Research Program in Finance
WORKING PAPER SERIES

WORKING PAPER NO. 104

Market Value Maximization and Markov Dynamic Programming

BY

RICHARD C. GRINOLD

Research Program in Finance Working Papers are preliminary in nature; their purpose is to stimulate discussion and comment. Therefore, they should not be cited or quoted in any publication without the permission of the author. Single copies of a paper may be requested from the Institute of Business and Economic Research.
RESEARCH PROGRAM IN FINANCE AT THE
WALTER A. HAAS SCHOOL OF BUSINESS,
UNIVERSITY OF CALIFORNIA, BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH
Carl Shapiro, Director

The Institute of Business and Economic Research is an organized research unit at the University of California, Berkeley, whose mission is to promote research by faculty and graduate students in the fields of business and economics. The Institute carries out its mission by organizing programs and activities that enrich the research environment, administering extramural research awards, publishing working papers, and making direct grants for research.
MARKET VALUE MAXIMIZATION AND
MARKOV DYNAMIC PROGRAMMING

by

Richard C. Grinold

November 1980
ABSTRACT

This paper shows how an operational method for solving dynamic programs can be used, in some cases, to solve the problem of maximizing a firm's market value. The problem is formulated as a Markov decision problem that can be solved via linear programming. The paper shows how to calculate (or estimate) the state-contingent prices that are used to value the firm. In addition, the paper points out how states can be aggregated to make the solution technique more practical. The paper's final section contains a specific example.
1. Introduction

This paper shows how to harness an operational technique, Markov dynamic programming, to the theoretical problem of choosing a firm's decisions in order to maximize the market value of its future cash flows. There are several key steps that allow us to do this. The first is the use of modern financial theory to calculate state contingent claims prices that can be used to value alternative cash flows. The second step is to structure the state space of the problem in a sensible way in order to separate the macroeconomic state variables that determine value from the microeconomic variables that characterize the state and actions available to the firm. The third step is to identify a value law of motion that allows us to relate values in one time period to subsequent values that are conditioned on the uncertain evolution of the system. In this way we can obtain a modified Markov decision problem with the values playing the role of probabilities. Thus the result blends the theory of Markov dynamic programming with a theory of market valuation under uncertainty.

The theory of Markov's decision processes, as developed by Howard [11], offers a context in which one can study the evolution of a system under uncertainty, and where that uncertainty is to some extent under the control of the decision maker.

The theory can also be used to examine the effect of risk aversion in a decision process. Suppose, for example, a transition
from state \( y \) to state \( y' \) after selection of decision \( a \) leads to consumption \( c(y,a) \). Suppose further, that one starts with a utility for consumption \( U(c) \), and that the objective is to maximize

\[
E \left[ \sum_{t=1}^{\infty} \rho^t U[c(t)] \right] \quad \text{where} \quad 0 < \rho < 1 \quad \text{discounts the utility of future consumption, and} \quad c(t) \text{ is the uncertain level of consumption in period } t. \quad \text{In this case, we can set} \quad r(y,a) = U[c(y,a)] \quad \text{and maximize the expected value of the rewards, which is actually the expected utility of future consumption.}
\]

Howard and Matheson [11] and Denardo and Rothblum [6] consider a slightly different problem in which the consumer obtains the cumulative bundle earned at either the end of \( n \) steps or, more generally, when the system enters a set of stopping states. If we make the special assumption that the utility of the eventual consumption of \( c \) in period \( t \) is \( U(c,t) = \rho^t [a - be^{-\lambda c}] \), where \( \lambda > 0 \). Then, the problem can be changed into an equivalent problem of maximizing expected returns with altered rewards and altered transition probabilities.

When one make the transition from the decision problems of individuals to those of corporations, the problem of criteria is raised anew. One could maintain that corporations, like individuals, have a utility function for the cash they earn. However, this is not consistent with the widely held view that corporations certainly should (and actually might) act to maximize the market value of their firm; see Leland [13].
The valuation theory is based on the concept of complete markets and the state contingent claims price; i.e., the price today of a promise of $1.00 at some future time conditional on the system's evolution over that interval. This theory originated with Arrow [1], Debreu [4], and has seen more recent application by Kraus and Litzenberger [12], Rubinstein [16, 17, 18] Banz and Miller [2], and Breeden and Litzenberger [3]. This theory of market valuation is presented in section 2. It is based on the work of Arrow, and Debreu, and the specific models of Kraus and Litzenberger [12] and Rubinstein [16].

In section 3 we present a prototype decision model that details model structure and indicates the type of assumptions one must make in order to render the theory solvable by Markov dynamic programming. Section 4 looks at the problem in another way. We construct a dynamic decision model over an infinite horizon and then show how that problem could be solved using a linear program. In fact, one can solve the dynamic program via linear programming, and it turns out that these two linear programs are duals. This result lets us interpret the sensitivity results we will obtain by solving the Markov dynamic program.

In section 5 we show how to aggregate states. This procedure can be used to make the method more practical. We can aggregate states of the economy that may differ in terms of valuation but have the same impact on the state of and choice available to the firm.
Section 6 presents an example. The problem is to determine the optimal policy for term lending in banks. The model in section 6 shows how an optimal policy can be calculated using the market value criterion.

2. Market Valuation

In this section we present a model for determining the present market value of uncertain future cash flows. The model is based on the notion of a contingent claim price: i.e., the price today of a promise to pay $1.00 at a future time should some specified contingency be satisfied at that time.

There are several possible ways to do this. We have selected the logarithmic, or Kraus-Litzenberger [12], Rubinstein [16] model due to its simplicity and generality.

We divide time into periods of equal length. The interval of time after time \( t-1 \) and up to time \( t \) is called period \( t \) and is denoted \( [t-1, t] \). The economic state of the economy at time \( t \) is denoted \( y(t) \), where \( y(t) \) is an element of a state space \( Y \). The changes in state are governed by a stationary transaction function. If the state is \( y(t-1) \) at time \( t-1 \), then it will be \( y(t) \) at time \( t \) with probability \( \pi[y(t)|y(t-1)] \).

A history of the economy up to time \( t \) is denoted by \( \omega(t) \) and describes the infinite sequence of economic states that have obtained; thus \( \omega(t) = \{y(t), y(t-1), \ldots\} \). We start with an initial history \( \omega(0) \), and let \( \pi[\omega(t)|\omega(0)] \) be the probability of having history \( \omega(t) \) at time \( t \), given \( \omega(0) \) at time zero.
We wish to calculate \( q[\omega(t)|\omega(0)] \), the market value at time zero of a promise of $1.00 at time \( t \) if history \( \omega(t) \) obtains. The function \( q[\omega(t)|\omega(0)] \) is the state contingent claims price. We will show, following Krauz and Litzenberger [12], and Rubinstein [16], that it is possible to calculate an explicit formula for \( q[\omega(t)|\omega(0)] \) in terms of observable market values. To do this we need some additional notation.

\[
(2:1) \quad (i) \quad R_m[\omega(t)|\omega(0)] \text{ is the cumulative return on the aggregate (market) portfolio of all assets when history } \omega(t) \text{ obtains. A dollar invested in the the aggregate portfolio at time zero will be worth } R_m[\omega(t)|\omega(0)] \text{ at time } t \text{ if } \omega(t) \text{ occurs.}
\]

\[
(ii) \quad D[t|\omega(0)] = \sum_{\omega(t)} q[\omega(t)|\omega(0)] \text{ is the value at time zero of a promise of $1.00 at time } t \text{ regardless of the value of the history } \omega(t) \text{ at time } t.
\]

\[
(iii) \quad R_f[t|\omega(0)] = 1/D[t|\omega(0)] \text{ is one plus the } t \text{ period rate of interest.}
\]

\[
(iv) \quad K[t|\omega(0)] = \left\{ \frac{R_f[t|\omega(0)]}{R_m[t|\omega(0)]} \right\}^{\omega(0)} \text{ is a normalizing constant, depending only on } t.
\]

With this notation we can write
(2:2) \[ q[\omega(t)|\omega(0)] = \frac{\pi[\omega(t)|\omega(0)]}{R_m[\omega(t)|\omega(0)]K[t|\omega(0)]} \]

This form is not suitable for our purpose. We would like to calculate \( q[y(t)|y(t-1)] \); the market value, at time \( t-1 \) given Markov state \( y(t-1) \), of a promise of \$1.00 at time \( t \) if economic state \( y(t) \) should occur. That is, we want our values to be state dependent, but history independent. Fortunately, we can make this calculation; although a few additional assumptions are necessary.

First, let us write \( \omega(t) \) as \( \{y(t), \omega(t-1)\} \); i.e., the history at time \( t \) consists of the history up to time \( t-1 \) along with the economic state at time \( t \). By a simple arbitrage argument (based on buying and selling contingent claims in complete markets), we can establish that

(2:3) \[ q[\omega(t)|\omega(0)] = q[\omega(t)|\omega(t-1)] \cdot q[\omega(t-1)|\omega(0)] \]

or using (2:2)

(2:4) \[ q[\omega(t)|\omega(t-1)] = \frac{\pi[\omega(t)|\omega(0)]}{\pi[\omega(t)|\omega(0)]} \cdot \frac{R_m[\omega(t-1)|\omega(0)]}{R_m[\omega(t)|\omega(0)]} \cdot \frac{K[t|\omega(0)]}{K[t-1|\omega(0)]} \]

Since the underlying stochastic process is assumed to be Markovian, the first ratio on the right of (2:4) is equal to \( \pi[y(t)|y(t-1)] \). If we assume in addition the return on aggregate wealth is Markovian; i.e., increase in market wealth during period \( t \) depends only on \( y(t-1) \) and \( y(t) \). That increase is denoted \( r_m[y(t)|y(t-1)] \). Thus the second term on the right of (2:4) is
equal to one divided by \( 1 + r_m[y(t)\mid y(t-1)] \). We have to this point

\[
q[y(t)\mid y(t-1)] = \frac{\pi[y(t)\mid y(t-1)]}{1 + r_m[y(t)\mid y(t-1)]} \cdot \frac{k[t\mid \omega(0)]}{k[t-1\mid \omega(0)]}.
\] (2:5)

However, the sum of the right-hand side of (2:5) over all \( y(t) \) in \( \mathcal{Y} \) must give the value at time \( t-1 \) given history \( \omega(t-1) \) of a promise of $1.00 at time \( t \). Notice this sum will only depend on \( y(t-1) \) and \( t \); it is the inverse of \( 1 + r_f[y(t-1)] \), one plus the one-period risk-free rate of interest should \( y(t-1) \) be the economic state at time \( t-1 \). Let \( k[t\mid \omega(0)] \) be the third ratio on the right of (2:4), we have

\[
\frac{1}{1 + r_f[y(t-1)]} = \mathbb{E}\left\{ \frac{1}{1 + r_m[y(t)\mid y(t-1)]} \mid y(t-1) \right\} k[t\mid \omega(0)].
\] (2:6)

Our second assumption is that the one period risk free rate of interest depends on the current economic state \( y(t) \in \mathcal{Y} \) but is independent of time. Thus, the second term in (2:6) must be independent of \( t \), since we have already assumed that the first term in (2:6) is independent of \( t \). Thus we can write

\[
q[y(t)\mid y(t-1)] = \frac{\pi[y(t)\mid y(t-1)]k[y(t-1)]}{1 + r_m[y(t)\mid y(t-1)]}.
\] (2:7)

where

\[
\frac{1}{k[y(t-1)]} = \mathbb{E}\left\{ \frac{1 + r_f[y(t-1)]}{1 + r_m[y(t-1)]} \mid y(t-1) \right\}.
\]
Finally, we are assured that \( q[y(t)|y(t-1)] \) is nonnegative, and with \( r_f[y(t-1)] > 0 \), we have

\[
\sum_{y(t)} q[y(t)|y(t-1)] = \frac{1}{1 + r_f[y(t-1)]} < 1.
\]

We have presented one model for calculating market values from economic variables. Other exist; see Rubinstein [16, 17], Banz and Miller [2], and Breeden and Litzenberger [3]. The wide variety of models does not make selection of a particular model as ad hoc as it might appear. In Grauer [7] and Roll [15], we find that the particular specification of the model is not important as long as the time between decision periods is relatively short, i.e., one year or less.

The next section demonstrates use of these valuations in a prototype model of firm decision making.

3. Prototype Model

Consider a firm and suppose that the condition of the firm at any time \( t \) is described by an element \( x \) in the set \( X \), while the condition of the overall economy is described by an element \( y \) in the set \( Y \). The pair of descriptors \( (x,y) \in S = X \times Y \) will form the state space of our decision process.

In state \( (x,y) \) the firm can choose an action \( a \) from the set \( A(x,y) \). As a result of choosing action \( A \) the system will move to state \( (x',y') \) and the firm receives a cash flow \( \tau(x',y'|x,y,a) \) with probability \( \pi(x',y'|x,y,a) \).
The probabilities $\pi(x', y'| x, y, a)$ are calculated from more basic information. We presume to know $\pi(x'| x, y, y', a)$; this is the probability that the firm moves to state $x'$, given we start in $(x, y)$, choose action $a$, and that the economy moves to $y'$. We also presume to know $\pi(y'| y)$; this is the probability the economy moves from state $y$ to $y'$.

We also make the quite sensible assumption that the initial state of the firm $x$, and the actions the firm selects ($a$), will not influence the evolution of the economy; i.e.,

\begin{equation}
\pi(y'| y) = \pi(y'| x, y, a)
\end{equation}

This allows us to prove

**Proposition 1:**

\begin{equation}
\pi(x', y'| x, y, a) = \pi(x'| x, y, y', a) \cdot \pi(y'| y).
\end{equation}

**Proof:** Define the following events:

- $A \equiv \{x(t+1) = x'\}$, $C = \{x(t) = x, a(t) = a\}$
- $B \equiv \{y(t+1) = y'\}$, $D = \{y(t) = y\}$.

We wish to calculate

\[ \pi[A \cap B | C \cap D] = \frac{\pi[A \cap B \cap C \cap D]}{\pi[C \cap D]} \]

However,

\[ \pi[A | B \cap C \cap D] = \frac{\pi[A \cap B \cap C \cap D]}{\pi[B \cap C \cap D]} \]

and, by assumption,

\[ \pi[B | D] = \pi[B | C \cap D] = \frac{\pi[B \cap D \cap C]}{\pi[C \cap D]} \].
When these expressions are combined we obtain
\[ \pi(A \cap B \cap C \cap D) = \pi(A \cap B \cap C \cap D) \cdot \pi(B \mid D) \]

Actually, the knowledge of any two of the probabilities in (3:2) will allow us to calculate the third.

Recall, from section 2, that we know the present value of $1.00 should y' occur in the next period is given by \( q[y'|y] \). We now expand this notion to say that the present value of $1.00 should \((x', y')\) occur is

\[ q[x', y'|x, y, a] = q[y'|y] \pi[x'|y', x, y, a] \]

If we sum this expression over all possible values of \( x' \), we see that

\[ \sum_{x'} q[x', y'|x, y, a] = q[y'|y] . \]

Equation (3:3) simply says that the economy as a whole is not sensitive to the refinements in the state space caused by considering the explicit condition of one firm. Another way to say this is to say that the market is averse to the risk of changes in general economic conditions, but that it is risk neutral regarding the specific risks of firms.

Now consider the problem of operating the firm. We can proceed in two ways. The more theoretical and less realistic way is to assume the firm has an \( T \) period existence, plan over those \( T \) periods, and calculate \( V^T(x, y) \), the market value of the \( T \) period
firm with initial state \((x,y)\). Then we look at the limiting behavior of \(V^T(x,y)\). The second, more realistic and heuristic, is to assume there exists a function \(V(x,y)\) that gives the present market value of the firm as a function of the rate \((x,y)\). Then we indicate that \(V(x,y)\) should solve a dynamic programming functional equation. We shall follow the second tack here. In section 5 we follow the other approach.

Let \(V(x,y)\) be the market value of the firm given that we start in state \((x,y)\) and follow an optimal policy. Then for any decision we must have \(a \in A(X,Y)\).

\[
\begin{align*}
(3:5) \quad V(x,y) & \geq \sum_x \sum_y \left[ r(x',y',x,y,a) + V(x',y') q[x',y'|x,y,a] \right] \\
\end{align*}
\]

The expression on the right-hand side is just the present market value of the cash flows and residual firm value, \(V(x',y')\), if action \(a\) is taken.

Common sense indicates that if we maximize the right-hand side over all values of \(a\) then we should obtain equality: i.e., \(V(x,y)\) satisfies the functional equation.

\[
(3:6) \quad v(x,y) = \max_{a \in A(x,y)} \left\{ r(x,y,a) + \sum_x \sum_y V(x',y') q[x',y'|x,y,a] \right\}
\]

where

\[
r(x,y,a) = \sum_x \sum_y r(x',y'|x,y,a) q[x',y'|x,y,a]
\]
is the market value of the reward obtained by taking action \( a \) in state \((x,y)\).

For our model, the values \( q[x',y'|x,y,a] \) are nonnegative (see (3:4) and (2:8)), and they sum to a value less than 1. Thus with finite state space \( S = \times Y \), and finite action space \( A(x,y) \), we are assured (see Denardo [5]) that the functional equation above will have a unique solution and we can use all the standard policy iteration, value iteration, and linear programming techniques to find that solution. Moreover, the optimal policy, i.e., the rule that chooses an optimal \( a(x,y) \in A(a,y) \) for each \( (x,y) \in S \), will be stationary and independent of earlier values of \( x \) and \( y \).

In more complicated cases, where say \( X \), or \( Y \), or \( A(x,y) \) is infinite, we still retain the Markov decision process structure, so we can appeal to any of the special results that have been obtained; for more general classes of Markov decision problems, or as in section 5, we can aggregate economic states to make a more easily solved problem.

Section 4 shows how a solution of (3:6) can be obtained via linear programming, section 5 shows how to aggregate the state space, and the final section presents a specific example.

4. Linear Programming and Duality

As mentioned in the previous section, problem (3:6) can be solved as a linear program (problem (4:2) below) in case both the state space \( X \times Y \) and action space \( A(x,y) \) are finite. There is,
in general, a duality among decision problems in which the primal (action) problem is concerned with decisions, resources, and rewards and a dual (valuation) problem that is concerned with valuation of resources. Problem (3:6) deals very much with values. This section will uncover the primal action problems (4:14) and (4:15) below that leads to (4:2) as a dual.

Let $A$ be a matrix with one row for each state $(x,y)$ and one column for each state action combination $(x,y,a)$ with $a \in A(x,y)$. Thus, if $y$ takes on 3 possible values, $x$ 4 values, and there are 5 actions possible in each state, then $A$ will have 12 rows ($3 \times 4$) and 60 columns ($3 \times 5 \times 4$). Each element of $A$ is either 0 or 1; element $A(x,y; x',y',a')$ is equal to 1 if $(x',y') = (x,y)$ and $a' \in A(x,y)$; it is equal to zero otherwise. Let $Q$ be a matrix with the same dimensions as $A$; and elements $q(x,y|x',y',a')$, let $v$ be the vector $V(x,y)$, $r$ the vector $r(x,y,a)$, and $\pi(0)$. The vector $\pi(x,y,0)$; $\pi(x,y,0)$ is the probability of being in state $(x,y)$ at time 0.

Then we can find the solution of (3:6) by solving

\[(4:1) \quad \text{minimize} \quad v\pi(0) \]
\[\text{subject to} \quad v(A-Q) \geq r.\]

and, indeed, there will exist an optimal solution $v$ that is independent of $\pi(0)$ and we can find it by artificially making $\pi(0)$ strictly positive. The dual of (4:1) is
We summarize our results with proposition 2:

**Proposition 2:**

The solution of the functional equation (3:6) can be found by solving (4:2).

**Proof:**

The matrix $A-Q$ is a Leontief substitution matrix: each column has exactly one positive element, and $(A-Q)e$ (where $e$ is a vector of ones) is strictly positive; see Vienott [19]. If $\pi(0)$ is strictly positive, then we can be assured that the optimal basis in (4:2) will have one positive element in each row and column; thus it will be a square Leontief matrix with a nonnegative inverse.

This means the basis will remain optimal for any non-negative $\pi(0)$. It also means that the optimal dual solution will solve (3:6), since for each $(x,y)$ there will be an $a \in A(x,y)$ such that column $(x,y,a)$ is in the optimal basis; call this $a^*(x,y)$. For $a^*(x,y)$ we have

\begin{equation}
(4:3) \quad V(x,y) = r(x,y, a^*(x,y)) + \sum_{x'} \sum_{y'} V(x',y')q[x',y'|x,y, a^*(x,y)]
\end{equation}

and for any $a \in A(x,y)$

\begin{equation}
(4:4) \quad V(x,y) \geq r(x,y,a) + \sum_{x'} \sum_{y'} V(x',y')q[x',y'|x,y,a]
\end{equation}

When (4:3) and (4:4) are combined, we obtain (3:6).
What is more interesting than this is an attempt to produce problem (4:2) directly; i.e., to devise a decision problem that can be solved with (4:2) and whose dual (4:1) will naturally produce the market value \( V(x,y) \) of starting in state \((x,y)\).

We can define a policy as follows: for each \( t \) and each \((x,y)\), let \( \pi(a|x,y,t) \) be the conditional probability of using action \( a \in A(x,y) \) given that we are in state \((x,y)\) at time \( t \). The policy is deterministic if \( \pi(a|x,y,t) \) equals 0 or 1, and it is called stationary if \( \pi(a|x,y,t) \) is independent of \( t \). Given a strategy, we can calculate \( \pi(x,y,t) \), the probability of being in state \((x,y)\) at time \( t \). This will, of course, depend on the initial conditions \( \pi(x,y,0) \) and the policy \( \pi(a|x,y,z) \) for \( z = 0,1,\ldots,t-1 \). The joint probability of using action \( a \) in state \((x,y)\) at time \( t \) is given by

\[
(4:5) \quad \pi(x,y,a,t) = \pi[a|x,y,t] \pi[x,y,t]
\]

The \( \pi(x,y,a,t) \) must satisfy the following constraints.

\[
(4:6) \quad \sum_{a \in A(x,y)} \pi(x,y,a,t) = \pi(x,y,t) \quad \text{for } t = 0,1,2,\ldots
\]

and

\[
(4:7) \quad \pi(x,y,t) = \sum_{x'} \sum_{y'} \sum_{a' \in A(x',y')} \pi(x,y|x',y',a') \pi(x',y',a',t-1) \quad \text{for } t = 1,2,\ldots
\]

with \( \pi(x,y,a,t) \geq 0 \) for all \((x,y,a,t)\).
Now we shall show that there are conditional values, \( q(x,y,a,t) \) being the value at time zero of \$1.00 at time \( t \) should we be using action \( a \) in state \((x,y)\), that satisfy a system of equations similar to (4:6) and (4:7).

Let \( q[x,y,a,t] \) be the market value at time zero of a promise of \$1.00 should we land in state \((x,y)\) at time \( t \) and use action \( a \).

The market value is controlled by the \( y \) variable. Thus if \( q(y,t) \) is the market value of a promise of \$1.00 if state \( y \) obtains at time \( t \), then

\[
(4:8) \quad q(x,y,a,t) = q(y,t) \pi(x,a|y,t)
\]

and by the laws of conditional probability

\[
(4:9) \quad q(x,y,a,t) = \frac{q(y,t)}{\pi(y,t)} \pi(x,y,a,t)
\]

**Proposition 3:**

The values of \( q(x,y,a,t) \) satisfy the equations

\[
(4:10) \quad \sum_{A(x,y)} q(s,y,a,t) = q(x,y,t)
\]

\[\quad t = 0,1,2,\ldots\]

and

\[
(4:11) \quad q(x,y,t) = \sum_{x'} \sum_{y'} \sum_{A(x',y')} q(x,y|x',y',a') q(x',y',a',t-1).
\]

where \( q(x,y,0) = \pi(x,y,0) \): i.e., the probability of being in state \((x,y)\) at time zero.
Proof:

Let \( q(x,y,x',y',t-1) \) be the value, at time \( t-1 \) given state \((x',y')\), of $1.00 at time \( t \) should state \((x,y)\) obtain. This implies that

\[
q(x,y,t) = \sum_{x} \sum_{y} q(x,y|x',y',t-1)q(x',y',t-1),
\]

since in complete markets we could obtain a dollar at \( x,y,t \) in two ways: first by buying a certificate for \( q(x,y,t) \) and second by buying \( q(x|y|x',y',t-1) \) certificates for each possible \((x',y')\) at time \( t-1 \) and then whatever \( x',y' \) obtains, investing the proceeds \( q(x,y|x',y',t-1) \) in a certificate that will be worth $1.00 should \((x,y)\) obtain at time \( t \). Thus both sides of equation (4:12) give us $1.00 at time \( t \) if \((x,y)\) obtains at time \( t \).

They must be equal; if not we could go short on one scheme and long on the other and produce an arbitrage product.

It is obvious that

\[
q(x,y,t) = \sum_{a(x,y)} q(x,y,a,t).
\]

We must show three slightly more subtle relations.

(4:13) (i) \( q(x,y|x',y',t-1) = q(y'|y) \frac{\pi(x,y|x',y',t-1)}{\pi(y|y')} \)

(ii) \( q(x,y|x',y',a') = \frac{q(y'|y')}{\pi(y|y')} \pi(x,y|x',y',a') \)

(iii) \( q(x',y',a',t-1) = \pi(a'|x',y',t-1)q(x',y',t-1) \).
Item (iii) follows from our assumption, see the discussion following equation (3.4), that risk is associated with the y variable only. Item (ii) is merely a rearrangement of (3.2) and (3.3). To prove (i), note that

\[ q(x, y|x', y', t-1) = q(y|y') \pi(x|y, x', y', t-1) \]

by an argument similar to that following equation (3.4). Then, following (3.2) and proposition 1, we can write

\[ \pi(x|y, x', y', t-1) = \frac{\pi(x, y|x', y', t-1)}{\pi(y|y')} \]

(4.14) \[ q(x, y, t) = \sum_{x'} \sum_{y'} \frac{q(y|y') \pi(x, y|x', y', t-1)}{\pi(y|y')} q(x', y', t-1) \]

However, we can write

(4.15) \[ \pi(x, y|x', y', t-1) = \sum_{a(x', y')} \pi(x, y|x', y', a') \pi(x', y', t-1) \]

When (4.15) is substituted into (4.14), we get

\[ q(x, y, t) = \sum_{x'} \sum_{y'} \sum_{a(x', y')} \frac{q(y|y') \pi(x, y|x', y', a') \pi(a', x', y', t-1) q(x', y', t-1)}{\pi(y|y')} \]

However, from (4.13) items (ii) and (iii), this is just

\[ q(x, y, t) = \sum_{x'} \sum_{y'} \sum_{a(x', y')} q(x, y|x', y', a') q(x', y', a', t-1) \]

which is what we hoped to prove. □

If we let \( q(t) \) represent the vector \( q(x, y, a, t) \), then we have
\begin{equation}
Aq(0) = \pi(0) \tag{4.16}
\end{equation}
\begin{equation}
Aq(t) - Qq(t-1) = 0 \quad t = 1, 2, \ldots
\end{equation}
\begin{equation}
q(t) \geq 0
\end{equation}

where the matrices A and Q were defined earlier in this section.

Given $q(t)$ for all $t$, the market value of the rewards is simply

\begin{equation}
\sum_{t=0}^{\infty} \sum_{x} \sum_{y} \sum_{A} \tau(x, y, a)q(x, y, a, t) = r \sum_{t=0}^{\infty} q(t) \tag{4.17}
\end{equation}

Because of the special structure of this problem, see Grinold and Hopkins [9], and Grinold [8], we can solve the infinite programming problem of maximizing (4.17) subject to the constraints of (4.16) by solving the linear program obtained by adding the constraints and letting the sum $q = \sum_{t=0}^{\infty} q(t)$ be the decision variable.

\begin{equation}
\text{maximize} \quad rq
\end{equation}
\begin{equation}
(A-Q)q = \pi(0)
\end{equation}
\begin{equation}
q \geq 0
\end{equation}

This is (4.2), the dual of (4.1), obtained by another route, and it lets us see that $q(x, y, a) = \sum_{t=0}^{\infty} q(x, y, a, t)$ is just time zero market value of using action $a$ in state $(x, y)$ over the entire future.
5. Aggregation

On most cases the actual economic state space $\mathcal{Y}$ may be too large to allow for solving the decision problem expressed in equation (3:6). In this section we describe a special case where it is possible to aggregate economic states and thus reduce the difficulty of the computations. Even if the conditions stated below are not satisfied, this section gives a lead on how to approximate the problem with a simpler solvable problem and also gives some backing to those who heuristically (and perhaps not explicitly) have made these assumption from the start.

Let $Y_1, Y_2, \ldots, Y_n$ be a partition of $\mathcal{Y}$; the sets cover $\mathcal{Y}$ and are pairwise disjoint. Now assume that $r(x, y, a)$ is constant on each $Y_j$, that the set $a(x, y)$ is the same for all $y \in Y_j$, and $\pi(x'|x, y, y', a)$ is constant for $y \in Y_j$ and $y' \in Y_i$. Suppose further that

$$\sum_{y' \in Y_i} q(y'|y) = E\left[\frac{k(y)}{1 + r_m(y')|y)} | y \in Y_j\right]$$

depends only on $i$ and $j$.

Consistent with the assumptions above, we can use the following notation.

\begin{align*}
(5:1) & \\
(i) & r(x, j, a) = r(x, y, a) \quad \text{for } y \in Y_j \\
(ii) & \pi(x'|x, j, i, a) = \pi(x'|x, y, y', a) \quad \text{for } y \in Y_j \text{ and } y' \in Y_i \\
(iii) & q(i, j) = \sum_{y' \in Y_i} q(y'|y) \quad \text{for } y \in Y_j \\
(iv) & a(x, j) = a(x, y) \quad \text{for } y \in Y_j
\end{align*}
Note that (5:1) and (3:3) combine to give us

\[(5:2) \quad q(x', y' | x, y, a) = q(y' | y) \pi(x' | x, j, i, a) \quad \text{for } y \in Y_j \text{ and } y' \in Y_{\bar{j}}.\]

Now let \( V^T(x, y) \) be the optimal market value over \( T \) periods and suppose \( V^0(x, y) = 0 \). Then we can calculate \( V^{T+1}(x, y) \) by

\[(5:3) \quad V^{T+1}(x, y) = \max_{a \in A(x, y)} \left\{ r(x, y, a) + \sum_{x', y'} \sum_{y' \in Y_{\bar{j}}} V^T(x', y') q(x', y' | x, y, a) \right\} \]

We shall assume \( V^T(x', y') \) is constant for \( y' \in Y_{\bar{j}} \); thus \( V^T(x', i) = V^T(x', y') \) for any \( y' \in Y_{\bar{j}} \), and then prove that \( V^{T+1}(x, y) \) has the same property. To do this we first substitute (5:1) and (5:2) into (5:3). This yields, for any \( y \in Y_j \),

\[(5:4) \quad V^{T+1}(x, y) = \max_{a \in A(x, y)} \left\{ r(x, j, a) \right. \]

\[+ \left. \sum_{x', i=1}^n \sum_{y' \in Y_{\bar{j}}} V^T(x', i) \pi(x' | x, j, i, a) q(y' | y) \right\} \]

Equation (5:4) can be simplified, using our assumptions, to

\[(5:5) \quad V^{T+1}(x, y) = \max_{a \in A(x, j)} \left\{ r(x, j, a) + \sum_{x', i=1}^n V^T(x, i) \pi(x' | x, j, i, a) q(i, j) \right\}. \]

From equation (5:5) we can see that \( V^{T+1}(x, y) \) will be constant for \( y \in Y_j \) since the right side depends only on \( j \) and not on the particular \( y \in Y_j \). If we let \( q(x', i | x, j, a) = \pi(x' | x, j, i, a) q(i, j) \), then (5:5) becomes
\[(5.6) \quad V^{T+1}(x,j) = \max_{a \in A(x,j)} \left\{ r(x,j,a) + \sum_{x} \sum_{i=1}^{n} V^{T}(x',i)q[x',i|x,j,a] \right\} \]

and as \( T \to \infty \) if followed by exact analogy with the finite state discounted Markov decision process that \( V^{T}(x,j) \to V(x,j) \) which is the unique solution of

\[(5.7) \quad V(x,j) = \max_{a \in A(x,j)} \left\{ r(x,j,a) + \sum_{x} \sum_{i=1}^{n} V^{T}(x',i)q[x',i|x,j,a] \right\} \]

Equation (5.7) tells us that, under the assumptions given by (5.1), we can solve an aggregated problem and reduce an infinite state space to a finite state space. What (5.7) suggests is perhaps more important. If we are reasonably clever in the context of any particular problem about selecting our partition of the state space \( \mathcal{X} \), then the aggregation rules (5.1), although not exact, will give us a reasonable way in which to reduce the problem to a more manageable size.

6. An Example

This section contains a specific example of a decision problem that can be solved using the market value criterion. The problem is one of term lending in banks; see Pyle [14] for a review of this topic. The banks borrow for one period and can lend for two periods at higher rates of interest, however, they lose flexibility (liquidity) in adjusting to short-run (one-period) changes in the level of deposits and market rates of interests on loans and deposits.
Let $t$ indicate time, with period $t$ running from after time $t-1$ up to and including time $t$. The Markov state of the system at time $t$ is denoted $y(t)$. In what follows the parameters $D$, $r(1)$, $r(2)$, $\mu(1)$, $\mu(2)$, $\phi$, and $\psi$ are all permitted to vary with the Markov state $y(t)$.

At time $t-1$, the bank's balance sheet is as follows.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>reserves</td>
<td>$\rho D(t-1)$</td>
</tr>
<tr>
<td>short loans</td>
<td>$S(t-1)$</td>
</tr>
<tr>
<td>long loans</td>
<td>$L(t-1)$</td>
</tr>
</tbody>
</table>

We shall assume for the moment that $K$ is fixed, although the model is quite capable of either parametrically determining an optimal level of $K$ subject to market constraints or of allowing $K$ to be a decision variable that varies with time. This, however, is a distraction from the primary goal of the model: to measure the long/short loan tradeoff and thus get some insights into the benefits and costs of liquidity.

The level of deposits at time $t-1$ is $D(t-1)$, and corresponding level of required reserves is $\rho D(t-1)$ where $\rho$ is of the order $(0.15-0.2)$. $S(t-1)$ is the face amount of short loans due at time $t$, and $L(t-1)$ is the face amount of loans due at $t+1$.

At time $t$, the new Markov state $i(t)$ is known, thus the new level of deposits $D(t)$, and the new rates of interests, etc. We assume the bank can do three things: make new two-period loans
with face amount $L(t)$ due at time $t+2$, make new one-period loans
with face amount $H(t)$ due at time $t+1$, and finally sell an existing
amount of one-period loans $J(t)$. The short-loan portfolio after
these decisions will have value (face value due at $t+1$) $S(t)$ where

(6:1) $$S(t) = L(t-1) + H(t) - J(t)$$

Deposits and loans are discounted. Let $r(1)$ and $r(2)$
represent the market rates of interest for one and two periods, and
$\mu(1)$ and $\mu(2)$ the profit rates on one- and two-period loans.
The bank gives one-period borrowers

(6:2) $$\frac{[1-\mu(1)]}{(1+r(1))} H(t) = (1-h) \cdot H(t)$$

at time $t$ in return for a promise of a repayment of $H(t)$ at time
$t+1$. Similarly, the bank gives two-period borrowers. The parameter
$h$ defined by (6:2) measures the profitability of these loans.

(6:3) $$\frac{[1-\mu(2)]^2}{(1+r(2))^2} L(t) = (1-\xi)L(t)$$

for a promised repayment of $L(t)$ at time $t+1$, and $\xi$ is the
profitability.

The bank takes in an amount

(6:4) $$\frac{1+\psi}{(1+r(1))} D(t) = (1-d) \cdot D(t)$$

of deposits at time $t$, and promised to repay $D(t)$ at time $t+1$;
d is the cost of these loans. Finally, if \( Y(t) \) is the face value of the existing loans that are discounted, then the bank received

\[
(6.5) \quad \frac{[1-\psi]}{[1+r(1)]} J(t) = (1-j)J(t)
\]

for these loans; \( j \) is the cost of these premature sales.

Let \( P(t) \) be the profit realized at time \( t \), and recall that the time \( t \) balance sheet says

\[
(6.6) \quad \rho D(t) + S(t) + L(t) = D(t) + K
\]

The cash flow for \( t \) is

\[
(6.7) \quad \text{Cash Inflows} \quad \rho D(t-1) \quad \text{free reserves} \\
\quad S(t-1) \quad \text{repay short loan} \\
\quad (1-j)J(t) \quad \text{sale of existing loans} \\
\quad (1-d)D(t) \quad \text{new deposits}
\]

\[
\text{Cash Outflows} \\
\quad D(t-1) \quad \text{repay deposits} \\
\quad \rho D(t) \quad \text{new reserves} \\
\quad (1-\xi) L(t) \quad \text{two-period loans} \\
\quad (1-h)H(t) \quad \text{one-period loans} \\
\quad P(t) \quad \text{profit}
\]

When (6.7), (6.1), and (6.6) for \( t-1 \) and \( t \) are combined, we obtain
\( P(t) = L \cdot L(t) + h \cdot H(t) - d \cdot D(t) - j \cdot J(t) \)

The decision \( H(t), J(t), \) and \( L(t) \) are taken subject to the constraints

\[ H(t) - J(t) + L(t) = (1-p)D(t) + K - L(t-1) \]
\[ J(t) \leq L(t-1) \]
\[ L(t) \geq 0, \quad H(t) \geq 0, \quad J(t) \geq 0 \]

We could also impose lower bounds on \( L(t) \) and \( H(t) \), say \( L_0 \) and \( H_0 \). This would correspond to the practice of many banks of making loans to their best customers even if this leads to short-term losses. This added feature does not substantilly alter the form of the model, although it allows us to calculate the cost of this constraint.

Now our description of the model is complete. At time \( t \) the Markov state is \( y(t) \), and the bank's state is described by \( L(t-1) \). Given \( y(t) \) and \( L(t-1) \). We can choose, \( H(t), J(t), \) and \( L(t) \) in order to satisfy (6:9) and produce a profit \( P(t) \) (6:8).

Let \( V(y,x) \) be the market value of future profits giving \( y(t) = y \), and \( L(t-1) = x \). Then the function \( V(y,x) \) must satisfy.

\[ V(y,x) = \max_{H,J,L \in \mathcal{A}(y,x)} \left[ P(H,J,L,y) + \sum_{y'} q[y'|y] V(y',L) \right] \]

where

(i) \( P(H,J,L,y) \) is the profit: recall that \( D, d, j, h, \) and \( L \) in (6:9) may depend on \( y \).
(ii) \( A(y,x) = \begin{cases} H, J, L & H + L - J = (1-\rho)D + K - x \\ J \leq x & H \geq 0, \ J \geq 0, \ L \geq 0 \end{cases} \)

where \( D \) depends on \( y \).

(iii) \( q[y'|y] \) is the market value at time \( t \) in state \( y \) of a promise of \$1.00 \) at time \( t+1 \) contingent on state \( y' \) occurring. It is calculated in the manner of (2:7).

In order to establish the existence of a unique solution of (6:10) and to solve it, it is convenient to transform the problem.

Let \( G(t) = x - J(t) \). Then the constraints become

\[
(6:11) \quad A(y,x) = \begin{cases} G, H, L & G + H + L = (1-\rho)D + K \\ 0 \leq G \leq x & 0 \leq H, \ 0 \leq L \end{cases}
\]

and

\[
P(H,G,L,y) = h\cdot H + l\cdot L + jG - P_0(y,x)
\]

where

\[
P_0(y,x) = d\cdot D + j\cdot x
\]

If we assume these are a finite number of states \( y \), then it quickly follows that \( x \) is also bounded and that the per period profit is bounded above and below. Thus, any policy will lead to a finite market value.
If for each state $y$, the risk free rate of interest $r_f[y]$ is positive, and

$$\delta = \sup_y \left[ \frac{1}{1 + r_f[y]} \right] < 1$$

then the functional equation, see Denardo [5], will have a unique solution (6:10).

A most sensible way to solve (6:10) is by value iteration, with an initial guess of $V^0(y,x) = K + (1-j)x$; i.e., the existing loans are sold and capital returned to the owners.

A full data gathering, solution, use, and testing of this model will be the subject of a subsequent paper. The description here is merely to illustrate the use of the market value criteria in a specific problem.
REFERENCES


