AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

DYNAMIC PROGRAMMING MODELS WITH RISK ORIENTED CRITERION FUNCTIONS

by

Oscar R. Burt

University of California, Davis

Paper presented at the 1988 meeting of Southern Regional Project S-180 in Savannah, Georgia, March 21-23.

DYNAMIC PROGRAMING MODELS WITH RISK ORIENTED CRITERION FUNCTIONS

*By Oscar R. Burt, Professor, Department of Agricultural Economics, University of California-Davis

Dynamic programming (DP) is traditionally thought of being associated with an expected value objective funciton, although Bellman frequently illustrated the method with terminal control objective functions in his early research (Bellman, 1961). It was Ronald Howard's seminal book in 1960 which popularized the expected value criterion with periodic rewards in DP Markov processes. It appears that the most fruitful way to bring risk into the usual DP algorithm is by introducing a somewhat artifical state variable which compounds returns to th end of the planning horizon. The other alternative is to use the linear programming (LP) formulation of the Markov chain DP model (Hadley, p. 471). Both of these methods are explored below for various criteria which incorporate risk in an optimization setting. We begin with a brief summary of the Markov chain DP model.

The decision process can be in a finite number of states at the beginning of a series of discrete time stages, and the dicision agent controls the conditional probability of going from the ith to the jth state by choosing from a finite set of alternatives, k = 1, 2, ...K(i). Let pijk denote this probability under decision k. In general, an immediate return is associated with the transition from state i to state j under decision k, and this return is a random variable which may have a component of variation associated with the transition from state i to j under decision k, as well as independent variation with i, j, and k all conditionally given. This total variation is particularly important under risk considerations, although it is moot with an expected value criterion. The random variable return is denoted r_{ik} with the subscripts indicating that the given state i and decision k are parameters in the probability distribution. The conditional expectation of r_{ik} is denoted q_{ik}.

A policy for this multistage decision process is comprised of a decision rule which specifies a choice of k for each state i, denoted d(i). With a particular decision rule, the transition probabilities at each state are determined by $p_{ijd(i)}$ which in the discussion below are abbreviated p_{ij}. If there are M states, this Markovian process can be summarized by an MXM matrix P with (i,j) element equal to p_{ij}.

Define the M-component row vector $\pi(0)$ as the a priori probability distribution for the initial state of the process at time t = 0. Then it follows that the probability vector for t = 1 is $\pi(1) = \pi(0)P$, and inductively $\pi(t) = \pi(0)P^t$. In an ergodic Markov process, where every state is accessible from any given state after a finite number of transitions,

$$\lim_{t\to\infty} P^t = A,$$

where all rows of A are identical. It follows from this result that there exists a limiting probability vector π such that

$$\pi = \pi P, \tag{1}$$

and π is the limit of $\pi(t)$ as $t\to\infty$ as well as being the identical rows of A.

An obvious criterion for the decision process is to choose the decision rule d(i) such that long-run expected returns are a maximum. Each decision rule implies a vector of probabilities π given by (1), and an expected immediate return vector $q = (q_1\ q_2 \ldots q_M)$, so we could search over all decision rules d(i) and choose the maximum of $\pi_1 q_1 + \pi_2 q_2 + \ldots + \pi_M q_M$.

Although economists will note that the criterion should be one of maximum expected present value, it cannot be directly applied with the incorporation of risk without assuming additivity of expected utilities over the planning horizon. DP has long been recognized as the efficient algorithm for solving this problem using an expected value criterion (Bellman, Howard), with or without discounting.

The DP recursion formula in our notation is

(2) $$v_n(i) = \underset{k}{\text{Max}}\left[q_{ik} + \beta \sum_{j=1}^{M} p_{ijk} v_{n-1}(j)\right],$$

where $0 < \beta < 1$ is the discount factor and $v_n(i)$ is the present value of expected returns over an n-stage planning horizon under an optimal policy. An expected utility function (Von Neuman and Morgenstern) can be applied to the conditional variation in r_{ik} and thus replace q_{ik} with this conditional expected utility, but the variation in returns over states through time will be ignored and the tacit assumption is that the periodic conditional utilities are additive. Clearly this poses a serious problem.

On the other hand maximizing $\pi' q$, which is equivalent to maximizing the long-run expected gain per period, is readily adapted to introducing risk into the criterion. All the variation in returns can be incorporated by applying a utility function to r_{ik}. Denote the utility function $U(\cdot)$, and $U_{ik} = E(U(r_{ik})|i,k)$ denotes the conditional expected utility of random returns r_{ik}, given i and k fixed. For a given decision rule d(i) let the conditional expected utility be $\bar{U}_i$, then unconditional periodic expected utility in the long-run is $\pi_1\bar{U}_1 + \pi_2\bar{U}_2 + \ldots + \pi_M\bar{U}_M$. Fortunately an efficient computational algorithm is available for this criterion if the number of states M is relatively small.

LINEAR PROGRAMMING OF MARKOV CHAINS

The following LP formulation was first published by Manne in 1960 for an inventory problem which was for the nondiscounting case. Discounting formulations soon followed (d'Epenoux), but they do not seem to be any more adaptable to risk criteria than the classic DP recursion.

For any given policy, the steady-state probabilities must satisfy (1) and $\pi_1 + \pi_2 + \ldots + \pi_M = 1$, $\pi_i \geq 0$, i = 1, 2,...M. For clarification, we write out (1) in detail:

(1)' $$\pi_j - \sum_{i=1}^{M} \pi_i p_{ij} = 0, \quad j = 1, 2, \ldots M.$$

For the optimal policy, $\pi q'$ must be a maximum. This provides the basis for the LP formulation which introduces the set of activities $\{\pi_{ik}\}$, k = 1, 2,...K(i), where at most one π_{ik} can be positive over k, for given i. But also the value of k where $\pi_{ik} > 0$ must be associated with the optimal policy. Therefore, we can write the expected value criterion LP model as (see Hadley, p. 471)

(3) $$\max_{\{\pi_{ik}\}} \left[\sum_{i=1}^{M} \sum_{k=1}^{K(i)} \pi_{ik} q_{ik} \right], \text{ subject to}$$

(3-a) $$\sum_{k=1}^{K(j)} \pi_{jk} - \sum_{i=1}^{M} \sum_{k=1}^{K(i)} \pi_{ik} p_{ijk} = 0, \; j = 1, 2, \ldots M,$$

(3-b) $$\sum_{i=1}^{M} \sum_{k=1}^{K(i)} \pi_{ik} = 1, \; \pi_{ik} \geq 0, \text{ all } i,k.$$

The expected utility criterion can be applied in (3) by replacing q_{ik} in the objective function with U_{ik} as defined earlier. The summations over k in (3-a), and the fact that one and only one π_{ik} can be positive for given i, impose the constraints in (1)' on all basic feasible solutions associated with the LP problem.

The mean/variance criterion can be introduced into the objective function of (3) to give a quadratic programming model which would be feasible for only small problems. Note that the number of activities in (3) is K(1) + K(2) + ... K(M), or KM if K(i) = K, i = 1, 2,...M. The number of constraints is M + 1. Applications can easily have over 1,000 states which is a large LP problem by any standard. The dimension gets out of reason very quickly with a quadratic objective function or some other nonlinear form of risk such as semivariance.

But Fishburn's "risk associated with below-target returns" is easily applied in this expected utility framework by using the results of his Theorem 2 (Fishburn, p. 120). His α-t model defines risk by the two-parameter function

$$F_{\alpha}(t) = \int_{-\infty}^{t} (t-y)^{\alpha} dF(y) \qquad (4)$$

where $F(\cdot)$ is the cumulative distribution function for a portfolio's return. Using Fishburn's notation, let F and G be two distribution functions while $\mu(F)$ and $\mu(G)$ denote the respective means of F and G; $\rho(F)$ and $\rho(G)$ denote the risk measure given by (4). Then by the α-t risk criterion, F Dominates G if and only if $\mu(F) \geq \mu(G)$ and $\rho(F) \leq \rho(G)$ with at least one strict inequality.

Fishburn proves that when the α-t risk criterion is congruent with the expected utility model, the utility function can be written in the form

$$\begin{aligned} U(y) &= y && \text{for } y \geq t \\ U(y) &= y - \lambda(t-y)^{\alpha} && \text{for } y \leq t. \end{aligned} \qquad (5)$$

Of course, the problems of choosing a specific utility function from this three-parameter family to use in research applications in agricultural economics are not trivial.

It is demonstrated below that the classic DP model can be modified to a terminal control process with a risk criterion applied to total compounded returns at the end of a finite planning horizon. Although this reformulation requires another state variable for Markov processes with periodic rewards, it appears to be more manageable computationally than the LP approach.

A TERMINAL CONTROL FORMULATION

Although stochastic firm growth models are logically formulated with an objective function equal to net wealth at the end of a finite planning horizon (Larson, Stauber, and Burt; Schnitkey), many other applications focus on a small segment of the firm such that discounted value of returns is the most common criterion. These problems are typically modeled with an expected present value criterion, and there is a stochastic stream of returns over the planning horizon. The task now is to convert these problems to a criterion of maximum compounded returns at the end of a finite horizon, which gives the same decision rule as discounting under an expected value criterion, but risk models can be applied to the terminal value problem, e.g., expected utility of total compounded returns.

Expected Value Criterion

The ideas are best illustrated with an example using one state variable where we can use continuous variables and functional relationships. We denote the decision and state variables by u and x, respectively, and immediate returns are $R(u,x,\epsilon)$, where ϵ is a random variable. The state variable x obeys the stochastic difference equation

(6) $$x_{n-1} = x_{nt} + g(u_n, x_n, \delta),$$

where δ is a random variable and n is the number of periods remaining in the planning horizon. The expected present value criterion DP equation for this problem is

$$f_n(x) = \underset{u}{\text{Max}}\ E\left[R(u,x,\epsilon) + \beta f_{n-1}(x+g(u,x,\delta))\right] \tag{7}$$

where $f_n(x)$ is the expected discounted value of net returns over an n-stage planning horizon when following an optimal policy and the initial value of the state variable is x.

To convert this to a terminal control process, we introduce a second state variable which systematically compounds returns to the terminal period, n = 0 in this notation. The stochastic difference equation for this state variable y is

$$y_{n-1} = (1+\rho)y_n + R(u_n,x_n,\epsilon_n), \tag{8}$$

where y_n is defined as the amount of money on account at the beginning of period n and which earns interest at the rate ρ. Not compounding $R(\cdot)$ in (8) implies that returns are received at the beginning of the stage, and that the process is terminated at the beginning of the last period of the planning horizon. These assumptions are the same as implied by the maximum expected present value model in (7). We note that $\beta = 1/(1+\rho)$ in this comparison.

The DP equation for this model is

$$f_n(x,y) = \underset{u}{\text{Max}}\ E\{f_{n-1}(x+g(u,x,\epsilon),(1+\rho)y + R(u,x,\delta))\}, \tag{9}$$

and the initial condition is $f_0(x,y) = y$, unless x has some salvage value to be added to y. The iterative solution of (9) starting at n=1 will make $f_n(x,y)$ the expected compounded value of returns over the n-period planning horizon. It should be clear that the interest rate ρ could be a random variable with known distribution and the problem would not be much more complex.

<u>Maximum Expected Utility</u>

A utility function to account for risk preferences can be imposed on this formulation with ease. This is accomplished by defining $f_0(x,y) = U(y)$; then solution of (9) for n=T yields the optimal decision rule to maximize expected utility of wealth at the end of the T-stage planning horizon.

Numerical solution of this problem is achieved by using discrete valued approximations for the variables u,x, and y, as well as discrete distributions for ϵ and δ. Thus, the problem is reduced to the discrete Markov process discussed above. In that notation the state i designates a pair of levels for x and y within the discrete approximation (actually a rectangular area in the (x,y) plane). Again, Fishburn's α-t model is conveniently used to obtain a three-parameter utility function which is congruent with a large family of targeted risk criteria.

<u>Safety-First Criteria</u>

A two-parameter safety first criterion which can be applied to (9) is to maximize the mean of terminal wealth subject to a constraint that the probability of wealth falling below y^* is less than a specified probability denoted γ (Telser). First, we will examine Roy's safety-first criterion of simply minimizing the probability of falling below the specific level y^* (Roy). This can be done by using the model in (9) after it has been approximated as a finite Markov chain. Since the terminal state collapses the two state variables x and y into y only by setting the terminal condition $f_0(x,y) = y$, we do the same thing in the discrete case. The transition matrices in stage 1 are defined such that the first subscript on p_{ijk} goes from 1 to M, but the second subscript (associated with a state experienced in stage 0) goes from 1 to m, where m is the number of discrete values on y. Technically, all of the p_{ijk} associated with a given value of y are added over all possible values of x on the subscript j. Without loss of generality, we can take the definition of subscripts for $j = 1, 2, \ldots m$ such that y is nondecreasing in the counting integers j. Then for the states defined in stage 0, $i < i^*$ implies $y < y^*$.

The terminal control Markov chain DP equation is

$$(10) \qquad v_n(i) = \underset{k}{\text{Max}}\left\{\sum_{j=1}^{M} p_{ijk} v_{n-1}(j)\right\}$$

Suppose we were to assign $v_0(i) = 0$ for $i = 1, 2, \ldots i^*-1$ and $v_0(i) = 1$ for $i \geq i^*$. Then

$$v_1(i) = \underset{k}{\text{Max}}\left[\sum_{j=1}^{M} p_{ijk} v_0(j)\right] = \underset{k}{\text{Max}}\left[\sum_{j=i^*}^{M} p_{ijk}\right],$$

which is maximization of the probability that $i \geq i^*$ ($y \geq y^*$) in the terminal stage. Clearly $v_1(i)$ is the maximum probability of reaching a state where $y \geq y^*$ in the last stage when starting from state i at the beginning of the first stage. If w_i is the probability of going to state i in the first stage, then it follows from elementary probability theory that the probability of $y \geq y^*$ in a two-stage process is $w_1v_1(1) + \ldots + w_Mv_1(M)$.

But we want to maximize the above probability for a given state i at the beginning of a two-stage process. Therefore, we use the recursion formula (10) again,

$$v_2(i) = \underset{k}{\text{Max}}\left[\sum_{j=1}^{M} p_{ijk}v_1(j)\right], \quad i = 1, 2, \ldots M,$$

which makes $v_2(i)$ the maximum probability of $y \geq y^*$ in the terminal stage starting from the beginning of a two-stage process. Inductively, we see that our choice of $v_0(i)$ makes $v_n(i)$ in (10) the maximum probability of $i \geq i^*$, or $y \geq y^*$, and the safety first criterion of minimizing the probability that $y < y^*$ can be calculated efficiently with DP.

The less conservative safety first criterion of maximizing mean returns subject to a constraint that actual returns are at least y^* with probability $(1-\gamma)$ is now presented. Under this criterion we calculate (10) as an expected value criterion by setting $v_0(i) = y(i)$, where $y(i)$ denotes the value of returns associated with the discrete valued state i (midpoint of the interval associated with i). But we also use the recursion in (10) without the optimization operation, and $v_0(i)$ is defined according to Roy's safety first criterion. This second sequence is denoted $u_n(i)$ and calculated by the recursion

(11) $$u_n(i) = \sum_{j=i^*}^{M} p_{ijk}u_{n-1}(j),$$

for the optimal value of k given by (10). For each trial k in (10) the following inequality must be met

$$\sum_{j=i^*}^{M} p_{ijk} u_{n-1}(j) \geq (1-\gamma) \quad (12)$$

The initial conditions $u_0(i)$ are for Roy's safety first criterion: $u_0(i) = 0$ for $i < i^*$ and $u_0(i) = 1$ for $i \geq i^*$. At each iteration in (10) as the optimal k for given i is determined, subject to the constraint (12), (11) is used to update $u_n(i)$ by using the optimal k on the right hand side. The two sequences (10) and (11) are updated simultaneously for n = 1, 2,... and i = 1, 2,...M for given n. After T iterations (n=T), we have the optimal decision rule for maximizing the expected value of returns at the end of a T-period process, subject to the constraint that the returns are less than y^* with probability γ.

CONCLUDING REMARKS

Methods have been presented for the incorporation of several risk criteria into stochastic dynamic optimization models, but the application of these methods will require a good deal of ingenuity, particularly with respect to the definition of "returns." Many partial farm analyses such as in integrated pest management (IPM) and cropping systems research, are done on a gross margins per acre basis. This measure is not very amenable to incorporation of risk analysis because the returns need to reflect the impact of various random outcomes on the financial viability of the farm operation. If one were to use gross margins per acre, I would recommend an expected value criterion and justify it by the assumption that the differences in the distributions of returns over various decision rules would not involve substantial changes in the risk exposure of the total firm.

If one is bent on using a risk sensitive criterion, then the burden is on them to reformulate the return measure so that it accounts for the risk exposure of the firm across the many possible decision rules. In general, this would require various scenarios with respect to the debt burden of the operator (maybe on a per acre basis) and reflect his fixed payments per period which would be subtracted from the returns denoted r_{ik} in the Markov chain model. One might even add other sources of random returns from the farm operation to the $\{r_{ik}\}$ to yield a composite measure of returns from which various cash payments would be deducted. The ancillary sources of returns would be modeled without use of stochastic optimization to keep the focus on the main problem at hand, e.g., IPM decision problems.

The other main alternative is to model the entire firm, i.e., use a full-fledged stochastic firm growth model (Larson, et al., Schnitkey). But then the size of the model is very demanding and focuses on the entire farming/financial operation and management.

REFERENCES

Bellman, R. Adaptive Control Processes: A Guided Tour, Princeton, Princeton: University Press, 1961.

d'Epenoux, F. "A Probabilistic Production and Inventory Problem." Management Science 10 (1963):98-108.

Fishburn, P. C. "Mean-Risk Analysis with Risk Associated with Below-Target Returns." American Economic Review 67 (1977):116-126.

Hadley, G. Nonlinear and Dynamic Programming. Reading: Addison-Wesley Publishing, 1964.

Howard R. Dynamic Programming and Markov Processes. New York: John Wiley and Sons, 1960.

Larson, D. K., O. Burt, and M. S. Stauber. Economic Analysis of Farm Firm Growth in Northcentral Montana. Research Report 62 (August, 1974). Montana Ag. Exp. Sta., Bozeman, Montana.

Manne, A. S. "Linear Programming and Sequential Decisions." Management Science 6 (1960):259-267.

Roy, A. D. "Safety-First and the Holding of Assets." Econometrica 20 (1952):431-439.

Schnitkey, G. D. "A Dynamic Programming Analysis of Optimal Farmland Investment Decisions: An Application to Central Illinois High-Quality Farmland. Ph.D. thesis, University of Illinois, Urbana, 1987.

Telser, L. "Safety-First and Hedging. Review of Economic Studies 23 (55-56):1-6.

von Neumann, J. and O. Morgenstern. Theory of Games and Economic Behavior. Princeton: Princeton University Press, 1947.