Theory of dynamic portfolio choice for survival under uncertainty

Santanu Roy

Econometric Institute, Erasmus University, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands

Received March 1993; revised December 1994

Abstract

This paper develops a discrete time theory of dynamic optimization where the objective is to maximize the long-run probability of survival through risk portfolio choice over time. There is an exogenously given minimum withdrawal (subsistence consumption) requirement and the agent survives only if his wealth is large enough to meet this requirement every time period over an infinite horizon. The agent is endowed with an initial wealth. Every time period, he withdraws a part of the current wealth and allocates the rest between a risky and a risk-less asset. Assuming the returns on the risky asset to be i.i.d. with continuous density, the existence of a stationary optimal policy and the functional equation of dynamic programming are established. This is used to characterize the maximum survival probability and the stationary optimal policies. The stationary optimal policies exhibit variable risk preference ranging from extreme 'risk-loving' behavior for low levels of wealth to 'risk-averse' behavior for high levels of wealth.

Keywords: Survival; Risk; Subsistence consumption; Ruin; Dynamic portfolio choice

1. Introduction

One of the basic models of decision making under uncertainty that has been used to study the nature of risk preference and its implications for resource allocation is the model of investment portfolio choice. In a one-period framework, the problem is one of choosing an 'optimal' allocation of wealth over alternative assets that have random yields. While various criteria have been used to rank alternative portfolios, much of the focus in the literature has been on the mean-variance and the expected utility criteria. Models of intertemporal consumption, investment and portfolio choice have mostly used dynamic versions of
these criteria (for example, Samuelson, 1969; Hakansson, 1970; and Merton, 1971).

Most of these models are based on an assumption of some kind of risk aversion or uncertainty avoidance on the part of the agent. However, empirical and experimental studies indicate that risk preference is not fixed but depends on the context of choice. Given a choice between two alternatives of equal expected return in an experimental setting, human subjects are more likely to select the riskier alternative if the outcomes involve losses or are below some target level, than if the outcomes involve gains or are above some target level. Empirical observations on households and business enterprises reveal risk-averse behavior at high levels of wealth (or similar attribute) and risk-loving behavior at low levels (see, among others, March, 1988; Selten, 1990).

It is possible to accommodate these observations within a general analytical model by assuming, for example, that utility functions are convex-concave. However, making such an assumption on the preferences is equivalent to assuming the observed behavior.

There has been the gradual development of a theory of decision making that has focused on agents who seek to attain some aspiration or target level of outcome through their actions. Various alternative formulations of aspiration-level guided behavior have been used in the literature in different contexts (see, among others, Simon, 1955; and Radner, 1975).

In the portfolio literature, since the early 1950s there have been a number of static models that have considered 'safety first' as a criterion for portfolio selection (see Roy, 1952). In such models, a subsistence or disaster level of returns is identified. The objective is taken to be the maximization of the probability that the returns are above the disaster level or, alternatively, the maximization of mean return while constraining the probability of avoiding disaster to be above a certain exogenous level (or some other variation on this theme). Most of the early models in this tradition approximate the probability of avoiding disaster by some upper bound; for example, that given by Chebyshev's inequality (see Pyle and Turnovsky, 1970). Such bounds are 'loose' approximations of the theoretical probability of interest and it is not at all clear that the behavior obtained by using these bounds in the maximization exercise is an approximation, in any sense, of the one that would be obtained by using the actual probability of avoiding disaster.

We consider an infinite horizon model of intertemporal investment portfolio choice in which the agent's objective is to maximize the probability of survival. Here, survival is defined as the event that the agent is able to meet a minimum withdrawal or 'consumption' requirement out of his current wealth in every time period. The minimum withdrawal can be given different interpretations (see Majumdar and Radner, 1992). It can be interpreted as an exogenously given subsistence consumption standard for low-income households. One may also
interpret the consumption constraint as a minimal level of debt service obligation imposed by creditors, where inability to meet this obligation leads to bankruptcy.¹

The agent is endowed with some initial wealth and, in every time period, he decides how to allocate his wealth remaining after consumption between two assets, one risky and the other risk-less. Specifically, we assume that the agent fixes his consumption at the subsistence level. Note that there is no loss of generality in assuming this. Even if we allow the agent to choose his consumption path, it is always optimal for him to set his consumption at the subsistence level in order to maximize his chances of survival. The agent is said to be ruined if his wealth level falls below the subsistence requirement. If the agent is not ruined in finite time, he is said to survive. For any sequence of adapted portfolio decision rules that the agent decides on, there is an associated stochastic process of wealth and a probability of survival. The dynamic optimization problem is to design a sequence of decision rules so as to maximize the probability of survival. The returns on the risky asset are assumed to be independent and identically distributed (i.i.d.) over time with compact support and a continuous density function.

The structure of our survival problem is similar to the 'gambler's ruin' problem (Dubins and Savage, 1965). In the classical version of this problem, the agent attempts to attain a target level of wealth by placing successive bets. The subsequent literature has considered variations of the gambler's ruin problem where agents control wealth processes in order to maximize objectives such as the probability that the wealth reaches a target level by a terminal date, the expected time to 'failure' and the expected discounted time to bankruptcy (see, among others, Heath et al., 1987; and Dutta, 1994). The analysis in these models is in continuous time and the wealth process is modelled as a diffusion or Ito-process. There is no withdrawal or consumption requirement over time.

Our specific model is a discrete-time variation of the dynamic portfolio choice model contained in Majumdar and Radner (1991). They analyze the problem of survival probability maximization in a continuous-time framework, where the risky asset is assumed to follow a two-parameter diffusion process and where the agent has a strictly positive consumption requirement every period. In discrete-time models, it is difficult to arrive at any explicit solution for the probability of survival even if one makes specific distributional assumptions. Deriving qualitative properties is also more difficult compared with continuous-time models.

¹The literature on moral hazard in a dynamic principal–agent framework has demonstrated, for a certain class of situations, the 'optimality' of incentive schemes that contain a minimum performance requirement (see Radner, 1986; Dutta and Radner, 1994). Contracts with 'target' performance requirements can be used as pre-commitment devices by firms in order to soften competition in the market (see Fershtman et al., 1991). These, in turn, have definite implications for the risk preference of agents who decide on capital structure, technology and R&D portfolio choice within a firm. The dynamic optimization problem for agents in such environments can be qualitatively similar to ours.
This paper is an attempt to build the foundations for a discrete-time theory of dynamic optimization with survival as the objective, and to derive a few interesting properties of the optimal portfolio policy. The standard results derived in the theory of discounted dynamic programming are not directly applicable here. Our objective function is not additively separable in time.

We prove the existence of an optimal policy that is stationary and develop an optimality equation for the value function, similar to that developed in the theory of stationary dynamic programming (see Blackwell, 1965; Bertsekas and Shreve, 1978). The functional equation generates a stationary optimal policy, not necessarily unique. The potential non-uniqueness of the optimal policy reflects the non-convex nature of our optimization problem.

We show that there exists a critical level of initial wealth below which survival is impossible, independent of what actions are chosen. There is another critical level of initial wealth above which survival can be ensured with probability one by choosing to concentrate all investment on the risk-less asset in every period. Between these two critical levels, the maximum survival probability is continuous and strictly increasing in current wealth. In particular, both ruin and survival may occur with positive probability.

If the current wealth of the agent lies in the interior of this interval, the agent always invests a strictly positive fraction of his investible wealth in the risky asset. We identify a lower bound (which is decreasing in wealth level) for the fraction of investment devoted to the risky asset. If the mean return on the risky asset is less than the return on the risk-less asset, then our results imply that the agent displays risk-loving behavior on this interval; a risk-averse agent would have never invested in the risky asset in this situation. In fact, there exists a range of values at the lower end of this interval such that if the current wealth of the agent lies there, then the unique optimal action is to put all investment in the risky asset (i.e. bold play is optimal).

We characterize fully the stationary optimal policy correspondence for the range where the maximum survival probability is equal to one. Here, investing all wealth in the risk-less asset (i.e. safe play) is an optimal action. We identify an upper bound to the fraction of wealth that is invested in the risky asset under any optimal policy. For a certain range of wealth, the agent necessarily concentrates almost his entire investment on the risk-less asset, independent of how good the risky asset is.

Our results, therefore, indicate that a highly variable risk preference can be demonstrated by agents who maximize their chances of survival. In particular, the agent can exhibit extreme risk-loving behavior at low levels of wealth and equally extreme risk-averse behavior at high levels of wealth. These results are a step towards providing a theoretical explanation of empirical and experimental observations. In particular, our results indicate why we might expect risk-loving behavior among low-income households, as observed, for example, in the
preference for high-risk cash crops by destitute farmers, as well as high fertility in such households. It also explains why firms and financial institutions tend to go for high (and often bad) risk options when faced with the prospect of bankruptcy. Our results also predict extreme aversion to risk among middle-income households who would rather ensure survival by investing in a deterministic asset than expose themselves to any uncertainty.

Qualitatively similar results have been derived by Majumdar and Radner (1991) in their continuous-time model. They show that optimal portfolios may be ‘inefficient’ in the mean-variance sense. In models of survival under production uncertainty, where the agent is ‘passive’, an increase in the mean-preserving spread of the random shock has been shown to increase the probability of survival for low levels of stock and the opposite for high stock levels (see Mitra and Roy, 1993).

If the wealth of the agent is such that survival is impossible, our theory yields no prediction. Perhaps a model where agents maximize expected time to ruin is appropriate there. If the wealth of the agent is very high, survival is ensured no matter what the agent does and, again, our theory yields no prediction. It appears that standard utilitarian models, or models where the aspiration level increases with wealth, are much more applicable to such situations. A limitation of our model is that it does not allow for borrowing or short-sales by the agent. However, it appears that qualitatively similar results on risk choice might obtain if there is a binding constraint on the total volume of net borrowing or debt.

The model is formally outlined in Section 2 and the main results are discussed in Section 3. Section 4 contains the proofs.

2. The model

Consider the following problem of resource allocation under uncertainty. An agent is endowed with positive initial wealth in period zero. He is ruined in the...
first time period in which the current wealth is less than some fixed subsistence consumption level, \( c > 0 \). In each period \( t \geq 0 \), the agent observes the current wealth \( W_t \). He consumes \( c \) and allocates the remaining wealth \((W_t - c)\) between investment in a risky and a risk-less asset. Let \( \alpha_t \) denote the fraction of current wealth after consumption, which the agent invests in the risky asset. The gross rate of return on the risky asset available in period \( t \), and yielding a return in period \( t + 1 \), is denoted by \( \rho_t \). The (time-invariant) gross return on the risk-less asset is denoted by \( r \). A current portfolio decision by the agent determines his next period's wealth through the relation

\[
W_{t+1} = (\rho_t \alpha_t + (1 - \alpha_t)r)(W_t - c).
\]

We make the following assumption:

**Assumption 1.** \( \{\rho_t\}_{t=0,1,2,...} \) is a sequence of i.i.d. random variables defined on some probability space \((\Omega, \mathcal{F}, P)\). The support of the distribution of \( \rho_t \) is a closed interval \([a, b] \) where \( 0 < a < r < b < \infty \). Further, \( \rho_t \) has a density function \( g \) that is continuous on \([a, b]\). Lastly, \( r > 1 \).

For \( t \geq 0 \), let \( h_t = (\alpha_0, \alpha_1, \ldots, \alpha_{t-1}, \rho_0, \rho_1, \ldots, \rho_{t-1}, W_0, W_1, \ldots, W_{t-1}, W_t) \) denote the history of the process as observable at the beginning of period \( t \). A policy \( \pi \) is a sequence of decision rules \( \{\pi_t\} \), where \( \pi_t \) associates (Borel measurably) with each observable history \( h_t \), an element of the interval \([0, 1]\), indicating the fraction of investible wealth that is allocated to the risky asset in period \( t \). Any policy \( \pi \) and initial wealth \( w \) generates a stochastic process of \( t \)-period history \( \{h_t(w, \pi)\} \) and a stochastic process of wealth given by

\[
W_0(\pi)(w) = w,
W_{t+1}(\pi)(w) = (\pi_t \rho_t + (1 - \pi_t)r)(W_t(\pi)(w) - c), \quad t \geq 0, \tag{2}
\]

where \( \pi_t = \pi_t(h_t(w, \pi)) \).

The probability of survival for a given policy \( \pi \) and initial wealth \( w \), denoted by \( P(w, \pi) \), is defined by

\[
P(w, \pi) = \text{Prob}[W_t(\pi)(w) \geq c \text{ for all } t \geq 0]. \tag{3}
\]

The dynamic optimization problem is to choose a policy so as to maximize the probability of survival. Let \( \Pi = \{\pi: \pi \text{ is a policy}\} \). Also, let the value function \( V(w) \) for the dynamic optimization problem be defined by

\[
V(w) = \sup_{\pi \in \Pi} P(w, \pi). \tag{4}
\]

\( V(w) \) is the maximum survival probability from initial wealth \( w \). A policy \( \pi^* \) is said to be an optimal policy if \( P(w, \pi^*) = V(w) \) for all \( w \geq 0 \).

\(^5\) Note that we make no assumption on the relationship of \( \text{E}(\rho_t) \) and \( r \).
We are particularly interested in the class of policies that are stationary. A policy \( \pi = \{ \pi_t \} \) is stationary if the decision rules \( \pi_t \) are identical in all time periods and, further, depend only on the current state (wealth) \( W_t \). A stationary policy can, therefore, be described by a measurable function \( f : \mathbb{R} \to [0, 1] \). If the current wealth in any period is \( w \), then \( f(w) \) is the fraction of \( (w - c) \) that is invested in the risky asset. \( f \) is called the policy function generating the stationary policy and we denote such a policy by \( f^{(\omega)} \).

3. Main results

In this section we describe the main results of the paper. Our primary objective is to lay down the foundations of dynamic optimization for this problem. As our objective function is not additively separable over time, we cannot appeal to standard results on dynamic programming. In some respects, our problem is akin to dynamic optimization problems where the objective is maximization of expected long-run average reward.\(^6\) The use of a discrete time structure differentiates the methodology of our work from the analysis in continuous time by Majumdar and Radner (1991).

Recall that \( a \) and \( b \) are the lowest and highest possible returns on the risky asset and \( r \) is the return on the risk-less asset, \( 0 < a < r < b \) and \( r > 1 \). Define the critical levels of wealth \( A \) and \( B \) by

\[
b(A - c) = A, \tag{5}
\]

\[
r(B - c) = B. \tag{6}
\]

It is easy to check that \( 0 < A < B < \infty \) (see Fig. 1).

Now, if the initial wealth is equal to \( A \), then it is possible to sustain consumption equal to \( c \) over an infinite horizon only if, in every period, the gross rate of return on the total investment is equal to \( b \), which is a null event (under any portfolio policy). Therefore, the maximum probability of survival \( V(w) \) is zero for initial wealth \( w = A \) and, in fact, for all \( w \leq A \). However, one can check from (6) that if the initial wealth is equal to \( B \), and the agent follows the simple policy of concentrating all investment in the risk-less asset in every time period, survival is ensured. Therefore, for \( w \geq B \), \( V(w) = 1 \). We show that for \( w \in (A, B) \),

\(^6\) Consider a one-period utility function of consumption of the form \( u(x) = 0 \) for \( x < c \) and \( u(x) = 1 \) for \( x \geq c \), and further, assume that utility is zero for all time periods following any time period in which \( x < c \). Then the expected long-run average reward from consumption is, in fact, the probability of survival in our sense. See, also, Dutta (1994).
$0 < V(w) < 1$; that is, under optimal policy, both survival and ruin occur with positive probability. Further, $V(w)$ can be shown to be strictly increasing on $[A, B]$ (see Fig. 2).
Using recursive arguments and the stationary structure of the model, we establish the following functional equation of dynamic programming (or optimality equation) which is satisfied by the value function $V$:

$$V(w) = \max_{0 \leq x \leq 1} E[V((xp + (1 - x)r)(w - c))].$$  \hspace{1cm} (7)

The functional equation states that the maximum probability of survival from any stock today is equal to the expected maximum probability of survival from the stock to be attained tomorrow (when the current portfolio is chosen so as to maximize the latter). This illustrates the fundamental nature of our optimization problem: from any current state we are concerned only about the state that we move into.

We derive certain qualitative properties of the value function. In particular, the value function can be shown to be continuous on $\mathbb{R}^+$. This continuity property is obtained by using the functional equation and by exploiting the continuity of the density function for the random shock. In particular, we consider the parametric maximization problem on the right-hand side of (7). The continuity of the density function leads to the continuity of the maximand in both the parameter $w$ as well as the action $\alpha$. Berge's Maximum Theorem is then applied to establish parametric continuity of the maximum.

One of the implications of the continuity of $V(w)$ is that the maximization problem on the right-hand side of (7) has a solution for every $w$. Using this and the optimality equation, we establish the existence of an optimal policy that is stationary. The proof of existence is based on the property that, under any policy, the $T$-period survival probability decreases to the infinite horizon survival probability (as $T \to \infty$), as well as on the Markov structure of the stochastic process of wealth resulting from any policy. We show that the class of all policy functions that generates stationary optimal policies is identical to the class of all measurable functions $f(w)$, such that for each $w$, $f(w)$ solves the maximization problem on the right-hand side of (7).

In our framework, one cannot, in general, obtain any kind of concavity property for the maximand. In fact, under certain conditions, and for the case where the portfolio choice rule is fixed as one of investing all wealth in the risky asset, the survival probability as a function of initial wealth has been shown to be S-shaped (see Mitra and Roy, 1993). Thus, the value function is typically non-concave and there is always the possibility of multiple solutions to the maximization problem on the right-hand side of (7). There may, therefore, be

---

7 Note that $V$ need not be a unique solution to the functional equation (in the appropriate function space). The operator defined by the right-hand side of the functional equation is not a contraction. However, along with the boundary conditions $V(w) = 0$, for $w \leq A$ and $V(w) = 1$ for $w \geq B$, the functional equation does provide a rather strong characterization of the value function on the interval $[A, B]$. 
multiple stationary optimal policies. We need to characterize the stationary optimal policy correspondence \( \phi(w) \) where

\[
\phi(w) = \{ x \in [0, 1] : x \text{ solves the maximization problem on the right-hand side of (7)} \}.
\]

The (stationary) optimal policy correspondence is upper semi-continuous. It is not possible to ensure, in general, the existence of a continuous measurable selection from this correspondence.

From the earlier discussion it is clear that if the initial wealth is at least as large as \( B \), then choosing the safe strategy of investing only in the risk-less asset is an optimal policy. To say something more about the (stationary) optimal policy correspondence \( \phi(w) \), define the critical points \( G \) and \( K \) by

\[
r(G - c) = A,
\]

\[
a(K - c) = B.
\]

The critical wealth levels \( G \) and \( K \) are illustrated in Fig. 1. Note that \( A < G < B < K < \infty \), \( 0 < V(G) < 1 \) and \( V(K) = 1 \) (see Fig. 2). To understand the significance of the critical level \( G \), suppose that we call the event that the return on the risky asset is greater than the safe return \( r \), a 'good return', and a return below \( r \), a 'bad return'. \( G \) is the highest level of current wealth such that no matter what portfolio allocation is chosen in the current period, if a bad return occurs, then the next period's wealth is no greater than \( A \), from where ruin occurs almost surely. Thus, for wealth levels below \( G \), survival is impossible unless a good return is obtained in the next period. On the other hand, \( K \) is the smallest level of current wealth such that no matter what portfolio allocation is chosen in the current period, the next period's wealth is at least as large as \( B \) almost surely, from where it is possible to survive with probability one.

Consider \( w \geq K \). The expected survival probability from tomorrow's wealth is maximized (equal to one) by every feasible portfolio choice. The image of the optimal policy correspondence is, therefore, the entire \([0, 1]\) interval. Similarly, for \( w \leq A \) the maximum survival probability is equal to zero and so the image of \( \phi(w) \) is the entire \([0, 1]\) interval. Therefore, the focus of interest is on characterizing the correspondence \( \phi \) on the interval \((A, K)\).

Here, the first interesting result is that if the current wealth \( w \) lies in \((A, G]\), then the unique optimal portfolio choice under any stationary optimal policy is to concentrate all current investment on the risky asset. In other words, the correspondence \( \phi \) is single and constant valued on \((A, G]\) and, in fact, assumes the value 1. To understand the implications of this result, note that we allow for the possibility that the mean return from the risky asset is less than the return on the safe asset. By definition of risk aversion, no risk-averse agent would, in such an event, invest any wealth in the risky asset (see Arrow, 1965). However, our results indicate that agents maximizing survival probability will choose to invest
everything in the risky asset if their current wealth levels are low enough (more precisely, less than \( G \)) independent of whether the mean return from the risky asset is greater or less than \( r \). Thus, maximization of survival probability can lead to extreme risk-loving behavior.\(^8\) Note that this is not a trivial choice; the strategy of bold play ensures strictly positive survival probability.

To understand the factors leading to such a strong preference for risk, consider an agent whose current wealth lies in this region. Then, no matter what portfolio is chosen, if the risky asset yields a 'bad return' in the sense described earlier, the wealth next period will reach a level from where ruin occurs almost surely. Therefore, the choice of portfolio makes a difference to the agent's survival chances only in the event that a good return is realized from the risky asset; that is, return on the risky asset is greater than the safe return. But in the latter event, the wealth next period is maximized by concentrating all investment in the risky asset.

Unfortunately, we do not have any such strong characterization of the optimal policy correspondence on the rest of the interval between \( A \) and \( B \), namely \((G, B)\). We are, however, able to identify a strictly positive lower bound for the fraction of wealth invested in the risky asset. Let the function \( \lambda(w) \) be defined on \([A, B]\) by

\[
(\lambda(w)b + (1 - \lambda(w))r)(w - c) = w.
\]

Note that \( \lambda(w) \) is strictly decreasing, \( \lambda(w) \in (0, 1) \) for \( w \in (G, B) \), \( \lambda(G) = (1/b) \), and \( \lambda(B) = 0 \). We show that under any stationary optimal policy, the fraction of total investment devoted to the risky asset is bounded below by \( \lambda(w) \) on the interval \((G, B)\). Again, if \( E(\rho) \leq r \), this kind of portfolio choice implies risk-loving behavior.\(^9\)

Lastly, we consider the segment \([B, K)\). As noted earlier, it is always optimal to concentrate all investment on the risk-less asset for this range. In other words, \( 0 \in \phi(w) \) for \( w \in [B, K) \) and the agent can exhibit extreme risk-averse behavior in this range. However, there can be other optimal portfolio rules. Let \( \eta(w) \) be defined by

\[
\eta(w) = [(w - c)r - B]/[(r - a)(w - c)].
\]

Note that \( \eta(w) \in [0, 1] \) for \( w \in [B, K) \). Further, \( \eta(B) = 0 \), \( \eta(w) \) is increasing on \([B, K)\) and \( \eta(w) \uparrow 1 \) as \( w \downarrow K \). We show that for \( w \in [B, K) \), \( \phi(w) \) is, in fact, equal to the interval \([0, \eta(w)]\). For wealth levels in this range that are close

\(^8\) Majumdar and Radner (1991) derive qualitatively similar results about optimal portfolio choice that are inefficient in a mean–variance sense if the current wealth is low.

\(^9\) One might want to investigate whether the optimal policy correspondence is ascending (or expanding) in this interval. However, since the value function for this problem is typically non-concave, it is very difficult to derive any general result on this aspect. It appears to be dependent on the curvature of the value function, which, in turn, is closely related to the specific distribution of returns from the risky asset.
enough to $B$, optimal decision rules necessarily imply a high degree of risk
aversion, particularly if $E(p)$ is significantly greater than $r$. (In this latter event,
the standard expected utility theory of portfolio choice predicts that, under
regularity conditions, the agent always invests some positive wealth in the
risky asset, however concave his utility function may be; see Arrow, 1965). To
understand the motivation behind such safe play, note that in this range the agent
can ensure survival by investing only in the risk-less asset and excessive risk taking
is not only unnecessary, but might expose him to chances of ruin. The optimal
policy correspondence is illustrated in Fig. 3. We summarize the main results of
the paper in the following theorem:

**Theorem 1** (The main result).

(i) The value function $V$ is continuous and non-decreasing on $R_+$.

(ii) Let $A = [cb/(b - 1)]$ and $B = [cr/(r - 1)]$. Then $V(w) = 0$ for $w \leq A$, $V(w) = 1$ for $w \geq B$ and $V(w) \in (0, 1)$ for $w \in (A, B)$. Further, $V$ is strictly increasing on
$[A, B]$.

(iii) (The optimality equation) For all $w \geq 0$,

$$V(w) = \max_{0 \leq x \leq 1} E[V((xp + (1 - x)r)(w - c))].$$

Let $\phi(w) = \{x: x \text{ solves the maximization problem on the right-hand side of (*)}\}$.

(iv) $\phi$ is upper semi-continuous.

(v) There exists an optimal policy that is stationary. In fact, every measurable
selection $f$ from $\phi$ generates a stationary policy $f^{(\alpha)}$ that is optimal. Further, if $\alpha^{(\alpha)}$

![Fig. 3. Optimal policy correspondence $\phi(w)$. (Note: for $w \in (G, B)$, $\lambda(w)$ is a lower bound of $\phi(w)$.)](image.png)
is a stationary optimal policy generated by a measurable function \( \alpha \), then \( \alpha \) is a selection from \( \phi \); that is, \( \alpha(w) \) solves the maximization problem on the right-hand side of (\(*\)). Thus, the class of all stationary optimal policy functions is exactly identical to the class of all measurable selections from the correspondence \( \phi \).

(vi) Let \( G = \lfloor (A/r) + c \rfloor \) and \( K = (B/a) + c \). Then, \( A < G < B < K < \infty \) and the following hold:

(a) If \( w < A \) or \( w > K \), then \( \phi(w) = \{ \alpha : 0 \leq \alpha \leq 1 \} \),

(b) If \( w \in (A, G] \), then \( \phi(w) = \{1\} \),

(c) If \( w \in (G, B) \), then \( \alpha \in \phi(w) \) implies \( \alpha \geq \lambda(w) \), where \( \lambda(w) = \lfloor (w - r(w - c)) / ((b - r)(w - c)) \rfloor \), \( \lambda(G) = 1/b \), \( \lambda(B) = 0 \), \( \lambda(w) > 0 \) for \( w \in (G, B) \) and \( \lambda(w) \) is strictly decreasing on \( [G, B] \),

(d) If \( w \in [B, K) \), then \( \phi(w) = \{ \alpha : 0 \leq \alpha \leq \eta(w) \} \), where \( \eta(w) = [ (w - c)r - B ] / [(r - a)(w - c)] \), \( \eta(w) \) is strictly increasing on \( [B, K] \), \( \eta(B) = 0 \) and \( \eta(w) \uparrow 1 \) as \( w \uparrow K \).

4. Proofs

The proof of Theorem 1 is accomplished through the following steps:

(1) We show that \( V(w) = 0 \), for \( w < A \), \( V(w) = 1 \), for \( w \geq B \), \( V(w) \in (0, 1] \) for \( w \in (A, B) \) and that \( V \) is non-decreasing (Lemmas 1 and 2).

(2) We prove a recursive functional equation for the probability of survival \( P(w, \pi) \), under any given policy \( \pi \) (Proposition 1); this is then used to establish the functional equation of dynamic programming (Proposition 2).

(3) We show that \( V \) is continuous (Proposition 3) and that the maximization problem on the right-hand side of functional equation (\(*\)) has a solution.

(4) We show the existence of an optimal policy that is stationary, and that any measurable selection from \( \phi \) generates such a policy (Proposition 4). Further, any measurable function generating a stationary optimal policy is a selection from \( \phi \) (Proposition 5).

(5) We show that for \( w \in (A, B) \), \( V(w) \in (0, 1) \) (Proposition 6), and that \( V \) is strictly increasing on \( [A, B] \) (Proposition 7).

(6) Lastly, we prove part (vi) of Theorem 1, which characterizes \( \phi \) (Proposition 8).

Recall the definition of the critical wealth levels \( A \) and \( B \) in (5) and (6). Note that \( 0 < A < B < \infty \). Our first result is a characterization of the value function \( V(w) \).

**Lemma 1.** \( V(w) = 0 \), for \( w \leq A \), \( V(w) = 1 \) for \( w \geq B \) and \( V(w) \in (0, 1] \) for \( A < w < B \). For, \( w \geq B \), the policy \( \alpha = \{ a_t \} \), where \( a_t = 0 \) for all \( t \) and \( h_t \), is optimal. For \( w \leq A \), every policy is optimal.
Proof. First consider \( w < A \). Then, \( b(w - c) < w \). Let \( Y_t \) be the sequence defined by \( Y_0 = w \), \( Y_{t+1} = b(Y_t - c) \), \( t \geq 0 \). Then \( Y_t \) is monotonically decreasing. Suppose \( Y_t \) is bounded below. Then it converges to, say, \( y \). It must be true that \( b(y - c) = y \). But \( y < Y_0 = w < A \). Since \( A \) is the unique solution to the equation \( b(w - c) = w \), we have a contradiction. Hence \( Y_t \downarrow -\infty \). For any policy \( \pi = \{ \pi_t \} \), one can check by induction that \( W_t(\pi)(w) \leq Y_t \) almost surely. So there exists \( T \geq 0 \), such that \( W_T(\pi)(w) < c \) almost surely. Thus, \( P(w, \pi) = 0 \) for every policy \( \pi \). Hence, \( V(w) = 0 \). At \( w = A \), note that for any policy \( \pi \), \( W_1(\pi)(w) < A \) with probability one and, as seen earlier, ruin occurs almost surely from there. Thus \( P(A, \pi) = 0 \) for every policy \( \pi \) which implies \( V(A) = 0 \).

For \( w \geq B \), let \( \alpha = \{ \alpha_t \} \) be the policy where \( \alpha_t(h_t) = 0 \) for all \( t \) and \( h_t \). Since \( w > B = (c + (c/r) + (c/r^2) + (c/r^3) + \cdots) \), it is possible to sustain a consumption equal to \( c \) over an infinite horizon. Thus, \( P(w, \pi) = 1 \). Hence, \( V(w) = 1 \) and the policy \( \alpha \) is optimal.

Lastly, choose \( w \in (A, B) \). Consider the policy \( \pi \) where \( \pi_t(h_t) = 1 \) for all \( t \) and \( h_t \). We shall show that \( P(w, \pi) > 0 \) so that \( V(w) > 0 \). Since \( w > A \), \( b(w - c) > w \). There exists \( \rho^* \in (a, b) \) such that

\[ \rho^*(w - c) > w. \]

Let \( \{ y_t \} \) be the sequence defined by \( y_0 = w \), \( y_t = \rho^*(y_{t-1} - c) \). It is easy to check that \( \{ y_t \} \uparrow +\infty \). There exists finite \( T \geq 0 \), such that \( y_T > B \). This, in turn, implies that \( W_T(\pi)(w) > B \) almost surely on the set \( \{ \omega \in \Omega: \rho_t \in [\rho^*, b] \text{ for } t = 0, 1, \ldots, T - 1 \} \). As survival occurs almost surely from any wealth level lying in \([B, \infty)\):

\[ P(w, \pi) \geq \text{Prob}\{ \rho_t \in [\rho^*, b] \text{ for } t = 0, 1, \ldots, T - 1 \} > 0. \]

Next, we state a weak monotonicity property of \( V \). The proof is obvious and, hence, is omitted.

**Lemma 2.** For any policy \( \pi \), \( P(w, \pi) \) is non-decreasing in \( w \). Further \( V(w) \) is non-decreasing on \( R_+ \).

We now proceed to state a functional equation for the survival probability \( P(w, \pi) \) as a function of \( w \), for any given policy \( \pi \).

**Proposition 1.** For any policy \( \pi \),

\[ P(w, \pi) = E[P([\pi_0\rho + (1 - \pi_0)r](w - c)], \pi')] , \tag{12} \]

where \( \pi' \) is the policy defined by \( \pi'_t = \pi_{t+1} \) for all \( t \geq 0 \).
Proof. Let $S$ be the event that survival occurs from $w$ under policy $\pi$. Then

$$P(S \mid \rho_0) = \text{Prob}(W_t(\pi)(w) \geq c \text{ for all } t \geq 1 \mid \rho_0) = P\left(\left(\pi_0 \rho_0 + (1 - \pi_0) r\right)(w - c), \pi'\right).$$

Taking expectation with respect to $\rho_0$, and using the law of iterated expectations:

$$E[(P(\left(\pi_0 \rho_0 + (1 - \pi_0) r\right)(w - c)), \pi')] = E[P(S \mid \rho_0)] = P(S) = P(w, \pi). \square$$

Next, we establish the optimality equation for our dynamic optimization problem.

**Proposition 2** (The optimality equation). For any $w > 0$,

$$V(w) = \sup_{0 \leq x \leq 1} E[V((x \rho + (1 - x)r)(w - c))]. \quad (13)$$

Proof. For any policy $\pi$, we have from (12):

$$P[w, \pi] = E\{P(\left(\pi_0 \rho + (1 - \pi_0) r\right)(w - c)), \pi')\} \leq E[V(\left(\pi_0 \rho + (1 - \pi_0) r\right)(w - c))] \leq \sup_{0 \leq x \leq 1} E[V((x \rho + (1 - x)r)(w - c))].$$

Hence,

$$V(w) \leq \sup_{0 \leq x \leq 1} E[V((x \rho + (1 - x)r)(w - c))]. \quad (14)$$

Fix any $\epsilon > 0$. Consider any $x \in [0, 1]$. Suppose $x$ is the action chosen in period 0. The wealth at the beginning of period 1 is therefore $[x \rho_0 + (1 - x)r][w - c]$. From the definition of $V$ we know that for each possible realization $\rho \in [a, b]$, there exists policy $\pi(\rho)$ such that

$$P(\left(\pi_0 \rho + (1 - \pi_0) r\right)(w - c), \pi(\rho)] \geq V[\left(\pi_0 \rho + (1 - \pi_0) r\right)(w - c)] - \epsilon.$$

Define the policy $\alpha = \{\alpha_t\}$ by: $\alpha_0 = x$ for all levels of initial wealth, and from period 1 onwards follow the policy $\pi(\rho)$ for each possible realization $\rho$ of $\rho_0$. Then (using (12)):

$$P(w, \alpha) = EP[\left(x \rho + (1 - x)r\right)(w - c), \pi(\rho)] \geq E\{V[\left(x \rho + (1 - x)r\right)(w - c)] - \epsilon\}.$$

Thus,

$$V(w) \geq EV[\left(x \rho + (1 - x)r\right)(w - c)] - \epsilon.$$

As this can be done for every $x \in [0, 1]$,

$$V(w) \geq \sup_{0 \leq x \leq 1} E[V((x \rho + (1 - x)r)(w - c))] - \epsilon.$$

Since $\epsilon$ is arbitrary, we have

$$V(w) \geq \sup_{0 \leq x \leq 1} E[V((x \rho + (1 - x)r)(w - c))]. \quad (15)$$

Combining (14) and (15), we obtain (13). \square
Next, we use the optimality equation and the continuity of the density function \( g \), to derive the continuity of \( V \).

**Proposition 3.** \( V \) is continuous on \( \mathbb{R}_+ \).

**Proof.** As \( V \) is constant valued on \([0, A]\) and \([B, \infty)\), it is sufficient to show continuity of \( V \) on \([A, B]\). Let \( f(x, w) \) be defined by

\[
f(x, w) = \mathbb{E}[V((xp + (1 - x)r)(w - c))] = \int_{a}^{b} V((xp + (1 - x)r)(w - c))g(\rho) \, d\rho.
\]

For \( x \in (0, 1] \), after a change of variables we can write

\[
f(x, w) = \int_{[a/(x(w - c))]^{(1/x)}{(t/(w - c)) - (r(1 - x))}^{-1}/(x(w - c))}^{b} V([t]g) \, dt.
\]

(16)

Choose any \((x, w) \in (0, 1] \times [A, B]\) and any sequence \( \{x_n, w_n\} \) converging to \((x, w)\). As \( x > 0 \), there exists \( N \) such that \( x_n > 0 \) for \( n > N \). Using the dominated convergence theorem and the continuity of \( g \) in (16), we can see that \( f(x_n, w_n) \to f(x, w) \) as \( n \to \infty \). Thus, \( f(x, w) \) is continuous at each point in \((0, 1] \times [A \times B]\).

We will now define a sequence of intervals covering \([A, B]\) and show by an inductive argument that \( f(x, w) \) is continuous at all points \((x, w)\) where \( x = 0 \). Using Berge’s Maximum Theorem (see Berge, 1963), we will simultaneously establish continuity of \( V \) on each of these intervals. A separate argument will be used to establish the continuity of \( V \) at \( B \).

Define a sequence of points \( \{z_t\} \) in \([A, B]\) by the following recursion:

\[
z_{1} = A, \quad z_{t} = r(z_{t+1} - c), \quad \text{for } t = 1, 2, \ldots.
\]

It is easy to check that \( z_t < z_{t+1} \) and \( z_t \uparrow B \) as \( t \to \infty \).

Choose \( w \in [z_1, z_2] \). Then, \( r(w - c) \in (0, A) \). Consider a sequence \( \{(x_n, w_n)\} \) converging to \((0, w)\). For each \( \rho \in [a, b] \), there exists \( N_0(\rho) \) such that \([x_n \rho + (1 - x_n) r)(w_n - c)] \in (0, A) \) for \( n \geq N_0(\rho) \). This implies that \( V((x_n \rho + (1 - x_n) r)(w_n - c)) = 0 \) for \( n \geq N_0(\rho) \). Hence, using the dominated convergence theorem:

\[
f(x_n, w_n) = \mathbb{E}[V((x_n \rho + (1 - x_n) r)(w_n - c))] = \int_{a}^{b} V((x_n \rho + (1 - x_n) r)(w_n - c))g(\rho) \, d\rho \to 0 = f(0, w) \quad \text{as } n \to \infty.
\]

Therefore, \( f \) is continuous at points \((x, w)\) where \( x = 0 \) and \( w \in [z_1, z_2] \). Combined with the fact that \( f \) is continuous on \((0, 1] \times [A, B]\), we have that \( f \) is...
continuous on \([0, 1] \times [z_1, z_2)\). Using Berge’s Maximum Theorem we have the continuity of \(V\) each point in \([z_1, z_2)\). Thus, \(V\) is continuous on \((0, z_2)\).

Suppose \(V\) is continuous on \([0, z_1)\). Choose \(w \in [z_1, z_{1+1})\). Then, \(r(w - c) \in \[z_1, z_2)\). Again, consider a sequence \(\{(x_n, w_n)\}\) converging to \((0, w)\). For each \(\rho \in [a, b]\), there exists \(N(\rho) > 0\) such that for \(n \geq N(\rho)\), \([(x_n, \rho + (1 - x_n)r)(w_n - c)] \in (0, z_1)\). Using the continuity of \(V\) on \((0, z_1)\) we have that \(V((x_n, \rho + (1 - x_n)r)(w_n - c)) \to V(r(w - c))\) as \(n \to \infty\). Using the dominated convergence theorem, it follows that \(f(x_n, w_n) = EV[(x_n, \rho + (1 - x_n)r)(w_n - c)] \to f(0, w) = V(r(w - c))\). Thus, \(f\) is continuous at each point \((x, w)\) where \(x = 0\) and \(w \in [z_1, z_{1+1})\). Combined with the fact that \(f\) is continuous on \((0, 1\) \times \([A, B]\), this implies that \(f\) is continuous on \([0, 1] \times [z_1, z_{1+1})\). Using the maximum theorem again, we have that \(V\) is continuous on \([z_1, z_{1+1})\) and hence, on \((0, z_{1+1})\). By induction, \(V\) is continuous on \((0, B)\) and, in particular, on \([A, B]\).

To see continuity at the point \(B\), choose any \(w \uparrow B\). Define a sequence \(\{R_n\}\) by \(R_n(w - c) = B\). Note that \(R_n \downarrow a\) as \(n \to \infty\). There exists \(N^\uparrow > 0\) such that for \(n > N^\uparrow\), \(R_n \in [a, b]\). Let \(\pi\) be the following policy: invest all wealth in the risky asset in the first time period and invest all wealth in the risk-less asset for all subsequent periods. Then, for \(n > N^\uparrow\), \(V(w_n) \geq P(w_n, \pi) \geq \text{Prob}\{\rho_0 \in [R_n, b]\}\). As the \(\text{Prob}\{\rho_0 \in [R_n, b]\}\) \(\uparrow 1\) as \(n \to \infty\) it follows that \(V(w_n) \to 1 = V(B)\) as \(n \to \infty\). The proof is complete. \(\square\)

In view of the above proposition, we know that the supremum on the right-hand side of (13) is actually attained for all \(w \geq 0\). Thus, one can rewrite the optimality equation as:

**Corollary 1.**

\[V(w) = \max_{0 \leq x \leq 1} E[V((x, \rho + (1 - x)r)(w - c))]. \tag{17}\]

Let \(\phi(w) = \{x: x\ \text{is a solution to the maximization on the right-hand side of (17)}\}\). Using Berge’s Maximum Theorem, we have that \(\phi(w)\) is an upper semi-continuous correspondence. Let \(f(w)\) be a measurable selection from the correspondence \(\phi(w)\), and \(f^{(\infty)}\) the stationary policy generated by \(f\).

**Proposition 4.** There exists an optimal policy that is stationary. In particular, if \(f\) is any measurable selection from \(\phi(w)\), then the stationary policy generated by the function \(f\) (denoted by \(f^{(\infty)}\)) is an optimal policy.\(^{10}\)

**Proof.** Consider any \(w \geq 0\). For any \(T \geq 0\), let \(A_T = \{\omega \in \Omega: W_t(f^{(\infty)})(w) \geq c, 0 \leq t \leq T\}\), and \(A = \{\omega \in \Omega: W_t(f^{(\infty)})(w) \geq c \text{ for all } t \geq 0\}\). Then,

\(^{10}\)The existence of a measurable selection \(f\) from \(\phi\) follows from a selection theorem due to Kuratowski and Ryll-Nardzewski, see Hildenbrand (1974, Part I, Section D, Lemma 1).
At \supseteq A_{t+1}, A = \bigcap_{t=0}^{\infty} A_t.

Hence, \( P(A_t) \downarrow P(A) = P(w, f^{(\omega)}). \)

Fix any \( \varepsilon > 0 \) small enough. Then, there exists \( T_1 \geq 0 \) such that for all \( t \geq T_1, \)

\[
P(w, f^{(\omega)}) \geq \text{Prob}(A_t) - (\varepsilon/2). \tag{18}
\]

Let \( \pi^* \) be a policy such that \( P(w, \pi^*) \geq [V(w) - (\varepsilon/2)] \) for all \( w \geq 0 \). The policy \( \pi^* \) can be constructed by choosing separately for each \( w \) a policy that ensures survival probability \( \geq [V(w) - (\varepsilon/2)] \) (the latter exists for each \( w \) by definition of \( V(w) \)). For each \( T \geq 0 \) and for \( i = 0, 1, 2, \ldots \), let \( h^T_i \) be the truncated history from period \((T + 1) \) till period \((T + 1 + i) \) (as observable at the beginning of period \((T + 1 + i) \)). More precisely, \( h^T_0 = (W_{T+1}) \) and for \( i > 0, h^T_i = (\alpha_{T+1}, \ldots, \alpha_{T+i}, \rho_{T+1}, \ldots, \rho_{T+i}, W_{T+1} \ldots, W_{T+i}, W_{T+i+1}) \). Also for any such \( T \geq 0 \), let \( f^{(T)} \) be the (possibly non-stationary) policy such that

\[
f^{(T)}_t(h_t) = f(W_t), \quad t = 0, 1, 2, \ldots, T
\]

\[
= \pi^*_t(h^T_t), \quad t = T + 1 + i, \quad i = 0, 1, 2, \ldots
\]

Recall that by definition of a policy, \( \pi^*_t \) associates with each \( i \)-period history an action in \([0, 1]\). So \( f^{(T)} \) is a well-defined policy. As in Proposition 1, for any policy \( \pi \) let \( \pi' \) denote the one-period shifted policy, that is, \( \pi'_t = \pi_{t+1} \) for all \( t \geq 0 \). Observe that, by construction, \( f^{(0)} = \pi^* \) and \( f^{(T)} = f^{(T-1)} \). We claim that for all \( T \geq 0 \) and for all \( w \):

\[
P(w, f^{(T)}) \geq V(w) - (\varepsilon/2). \tag{19}
\]

To see (19), proceed by induction. From (12), for \( T = 0 \):

\[
P(w, f^{(0)}) = E\{P([f(w)\rho_0 + (1 - f(w))r](w - c)], \pi^*)\}
\geq E\{V((f(w)\rho_0 + (1 - f(w))r)(w - c))\} - (\varepsilon/2)
= V(w) - (\varepsilon/2),
\]

using (17) and the definition of \( f(w) \).

Suppose that for all \( w \):

\[
P(w, f^{(T-1)}) \geq V(w) - (\varepsilon/2). \tag{20}
\]

Then, using (12) again, we have

\[
P(w, f^{(T)}) = E\{P([f(w)\rho_0 + (1 - f(w))r](w - c)], f^{(T-1)})\}
\geq E\{V((f(w)\rho_0 + (1 - f(w))r)(w - c))\} - (\varepsilon/2) \text{ (using (20))}
= V(w) - (\varepsilon/2), \text{ by definition of } f(w).
\]

We have established that (19) holds for all \( T \).
Observe that, by construction, for all \( T \geq 0 \),
\[
W_t(f^{(T)})(w) = W_t(f^{(\omega)})(w), \quad t = 0, 1, \ldots, T \text{ a.s.}
\]  
(21)

Choose \( T \geq T_1 \). From (18) and (21):
\[
P(w, f^{(\omega)}) \geq \Pr(A_T) - (\epsilon/2)
= \Pr\{W_t(f^{(\omega)})(w) \geq c, t = 0, 1, \ldots, T\} - (\epsilon/2)
= \Pr\{W_t(f^{(T)})(w) \geq c, t = 0, 1, \ldots, T\} - (\epsilon/2)
\geq \Pr\{W_t(f^{(T)})(w) \geq c \text{ for all } t \geq 0\} - (\epsilon/2)
= P(w, f^{(T)}) - (\epsilon/2).
\]  
(22)

Using (19) and (22) we have
\[
P(w, f^{(\omega)}) \geq P(w, f^{(T)}) - (\epsilon/2) \geq [V(w) - (\epsilon/2)] - (\epsilon/2) = V(w) - \epsilon.
\]

As \( \epsilon \) is arbitrary, \( P(w, f^{(\omega)}) \geq V(w) \); that is, \( f^{(\omega)} \) is an optimal policy. The proof is complete. \( \square \)

In the previous proposition we have shown that if \( f \) is any measurable selection from the correspondence \( \phi(w) \), then \( f \) generates a stationary optimal policy \( f^{(\omega)} \). Now, we show that any measurable function \( f \) that generates a stationary optimal policy \( f^{(\omega)} \) must, in fact, be a selection from \( \phi(w) \). In other words, the function generating such a policy must be a solution to the maximization problem on the right-hand side of the optimality equation (17).

**Proposition 5.** Let \( f^{(\omega)} \) be any stationary optimal policy generated by a measurable function \( f \). Then, \( f(w) \in \phi(w) \) for all \( w \geq 0 \).

**Proof.** Suppose not. Then there exists some \( w_0 \) such that \( f(w_0) \not\in \phi(w_0) \). Then,
\[
V(w_0) > E\{V((f(w_0)\rho + (1 - f(w_0))\tau)(w_0 - c))\}.
\]  
(23)

As \( f^{(\omega)} \) is a stationary optimal policy, for each realization of \( \rho \):
\[
V((f(w_0)\rho + (1 - f(w_0))\tau)(w_0 - c))
= P([\{f(w_0)\rho + (1 - f(w_0))\tau](w_0 - c)], f^{(\omega)}).
\]

Thus,
\[
E\{V((f(w_0)\rho + (1 - f(w_0))\tau)(w_0 - c))\}
= EP([\{f(w_0)\rho + (1 - f(w_0))\tau](w_0 - c)], f^{(\omega)})
= P(w_0, f^{(\omega)}),
\]

using (12).

Using (23) we have, therefore:
\[
V(w_0) > P(w_0, f^{(\omega)}).
\]
This is a contradiction as \( f^{(w)} \) is an optimal policy. \( \square \)

From Propositions 4 and 5 it follows that:

**Corollary 2.** A measurable function \( f : \mathbb{R} \rightarrow [0, 1] \) generates an optimal stationary policy \( f^{(w)} \) if and only if \( f(w) \in \phi(w) \) for all \( w \geq 0 \).

Our next proposition is based on Corollaries 1 and 2.

**Proposition 6.** \( 0 < V(w) < 1 \) for all \( w \in (A, B) \).

**Proof.** As noted in Lemma 1, \( V(w) > 0 \) for \( w > A \) and \( V(w) = 0 \) for \( w \leq A \). Suppose \( V(w) = 1 \) for some \( w \in (A, B) \). Consider a stationary optimal policy \( \alpha^{(w)} \). It must be true that \( V[(\alpha w) + (1 - \alpha w) r (w - c)] = 1 \). (To see this, suppose not. Then \( V[(\alpha w) + (1 - \alpha w) r (w - c)] < 1 \) with positive probability, which contradicts (17).) Since \( V \) is non-decreasing, we must have \( V(r(w - c)) = 1 \). Define a sequence \( w_n \) by \( w_0 = w \), and \( w_n = r(w_{n-1} - c) \). Then \( w_n \downarrow A \). Repeating the arguments made above, we have that \( V(w_n) = 1 \) for all \( n \). Since \( V \) is continuous, we have that \( V(A) = 1 \), a contradiction. \( \square \)

We now state a useful lemma. The proof of this lemma follows from the fact that \( \rho_i \)'s have a density \( g \) whose support is a closed interval (use induction on \( t \)).

**Lemma 3.** Let \( \{W_i(\alpha)(w)\} \) be the stochastic process of wealth generated by any stationary policy \( \alpha^{(w)} \) from initial stock \( w \). Then, for any \( t \), the support of \( W_i(\alpha)(w) \) is a closed interval on \( \mathbb{R} \).

Our next result shows that \( V \) is, in fact, strictly increasing on \( [A, B] \).

**Proposition 7.** For any \( w_1, w_2 \in [A, B] \), \( w_1 < w_2 \) implies \( V(w_1) < V(w_2) \).

**Proof.** Since \( V(A) = 0 \), \( V(B) = 1 \) and \( V(w) \in (0, 1) \) for all \( w \in (A, B) \), the proposition must hold if \( w_1 = A \) or \( w_2 = B \). Therefore, consider the case where \( A < w_1 < w_2 < B \). From Lemma 2, we have that \( V(w_1) \leq V(w_2) \). Suppose the proposition does not hold. Then \( 0 < V(w_1) = V(w_2) < 1 \). Let \( f^{(w)} \) be a stationary optimal policy and, further, let \( \{W_i(w_1)\} \) be the stochastic process of wealth generated from initial wealth \( w_1 \) under this policy. Then, \( V(w_1) = \text{Prob}\{W_i(w_1) \geq c \text{ for all } t \geq 0\} \). It is easy to check that for all finite \( T > 0 \), \( V(w_1) \leq \text{Prob}\{W_i(w_1) \geq c, 0 \leq t \leq T\} \).

We first show that for all finite \( T > 0 \):

\[
V(w_1) < \text{Prob}\{W_i(w_1) \geq c, 0 \leq t \leq T\}. \tag{24}
\]
Suppose not. Then there exists some \( T > 0 \) such that
\[
V(w_1) = \text{Prob}\{W_t(w_1) \geq c \text{ for all } t \geq 0\}
= \text{Prob}\{W_t(w_1) \geq c, 0 \leq t \leq T\}
= \text{Prob}\{W_t(w_1) \geq c, 0 \leq t \leq \tau\}, \quad \text{for all } \tau > T. \tag{25}
\]
Since, \( A < w_1 < B \), one can use Proposition 6 to check that on the set \( \{\omega \in \Omega: W_t(w_1) \in (A, B)\} \), \( W_t(w_1) \leq c \) with positive probability for some \( \tau > T \), i.e. ruin occurs with positive probability in some future period. The last equality in (25), therefore, implies that
\[
\text{Prob}\{W_T(w_1) \in (A, B)\} = 0. \tag{26}
\]
Next, observe that
\[
\text{Prob}\{W_t(w_1) \geq B\} > 0, \quad \text{Prob}\{W_t(w_1) \leq A\} > 0. \tag{27}
\]
If \( \text{Prob}\{W_T(w_1) \geq B\} = 0 \), then (26) implies that \( W_T(w_1) \leq A \) almost surely, which contradicts the fact that \( V(w_1) > 0 \). Similarly, if \( \text{Prob}\{W_T(w_1) \leq A\} = 0 \), then (26) implies that \( W_T(w_1) \geq B \) almost surely, which contradicts the fact that \( V(w_1) < 1 \). (26) and (27) jointly contradict Lemma 3. Therefore, (24) holds.

Define a policy \( \alpha = \{\alpha_t\} \) to be followed from initial wealth \( w_2 \) by the following: for each \( t \), observe the history \( h_t \). Let \( W_t \) be the current wealth. Using the (stationary optimal) policy \( f^{(w)} \) and realizations of the random variables \( (\rho_0, \ldots, \rho_{t-1}) \), one can calculate \( W_t(w_1) \) at the beginning of period \( t \). If \( W_t(w_1) \geq c \), set \( \alpha_t \) such that
\[
\alpha_t(W_t - c) = f(W_t(w_1))[W_t(w_1) - c]; \quad \text{otherwise}, \quad \alpha_t = f(W_t).
\]
It can be checked that this is a well-defined policy and that
\[
W_t(\alpha)(w_2) = W_t(w_1) + r'(w_2 - w_1) \tag{28}
\]
almost surely on the set \( \{\omega \in \Omega: W_t(w_1) \geq c\} \) (use induction on \( t \)).

Thus, on the set \( \{\omega \in \Omega: W_t(w_1) \geq c \text{ for all } t \geq 0\} \), \( W_t(\alpha)(w_2) \geq W_t(w_1) \) almost surely, which implies that
\[
P(w_2, \alpha) > P(w_1, f^{(w)}). \tag{29}
\]
Let \( T^* \) be large enough such that \( [r^{T^*}(w_1 - w_0)] > B \). Let \( X \) be the event:
\[
X = \{\omega \in \Omega: W_t(w_1) \geq c, 0 \leq t \leq T^*, W_t(w_1) < c \text{ for some } \tau > T^*\}.
\]
From (28), we have that on the set \( X \) it is almost surely true that
\[
W_{T^*}(\alpha)(w_2) = W_{T^*}(w_1) + r^{T^*}(w_2 - w_1) > B.
\]
Thus, survival occurs almost surely on the set \( X \) if initial stock is \( w_2 \) and the policy followed is \( \alpha \). Note that, by construction, ruin occurs almost surely on \( X \), if the initial stock is \( w_1 \) and the stationary optimal policy \( f^{(w)} \) is followed. Since \( V(w_1) < \text{Prob}\{W_T(w_1) \geq c, 0 \leq t \leq T^*\} \), the event \( X \) occurs with strictly positive probability. Therefore, using (29), we have
\[
P(w_2, \alpha) - P(w_1, f^{(w)}) \geq P(X) > 0,
\]
which implies
\[ V(w_2) \geq P(w_2, \alpha) > P(w_1, f(\alpha)) = V(w_1), \]
which contradicts our initial supposition. □

Recall the definition of critical stocks \( K \) and \( G \) and that \( A < G < B < K \). For \( w \leq A \), the maximand on the right-hand side of (17) is equal to zero, whatever be the portfolio choice. Similarly, for \( w \geq K \), \( ((x\rho + (1 - x)r)(w - c)) \geq B \), no matter what \( x \in [0, 1] \) is chosen. So, the maximand on the right-hand side of (17) is equal to 1, whatever be the portfolio chosen. Thus, \( \phi(w) = \{ \alpha: 0 \leq \alpha \leq 1 \} \) for \( w \leq A \) and for \( w \geq K \). The next proposition characterizes the optimal policy correspondence \( \phi \) on \( (A, K) \). This, in turn, completes the proof of Theorem 1.

Recall the definitions (10) and (11) of the functions \( \lambda(w) \) and \( \eta(w) \) defined, respectively, on the intervals \([A, B]\) and \([B, K]\). It is easy to check that \( \lambda(w) \) is strictly decreasing on \([A, B]\), \( \lambda(w) > 0 \) on \([A, B]\), \( \lambda(A) = 1 \), \( \lambda(G) = 1/b \) and \( \lambda(B) = 0 \). Also, \( \eta(w) \) is strictly increasing on \([B, K]\), \( \eta(B) = 0 \) and \( \eta(w) \uparrow 1 \) as \( w \uparrow K \).

**Proposition 8.** The (stationary) optimal policy correspondence \( \phi \) satisfies the following on \((A, K)\):

(a) If \( w \in (A, B) \), then \( x \in \phi(w) \) implies \( x > \lambda(w) > 0 \).

(b) If \( w \in (A, G) \), then \( \phi(w) = \{1\} \).

(c) If \( w \in [B, K) \), then \( \phi(w) = \{x: 0 \leq x \leq \eta(w) < 1\} \).

**Proof.** (a) Suppose that for some \( w \in (A, B) \), there exists \( x \in \phi(w) \) such that \( x < \lambda(w) \). Then,
\[
E[V((x\rho + (1-x)r)(w - c))] \leq V((xb + (1-x)r)(w - c)) < V(w),
\]
using the definition of \( \lambda(w) \) and Proposition 7. This contradicts (17).

(b) Choose any \( w \in (A, G) \). For all realizations of \( \rho \) that lie in \([a, r]\):
\[
[(x\rho + (1-x)r)(w - c)] \leq r(w - c) \leq A.
\]
Since \( V(w) = 0 \) for \( w \leq A \):
\[
V((x\rho + (1-x)r)(w - c)) = 0, \quad \text{for } \rho \in [a, r].
\]
Therefore, the maximand on the right-hand side of (17) given by
\[
E[V((x\rho + (1-x)r)(w - c))] = E[V((x\rho + (1-x)r)(w - c))I_{(\rho \in [r,b])}],
\]
where \( I_{(\cdot)} \) is the indicator function.
However, for any $\rho \in [r, b]$, $((x \rho + (1 - x)r)(w - c))$ is maximized at a unique $x$, namely $x = 1$. Since $V$ is strictly increasing on $[A, B]$, we have that $x = 1$ is the unique solution to the maximization problem on the right-hand side of (17).

(c) Choose any $w \in [B, K)$. Note that $V(w) = 1$. Consider the maximization problem on the right-hand side of (17). We first show that if $x > \eta(w)$, then $x \not\in \phi(w)$. For any $x > \eta(w)$, there exists $\rho_0 \in (a, r)$ such that for all realizations of $\rho \in (a, \rho_0)$, we have

$$(x \rho + (1 - x)r)(w - c) < B,$$

so that

$$V[(x \rho + (1 - x)r)(w - c)] < V(B) = 1.$$

Thus, for $x > \eta(w)$:

$$\mathbb{E}[V((x \rho + (1 - x)r)(w - c))] < 1 = V(w).$$

Next, we show that $x \in [0, \eta(w)]$ implies $x \in \phi(w)$. Observe that $0 \leq x \leq \eta(w)$ implies that for every possible realization of $\rho$:

$$(x \rho + (1 - x)r)(w - c) \geq a(w - c) \geq B.$$

Thus,

$$\mathbb{E}[V((x \rho + (1 - x)r)(w - c))] = 1 = V(w).$$

Therefore, the maximization on the right-hand side of (17) is solved by any $x \in [0, \eta(w)]$. The proof is complete. □

Acknowledgements

I am grateful to T. Mitra for many useful suggestions. The current version of this paper has considerably benefited from detailed comments by M. Majumdar, an anonymous referee and participants at the 1993 European Meetings of the Econometric Society held in Uppsala (Sweden). The remaining errors are mine.

References


H.A. Simon, A behavioral model of rational choice, Q. J. Econ. 69 (1955) 99–118.