Dynamic Probabilistic Decision Processes

By Photios G. Ioannou,* A.M. ASCE

Abstract: This paper presents a general model for the formulation and solution of the risk-sensitive dynamic decision problem that maximizes the certain equivalent of the discounted rewards of a time-varying Markov decision process. The problem is solved by applying the principle of optimality and stochastic dynamic programming to the immediate rewards and the certain equivalent associated with the remaining transitions of a time-varying Markov process over a finite or infinite time horizon, under the assumptions of constant risk aversion and discounting of future cash flows. The solution provides transient and stationary optimal decision policies which depend on the presence or absence of discounting. The construction equipment replacement problem serves as an example application of the model to illustrate the solution methodology, and the sensitivity of the optimal policy to the discount factor and the degree of risk aversion.

Introduction

Many construction management decisions are dynamic in nature and must be re-evaluated over time based on the current state of some crucial underlying factor, such as the state of company finances, the foreign exchange rate, the current market penetration, the current outlook for the type and number of available future projects, the cost and schedule status of a project, the age and mix of the available spread of equipment, etc.

Even though several models exist for the solution of construction decision problems, none are

*Assistant Professor of Civil Engineering, University of Michigan, Ann Arbor, MI 48109-2125.

Note. This manuscript was published in the ASCE *Journal of Construction Engineering and Management*, Vol. 115, No. 2, June 1989, pp. 237-257. Paper No. 23553.

particularly well suited for addressing dynamic situations that are characterized by uncertainty and risk sensitivity, and which involve a large number of decision alternatives over an extended period of time. The notable exception is the model for analyzing risk-sensitive Markov decision processes proposed by Howard (1972). This model, however, is also limited because it ignores the importance of discounting future cash flows, a fact that may be crucial in determining optimal decision policies over time. This paper extends Howard's model and presents a general framework for the formulation and solution of the risk-sensitive dynamic decision problem that maximizes the certain equivalent of the discounted rewards of a time-varying Markov decision process. An example application of the model to the construction equipment replacement problem illustrates the effects of discounting and risk aversion.

Risk Preference

The development of the model is based on the assumption that the decision maker's behavior, when faced with an uncertain proposition, can be described by a utility curve $u(v)$ that assigns a real number u (in an ordinal scale) to each of the possible outcomes v. The technical term for an uncertain proposition is "lottery." If the decision maker accepts the axioms of utility theory, then his preference in ranking alternative lotteries with uncertain outcomes can be quantified by the expected utility $E[u(v)]$ of each lottery (Luce 1957). If larger values of v are preferred, then $u(v)$ is a monotonically increasing function of v. Whenever the decision maker has to chose between several lotteries, the one with the largest $E[u(v)]$ represents the most desirable alternative.

The certain equivalent of a lottery, $\tilde{v}$, is defined to be that value of the outcome which has the same utility as the expected utility of the lottery.

$$u(\tilde{v}) = E[u(v)] \tag{1}$$

By definition, the decision maker is indifferent between facing the uncertain outcomes of the lottery and receiving the certain equivalent. For this reason, the certain equivalent is also called the "selling price" of the lottery. It must be pointed out that the "selling price" of a lottery is a technical term that should be interpreted algebraically. For example, the certain equivalent for a lottery that involves monetary losses is negative, and represents the amount of money the decision maker is willing to "pay" in order to sell the risk of the lottery to some other party.

The risk premium v_p of a lottery represents the difference between the expected value of the outcomes and the certain equivalent:

$$v_p = E[v] - \tilde{v} \tag{2}$$

For example, the risk premium for a lottery involving monetary losses is the amount that the decision maker is willing to pay in excess of the expected loss in order to avoid facing the risk of the lottery.

The second assumption of the proposed model is that the decision maker is willing to accept what is known as "the delta property": if all the outcomes of a lottery are increased (decreased) by a constant amount Δ, then the certain equivalent of the lottery is also increased (decreased) by Δ:

$$\widetilde{(v + \Delta)} = \tilde{v} + \Delta \tag{3}$$

The principal implication of the delta property is that the certain equivalent of a lottery is independent of the decision maker's present asset position (or wealth). Consequently, if the decision maker is involved in a series of lotteries, then at any stage in this process the certain equivalent for the remaining lotteries does not depend on the specific outcomes of the previous lotteries. The reason for this is that the change Δ in the decision maker's wealth as a result of the outcomes of the first set of lotteries can simply be added to the certain equivalent for the second. Furthermore, if the series of lotteries are independent then the certain equivalent for the combination of sequential independent lotteries equals the sum of the expected equivalents of the lotteries when each is considered separately.

It is easy to prove that the utility function of a decision maker who accepts the delta property must be either linear or exponential (Raiffa 1968, Howard 1977). A linear utility function means that the decision maker is "risk-neutral". Such as person makes decisions by comparing the expected value of the outcomes of each possible alternative. The exponential utility function defines the case of "constant risk aversion" where the risk premium of a lottery does not change if all the outcomes of the lottery are increased by the same amount. The exponential utility function includes the linear case and its functional form can, in general, be written as follows:

$$u(v) = -(\text{sgn}\ \gamma) e^{-\gamma v} \tag{4}$$

Its inverse is given by:

$$v = -\frac{1}{\gamma} \ln[-(\text{sgn}\ \gamma)\ u] \tag{5}$$

The parameter γ is the “risk aversion coefficient” and (sgn γ) is the sign of γ. A positive γ indicates that the decision maker is risk averse: the certain equivalent of a lottery is smaller than the expected value of the outcomes, and thus, the risk premium for the lottery is positive. A negative γ indicates that the decision maker prefers risk, the opposite situation. In the limit, as γ approaches zero, the behavior implied by the exponential function tends to be risk neutral, the same as for a linear utility curve.

The delta property can now be expressed in terms of utility:

$$u(v + \Delta) = -(\text{sgn}\ \gamma) e^{-\gamma(v+\Delta)} = e^{-\gamma\Delta} u(v) \tag{6}$$

Increasing the outcomes of a lottery by Δ causes their utilities to be multiplied by the same quantity $e^{-\gamma\Delta}$. As a result, the expected utility of the lottery is also multiplied by $e^{-\gamma\Delta}$ and the certain equivalent $\tilde{v}$ is increased by Δ.

Time-Varying Markov Reward Process

The general dynamic decision problem is based on a time-varying, discrete-state, discrete-stage Markov process with N states in each stage. The process is described by the transition probability matrix $\mathbf{P}(n)$, which, in general, is a function of n, the number of transitions (or stages) to the end of the process. Each element $p_{ij}(n)$ in this matrix represents the probability that the process will make a transition into state j given that it currently occupies state i with n transitions remaining. A transition from state i to state j with n transitions remaining is associated with a reward $r_{ij}(n)$, which can be positive or negative. As a result, the reward matrix $\mathbf{R}(n)$ is also a function of n.

Consider that the process is currently in state i of stage $n + 1$ (i.e., it has $n + 1$ transitions remaining). The total value of the reward from the remaining transitions of the process is given by the random variable $v_i(n + 1)$. This random variable represents the outcomes of a complex lottery (event tree) which combines the transition probabilities and outcomes of the $n+1$ sequential simple lotteries that occur over the remaining $n + 1$ transitions. Notice that the $n + 1$ sequential simple lotteries are not independent because the transition probabilities and the reward of each transition depends on the state occupied in the previous stage.

For a decision maker that satisfies the delta property the certain equivalent for the future rewards of this process is independent of the changes in the decision maker's asset position due to the rewards received in previous transitions. The certain equivalent $\tilde{v}_i(n+1)$ is the amount the decision maker is willing to take for certain, instead of facing the remaining $n+1$ rewards of the process. Similarly, $\tilde{v}_i(n)$ is the certain equivalent with n transitions remaining. The transition from state i in stage $n+1$ to state j in stage n will pay an immediate reward $r_{ij}(n+1)$ and will lead to a certain equivalent $\tilde{v}_j(n)$ for the remaining n transitions. Thus, we can write the recursive relationship (Howard 1972):

$$u(\tilde{v}_i(n+1)) = \sum_{j=1}^{N} p_{ij}(n+1)\, u(r_{ij}(n+1) + \tilde{v}_j(n)) \quad n = 0, 1, 2, \ldots \tag{7}$$

The term $u(\tilde{v}_i(n+1))$ represents the expected utility of the reward process when it currently occupies state i in stage $n+1$ (i.e., with $n+1$ transitions remaining). Making a transition from state i in stage $n+1$ to state j in stage n, leads to a reward $r_{ij}(n+1)$ and the prospect of facing n more transitions whose certain equivalent is $\tilde{v}_j(n)$. The transition from i to j increases the end values for the remaining n transitions by the constant value $r_{ij}(n+1)$. Because of the delta property, the certain equivalent for the prospect of making the next transition into state j and facing the remaining n transitions is given by the sum $r_{ij}(n+1) + \tilde{v}_j(n)$. As a result, the original multistage process is equivalent to the single-transition process from state i to state j where the rewards are increased to $r_{ij}(n+1) + \tilde{v}_j(n)$. This equivalence allows the recursive solution of the problem starting at stage 0 and working backwards in time. The certain equivalents $\tilde{v}_j(0)$ in stage 0 must be assigned directly by the decision maker.

Time-Varying Markov Reward Process with Discounting

The above formulation is inappropriate for many business problems because it ignores the time value of money. To introduce the concept of discounting future cash flows we need to define two new parameters, $\alpha(n+1)$ and $\beta(n+1)$, so that $\alpha(n+1)r_{ij}(n+1)$ and $\beta(n+1)\tilde{v}_j(n)$ are the respective present values of $r_{ij}(n+1)$ and $\tilde{v}_j(n)$ at stage $n+1$. If the reward $r_{ij}(n+1)$ is paid at the beginning of the time span between stages $n+1$ and n, then $\alpha(n+1) = 1$. Similarly, if the reward $r_{ij}(n+1)$ is paid at the end of the time span between stages $n+1$ and n, then $\alpha(n+1) = \beta(n+1)$. Thus, parameters $\alpha(n+1)$ and $\beta(n+1)$ are discount factors, where $\beta(n+1)$ equals:

$$\beta(n+1) = \frac{1}{1+\rho(n+1)} \tag{8}$$

The parameter $\rho(n+1)$ is the discount rate for the time period from stage $n+1$ to stage n. For the ease of notation, and with no loss of generality, we shall assume that the time interval between stages is constant and thus use the simplified notation α and β.

Considering discounting we can now write:

$$u(\tilde{v}_i(n+1)) = \sum_{j=1}^{N} p_{ij}(n+1)\, u(\alpha r_{ij}(n+1) + \beta \tilde{v}_j(n)) \quad n = 0, 1, 2, \ldots \tag{9}$$

where the certain equivalent $\tilde{v}_i(n+1)$ for the remaining $n+1$ transitions is expressed in present value terms in stage $n+1$ and the certain equivalents $\tilde{v}_j(n)$ are expressed in present value terms in stage n. Substituting the analytical form of the exponential utility function gives:

$$u(\tilde{v}_i(n+1)) = \sum_{j=1}^{N} p_{ij}(n+1)\, [-(\operatorname{sgn}\gamma) e^{-\gamma(\alpha r_{ij}(n+1) + \beta \tilde{v}_j(n))}] \quad n = 0, 1, 2, \ldots \tag{10}$$

$$u(\tilde{v}_i(n+1)) = \sum_{j=1}^{N} p_{ij}(n+1)\, [-(\operatorname{sgn}\gamma)]\, e^{-\gamma\alpha r_{ij}(n+1)}\, e^{-\gamma\beta \tilde{v}_j(n)} \quad n = 0, 1, 2, \ldots \tag{11}$$

$$u(\tilde{v}_i(n+1)) = \sum_{j=1}^{N} p_{ij}(n+1)\, e^{-\gamma\alpha r_{ij}(n+1)}\, u(\beta \tilde{v}_j(n)) \quad n = 0, 1, 2, \ldots \tag{12}$$

The utility $u(\beta\tilde{v}_j(n))$ of the certain equivalent at stage n discounted to stage $n+1$ can be computed as follows:

$$\begin{aligned} u(\beta\tilde{v}_j(n)) &= -(\operatorname{sgn}\gamma)\, e^{-\gamma\beta\tilde{v}_j(n)} \\ &= -(\operatorname{sgn}\gamma)\, [e^{-\gamma\tilde{v}_j(n)}]^{\beta} \\ &= -(\operatorname{sgn}\gamma)\, [-(\operatorname{sgn}\gamma)\, u(\tilde{v}_j(n))]^{\beta} \end{aligned} \tag{13}$$

Substituting into the previous equation yields:

$$u(\tilde{v}_i(n+1)) = \sum_{j=1}^{N} p_{ij}(n+1)\, e^{-\gamma\alpha r_{ij}(n+1)}\, [-(\operatorname{sgn}\gamma)]\, [-(\operatorname{sgn}\gamma)\, u(\tilde{v}_j(n))]^{\beta} \quad n = 0, 1, 2, \ldots \tag{14}$$

This equation can be simplified by observing that:

$$u(\alpha r_{ij}(n+1)) = -(\text{sgn}\,\gamma)\, e^{-\gamma\alpha r_{ij}(n+1)} \tag{15}$$

where $u(\alpha r_{ij}(n+1))$ is the utility associated with the present value of the reward of the next transition to state j when currently in state i with $n+1$ transitions remaining. The notation is simplified further, if we write the expected utility $u(\tilde{v}_i(n))$ with n transitions remaining in the process as $u_i(n)$:

$$u_i(n+1) = \sum_{j=1}^{N} p_{ij}(n+1)\, u(\alpha r_{ij}(n+1))\, [-(\text{sgn}\,\gamma)\, u_j(n)]^{\beta} \quad n = 0, 1, 2, \ldots \tag{16}$$

This is the fundamental recursive equation that relates the expected utility $u_i(n+1)$ of state i in stage $n+1$ to the expected utility $u_j(n)$ of state j in stage n. The parameters $p_{ij}(n+1)$ and $r_{ij}(n+1)$ on the right-hand side describe the transition from stage $n+1$ to stage n. The effect of these parameters can be summarized by defining $q_{ij}(n+1)$ to be equal to the product of the utility of the present value of the immediate reward times the associated transition probability:

$$\begin{aligned} q_{ij}(n+1) &= p_{ij}(n+1)\, [-(\text{sgn}\,\gamma)]\, e^{-\gamma\alpha r_{ij}(n+1)} \\ &= p_{ij}(n+1)\, u(\alpha r_{ij}(n+1)) \end{aligned} \tag{17}$$

Thus, the recursive relationship can be written as:

$$u_i(n+1) = \sum_{j=1}^{N} q_{ij}(n+1)\, [-(\text{sgn}\,\gamma)\, u_j(n)]^{\beta} \quad n = 0, 1, 2, \ldots \tag{18}$$

In matrix form we can write:

$$\mathbf{U}(n+1) = \mathbf{Q}(n+1)\, [-(\text{sgn}\,\gamma)\, \mathbf{U}(n)]^{\beta} \quad n = 0, 1, 2, \ldots \tag{19}$$

It is understood that the exponent on the right-hand side applies to each individual element of the vector $[-(\text{sgn}\,\gamma)\, \mathbf{U}(n)]$ and not to the vector itself (the matrix power of a vector does not exist).

In the case of no discounting we have $\alpha = \beta = 1$ and we can write:

$$u_i(n+1) = \sum_{j=1}^{N} p_{ij}(n+1)\, e^{-\gamma r_{ij}(n+1)}\, u_j(n) \quad n = 0, 1, 2, \ldots \tag{20}$$

$$u_i(n+1) = -(\text{sgn}\,\gamma) \sum_{j=1}^{N} q_{ij}(n+1)\, u_j(n) \quad n = 0, 1, 2, \ldots \tag{21}$$

$$\mathbf{U}(n+1) = -(\text{sgn}\,\gamma)\, \mathbf{Q}(n+1)\, \mathbf{U}(n) \quad n = 0, 1, 2, \ldots \tag{22}$$

Using the above equations we can compute the expected utility at any state i and any stage n by starting at stage 0 and continuing backwards in time. At any stage n the certain equivalent of the process can be computed from:

$$\tilde{v}_i(n) = -\frac{1}{\gamma} \ln[-(\text{sgn}\,\gamma)\, u_i(n)] \tag{23}$$

In the case where we only compute and store the discounted form of utilities $u(\beta \tilde{v}_i(n))$, we can use the formula:

$$\tilde{v}_i(n) = -\frac{1}{\gamma\beta} \ln[-(\text{sgn}\,\gamma)\, u(\beta \tilde{v}_i(n))] \tag{24}$$

Stationary Markov Reward Process with Discounting

The classical Markov reward process is a special case of the time-varying process presented above in which the transition probability matrix **P**, the reward matrix **R**, and the discount factors α and β, are the same for all transitions:

$$\begin{aligned} p_{ij}(n) &= p_{ij} \\ r_{ij}(n) &= r_{ij} \\ \alpha(n) &= \alpha \\ \beta(n) &= \beta \end{aligned} \qquad n = 1, 2, \ldots \tag{25}$$

As a result, the matrix **Q** is also independent of n:

$$\begin{aligned} q_{ij} &= p_{ij}\, [-(\text{sgn}\,\gamma)]\, e^{-\gamma\alpha r_{ij}} \\ &= p_{ij}\, u(\alpha r_{ij}) \end{aligned} \tag{26}$$

The recursive equations for computing the expected utility $u_i(n+1)$ associated with state i with $n+1$ transitions remaining can be written as follows:

$$u_i(n+1) = \sum_{j=1}^{N} q_{ij}\, [-(\text{sgn}\,\gamma)\, u_j(n)]^{\beta} \quad n = 0, 1, 2, \ldots \tag{27}$$

When the matrix $\mathbf{Q}$ is not a function of time, the Markov reward process exhibits stationary behavior for large n. As the number of remaining transitions n increases, the expected utility $u_i(n)$ and the certain equivalent $\tilde{v}_i(n)$ associated each state i approach stationary (limiting) values u_i and $\tilde{v}_i$. The stationary (limiting) certain equivalent $\tilde{v}_i$ for each state i is analogous to the finite present value of a discounted perpetuity. The cause of this stationary behavior is the fact that the utility contribution of future rewards is discounted over time and approaches zero as the interval from the present stage to the future reward becomes very large. This, of course, is true for $\beta < 1$.

The computation of the stationary certain equivalents $\tilde{v}_i$ is of particular interest for ongoing processes that are not expected to end in the near future. The $\tilde{v}_i$ can be computed by repeatedly applying the above recursive equations until the change in expected utilities or certain equivalents from stage n to stage $n+1$ are small enough to be considered insignificant. Obviously, the assumed values for $\tilde{v}_i(0)$ or $u_i(0)$ influence the number of iterations needed for convergence. Any choice of starting values, however, will lead to the same stationary values $\tilde{v}_i$ and u_i because their contribution will eventually become zero.

Alternatively, the stationary values can be computed directly by solving the following set of N nonlinear simultaneous equations:

$$u_i = \sum_{j=1}^{N} q_{ij} \, [-(\text{sgn}\,\gamma)\, u_j]^{\beta} \quad i = 1, 2, \ldots, N \tag{28}$$

or, in matrix form:

$$\mathbf{U} = \mathbf{Q}\, [-(\text{sgn}\,\gamma)\, \mathbf{U}]^{\beta} \tag{29}$$

Indeed, the recursive equations (27) represent a simple algorithm for the numerical solution of the above nonlinear set of stationary equations.

Stationary Markov Reward Process without Discounting

In the case where the discount rate is zero (i.e. $\beta = 1$), the process assumes a slightly different behavior for large n. Since future rewards are not discounted, their contribution to expected utility does not depend on when the reward is received. Thus, the expected utility for each state of the process may not converge to a finite value. For example, the expected utilities $u_i(n)$ for a process in which the all the transition rewards represent profit increase indefinitely as the number of remaining transitions n becomes very large. As a result, $\lim_{n\to\infty} \tilde{v}_i(n) = \infty$ $(i = 1, 2, \ldots, N)$. The limiting behavior, in this case, is reflected in the way the expected utilities and certain equivalents change from stage to stage.

For the nondiscounted Markov reward process to exhibit stationary behavior the transition probability matrix $\mathbf{P}$ must satisfy the following two conditions (Howard 1972): (1) all states communicate (i.e. there are no trapping states), and (2) the states do not follow each other in a repeating cycle (the process is acyclic). (These conditions are not required if the future cash flows are discounted.)

If these two conditions are satisfied, then for large n:

$$u_i(n+1) = \lambda u_i(n) \tag{30}$$

This means that, for large n, the expected utility of state i will be multiplied by a constant λ at each preceding stage. The constant multiplier λ is the same for all states i.

Using this relationship and the recursive equations (27) leads to the following eigenvalue problem (system of N linear equations):

$$\lambda\, \mathbf{U} = -(\operatorname{sgn} \gamma)\, \mathbf{Q}\, \mathbf{U} \tag{31}$$

Under the two conditions posed above, it can be proven that the constant λ exists and is a positive real number equal to the largest eigenvalue of the matrix $-(\operatorname{sgn} \gamma)\, \mathbf{Q}$. The corresponding eigenvector $\mathbf{U}$ is the vector of relative expected utilities. Solving the eigenvalue problem is a computational alternative to the repeated application of the recursive equations (27), which allows the direct determination of λ and the relative expected utilities u_i for each state i. The values u_i are relative because for any constant c the vector $c\mathbf{U}$ will also satisfy the eigenvalue equations. As a result, we must assign an arbitrary value to the expected utility of one of the states i in order to compute the other $N - 1$ relative expected utilities. A convenient assignment is $u_N = -(\operatorname{sgn} \gamma)$ because it simplifies the interpretation of relative certain equivalents as explained below.

The stationary behavior of certain equivalents can be determined by substituting the explicit form of the exponential utility curve into equation (30):

$$-(\operatorname{sgn} \gamma) e^{-\gamma \tilde{v}_i(n+1)} = \lambda \left[-(\operatorname{sgn} \gamma) e^{-\gamma \tilde{v}_i(n)}\right] \tag{32}$$

$$\tilde{v}_i(n+1) = -\frac{1}{\gamma} \ln \lambda + \tilde{v}_i(n) \tag{33}$$

$$\tilde{v}_i(n+1) - \tilde{v}_i(n) = -\frac{1}{\gamma}\ln\lambda = \tilde{g} \tag{34}$$

Thus, for large n, the certain equivalent for state i increases by a constant $\tilde{g}$ at each preceding stage. The constant $\tilde{g}$ is called the "certain equivalent gain" of the process and is the same for all states i.

Because the certain equivalent gain $\tilde{g}$ is the same for all states i, it follows that the difference $\tilde{v}_i(n) - \tilde{v}_j(n)$ between the certain equivalents of any two states i and j in the same stage, will also approach a limiting value $\tilde{v}_i - \tilde{v}_j$ as n becomes very large. The limiting $\tilde{v}_i$ and $\tilde{v}_j$ are computed from the relative expected utilities u_i and u_j, and thus their values depend on the arbitrarily chosen value for u_N. As pointed out above, a convenient value is $u_N = -(\text{sgn}\ \gamma)$, which implies that $\tilde{v}_N = 0$. As a result, the computed limiting certain equivalents $\tilde{v}_i$ are "relative" to $\tilde{v}_N(n)$. In other words, they represent the following limit:

$$\tilde{v}_i = \lim_{n\to\infty}(\tilde{v}_i(n) - \tilde{v}_N(n)) \tag{35}$$

This means that the relative certain equivalent $\tilde{v}_i$ expresses the difference in value that the decision maker attributes to being in state i, as opposed to state N, with an infinite number of transitions remaining.

Time-Varying Markov Decision Process

The reward processes presented above have been based on the implicit assumption that the decision maker does not have any control over the transition and reward matrices $\mathbf{P}(n)$ and $\mathbf{R}(n)$. Let us now consider the decision process where the decision maker can influence the transition probabilities and future rewards of the process by selecting an alternative from a set of possible actions at each state and stage. Alternative action k is associated with its own transition matrix $\mathbf{P}^k(n)$ and reward matrix $\mathbf{R}^k(n)$. Thus, in any state i of stage n the decision maker has to decide whether to adopt alternative k, which in turn determines the $p_{ij}^k(n)$ and $r_{ij}^k(n)$ for the next transition.

The number of possible alternative actions in each state i may be different from the number of alternatives for some other state j, and may also vary from stage to stage as a function of the remaining transitions n. Furthermore, the nature of the alternatives can also vary from state to state and from stage to stage. For example, the alternatives might be whether to conduct a marketing study looking at new business areas in the case of a decline in the decision maker's current market, or whether to expand in new geographical regions if the demand for the decision maker's specialty is on the rise. From this definition it is clear that the generalized time-varying Markov decision process has the structure of a decision tree where the probabilities of chance nodes and the associated rewards are conditional on which alternative action is taken.

The problem is to find the optimal decision policy that will produce the maximum expected utility $u_i^*(n)$ for the rewards generated by the remaining n transitions of the process. An optimal policy would indicate the optimal alternative action k^* that should be chosen for a particular combination of state i and stage n. To facilitate the notation we will assign a number $k = 1, 2, \ldots,$ to each of the possible alternative actions in each state, and let $d_i(n)$ be the number of the optimum alternative k^* that we should follow when the process is in state i with n transitions remaining. Thus, the optimal policy is established, and the problem is solved, when we have determined the N values of the column vector $\mathbf{D}(n)$ for every n.

The procedure for solving this problem is recursive and is based on stochastic dynamic programming. Given that we know the optimal policies for stages $1, 2, \ldots, n$, and thus the maximum expected utilities $u_j^*(n)$ at stage n, the maximum expected utility $u_i^*(n+1)$ and the optimal policy $d_i(n+1)$ for state i at stage $n+1$ can be computed by using the principle of optimality:

$$u_i^*(n+1) = \max_k \sum_{j=1}^{N} p_{ij}^k(n+1)\, e^{-\gamma\alpha\, r_{ij}^k(n+1)}\, [-(\operatorname{sgn}\gamma)][-(\operatorname{sgn}\gamma)\, u_j^*(n)]^{\beta} \quad n = 0, 1, 2, \ldots \tag{36}$$

or,

$$u_i^*(n+1) = \max_k \sum_{j=1}^{N} q_{ij}^k(n+1)\, [-(\operatorname{sgn}\gamma)\, u_j^*(n)]^{\beta} \quad n = 0, 1, 2, \ldots \tag{37}$$

Thus, for each state i in stage $n+1$ we search for the alternative k^* that maximizes the expected utility $u_i(n+1)$ of the future rewards. The alternative k^* that maximizes $u_i(n+1)$ is the optimal policy $d_i(n+1)$. We can determine the N values of the vector $\mathbf{D}(n+1)$ by repeating this process over all states i. Once the vector $\mathbf{D}(n+1)$ has been established, we can repeat the process to compute $\mathbf{D}(n+2)$, etc. Using this method, the problem is solved backwards, starting at the end of the process (stage 0) and solving the decision problem in stage 1, stage 2, etc.

Example: The Equipment Replacement Problem

To illustrate the application of this model to a construction problem we shall examine the replacement problem for construction equipment — in this case, a scraper. The example is adapted from (Howard 1960) and can be easily modified and extended for other types of equipment. This application will provide additional insight into the effects of risk aversion and the discount factor on the development of an optimal policy.

The problem is whether to keep an existing scraper, or whether to trade it in and buy another one, and if so, to determine how old the replacement should be. We will assume that these decisions have to be made every 3 months. The states of the system i, are defined to be equal to the age of the existing scraper in 3-month units. For example, a brand new scraper is in state 0, a 4-year-old scraper is in state 16, and a 10-year-old scraper is in state 40. In order to keep the number of states reasonably small, we will assume that the maximum useful life for a scraper is 10 years. When a scraper reaches state 40 it is essentially worn out and thus it remains in state 40 forever.

The decision alternatives for any state i are defined as follows: For $k = 0, 1, 2, \ldots, 39$, alternative k is to trade the existing scraper for a $3k$-month-old one. Alternative 40 is to keep the existing scraper for another three months (this alternative is also indicated by the capital letter "K" which stands for "Keep"). Thus, we have 40 states, with 41 decision alternatives in each state, for a total of $41^{40} \approx 3 \times 10^{64}$ possible policies.

The data for this problem are shown in Table 1 and plotted in Fig. 1. They are defined as follows:

C_i = The cost of buying a scraper of age i.

T_i = The trade-in value (salvage value) of a scraper of age i.

E_i = The expected cost of operating a scraper of age i until it reaches age $i + 1$. (Reductions in productivity that may occur because of age can be modeled by an appropriate increase in operating costs.)

p_i = The probability that a scraper of age i will survive to reach age $i + 1$ without incurring a prohibitively expensive repair.

Table 1: Data for the Scraper Replacement Example

i	C_i	T_i	E_i	p_i		i	C_i	T_i	E_i	p_i
(1)	(2)	(3)	(4)	(5)		(6)	(7)	(8)	(9)	(10)
0	$200,000	$160,000	$5,000	1.000						
1	184,000	146,000	5,300	0.999		21	$34,500	$24,000	$11,500	0.925
2	168,000	134,000	5,600	0.998		22	33,000	22,500	11,800	0.919
3	156,000	123,000	5,900	0.997		23	31,500	21,000	12,100	0.910
4	130,000	105,000	6,200	0.996		24	30,000	20,000	12,500	0.900
5	122,000	98,000	6,500	0.994		25	29,000	19,000	12,900	0.890
6	115,000	91,000	6,800	0.991		26	28,000	18,000	13,300	0.880
7	108,000	84,000	7,100	0.988		27	26,500	17,000	13,700	0.865
8	90,000	71,000	7,500	0.985		28	25,000	16,000	14,100	0.850
9	84,000	65,000	7,800	0.983		29	24,000	15,000	14,500	0.820
10	78,000	60,000	8,100	0.980		30	23,000	14,500	15,000	0.790
11	73,000	55,000	8,400	0.975		31	22,000	14,000	15,500	0.760
12	60,000	48,000	8,700	0.970		32	21,000	13,500	16,000	0.730
13	56,000	43,000	9,000	0.965		33	20,000	13,000	16,700	0.660
14	52,000	39,000	9,300	0.960		34	19,000	12,000	17,500	0.590
15	48,000	36,000	9,600	0.955		35	18,000	11,500	18,200	0.510
16	44,000	33,000	10,000	0.950		36	17,000	11,000	19,000	0.430
17	42,000	31,000	10,300	0.945		37	16,000	10,500	20,500	0.300
18	40,000	29,000	10,600	0.940		38	15,000	9,500	22,000	0.200
19	38,000	27,000	10,900	0.935		39	14,000	8,700	23,500	0.100
20	36,000	25,500	11,200	0.930		40	13,000	8,000	25,000	0.000

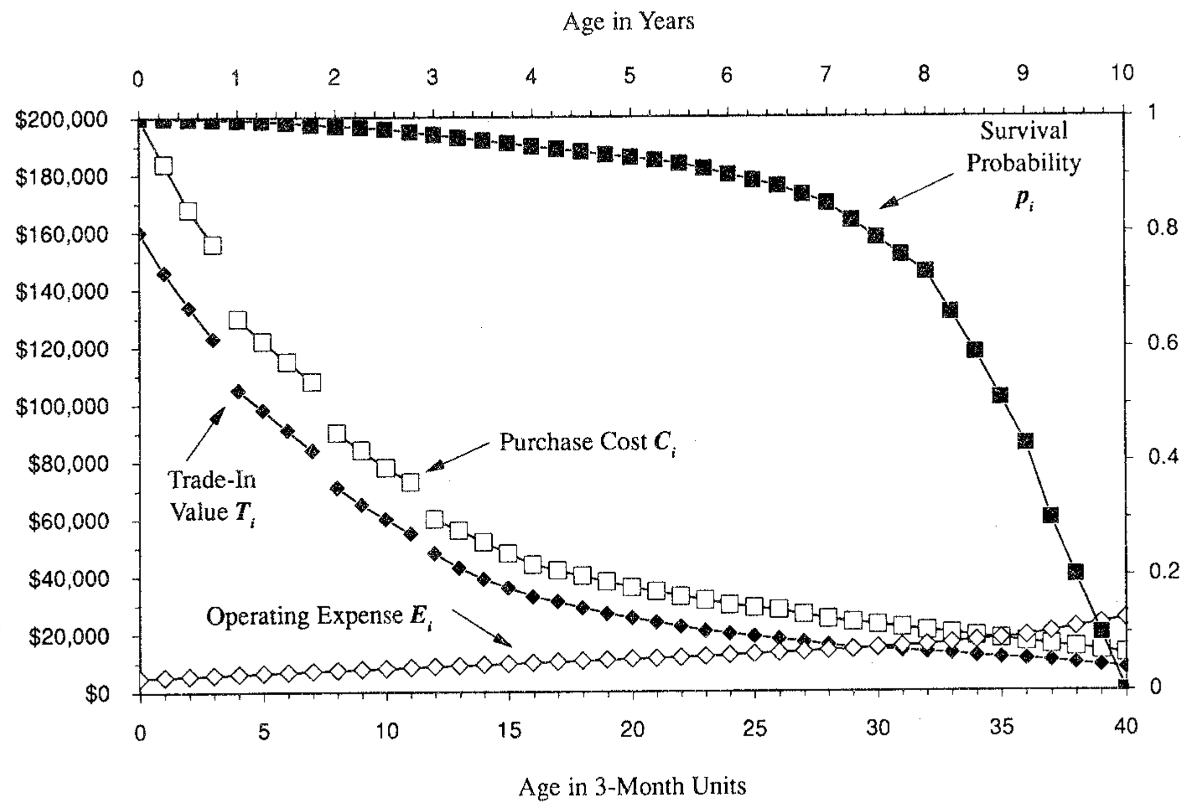

Figure 1: Data for the Scraper Replacement Example

The various costs are assumed to occur at the beginning of each stage so that $\alpha = 1$. Equivalently, we may assume that the cost data are given as present values at the beginning of each 3-month period and thus no further discounting is necessary. The survival probabilities defined above are necessary in order to limit the number of states. A scraper that incurs an expensive breakdown is immediately sent to state 40 and remains there forever (i.e., $p_{40} = 0$).

Problem Formulation

From the above data we can compute the probability transition matrices and the reward matrices for the different decision alternatives.

For $k = 0, 1, 2, \ldots, 39$ (trade-in for a $3k$-month-old scraper):

$$p_{ij}^k = \begin{cases} p_k & j = k+1 \\ 1 - p_k & j = 40 \\ 0 & \text{other } j \end{cases} \tag{38}$$

$$r_{ij}^k = \begin{cases} T_i - C_k - E_k & j = k+1 \\ T_i - C_k - E_k & j = 40 \\ 0 & \text{other } j \end{cases} \tag{39}$$

$$q_{ij}^k = \begin{cases} p_k\, u(T_i - C_k - E_k) & j = k+1 \\ (1 - p_k)\, u(T_i - C_k - E_k) & j = 40 \\ 0 & \text{other } j \end{cases} \tag{40}$$

For $k = 40$ (keep existing scraper for another 3 months):

$$p_{ij}^k = \begin{cases} p_i & j = i+1 \\ 1 - p_i & j = 40 \\ 0 & \text{other } j \end{cases} \tag{41}$$

$$r_{ij}^k = \begin{cases} -E_i & j = i+1 \\ -E_i & j = 40 \\ 0 & \text{other } j \end{cases} \tag{42}$$

$$q_{ij}^{k} = \begin{cases} p_i \, u(-E_i) & j = i+1 \\ (1-p_i)\, u(-E_i) & j = 40 \\ 0 & \text{other } j \end{cases} \tag{43}$$

The recursive equations for solving the problem can now be written as follows:

$$u_i^*(n+1) = \max_k \begin{cases} \begin{pmatrix} p_k \, e^{-\gamma\,(T_i - C_k - E_k)}[-(\text{sgn}\,\gamma)]\,[-(\text{sgn}\,\gamma)\,u_{k+1}^*(n)]^{\beta} \\ +(1-p_k)\, e^{-\gamma\,(T_i - C_k - E_k)}[-(\text{sgn}\,\gamma)]\,[-(\text{sgn}\,\gamma)\,u_{40}^*(n)]^{\beta} \end{pmatrix}; & k = 0, 1, \ldots, 39 \\ \\ \begin{pmatrix} p_i \, e^{-\gamma\,(-E_i)}[-(\text{sgn}\,\gamma)]\,[-(\text{sgn}\,\gamma)\,u_{i+1}^*(n)]^{\beta} \\ +(1-p_i)\, e^{-\gamma\,(-E_i)}[-(\text{sgn}\,\gamma)]\,[-(\text{sgn}\,\gamma)\,u_{40}^*(n)]^{\beta} \end{pmatrix}; & k = 40 \end{cases} \tag{44}$$

Solution Methodology

Probably the easiest and most accessible method for solving the problem is to program these equations in a microcomputer spreadsheet. (The spreadsheet for this problem can be obtained from the author.) The general organization of the spreadsheet can be described as follows: For each of the 41 decision alternatives k we define 40 cells, one for each state i, each containing a formula that evaluates the right-hand side of the expected utility equations (44) for the corresponding alternative k and state i. The $41 \times 40 = 1640$ formulas are arranged in a single column, sorted by decision alternative and then by state within each alternative. The maximum expected utilities $u_i^*(n+1)$ for each state i are computed in 40 separate cells at the top of the column. The formula in the ith cell compares the expected utilities $u_i^k(n+1)$ of the 41 decision alternatives below and selects the maximum.

The 1640 formulas for the alternative-state combinations in stage $n+1$ are, of course, a function of the maximum expected utilities $u_j^*(n)$ in stage n. Since the $u_j^*(n)$ are common input to all 1640 formulas they are stored to the left of the 40 cells used for storing $u_i^*(n+1)$. Between the two columns for $u_j^*(n)$ and $u_i^*(n+1)$ there is one more column that computes which alternative k^* produces the maximum $u_i^*(n+1)$. Thus, the 40 cells in this column contain the optimal policies $d_i(n+1)$ for the corresponding states i. To perform iteration over stages we need a simple program (macro) that copies the values (not the formulas) of $u_i^*(n+1)$ into the cells for $u_i^*(n)$ and recalculates the worksheet. The 40 cells that contained $u_i^*(n)$ now hold the values $u_i^*(n+1)$ and the columns $d_i(n+1)$ and $u_i^*(n+1)$ now show the solution for $d_i(n+2)$ and $u_i^*(n+2)$. By repeating this process we can solve the problem for any number of stages. Obviously, in order to start the iteration process we have to initially define $u_i(0)$ in the $u_i^*(n)$ column. For this example, the values $u_i(0)$ were computed from the certain equivalents $\tilde{v}_i(0)$ which were set equal to the corresponding trade-in values T_i. In other words, at the end of the process the scraper is no longer needed and it is sold.

The optimal policies determined by this approach are a function of the risk aversion coefficient γ and the discount factor β. Sensitivity analysis over these parameters can be performed by changing these parameters and resolving the problem for the optimal policies. A similar sensitivity analysis can be performed for all the cost and probability data shown in Table 1.

Interpretation of the Results

This problem has been solved for $\beta = 1$ (no discounting), $\beta = 0.97$ (annual discount rate $\approx$ 12.96%) and $\beta = 0.95$ (annual discount rate $\approx 22.77\%$), and for various values of the risk aversion coefficient γ. The results for the first 10 iterations (10 stages) for $\beta = 0.97$ and $\gamma = 0.00005$ (risk averse behavior) are shown in Fig. 2. This figure also shows the stationary optimal policy when the number of remaining stages is infinite.

The abscissa in this figure represents the states i of the process (the age of the existing scraper in 3-month units). Each column represents the certain equivalent in dollars of owning a scraper of that age with $n = 0, 1, 2, \ldots, 10$, transitions remaining (these amounts are measured using the left dollar scale). The optimal decision alternative $d_i(n) = k^*$ for each state i and each stage n is shown above each column. For example, if the existing scraper is 1 year old ($i = 4$) and we need a scraper for only 3 months ($n = 1$), then we should keep it ($d_4(1) = \text{K}$). Similarly, if we need a scraper for 2 more years ($n = 8$) then we should trade the existing 1-year-old scraper for a 4-year-old ($d_4(8) = 16$). Notice that as the number of remaining transitions increases, the optimal policy calls for a newer scraper as a replacement, which is the expected behavior.

The optimal policy reached in stage 10 remains unchanged as the number of remaining transitions increases (stationary decision behavior). The line graph at the bottom of Fig. 2 shows the

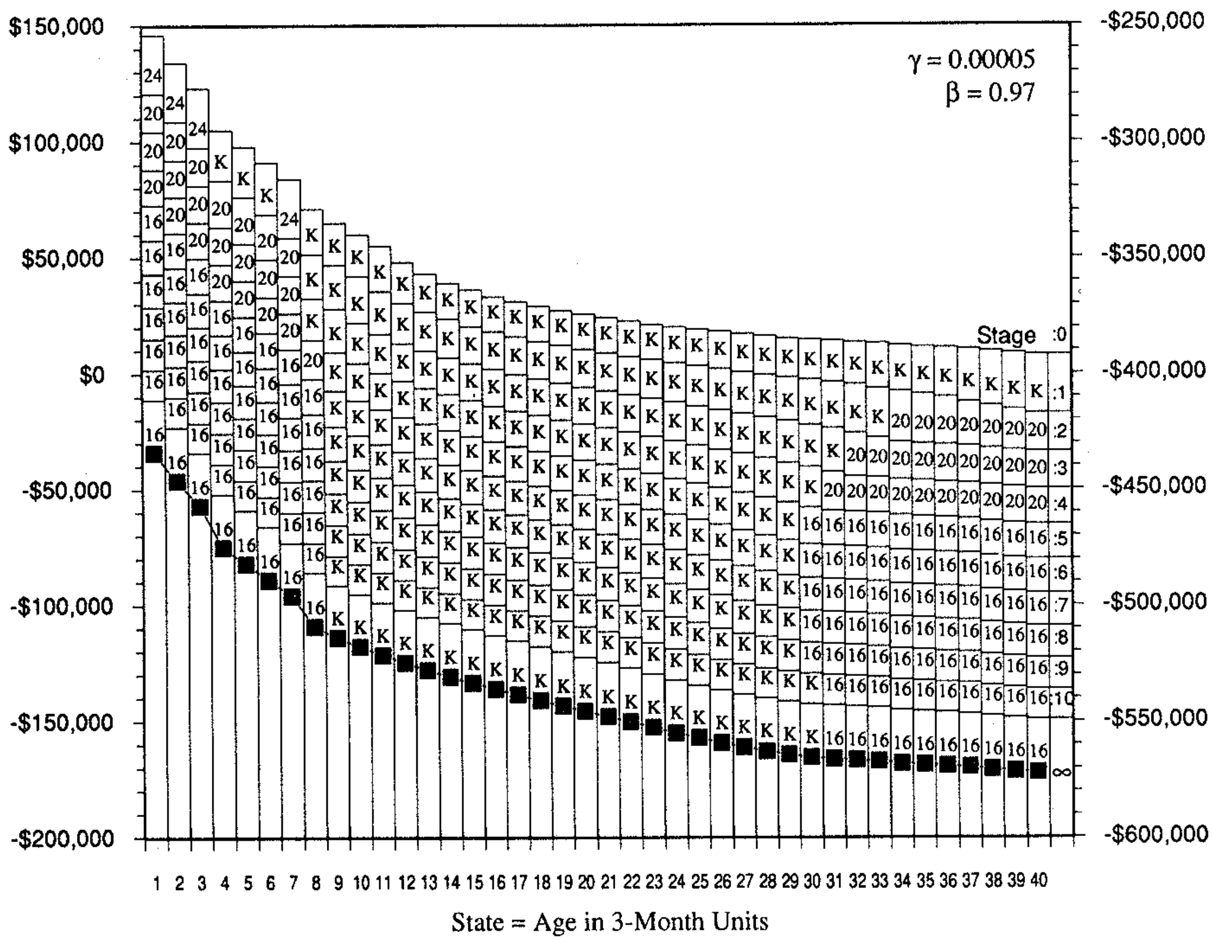

Figure 2: Certain Equivalents $\tilde{v}_i(n)$ and Optimal Policies $d_i(n) = k^*$ for $n = 1, 2, \ldots, 10, \infty$ ($\gamma = 0.00005$, $\beta = 0.97$)

stationary certain equivalents (in dollars measured on the right scale) and the optimal policy for an infinite number of transitions remaining. This policy is the same as for stage $n = 10$ and is probably the policy that interests an earth moving company the most.

The decision policy for an infinite number of remaining transitions is to keep the current scraper if it is between 2.25 and 7.5 years old and to trade it for a 4-year-old otherwise. This means that a company following this policy will always own a scraper that is between 4 and 7.5 years old. Furthermore, it is possible to sell to this company a special lease (or the rights) to use a cost-free scraper between these ages for a lump sum amount equal to the certain equivalent corresponding to the age of the scraper it currently owns. For example, if the company currently owns a 5-year-old scraper, it should be indifferent between following this policy and paying a one-time fee of $545,581 for the right to use a scraper between 4 and 7.5 years old over an infinite time horizon without incurring any additional ownership and operating costs.

The stationary optimal certain equivalents and the corresponding policies for $\beta = 0.97$ and $\beta = 0.95$ are shown in Tables 2 and 3, and are plotted in Fig. 3 and 4 for various risk aversion coefficients γ ranging from -0.001 (gambling behavior) to 0.0005 (risk averse behavior). The values for $\gamma = 0$ indicate the optimal policies for risk neutral behavior. Notice that as γ decreases, the implied behavior is less risk averse, and that the optimal policies involve trade-ins for newer equipment. The opposite is true for large values of γ. In the extreme, when $\gamma = 0.0005$, the optimal policy includes keeping a 10-year-old (worn out) scraper. As a result, a company following this policy will eventually end up with a 10-year-old scraper that will be kept forever at a high operating cost of $25,000 every three months. The certain equivalent in this case equals the net present value of a discounted perpetuity, a fact easily verified from the given data. For example, for $\beta = 0.97$ we have $\tilde{v}_{40} = \$25,000/(1 - 0.97) = \$833,333$ (Fig. 3).

Table 4 shows the optimal policies for $\beta = 1$ (no discounting) and various degrees of risk aversion. Notice that since there is no discounting, the notion of present value is substituted by the notion of the certain equivalent gain $\tilde{g}$, which represents the cost of owning and operating a scraper over a 3-month period for a company that follows the optimal policy. Fig. 5 shows how the certain equivalent gain $\tilde{g}$ varies as a function of the risk aversion coefficient γ. Obviously, the limit for $\tilde{g}$ as γ increases is $-E_{40} = -\$25,000$.

The certain equivalents $\tilde{v}_i$ for $\beta = 1$ are relative to $\tilde{v}_{40}$ which was set to $0. If we increase

Table 2: Stationary Optimal Policies $d_i = k^*$ and Certain Equivalents $\tilde{v}_i$ ($\beta = 0.97$)

	$\gamma = -0.0005$		$\gamma = -0.0001$		$\gamma = 0$		$\gamma = 0.00005$		$\gamma = 0.0001$		$\gamma = 0.0002$	
i	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	12	-350137	12	-363394	12	-392471	16	-434130	22	-510340	28	-622635
2	12	-362137	12	-375394	12	-404471	16	-446130	22	-522340	28	-634635
3	12	-373137	12	-386394	12	-415471	16	-457130	22	-533340	28	-645635
4	12	-391137	12	-404394	K	-433164	16	-475130	22	-551340	28	-663635
5	12	-398137	12	-411394	K	-439806	16	-482130	22	-558340	28	-670635
6	12	-405137	K	-417479	K	-446202	16	-489130	22	-565340	28	-677635
7	K	-411236	K	-423287	K	-452288	16	-496130	22	-572340	28	-684635
8	K	-416611	K	-428935	K	-458088	16	-509130	22	-585340	28	-697635
9	K	-421732	K	-434313	K	-463519	K	-514077	22	-591340	28	-703635
10	K	-426699	K	-439528	K	-468765	K	-517842	22	-596340	28	-708635
11	K	-431504	K	-444564	K	-473778	K	-521531	22	-601340	28	-713635
12	K	-436137	K	-449394	K	-478471	K	-524858	22	-608340	28	-720635
13	K	-440594	K	-454013	K	-482872	K	-527920	K	-610329	28	-725635
14	K	-444869	K	-458415	K	-487007	K	-530785	K	-612043	28	-729635
15	K	-448956	K	-462594	K	-490898	K	-533507	K	-613652	K	-732419
16	K	-452849	K	-466546	K	-494568	K	-536130	K	-615216	K	-733368
17	K	-456440	K	-470163	K	-497928	K	-538563	K	-616603	K	-734169
18	K	-459821	K	-473543	K	-501093	K	-540932	K	-617932	K	-734930
19	K	-462987	K	-476682	K	-504080	K	-543263	K	-619228	K	-735659
20	K	-465931	K	-479581	K	-506900	K	-545581	K	-620522	K	-736361
21	K	-468645	K	-482239	K	-509568	K	-547908	K	-621865	K	-737049
22	K	-471123	K	-484658	K	-512094	K	-550268	K	-623340	K	-737792
23	K	-473355	K	-486830	K	-514471	K	-552650	K	-624988	K	-738672
24	K	-475326	K	-488735	K	-516666	K	-554994	K	-626730	K	-739648
25	K	-476924	K	-490259	K	-518568	K	-557187	K	-628459	K	-740583
26	12	-478137	12	-491394	K	-520175	K	-559239	K	-630233	K	-741510
27	12	-479137	12	-492394	12	-521471	K	-561154	K	-632121	K	-742566
28	12	-480137	12	-493394	12	-522471	K	-562864	K	-634037	K	-743635
29	12	-481137	12	-494394	12	-523471	K	-564368	K	-636069	K	-744991
30	12	-481637	12	-494894	12	-523971	K	-565517	K	-638004	K	-746325
31	12	-482137	12	-495394	12	-524471	16	-566130	K	-639802	K	-747603
32	12	-482637	12	-495894	12	-524971	16	-566630	K	-641516	K	-748964
33	12	-483137	12	-496394	12	-525471	16	-567130	K	-643205	K	-750828
34	12	-484137	12	-497394	12	-526471	16	-568130	22	-644340	K	-752578
35	12	-484637	12	-497894	12	-526971	16	-568630	22	-644840	K	-754172
36	12	-485137	12	-498394	12	-527471	16	-569130	22	-645340	K	-755848
37	12	-485637	12	-498894	12	-527971	16	-569630	22	-645840	K	-757922
38	12	-486637	12	-499894	12	-528971	16	-570630	22	-646840	28	-759135
39	12	-487437	12	-500694	12	-529771	16	-571430	22	-647640	28	-759935
40	12	-488137	12	-501394	12	-530471	16	-572130	22	-648340	28	-760635

Table 3: Stationary Optimal Policies $d_i = k^*$ and Certain Equivalents $\tilde{v}_i$ ($\beta = 0.95$)

	$\gamma = -0.0005$		$\gamma = -0.0001$		$\gamma = 0$		$\gamma = 0.00005$		$\gamma = 0.0001$		$\gamma = 0.0002$	
i	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	16	-164588	16	-173659	16	-188742	16	-213809	22	-255747	28	-319733
2	16	-176588	16	-185659	16	-200742	16	-225809	22	-267747	28	-331733
3	16	-187588	16	-196659	16	-211742	16	-236809	22	-278747	28	-342733
4	16	-205588	16	-214659	16	-229742	16	-254809	22	-296747	28	-360733
5	16	-212588	16	-221659	16	-236742	16	-261809	22	-303747	28	-367733
6	16	-219588	16	-228659	16	-243742	16	-268809	22	-310747	28	-374733
7	16	-226588	16	-235659	16	-250742	16	-275809	22	-317747	28	-381733
8	K	-234245	K	-241721	K	-258058	16	-288809	22	-330747	28	-394733
9	K	-238647	K	-246390	K	-262786	16	-294809	22	-336747	28	-400733
10	K	-242960	K	-250968	K	-267397	K	-298542	22	-341747	28	-405733
11	K	-247179	K	-255438	K	-271847	K	-302014	22	-346747	28	-410733
12	K	-251293	K	-259776	K	-276045	K	-305154	22	-353747	28	-417733
13	K	-255297	K	-263973	K	-280014	K	-308046	K	-355892	28	-422733
14	K	-259185	K	-268024	K	-283774	K	-310755	K	-357587	28	-426733
15	K	-262951	K	-271921	K	-287344	K	-313328	K	-359180	K	-429503
16	K	-266588	K	-275659	K	-290742	K	-315809	K	-360729	K	-430452
17	K	-269985	K	-279127	K	-293874	K	-318103	K	-362101	K	-431254
18	K	-273233	K	-282418	K	-296856	K	-320336	K	-363416	K	-432016
19	K	-276326	K	-285529	K	-299699	K	-322535	K	-364698	K	-432746
20	K	-279254	K	-288456	K	-302418	K	-324727	K	-365975	K	-433449
21	K	-282009	K	-291197	K	-305025	K	-326938	K	-367298	K	-434139
22	K	-284583	K	-293752	K	-307533	K	-329197	K	-368747	K	-434884
23	K	-286962	K	-296112	K	-309937	K	-331501	K	-370360	K	-435765
24	K	-289130	K	-298255	K	-312204	K	-333792	K	-372063	K	-436742
25	K	-290969	K	-300067	K	-314226	K	-335961	K	-373749	K	-437678
26	K	-292461	K	-301541	K	-316004	K	-338025	K	-375475	K	-438607
27	16	-293588	16	-302659	K	-317531	K	-339996	K	-377313	K	-439663
28	16	-294588	16	-303659	16	-318742	K	-341808	K	-379182	K	-440733
29	16	-295588	16	-304659	16	-319742	K	-343476	K	-381176	K	-442084
30	16	-296088	16	-305159	16	-320242	K	-344858	K	-383088	K	-443414
31	16	-296588	16	-305659	16	-320742	16	-345809	K	-384885	K	-444687
32	16	-297088	16	-306159	16	-321242	16	-346309	K	-386631	K	-446042
33	16	-297588	16	-306659	16	-321742	16	-346809	K	-388412	K	-447893
34	16	-298588	16	-307659	16	-322742	16	-347809	K	-389733	K	-449630
35	16	-299088	16	-308159	16	-323242	16	-348309	22	-390247	K	-451211
36	16	-299588	16	-308659	16	-323742	16	-348809	22	-390747	K	-452878
37	16	-300088	16	-309159	16	-324242	16	-349309	22	-391247	K	-454960
38	16	-301088	16	-310159	16	-325242	16	-350309	22	-392247	28	-456233
39	16	-301888	16	-310959	16	-326042	16	-351109	22	-393047	28	-457033
40	16	-302588	16	-311659	16	-326742	16	-351809	22	-393747	28	-457733

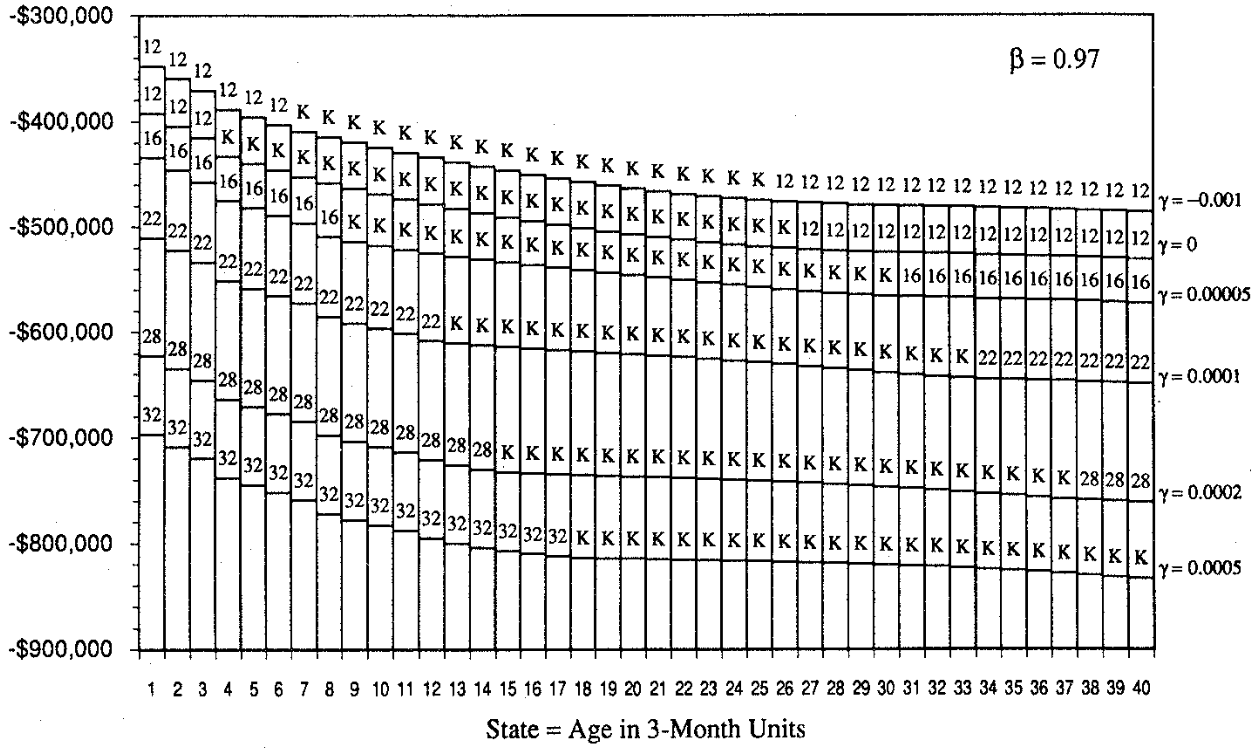

Figure 3: Stationary Optimal Policies $d_i = k^*$ and Certain Equivalents $\tilde{v}_i$ ($\beta = 0.97$)

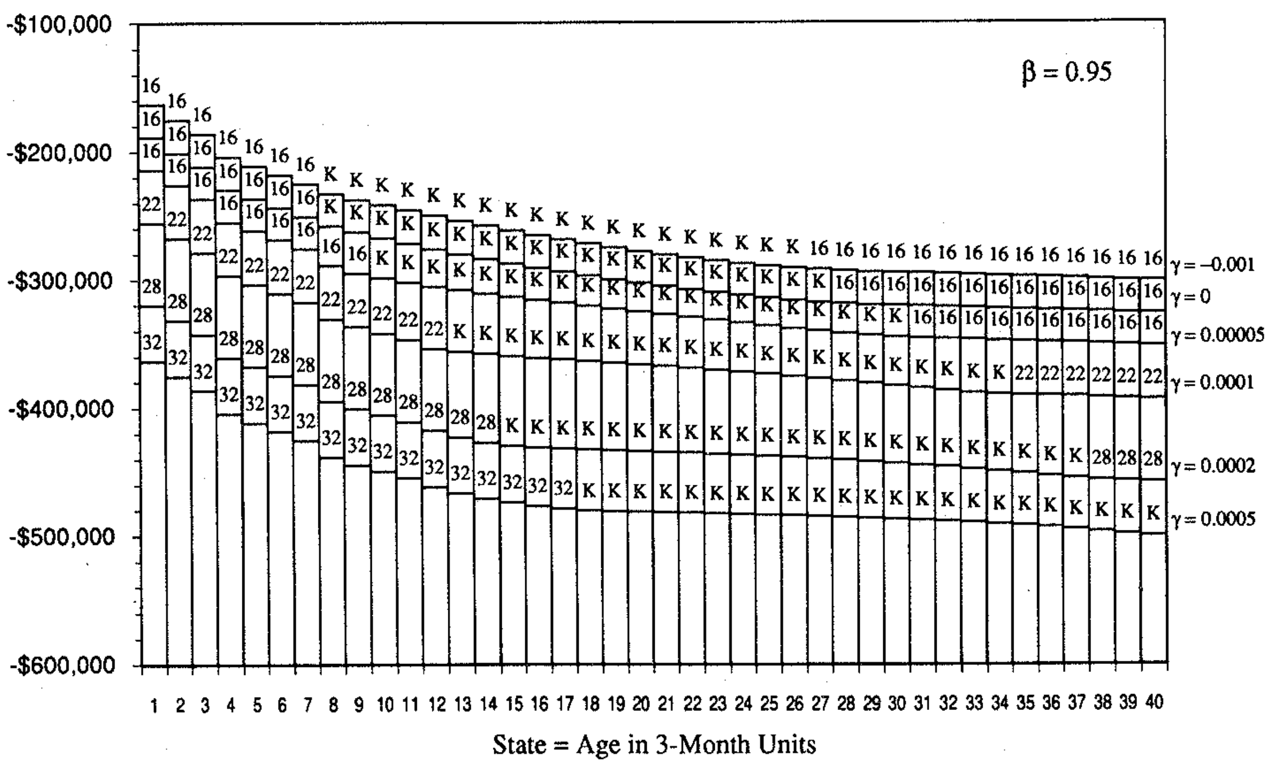

Figure 4: Stationary Optimal Policies $d_i = k^*$ and Certain Equivalents $\tilde{v}_i$ ($\beta = 0.95$)

Table 4: Certain Equiv. Gain $\tilde{g}$, Optim. Policies $d_i = k^*$ and Certain Equivalents $\tilde{v}_i$ ($\beta = 1.00$)

	$\gamma = -0.0005$		$\gamma = -0.0001$		$\gamma = 0$		$\gamma = 0.00005$		$\gamma = 0.0001$		$\gamma = 0.0002$	
	$\tilde{g} = -\$13862$		$\tilde{g} = -\$14252$		$\tilde{g} = -\$15095$		$\tilde{g} = -\$16551$		$\tilde{g} = -\$19098$		$\tilde{g} = -\$22723$	
i	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$	k^*	$\tilde{v}_i$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	12	138000	12	138000	12	138000	16	138000	21	138000	28	138000
2	12	126000	12	126000	12	126000	16	126000	21	126000	28	126000
3	12	115000	12	115000	K	116066	16	115000	21	115000	28	115000
4	K	103461	K	106565	K	107193	16	97000	21	97000	28	97000
5	K	95959	K	98546	K	98693	16	90000	21	90000	28	90000
6	K	88761	K	90847	K	90642	16	83000	21	83000	28	83000
7	K	81869	K	83477	K	83095	16	76000	21	76000	28	76000
8	K	75283	K	76439	K	76013	K	64031	21	63000	28	63000
9	K	69103	K	69833	K	69460	K	60019	21	57000	28	57000
10	K	63169	K	63549	K	63241	K	55910	21	52000	28	52000
11	K	57471	K	57597	K	57395	K	51886	21	47000	28	47000
12	K	52000	K	52000	K	52000	K	48265	K	40106	28	40000
13	K	46746	K	46760	K	47016	K	44936	K	38204	28	35000
14	K	41714	K	41864	K	42406	K	41819	K	36467	28	31000
15	K	36977	K	37312	K	38136	K	38856	K	34838	K	28192
16	K	32624	K	33102	K	34180	K	36000	K	33253	K	27244
17	K	28940	K	29330	K	30616	K	33340	K	31848	K	26445
18	K	25643	K	25894	K	27324	K	30751	K	30501	K	25685
19	K	22657	K	22788	K	24287	K	28209	K	29186	K	24957
20	K	19982	K	20007	K	21489	K	25696	K	27870	K	24257
21	K	17616	K	17545	K	18919	K	23196	K	26500	K	23572
22	K	15561	K	15395	K	16567	K	20697	K	24986	K	22832
23	K	13762	K	13560	K	14442	K	18220	K	23287	K	21955
24	K	12212	K	12058	K	12580	K	15830	K	21486	K	20980
25	12	11000	12	11000	K	11095	K	13646	K	19691	K	20047
26	12	10000	12	10000	12	10000	K	11664	K	17845	K	19123
27	12	9000	12	9000	12	9000	K	9892	K	15882	K	18067
28	12	8000	12	8000	12	8000	K	8404	K	13899	K	17000
29	12	7000	12	7000	12	7000	K	7215	K	11822	K	15638
30	12	6500	12	6500	12	6500	16	6500	K	9871	K	14298
31	12	6000	12	6000	12	6000	16	6000	K	8101	K	13015
32	12	5500	12	5500	12	5500	16	5500	K	6484	K	11646
33	12	5000	12	5000	12	5000	16	5000	21	5000	K	9762
34	12	4000	12	4000	12	4000	16	4000	21	4000	K	7993
35	12	3500	12	3500	12	3500	16	3500	21	3500	K	6382
36	12	3000	12	3000	12	3000	16	3000	21	3000	K	4690
37	12	2500	12	2500	12	2500	16	2500	21	2500	K	2628
38	12	1500	12	1500	12	1500	16	1500	21	1500	28	1500
39	12	700	12	700	12	700	16	700	21	700	28	700
40	12	0	12	0	12	0	16	0	21	0	28	0

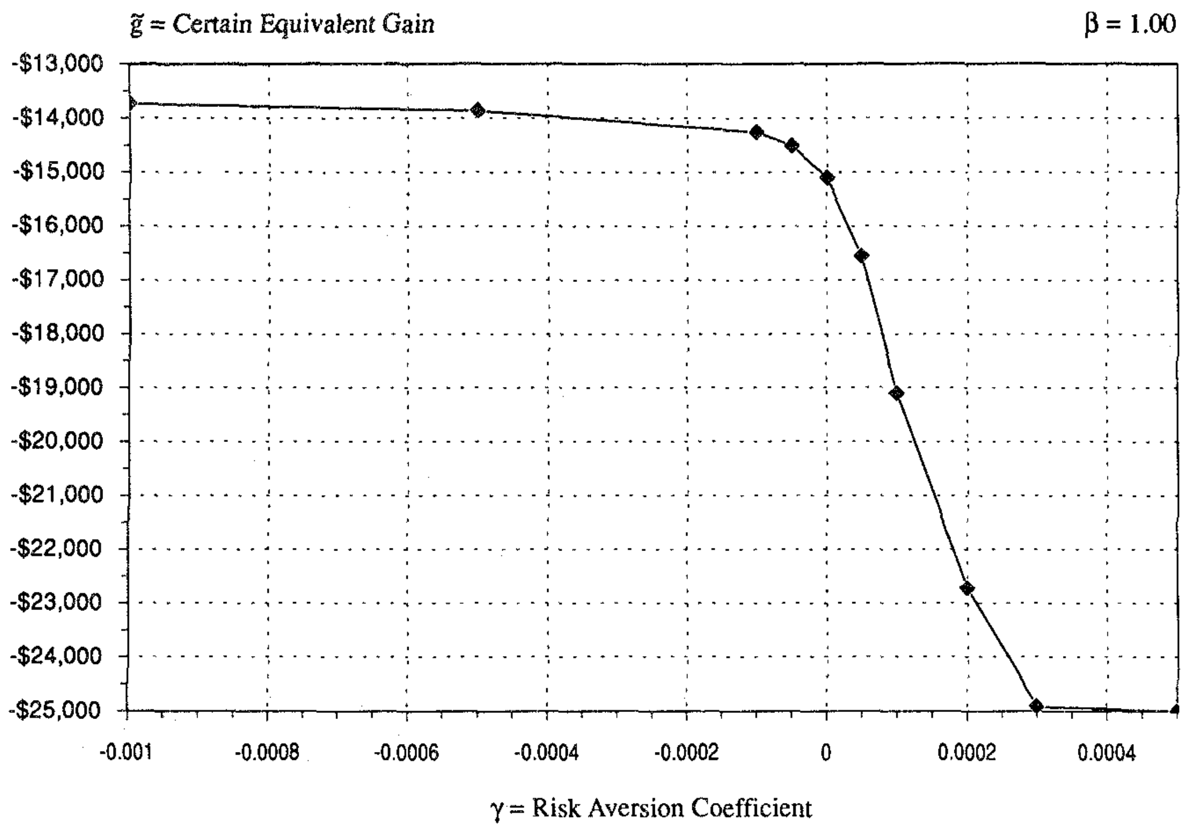

Figure 5: Impact of the Risk Aversion Coefficient γ on Certain Equivalent Gain $\tilde{g}$ ($\beta = 1.00$)

the values $\tilde{v}_i$ by $T_{40} = \$8,000$ (the trade-in value of a 10-year-old scraper), then the adjusted $\tilde{v}_i$ represent the dollar value of owning a scraper in state i to a company that follows the corresponding optimal policy. If the optimal policy for state i is to replace the existing scraper, the adjusted $\tilde{v}_i$ equals the trade-in value T_i. In other words, the market value of the scraper is more than the value that the company attributes to it, and as a result, the optimal decision is to sell the existing machine and replace it with another one.

To interpret the optimal policy of keeping the existing scraper let us consider a company that follows the optimal policy for $\gamma = 0.00005$, and which currently owns a 5-year-old scraper ($i = 20$). The value of this scraper to the company is \$25,696 more than the value of a 10-year-old (which has a trade-in value of \$8,000): $\$25,696 + \$8,000 = \$33,696$. This is more than its trade-in value $T_{20} = \$25,500$, but less than its purchase cost $C_{20} = \$36,000$. Thus, the company would prefer to keep a 5-year-old scraper rather than to trade it in, but it would not buy it as a replacement for another machine. This, of course, is true for all states i for which the optimal policy is to keep the current scraper. The only exception occurs for $i = 16$ which is the age of the replacement scraper under the optimal policy.

A 4-year-old scraper ($i = 16$) has a relative value $\tilde{v}_{16} = \$36,000$, which when adjusted by \$8,000 equals its purchase cost $C_{16} = \$44,000$. If we examine the difference between the adjusted value $\tilde{v}_i + \$8000$ and the purchase cost C_i for all states i we see that it is always negative except when $i = 16$ in which case it is \$0. This, of course, is the expected equilibrium condition. The value that the company attributes to the age (state) of the replacement scraper must be the maximum over all other states and it must equal its purchase price, otherwise, the company would not buy it. It is interesting to notice that the optimal age for the replacement scraper does not depend on the age of the machine being traded. This is as expected given the logic of the situation.

Similar interpretations apply to the optimal policies when future cash flows are discounted. To illustrate, let us consider a similar case, where the existing scraper is 5 years old ($i = 20$), and $\gamma = 0.0005$ and $\beta = 0.97$. The stationary certain equivalent under this scenario is $\tilde{v}_{20} = -\$545,581$. This amount is equivalent to a perpetuity of $-\$545,581/(1-0.97) = -\$16,367$, payable every 3 months beginning immediately. This perpetuity represents the equivalent cost of owning and operating a scraper under the optimal policy given the age of the existing equipment, and is analogous to the certain equivalent gain $\tilde{g} = \$16,551$ which applies for $\beta = 1$ (no discounting). Notice, however, that because future costs are discounted, the value of the perpetuity depends on the age of the existing scraper, whereas the certain equivalent gain $\tilde{g}$ is the same for all states (when future costs are not discounted, it makes no difference when the transitions implied by the stationary decision policy occur).

The amount $\tilde{v}_{20} - \tilde{v}_{40} + T_{40} = -\$545,581 + \$572,130 + \$8,000 = \$34,549$ represents the adjusted relative certain equivalent for a 5-year-old scraper. This represents the scraper's present value to the company, and since the optimal policy is to keep it, this value is between $T_{20} = \$25,500$ and $C_{20} = \$36,000$. For all states i for which the optimal policy is to replace the existing equipment, the adjusted relative certain equivalent equals the trade-in value T_i. For state $i = 16$, the age of the replacement scraper, the adjusted relative certain equivalent is $-\$536,130 + \$572,130 + \$8,000 = \$44,000$ which equals the replacement cost C_{16}. Thus, the proper interpretation of the results with and without discounting is exactly the same.

A comparison of Fig. 3 and 4, and Tables 2, 3, 4, shows that the optimal decision policies and the especially the values of the stationary certain equivalents depend on the adopted discount rate and the degree of risk aversion exhibited by the decision maker. For a given γ, the effect of a higher discount rate is to decrease the number of states for which the existing scraper is worth keeping and to increase the optimal age for the replacement scraper. Thus, higher discount rates produce the same qualitative effect as higher degrees of risk aversion. The stationary certain equivalents depend primarily on the discount rate whereas the optimal replacement policy is more sensitive to the risk aversion coefficient. This phenomenon is apparent in almost all management problems.

Conclusion

The model presented in this article is a powerful tool for the analysis of many construction management decision problems that exhibit dynamic behavior over time. Its ability to encompass risk aversion and discounting of future cash flows makes it particularly useful for financial problems at the strategic and operations management levels. Obviously, it can also be used to tackle simplified dynamic decision problems where discounting or risk aversion are not particularly important.

It is interesting to notice that the same model can be applied to problems whose stages extend over space and not over time. An example would be the determination of an optimal excavation and support sequence in tunneling. By utilizing a probabilistic geologic prediction model that reflects all available information about the project geology (Ioannou 1984, 1987), it is possible to subdivide the project's horizontal alignment into segments (stages), and to construct the space-varying Markov transition probability matrix $\mathbf{P}(n)$ from each geologic state i in segment n to any other possible state j in the next segment $n-1$. The length of each segment could be as short as the length of each round. The decision alternatives in each stage would be the particular excavation and support methods that would be used in the corresponding segment. Obviously, the choice of a particular excavation-support combination affects the cost of constructing the segment (the reward matrix $\mathbf{R}(n)$) but does not influence the transition probabilities since those are a function of geologic information. This problem also requires that the reward structure of the model be modified to account for the fixed costs of switching to new construction methods.

Tunneling decisions in practice are characterized by a considerable amount of conservatism that depends on the amount of available geologic information and the amount of risk born by each party to the contract (Ioannou 1984). Even though a stochastic dynamic programming approach to the tunneling problem has already been attempted (Kim 1984) it does not capture the effects of risk aversion since it is based on the minimization of expected cost (i.e. risk neutral decision making). The proposed formulation can provide an excellent vehicle for capturing the effects of conservatism and conducting sensitivity analysis of the project cost for different values of the risk aversion coefficient γ.

Acknowledgments

The dynamic decision model presented in this article is based on research partially supported by the National Science Foundation under Grant No. 85–04902. This support and the encouragement given by Dr. Gifford Albright, Director, Structures and Building Systems Program, are gratefully acknowledged.

Appendix I. References

Howard, R.A. (1960). *Dynamic programming and Markov processes*, The MIT Press, Cambridge, Mass.

Howard, R.A., and Matheson J.E. (1972). "Risk-sensitive Markov decision processes," *Management Science*, Vol. 18, No. 7, March, 1972.

Howard, R.A. (1977). "Risk preference," *Readings in Decision Analysis*, Howard, R.A., and Matheson, J.E., Editors, Stanford Research Institute, Menlo Park, Calif.

Howard, R.A. (1979). *Dynamic probabilistic systems*, Vols. I and II, John Wiley and Sons, Inc., New York, N.Y.

Ioannou, P.G. (1984). "The economic value of geologic exploration as a risk reduction strategy in underground construction," thesis presented to the Massachusetts Institute of Technology, at Cambridge, Mass., in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Ioannou, P.G. (1987). "Geologic prediction model for tunneling," *Journal of Construction Engineering and Management*, Vol. 113, No.4, Dec. 1987.

Kim, Y.W. (1984). "Decision support framework in adaptable tunneling," thesis presented to the Massachusetts Institute of Technology, at Cambridge, Mass., in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Luce, R.D., and Raiffa, H. (1957). *Games and decisions*, John Wiley and Sons, Inc., New York, N.Y.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty*, Addison Wesley, Reading, Mass.

Appendix II. Notation

The following symbols are used in this paper:

$\alpha(n)$ = Discount factor for discounting the reward $r_{ij}(n)$ to its present value at stage n.

$\beta(n)$ = Discount factor for discounting the certain equivalent $\tilde{v}_j(n-1)$ (which is expressed in present value terms at stage $n-1$) to its present value at stage n.

γ = Risk aversion coefficient.

$(\text{sgn}\ \gamma)$ = Algebraic sign of the risk aversion coefficient γ.

$d_i(n)$ = Optimum decision alternative k^* for state i in stage n.

$\tilde{g}$ = Certain equivalent gain from any state i in stage n to the same state i in stage $n+1$, for large n.

k = The kth decision alternative in state i and stage n.

k^* = The optimal decision alternative in state i and stage n.

λ = Constant ratio $u_i(n+1)/u_i(n)$ for large n and $\beta = 1$ (no discounting).

N = Number of discrete states of the Markov process.

n = The nth stage as measured from the end of the Markov process (the last and latest stage is stage 0). The number of transitions remaining until the Markov process ends (reaches stage 0).

$p_{ij}(n)$ = Markov transition probability for the transition from state i in stage n to state j in stage $n-1$.

$p_{ij}^{k}(n)$ = Markov transition probability for the transition from state i in stage n to state j in stage $n-1$, given decision alternative k.

$q_{ij}(n)$ = Product of the Markov transition probability $p_{ij}(n)$ times the associated utility of the discounted reward $\alpha(n)r_{ij}(n)$.

$\rho(n)$ = The discount rate for the time period between stage n and stage $n-1$.

$r_{ij}(n)$ = Reward associated with the transition from state i in stage n to state j in stage $n-1$.

$r_{ij}^{k}(n)$ = Reward associated with the transition from state i in stage n to state j in stage $n-1$, given decision alternative k.

v = The value outcomes of a lottery.

$\tilde{v}$ = Certain equivalent of a lottery.

v_p = Risk premium of a lottery.

$\tilde{v}_i(n)$ = Certain equivalent for state i in stage n.

$\tilde{v}_i$ = Stationary certain equivalent for state i.

$u(v)$ = The utility of the value outcome v.

$u(\tilde{v}_i(n))$ = Expected utility (= utility of the certain equivalent $\tilde{v}_i(n)$) of state i in stage n.

$u_i(n)$ = Expected utility of state i in stage n.

$u_i^*(n)$ = Maximum expected utility of state i in stage n.