots, m,$$

where y_{it} is the amount of the demand, D_t, satisfied by resource i at time t, x_{it} is the amount of resource i committed at t so it may be extracted later, c_i is the current cost of technology i, k_i is the capital cost of i, β is the discount factor, δ_t is the extraction rate, and R_i is the initial availability of the resource used in technology i.

Chao showed that, for this model, a myopic solution is optimal for all future demands and supplies. This solution implies that a family of bases, $\{B(i)\}$, exists that satisfies the conditions of the corollary. Therefore, the stochastic program for (29), in which, D_t and R_i are random, has a deterministic and myopic solution and VSS = 0. We note that this very simple model can be modified so that VSS grows. Chao explored the case of price–responsive demands and found that, with a "sufficiently high" discount rate, the optimal decisions are still insensitive to "distant–future" uncertainties. Our example in Chapter II shows how near future uncertainties can greatly affect current decisions, also making VSS high.

This chapter has described the value of information in the consideration of decisions made over time. We presented the program, SDLP as a method for incorporating uncertainties into a decision process and explored the possibilities for finding a solution to SDLP without solving the full problem. The value of the stochastic solution is, however, not always low.

Chapter II

The Nature of the Stochastic Solution

1. Introduction

In Chapter I, we presented conditions, under which, the optimal solution to a certain deterministic program is the solution to the stochastic dynamic linear program. The deterministic model used the expected values of the random right–hand sides. . Unfortunately, the great majority of stochastic problems do not meet the certainty equivalence criterion, ie. the sufficient conditions for deterministic optimality. When a model fails the conditions of Chapter I, it would be desirable to solve the stochastic problem directly. In this chapter, we explore the value of that solution and the costs that can arise from not finding the optimal stochastic solution.

We shall concentrate on decisions based upon "risk neutrality", meaning we wish to optimize the expected value of our policy decisions. Alternatively, the decision maker might want to minimize the probability of a catastrophic loss or, otherwise, reduce the variance of his expected utility. These attributes could be reflected in a carefully defined nonlinear utility function or in penalties placed on the less attractive scenarios. In this discussion, we do not consider such specifications. Our consideration of linear models should be, however, sufficiently general to allow for further analysis in this area.

This chapter begins with a discussion of bounds on the expected value of perfect information. It then proceeds in Section 3 to examine uses of the deterministic optimization of different scenarios. In Section 4, the possibilities for combining these solutions and the inherent difficulties in the stochastic program are discussed. Sections 5 then presents examples of these problems.

The chapter concludes in Section 6 with a suggestion for a general strategy to be applied in linear optimization under uncertainty.

2. Bounds on the Expected Value of Perfect Information

We discussed above the expected value of perfect information (EVPI) and value of the stochastic solution (VSS) and showed examples when these quantities might be zero. When the conditions for deterministic optimality are not met, we would like to have simple bounds on the EVPI that may help us determine the worth of solving the stochastic program. If the EVPI and VSS are bounded within a tight range, it may be adequate to use a deterministic approach to the problem instead of following an expensive stochastic method.

The expected value, $\overline{z}$, of the objective function, $z(\xi)$, can be bounded because of its convexity. Madansky [42] and later Huang, Ziemba and Ben–Tal [35] examined this property using the theory of moment spaces to bound the expectation. Their work rests on the following result.

Lemma 1. *The objective function, $z(\xi)$, in (1.3) is a convex and continuous function of ξ, the right–hand side.*

Proof. See Madansky [42]. ∎

This result then allows the application of Jensen's inequality for convex functions. Directly from this, we have

$$\overline{z}_1 = \mathrm{E}[z(\xi)] \geq z(\ \mathrm{E}(\xi)) = z_4, \tag{1}$$

giving a lower bound on the perfect information solution, $\overline{z}_1$. (In this analysis, we use the definitions of z_1, z_2, z_3, and z_4 from Chapter I.)

An upper bound also can be found for $\overline{z}_1$. We present here only the one dimensional case. The multidimensional case can be found in Madansky [42]. The bound is known as the *Edmundson–Madansky inequality*, and it states that, for $\xi \in [a, b] \subseteq \Re^1$, $\mu = \mathrm{E}[\xi]$,

$$\overline{z}_1 = \mathrm{E}[z(\xi)] \leq \left(\frac{b-\mu}{b-a}\right) z(a) + \left(\frac{\mu - a}{b-a}\right) z(b), \tag{2}$$

where by definition here $\mu = \mathrm{E}[\xi]$. This is shown easily, since for any $t \in [a, b]$, $t = (a(b-t) + (t-a)b)/(b-a) = \lambda a + \overline{\lambda} b$, and, for $z(t)$ convex, $z(t) \leq \lambda z(a) + \overline{\lambda} z(b)$, which by integrating yields (2). We can now state (1) and (2) as

Theorem 1. *For $\xi \in [a, b] \subseteq \Re^1$, $\mu = E(\xi)$, and z as defined above,*

$$(b-\mu)/(b-a) z(a) + (\mu - a)/(b-a) z(b) \geq \overline{z}_1 \geq z(\mu).$$

Huang, Ziemba and Ben–Tal carry these principles further. They subdivide the interval $[a, b]$ and apply successively finer approximations, which they show approach the expectation. This method makes possible the refinement of the EVPI to whatever level is desired.

Another method for computing bounds on the EVPI was presented by Avriel and Williams [5]. They showed that

$$0 \leq EVPI \leq z_2 - z_4 \leq \overline{z}_4 - z_4, \tag{3}$$

where z_2 is the best stochastic solution and z_4 is the expected value deterministic solution, as in Chapter I. They also show that, for $z(\xi)$ differentiable, the bound using the expected value of ξ (as in z_4) is the tightest possible.

These approaches give us methods for estimates of the benefit we may gain from knowledge about the random variables. We concern ourselves here with the VSS, the value of solving a stochastic program over solving a deterministic one. The EVPI is used to find the value of additional information, but, in looking at the VSS, we assume that no more information is available. We ask: *what given our present state of knowledge, is the value of solving a large stochastic program?*

Similar results to those above can be found for VSS. We can bound VSS as in the following theorem.

Theorem 2. *The value of the stochastic solution (VSS $= \overline{z}_4 - \overline{z}_3$, as defined in Chapter I), satisfies the inequalities*

$$0 \leq VSS \leq \overline{z}_4 - z_4. \tag{4}$$

Proof. We showed VSS ≥ 0 in Chapter I. It suffices to show $z_4 \leq \overline{z}_3$. We had $\overline{z}_1 \leq \overline{z}_2 \leq \overline{z}_3$. Now, $z_4 = z(\,\mathrm{E}(\xi))$ and $\overline{z}_1 = \mathrm{E}(z(\xi))$, so, for $z(\xi)$ convex, by Jensen's inequality, $z_4 \leq \overline{z}_1$. The result follows. ∎

This bound can prove useful in estimating the benefit of the stochastic program, but, if it remains high, further analysis may be required. As a first step in solving the stochastic program, we may find other deterministic solutions corresponding to different scenarios or outcomes of the random variables. We describe this approach in the next section.

3. The Scenario Approach

In evaluating the perfect information or "wait–and–see" solution, $\overline{z}_1$, a solution to the program, $z_1(\xi)$, must be found for each possible outcome of the random vector, ξ. The scenario approach, also known as "modified wait–and–see" in Gunderson, Morris, and Thompson [31], involves solving several of these deterministic programs, evaluating the expected cost of using the strategy that is optimal for each scenario, and choosing the strategy that minimizes this expected cost. In many cases, one of the first period bases dominates the others, and the choice for a decision is clear. As Gunderson, et al, emphasize, this method can be quite responsive to management concerns and may prove very useful since it presents alternative possibilities and risks in a compact and easily understandable form.

In the scenario approach, we first choose a set of possible outcomes for ξ, which we call, $\tilde{\Xi}$, where

$$\tilde{\Xi} = \{\xi_1, \ldots, \xi_k\}. \tag{5}$$

For each $\xi^0 \in \tilde{\Xi}$, we find the optimal solution, x^*, for $z(\xi^0)$. Then, we compute

$$\varsigma(\xi^0) \equiv \mathrm{E}[\varsigma(x^*(\xi^0), \xi)] \tag{6}$$

for each ξ^0, where $\varsigma(x^*(\xi^0), \xi)$ is the resulting objective value from using x^* when ξ actually occurs.

Next, we find

$$\varsigma^* \equiv \min_{\xi^0 \in \tilde{\Xi}} \varsigma(\xi^0). \tag{7}$$

This value represents the least expected cost from using the solution of a deterministic program.

This implies that a deterministic problem other than the expected value problem may result in a better solution. This is because less penalty may be incurred by following a piecewise linear section of the objective function other than that that covers the expected value. (See Figure 1.)

We can also bound ς^* as in the following theorem.

Theorem 3. *For Ξ discrete, the best scenario solution , ς^*, satisfies*

$$\overline{z}_3 \leq \varsigma^* \leq \overline{z}_4. \tag{8}$$

Proof. We know

$$\varsigma^* = \min_{\xi^0 \in \Xi} \mathrm{E}[\varsigma(x(\xi^0)), \xi] \leq \mathrm{E}[\varsigma(x(\overline{\xi}), \xi)] = \overline{z}_4$$

and, for x^* optimal in ς^*,

$$\overline{z}_3 = \min_{\xi \in \Xi} \mathrm{E}[\varsigma(x(\xi)), \xi] \leq \mathrm{E}[\varsigma(x^*), \xi] = \varsigma^*.$$

Hence, the result.∎

The inequalities in (8) show that the scenario approach may be useful in finding a closer approximation to the solution of the stochastic program. ς^* may be especially valuable when $\varsigma^* - z_4$ is small, since it also bounds $\overline{z}_3 - z_4$ and may show that solving the stochastic program is unnecessary.

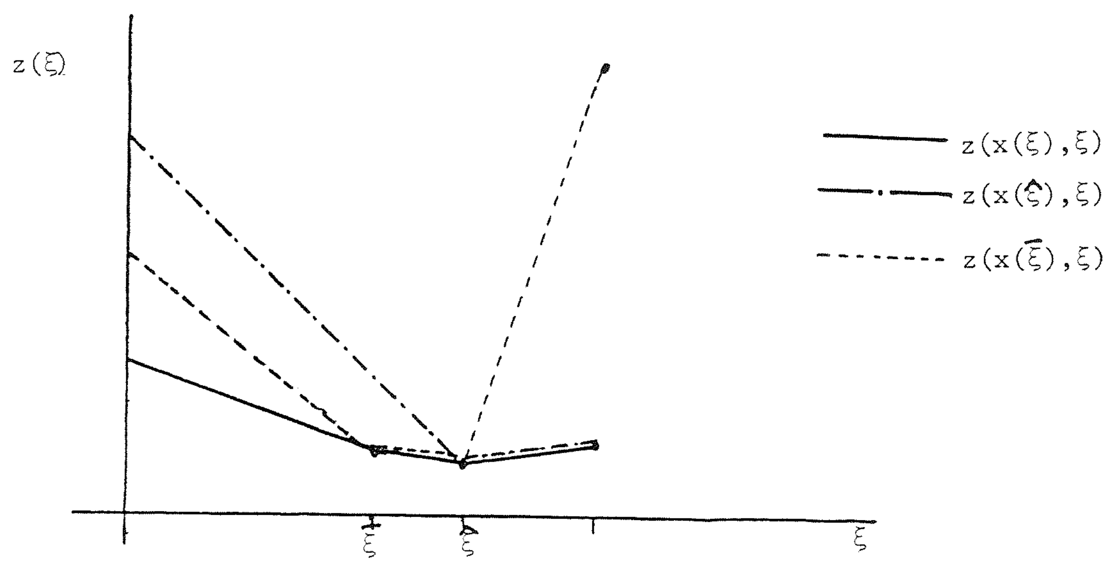

Figure 1. The expected value of $z(x(\overline{\xi}), \xi)$ is greater than that of $z(x(\hat{\xi}), \xi)$.

This outcome is most likely when the optimal decisions in the first period (those that are actually implemented) correspond to the same basis for most of the scenarios.

The scenario approach can also be used to find activities that may be basic in the optimal solution of the stochastic program. If the set of optimal basic activities remains fairly constant for the possible realizations of ξ, then this set of activities may be optimal in the stochastic program. We discuss this possibility in the next section.

4. Combining Scenarios

A natural approach to solving the stochastic program would be to use the optimal solutions of the different scenarios and to combine them in an appropriate manner. The proper combination may, however, be quite difficult to find and may lead to as much effort as solving the stochastic program directly. We present below the problems inherent in combining scenarios and some situations, in which, the optimal combination may be found directly.

In this analysis, we restrict ourselves to a two period case, for which there are only two scenarios considered. The difficulties involved in this example are typical of all stochastic programs, so we present this case because of its simplicity. The results may be easily generalized to more periods and scenarios.

We begin by defining two scenario problems as

$$\begin{aligned} \min z(\xi^1) &= c_1x_1 + c_2x_2 \\ \text{subject to} \quad A_1x_1 &\qquad\quad = b_1 \\ -B_1x_1 + A_2x_2 &= \xi_2^1 \\ x_1, x_2 &\geq 0, \end{aligned} \tag{S1}$$

and

$$
\begin{aligned}
\min\ z(\xi^2) &= c_1x_1 + c_2x_2 \\
\text{subject to}\quad & A_1x_1 \qquad\quad = b_1 \\
& -B_1x_1 + A_2x_2 = \xi_2^2 \\
& \qquad x_1, x_2 \ \geq 0.
\end{aligned}
\tag{$S2$}
$$

Solving S1 and S2 yields the optimal solutions, $x^{1,*}$ and $x^{2,*}$, and optimal dual prices $(\pi^{1,*};\sigma^{1,*})$ and $(\pi^{2,*};\sigma^{2,*})$, where $z^* = \pi^* \cdot b_1 + \sigma^* \cdot \xi$. We also define the following index sets:

$$\beta^1 \equiv \{j : A_1(*,j) \text{ is basic in S1 }\}$$

and

$$\beta^2 \equiv \{j : A_1(*,j) \text{ is basic in S2 }\}. \tag{9}$$

The complements of β^1 and β^2 are defined as $\overline{\beta}^1$ and $\overline{\beta}^2$, respectively.

Now, our actual goal is to solve the following two-period version of SDLP

:

$$
\begin{array}{llllll}
\min & c_1x_1 & +p^1c_2x_2^1 & +p^2c_2x_2^2 & & \\
\text{subject to} & A_1x_1 & & & = b_1 & \\
& -B_1x_1 & +A_2x_2^1 & & = \xi_2^1 & \qquad (SDLP2) \\
& -B_1x_1 & & +A_2x_2^2 & = \xi_2^2 & \\
& & & x_1, x_2^1, x_2^2 & \geq 0. &
\end{array}
$$

We would like to use the activities in β^1 and β^2 as the optimal basis of SDLP2. In other words, for $\beta = \{j : A_1(*,j) \text{ is basic in SDLP2 }\}$, we are looking for $\beta \subseteq \beta^1 \bigcup \beta^2$. Unfortunately, this is not always possible. The difficulty results from the properties of the basis in SDLP2. In order to maintain full rank of the basis in this program, the optimal basic activities from (S1) and (S2) cannot all be in the optimal basis of SDLP2, unless the square block situation we discussed in Chapter I holds.

In general, when we attempt to combine scenarios, we face a duality gap as the following lemma states.

Lemma 2. *For z^*, the optimal value of SDLP2,*

$$\begin{aligned} \underline{z} &= (p^1\sigma^{1,*} + p^2\sigma^{2,*})b + p^1\pi^{1,*}\xi_2^1 + p^2\pi^{2,*}\xi_2^2 \\ &\leq z^* \\ &\leq c_1(p^1x_1^{1,*} + p^2x_1^{2,*}) + pc_2\tilde{x}_2^{1,*} + p^2c_2\tilde{x}_2^{2,*} = \overline{z}, \end{aligned} \tag{10}$$

where $x_1^{1,*}$ and $x_1^{2,*}$ are the optimal primal solutions for S1 and S2, $\tilde{x}_2^{2,*}$ and $\tilde{x}_2^{1,*}$ are the values of the second period basic variables in S1 and S2 chosen to satisfy $A_2\tilde{x}_2^{1,*} = \xi_2^1 + B_1(p^1x_1^{1,*} + p^2x_1^{2,*})$ and $A_2\tilde{x}_2^{2,*} = \xi_2^2 + B_1(p^1x_2^{1,*} + p^2x_2^{2,*})$, [if $\tilde{x}_2^{1,*}$ or $\tilde{x}_2^{2,*} \leq 0$, set $\overline{z} = +\infty$], and $\pi^{1,*}$, $\pi^{2,*}$, $\sigma^{1,*}$, and $\sigma^{2,*}$ are the optimal dual values for S1 and S2. Furthermore, the duality gap, if $\tilde{x}_2^{1,*}$ and $\tilde{x}_2^{2,*}$ are feasible, is

$$\overline{z} - \underline{z} = (c_1^1 - \sigma^{2,*}A_1^1 + \pi^{2,*}B_1^1)p^2p^1x^{1,*} + (c_1^2 - \sigma^{1,*}A_1^2 + \pi^{1,*}B_1)p^1p^2x^{2,*}, \tag{11}$$

where $c_1^1 = \{c_j : j \in \beta^1 \bigcap \overline{\beta}^2\}$, $c_1^2 = \{c_j : j \in \overline{\beta}^1 \bigcap \beta_2\}$, $A_1^1 = \{a_j : a_j$ is a column of A_1 and $j \in \beta^1 \bigcap \overline{\beta}^2\}$, $A_2^1 = \{a_j : a_j$ is a column of A_1 and $j \in \overline{\beta}^1 \bigcap \beta^2\}$, and B_1^2 and B_1^1 are columns of B_1 corresponding to A_1^1 and A_1^2, respectively.

Proof. First, we show that $(p^1\sigma^{1,*} + p^2\sigma^{2,*}; p^1\pi^{1,*}; p^2\pi^{2,*})$ is a feasible solution to the dual of SDLP2. We have $\pi^{2,*}A_2 \leq c_2$ and $\pi^{1,*}A_2 \leq c_2$, since $\pi^{1,*}$ and $\pi^{2,*}$ solve S1 and S2. This also yields $-\pi^{1,*}B_1 + \sigma^{1,*}A_1 \leq c_1$ and $-\pi^{2,*}B_1 + \sigma^{2,*}A_1 \leq c_2$, so

$$-(p^1\pi^{1,*}+p^2\pi^{2,*})B_1+(p^1\sigma^{1,*}+p^2\sigma^{2,*})A_1 \le c_1, \tag{12}$$

since p^1, $p^2 \ge 0$ and $p^1+p^2=1$. Therefore, the solution is dual feasible, hence, by duality, we have $z^* \ge \underline{z}$.

Now, we consider $(p^1x^{1,*}+p^2x^{2,*};\tilde{x}_2^{1,*};\tilde{x}_2^{2,*})$ and observe that $A_1(p^1x^{1,*}+p^2x^{2,*})=b$, that, by definition, $\tilde{x}_2^{1,*}$ and $\tilde{x}_2^{2,*}$ satisfy the second set of inequalities in SDLP2, and, that, if $\tilde{x}_1^{2,*}$ and $\tilde{x}_2^{2,*}$ are not feasible, $\overline{z}=\infty$. Therefore, if $\tilde{x}_2^{1,*}\ge 0$ and $\tilde{x}_2^{2,*}\ge 0$, the solution is feasible and $z^*\le\overline{z}$.

Therefore, $\underline{z}\le z^*\le\overline{z}$. For the expression of the duality gap, we observe that

$$\begin{aligned}
\underline{z} \quad &= (p^1\sigma^{1,*}+p^2\sigma^{2,*})(p^1A_1x^{1,*}+p^2A_1x^{2,*})\\
&+p^1\pi^{1,*}(A_2\tilde{x}_2^{1,*}-p^1B_1x_1^{1,*}-p^2B_1x_1^{2,*})\\
&+p^2\pi^{2,*}(A_2\tilde{x}_2^{2,*}-p^1B_1x_1^{1,*}-p^2B_1x_1^{2,*})\\
\\
&=c_1^0(p^1x_1^{1,*}+p^2x_1^{2,*})+p^1c_1^1(p^1x_1^{1,*}+p^2x_1^{2,*})\\
&+p^2c_1^2(p^1x_1^{1,*}+p^2x_1^{2,*})+p^2\sigma^{2,*}(p^1A_1^1x_1^{1,*}+p^2A_1^1x_1^{2,*})\\
&+p^2\pi^{2,*}(-p^1B_1^1x_1^{1,*}-p^2B_1^1x_1^{2,*})+p^1\sigma^{1,*}(p^1A_1^2x_1^{1,*}+p^2A_1^2x_1^{2,*})\\
&+p^1\pi^{1,*}(-p^1B_1^2x_1^{1,*}-p^2B_1^2x_1^{2,*})+p^1c_2^1\tilde{x}_2^{1,*}\\
&+p^2c_2^2\tilde{x}_2^{2,*}
\end{aligned} \tag{13}$$

where $c_1^0=\{c_1(j):j\in\beta^1\bigcap\beta^2\}$. We then have

$$\begin{aligned}
\underline{z}=\overline{z}&-p^2c_1^1(p^1x_1^{1,*}+p^2x_1^{2,*})-p^1c_1^2(p^1x_1^{1,*}+p^2x_1^{2,*})\\
&+(p^2\sigma^{2,*}A_1^1-p^2\pi^{2,*}B_1^1+p^1\sigma^{1,*}A_1^2-p^1\pi^{1,*}B_1^2)(p^1x_1^{1,*}+p^2x_1^{2,*})
\end{aligned} \tag{14}$$

which, since $c_1^2\cdot x_1^{1,*}=0$ and $c_1^1\cdot x_1^{2,*}=0$, yields

$$\begin{aligned}
\underline{z}=\overline{z}&-(c_1^1-\sigma^{2,*}A_1^1+\pi^{2,*}B_1^1)(p^2p^1x_1^{1,*})\\
&-(c_1^2-\sigma^{1,*}A_1^2+\pi^{1,*}B_1^2)(p^1p^2x_1^{2,*})
\end{aligned} \tag{15}$$

Hence, the result in (11) follows. ∎

Measuring the duality gap in (11) is another way of finding the value of solving the stochastic problem directly. If the gap is small, then we can use the simple weighted average of the optimal scenario values without incurring a great penalty. A large gap signifies greater variance among the scenario solutions and should lead to a stochastic approach. A large difference can also result from infeasibility of the primal solutions since by definition in this case, $\bar{z} = \infty$. In some instances, the gap may be closed easily as we discuss below.

We first consider the case of $\beta = \beta^1 = \beta^2$ (same first period basic activities in S1 and S2), but where $|\beta| > m_1$, the rank of A_1. This situation, which we described in Chapter I, signifies that the same basic activities cannot exclusively be used to solve SDLP2. They will not span the row space of the matrix. If we have an inequality form, however, the weighted average of optimal first period values may be optimal in SDLP2.

We write the coefficient matrix of SDLP2 as

$$A = \begin{pmatrix} A_1^0 & A_1^1 & A_1^2 & A_1^N & & & & \\ -B_1^0 & -B_1^1 & -B_1^2 & -B_1^N & A_2^1 & A_2^{1,N} & & \\ -B_1^0 & -B_1^1 & -B_1^2 & -B_1^N & & & A_2^2 & A_2^{2,N} \end{pmatrix} \tag{16}$$

where $A_1^0 = (A_1(*,j) : j \in \beta^1 \bigcap \beta^2)$, A_1^1, A_1^2 are defined as in Lemma 2, $A_1^N = (A_1(*,j) : j \in \overline{\beta}^1 \bigcap \overline{\beta}^2)$, A_2^1, A_2^2 are the basic second period columns in S1 and S2, respectively, the columns of $A_2^{1,N}$ and $A_2^{2,N}$ are non–basic, and the B_j^i's are defined correspondingly.

Now, we assume that $\beta_1 = \beta_2$ and that we can replace the second equality in SDLP2 with an inequality. SDLP2 becomes

$$
\begin{array}{llllllllll}
\min & c_1^0 y_1^0 & +c_1^1 y_1^1 & +c_1^N y_1^N & +p^1 c_2^1 y_2^1 & +p^2 c_2^2 y_2^2 & +p^1 c_2^{1,N} y_2^{1,N} & +p^2 c_2^{2,N} y_2^{2,N} & \\
\text{s.t.} & A_1^0 y_1^0 & +A_1^1 y_1^1 & +A_1^N y_1^N & & & & & = b_1 \\
& -B_1^0 y_1^0 & -B_1^1 y_1^1 & -B_1^N y_1^N & +A_2^1 y_2^1 & & +A_2^{1,N} y_2^{1,N} & & \geq \xi_2^1 \\
& -B_1^0 y_1^0 & -B_1^1 y_1^1 & -B_1^N y_1^N & & +A_2^2 y_2^2 & +A_2^{2,N} y_2^{2,N} & & \geq \xi_2^2 \\
& & & & & & y_1, y_2 & & \geq 0,
\end{array}
$$

(SDLP2′)

where the variables (y_1, y_2) replace the variables (x_1, x_2) in SDLP, so that we can compare their values. First, we define a solution to SDLP2′ by the following. Let a solution $(y_1; y_1^1, y_1^2)$ be

$$
(y_1^{0,*}; y_1^{1,*}; y_2^{1,*}) = \begin{pmatrix} A_1^0 & A_1^1 & \\ -B_1^0 & -B_1^1 & A_2^1 \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ \xi_2^1 \end{pmatrix}. \tag{17}
$$

This solution is the same as $(x_1^{1,*}, x_2^{1,*})$. Next, define

$$
\nu = \{i : x_1^{1,*}(i) \text{is basic in row } j \text{ of S1 for } j > m_1\}, \tag{18}
$$

where m_1 is the number of rows in A_1. Now, partition A_2^1 as

$$
A_2^{1,\nu} = (A_2^1(j, *) : j \in \nu)
$$

and

$$
A_2^{1,\overline{\nu}} = (A_2^1(j, *) : j \not\in \nu). \tag{19}
$$

$A_2^{1,\overline{\nu}}$ is non–singular because the basis in (17) is non–singular.

We complete the definition of y^* by

$$
y_2^{2,\overline{\nu}^*} = (A_2^{1,\overline{\nu}})^{-1}(\xi_2^{2,\overline{\nu}} + B_1^{0,\overline{\nu}} y_1^{0,*} + B_1^{1,\overline{\nu}} y_1^{1,*})
$$

and

$$y_2^{2,\nu^*} = \xi_2^{2,\nu} + B_1^{0,\nu} y_1^{0,*} + B_1^{1,\nu} y_1^{1,*}. \tag{20}$$

Here, y_2^{2,ν^*} represents the slack variables. Now, we wish to show that y^* is an optimal solution to SDLP2$'$. We will have to restrict the solutions of S1 and S2 for this to be true. We first look at the dual of SDLP2$'$.

Let $\pi^{2,\nu} = (\pi_j^2 : j \in \nu) = 0$ and $\pi^{2,\overline{\nu}} = (\pi_j^2 : j \not\in \nu) = p^2 c_2^2 (A_2^{2,\overline{\nu}})^{-1}$, so we have $\pi^2 A_2^2 = (p^2 c_2^2; 0)$ where the objective row coefficients of the slacks y_2^{2,ν^*} are 0.

Next, let

$$\pi^{1,\overline{\nu}} = p^1 c_2^1 (A_2^{1,\overline{\nu}})^{-1} - \pi^{1,\nu}(A_2^{1,\nu})(A_2^{1,\overline{\nu}})^{-1}, \tag{21}$$

and

$$\pi^{1,\nu} = (c_1 - \sigma A_1^{\nu} - B_1^{\overline{\nu}} p^1 c_2^1 (A_2^{1,\overline{\nu}})^{-1})(B_1^{\overline{\nu}}(A_2^{1,\overline{\nu}})^{-1} A_2^{1,\nu} - B_1^{\nu})^{-1},$$

where $B_1^{\nu} = (B^1(*,i) : i$ is basic for row j such that $j \in \nu)$ and $B_1^{\overline{\nu}}$, A_1^{ν} are defined accordingly. Now, let $\hat{B} = (B_1^{1,\overline{\nu}}(A_2^{1,\overline{\nu}})^{-1} A_2^{1,\nu} - B_1^{\nu})^{-1}$ and compute σ as

$$\begin{aligned} \sigma = [c_1^{\overline{\nu}} + (B_1^{\nu} - B_1^{\overline{\nu}}(A_2^{1,\overline{\nu}})^{-1} A_2^{1,\nu})(\hat{B})^{-1}(c_1^{\nu} - B_1^{\nu}(A_2^{1,\overline{\nu}})^{-1} c_2^1) \\ + B_1^{\overline{\nu}}(A_2^{1,\overline{\nu}})^{-1} c_2^1](A_2^{1,\overline{\nu}} + (B_1^{\nu} - B_1^{\overline{\nu}}(A_2^{1,\overline{\nu}})^{-1} A_2^{1,\nu})(\hat{B})^{-1} A_1^{\nu})^{-1}. \end{aligned} \tag{22}$$

We substitute for $\pi^{1,\nu}$ in (21) and find that, π_{OLD} and σ_{OLD}, the optimal dual prices in S1 (or S2, since the bases and objective functions are the same), satisfy

(i) $\pi^{2,\nu} = 0$,

(ii) $\pi^{2,\overline{\nu}} = p^2(\pi^{\overline{\nu}}_{OLD} + \pi^{\nu}_{OLD} A_2^{2,\nu}(A_2^{2,\overline{\nu}})^{-1})$,

(iii) $\pi^{\nu} = \pi^{\nu}_{OLD}$,

(iv) $\pi^{1,\overline{\nu}} = p^1(\pi^{\overline{\nu}}_{OLD} - (p^2/p^1)\pi^{\nu}_{OLD} A_2^{2,\nu}(A_2^{2,\overline{\nu}})^{-1}$,

and (v) $\sigma = \sigma_{OLD}$.

Hence, we have

$$-\pi^{2,\overline{\nu}}B_1 - \pi^{2,\nu}B_1 - \pi^{1,\overline{\nu}}B_1 - \pi^{1,\nu}B_1 + \sigma A_1 \\ -\pi^{\overline{\nu}}_{OLD}B_1 - \pi^{\nu}_{OLD}B_1 + \sigma_{OLD}A_1 \leq c_1, \tag{23}$$

$$-\pi^{\overline{\nu}}_{OLD}B_1^{\overline{\nu}} - \pi^{\nu}_{OLD}B_1^{\overline{\nu}} + \sigma_{OLD}A_1^{\overline{\nu}} = c_1^{\overline{\nu}}, \tag{24}$$

and

$$-\pi^{\overline{\nu}}_{OLD}B_1^{\nu} - \pi^{\nu}_{OLD}B_1^{\nu} + \sigma_{OLD}A_1^{\nu} = c_1^{\nu}. \tag{25}$$

We also observe that

$$\begin{aligned}\pi^{1,\overline{\nu}}A_2^{1,\overline{\nu}} + \pi^{1,\nu}A_2^{1,\nu} &= p^1(\pi^{\overline{\nu}}_{OLD}A_2^{1,\overline{\nu}} + \pi^{\nu}_{OLD}A_2^{1,\nu}) \\ &\quad - p^2\pi^{\nu}_{OLD}A_2^{1,\nu} + p^2\pi^{\nu}_{OLD}A_2^{1,\nu} \\ &= p^1 c_2^1,\end{aligned} \tag{26}$$

and that, similarly,

$$\pi^{2,\overline{\nu}}A_2^{2,\overline{\nu}} + \pi^{2,\nu}A_2^{2,\nu} = p^2 c_2^2. \tag{27}$$

We need only that

$$\pi^{1,\overline{\nu}}A_2^{N,\overline{\nu}} + \pi^{1,\nu}A_2^{N,\nu} \leq p^1 c_2^{1,N} \tag{28}$$

and

$$\pi^{2,\overline{\nu}} A_2^{N,\overline{\nu}} + \pi^{2,\nu} A_2^{N,\nu} \leq p^2 c_2^{2,N} \tag{29}$$

in order for feasible y^* to be optimal, since we have already shown complementarity in (24), (25), (26), and (27). (23) also shows dual feasibility for the first period constraints.

To ensure that (28) and (29) hold, we need to have the following equality satisfied:

$$\pi_{OLD}^{\nu} A_2^{2,\nu} (A_2^{2,\overline{\nu}})^{-1} A_2^{N,\overline{\nu}} = \pi_{OLD}^{\nu} A_2^{N,\nu}. \tag{30}$$

By the definition of π^1 and π^2, (30) implies (28), since

$$\begin{aligned}\pi^{1,\nu} A_2^{N,\overline{\nu}} + \pi^{1,\nu} A_2^{N,\nu} &= p^1(\pi_{OLD}^{\overline{\nu}} A_2^{N,\overline{\nu}} + \pi_{OLD}^{\nu} A_2^{N,\nu}) \\ &\quad + p^2(\pi_{OLD}^{\nu} A_2^{N,\nu} - \pi_{OLD}^{\nu} A_2^{2,\nu} (A_2^{2,\overline{\nu}})^{-1} A_2^{N,\overline{\nu}}) \\ &\leq p^1 c_1^N,\end{aligned}$$

and (29) holds by

$$\begin{aligned}\pi^{2,\overline{\nu}} A_2^{N,\overline{\nu}} + \pi^{2,\nu} A_2^{N,\nu} &= p^2(\pi_{OLD}^{\overline{\nu}} A_2^{N,\overline{\nu}} + (\pi_{OLD}^{\nu} A_2^{2,\nu} (A_2^{2,\overline{\nu}})^{-1} A_2^{N,\overline{\nu}})) \\ &= p^2(\pi_{OLD}^{\overline{\nu}} A_2^{N,\overline{\nu}} + \pi_{OLD}^{\nu} A_2^{N,\nu}) \\ &\leq p^2 c_2^N.\end{aligned}$$

Hence, we have shown the following theorem.

Theorem 4. *For y^* as defined in (17) and (20), if y^* is a feasible solution to SDLP2′ and the optimal dual prices in S1 and S2 are such that (30) is satisfied, y^* is an optimal solution to SDLP2′.*

The restriction of the prices in (30) guarantees optimality, but (28) and (29) may be true even if (30) does not hold. In examining a problem of this type, if (30) fails, one may want to carry out the additional computations in (28) and (29) before solving the stochastic problem.

The difficulty in finding optimality conditions for combining scenarios for even a simple problem such as SDLP2$'$, shows the importance of the stochastic solution. The conditions in (30) can be generalized to allow for penalties in satisfying inequalities in the second period, but the results are more restrictive and direct computation of the dual feasibility conditions becomes more efficient than checking additional inequalities. Our development leads to the following method for combining scenarios.

(A.) Combine the first period scenario solutions, $x_1^1, \ldots, x_1^k$, by a simple weighted average, $y_1^* = \sum_{i=1}^k p^i x_1^i$.

(B.) Follow the branch of worst cases (what we shall call the *catastrophe* branch), that is, the set $\{\xi_t^*\}$ where $\xi_t^* = (\xi_t^*(1), \xi_t^*(2), \ldots, \xi_t^*(m_t))$ such that $\xi_t^*(i) = \sup_j \xi_t^j(i)$ for $i = 1, 2, \ldots, m_t$. Using these right–hand side values, we have that, if

$$-B_{t-1}y_{t-1}^* + A_t y_t^* \geq \xi_t^*,$$

then

$$-B_{t-1}y_{t-1}^* + A_t y_t^* \geq \xi_t^j$$

for all j. This procedure guarantees primal feasibility for y_t^*. We use the non–singular blocks from this scenario to determine new values based on y_1^*.

(C.) Use similar blocks for the other branches and maintain primal feasibility (by, perhaps, paying penalties).

(D.) Compute the dual prices and check for closure of the duality gap.

This approach would be most successful when the different scenarios have nearly identical bases. In these cases, the noncomplementarity would exist in only a few terms. The difficulty of using these optimal linear program solutions, however, is that they all correspond to extreme point solutions and may include very different sets of basic activities. This property could make their combination in a stochastic program most difficult. In the next section, we present small examples of this occurrence.

5. Examples

The extreme point properties of the basis in a linear program are crucial in understanding stochastic program solutions. Critical values of the parameters limit the use of different scenarios. In some cases, the implementation of any deterministic solution may lead to heavy penalties relative to the solution of the stochastic program. One example of this occurrence is the following linear program :

$$
\begin{array}{llllllll}
\min z = & x_1 & +4x_2 & + \mathrm{E}_\xi[\min y_1 & +10y_2 & \mid x_1 \text{ and } x_2] & & \\
\text{subject to} & x_1 & +x_2 & & & & & = 1 \\
 & -x_1 & +2x_2 & +y_1 & +y_2 & & & = \xi \\
 & & & & & & 0 \le y_1 & \le 2, \\
 & & & & & & x_1, x_2, y_2 & \ge 0, \\
 & & & & & & & \xi \text{ is Uniform } [0,4].
\end{array} \tag{EX1}
$$

We solve EX1 for $\overline{\xi} = (1,3)$ and find the expected value of using the optimal decisions for these scenarios as

$$\overline{z}_{\overline{\xi}} = \mathrm{E}[z_{\overline{\xi}}(\xi)], \tag{31}$$

where

$$z_{\overline{\xi}}(\xi) = c_1 x(\overline{\xi}) + [\ \min\ _y\ c_2 y | x(\overline{\xi}), \xi]. \tag{32}$$

We define

$$\overline{z}_I \equiv \overline{z}(1) \tag{33}$$

and

$$\overline{z}_{II} \equiv \overline{z}(3). \tag{34}$$

We want to consider also the perfect information solution

$$\overline{z}_p \equiv \ \mathrm{E}_{\xi}[\ \min\ c_1 \cdot x + c_2 \cdot y | \xi], \tag{35}$$

and the stochastic solution, $\overline{z}_s$, where we allow $\xi = 1$ or 3 with equal probability.

The function $z_{\overline{\xi}}(\xi)$ for z_I, z_{II}, z_s, and z_p appears in Figure 2. We observe that the optimal basis changes as ξ ranges over [0,4].

For $\xi \leq 1$, x_1 only is in the optimal basic set of variables, for $1 \leq \xi \leq 3$, $\{x_1, x_2\}$ is optimal, for $2 \leq \xi \leq 3$, an alternative optimal set includes x_2 alone, and, for $3 \leq \xi \leq 4$, the only optimal first period activity is x_2. In the stochastic solution, x_1 and x_2 must be in the optimal basis, reducing the expected loss relative to the perfect information solution. We find this from the figure as

$$\overline{z}_I - \overline{z}_p = 10.25, \tag{36}$$

$$\overline{z}_{II} - \overline{z}_p = 5.50, \text{ and} \tag{37}$$

$$\overline{z}_s - \overline{z}_p = 3.75. \tag{38}$$

The losses in (36), (37), and (38) associated with this problem demonstrate the usefulness of the stochastic solution. By assuming any deterministic value for the right–hand side, a large loss may result. The simple stochastic formulation with two possible ξ values, however, reduces the risk of this situation and enables us to approach the perfect information solution. The stochastic solution, therefore, lends resilience to the result. It provides a rationale for hedging strategies.

To demonstrate further the sensitivity of models to uncertainty, we return to the exhaustible resource model of Chapter I, (I.29). We stated that this model had an optimal deterministic solution, but, by adding an uncertain return from investment in exploration, we again arrive at a situation, in which, every deterministic solution will be associated with losses relative to

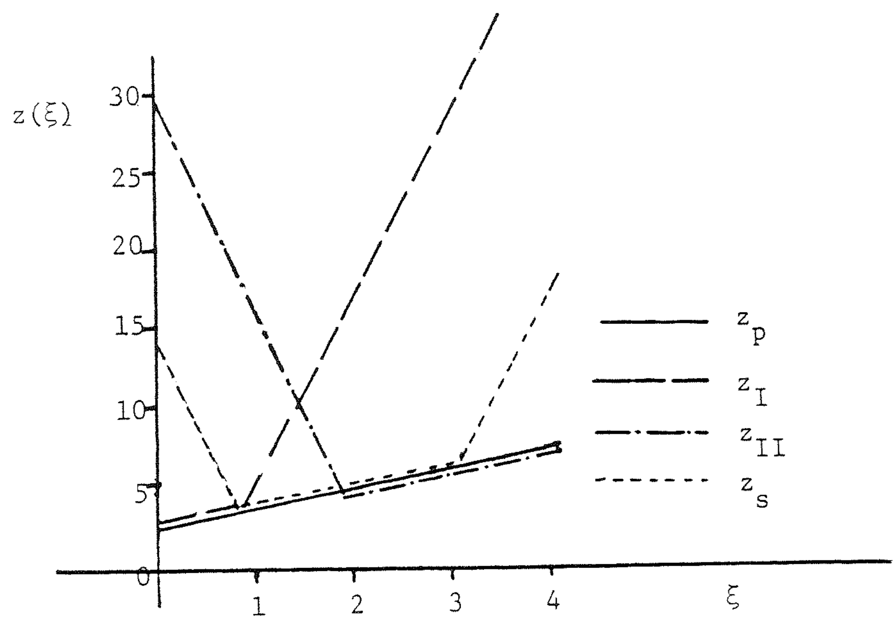

Figure 2. Example costs.

the stochastic solution.

The modified problem is

$$\min \sum_{i=1}^{m} c_i u_{i,0} + \sum_{i=1}^{n} k_i y_{i,0} + \sum_{t=1}^{T} \beta^t \sum_{i=1}^{m} \sum_{j=1}^{k} p^j [c_i u^j_{i,t} + k_i y^j_{i,t}]$$

$$\begin{aligned}
\text{subject to} \qquad \xi^0 + u_{i,0} &\le R_{i,0}; \text{ for } i = 1, \ldots, m, \\
\sum_{i=1}^{m} u_{i,0} &\ge D_0; \\
x^j_{i,t} + u^j_{i,t} &\le R_{i,j}; \text{ for all } i, j, \text{ and } t, \\
\sum_{i=1}^{m} u^j_{i,t} &\ge D_t; \text{ for all } j \text{ and } t, \\
x^j_{i,t-1} + u^j_{i,t-1} \qquad -\alpha^j_{i,t} y^j_{i,t-1} &= x^j_{i,t}; \text{ for all } i, j, \text{ and t}, \\
x^j_{i,t} \ge 0, u^j_{i,t} &\ge 0, \text{ for all } i, j, \text{ and } t,
\end{aligned} \tag{ERM}$$

where $R_{i,j}$ D_t, $x^j_{i,t}$, and $u^j_{i,t}$ are as defined in (1.29). $y^j_{i,t}$ represents the amount invested at time t under scenario j, and $\alpha^j_{i,t}$ is the return for investment.

ERM includes deterministic investment, but we can formulate an associated stochastic model by allowing $\alpha^j_{i,t}$ to take on several values, $\alpha^{j,1}_{i,t}, \ldots, \alpha^{i,l}_{i,t}$, and restricting investment in $y^j_{i,t}$ to be only one of the $\alpha^{j,q}_{i,t}$ for each scenario. By doing this, we maintain the random variable in the right–hand side as a constraint on $y^j_{i,t}$ for the different scenarios.

By solving ERM for fixed values, the investment decision may quickly swing from one resource to another as a different extreme point in the linear program becomes optimal. The stochastic solution has more basic activities, allowing for hedging against the different possible environments that one may face. To show this property, we consider a two period case of ERM with three alternative resources. We will call them oil, solar power, and some other high cost backstop technology. We use the following inputs in Table 1.

Resources

	Cost(c_i)	Availability(R_i)
Oil	5	25
Solar	10	10
Backstop	16	∞

Investment

	Cost (k_i)	Return(α_i)
Oil 1	1	1
Oil 2	1	0.1
Solar	1	1

Demand

Period 1	15
Period 2	25

Probabilities

Scenario 1	0.5
Scenario 2	0.5

Discount Rate

$\beta = .80$

Table 1. Two period model inputs.

The only uncertainty in this model is on the return for oil exploration. Investment in solar power can be interpreted as the relatively certain amount of capacity increase from investment. Each unit invested in oil in this model, however, results in either a full unit increase or a tenth of a unit increase in oil availability according to α. These two scenarios are assumed to occur with equal probability.

The deterministic models for "bad luck" ($\alpha = .1$), "good luck"($\alpha = 1$), and "myopic"(solving only given the first period availabilities and demands) were solved and compared to the stochastic program solution, in which both values of α were considered simultaneously. The results were :

Model	Expected Cost
Myopic	275
Good Luck	239
Bad Luck	245
Stochastic	231

Here, the stochastic solution represents a savings over the deterministic models because its solution involved investment in both oil and solar technologies, while the deterministic scenarios allowed for investment in only one. The relative savings would also increase with a higher cost backstop. It is interesting to note also that, with the addition of investment in exploration, the myopic solution is now far from optimal.

These examples have shown that models can be very sensitive to future uncertainties. The exhaustible resource model demonstrates the possibility of sharp changes in decision-making from near-term uncertainties. We, therefore, want to examine models with a strategy that considers their sensitivity to uncertain parameters. We discuss a general method for dealing with these problems in the following section.

6. A General Strategy

Problems modeled as dynamic linear programs often involve many uncertain assumptions about the parameters involved. In this case, a solution to the stochastic program is desired, but it may be quite costly. We have presented methods for checking whether deterministic solutions may be used.

The strategy resulting from our development in Chapters 1 and 2 follows.

I. Solve the expected value problem and check for basis feasibility and a possible optimal deterministic solution as in Chapter I.

II. If I fails, use the properties presented in 2.3 to bound the VSS and EVPI.

III. If the VSS bound is large, solve different deterministic scenarios and look for a dominant basis. Then, attempt to close the duality gap by combining scenarios.

IV. If the gap persists and if no single basis is indicated, proceed to solve SDLP.

This procedure outlines how we would evaluate the worth of solving successively more complicated problems. As our examples have shown, we may find that the stochastic program is worth solving, and that its solution may result in a substantial savings. We might attempt to solve such problems directly by brute–force methods. Alternatively, the next three chapters present methods that take advantage of the structure of SDLP to reduce the size of these possibly very large linear programs.