A Problem in Making Resources Last

J. MacQueen


Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28196411%2911%3A2%3C341%3AAPIMRL%3E2.0.CO%3B2-U

Management Science is currently published by INFORMS.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/informs.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
A PROBLEM IN MAKING RESOURCES LAST*1

J. MacQUEEN2

University of California, Los Angeles

An individual uses up a certain resource at a constant rate, but from time to time has a chance to gamble some of the resource in order to gain more. The expectation of each gamble is zero, i.e., the gambles are fair. This means that eventually the resource will run out. However, the form of the risk can be varied arbitrarily. The question is what form of risk to choose, when the object is maximize the expectation of the utility, \( u(T) \), of having the resources run out at time \( T \). This problem is solved for a more or less arbitrary function \( u(T) \).

Various possible interpretations of such a function \( u(T) \) are discussed briefly.

The model is mainly intended to provide a basis for empirical studies of individual decision making, where its complete mathematical tractability is convenient.

1. Introduction

Suppose that at a given time \( t \) an individual has certain resources \( x(t) \). These are used up at a specified rate, but from time to time opportunities arrive; at an ‘opportunity’ a decision is made and the resources are changed—increased or decreased—in a random manner depending on the decision. The situation is generally unfavorable; that is, on the average, resources are used up faster than they are increased, so that no matter what decisions are made, eventually the resources will run out. Nevertheless, the time at which they run out may be affected. The problem, then, is to make the decisions at each opportunity which will favorably effect the time when this happens.

This paper describes a simple model (Section 2) which to some extent typifies situations of this sort, and optimal policies are studied.3 More specifically, starting at time \( t = 0 \), the utility \( u(T) \) of having the resources run out at time \( T \) is assumed to be given, and the problem of maximizing \( Eu(T) \) is investigated. Optimal policies are obtained (Section 3) for the case where \( u \) is any continuous, bounded, increasing function. A number of alternative interpretations of the utility function \( u(T) \) are possible. These are considered briefly in Section 4.

The principal raison d’être of the model is to provide a well defined experimental situation in which decision making behavior related to the general topic of making resources last could be studied. Many more complicated models coming closer to being replicas of everyday situations can easily be imagined. However, this one has the advantage of admitting fairly complete and straight forward

* Received January 1964.
1 This research was supported in part by a grant from the Ford Foundation and in part under ONR contract 233(75).
2 Western Management Science Institute and Department of Public Health and Preventive Medicine.
3 Essentially the same model, but with a different optimality criterion, is treated in [4].
mathematical analysis. The knowledge of optimal behavior thus proved may prove to be a convenient starting point for useful qualitative and quantitative analysis of decision making behavior in more complicated situations.

Theorem 1 in Section 3 is essentially an extension to the continuous time case of a result of Blackwell’s [1]. The theorem provides a general way of obtaining upper bounds for the expected time until a process \( x(t) \) leaves a given region and may be useful in connection with other problems. Upper bounds obtained by the method extend to all (measurable) policies. This means that if a policy is found which actually achieves such an upper bound, then that policy is optimal over all policies and the bound itself is a least upper bound. Typically, such a policy will be stationary, and in this case the method provides a satisfactory way of dealing with the often technically difficult matter of whether or not an optimal stationary policy is optimal over all policies (See Derman [2] for a treatment of this problem.) The ability to handle the non-stationary case is essential in the following, in that optimality is established by first obtaining the least upper bound of \( ET \); a policy is then produced which maximizes \( Evu(T) \) subject to the constraint that \( ET \) be less than or equal to this least upper bound. But as it happens, this policy is non-stationary, so that it is necessary to the proof that the least upper bound on \( ET \) extend to the non-stationary policies.

2. The Model

The opportunities are assumed to arrive at random times \( t_1, t_2, \ldots \) with \( t_1, t_2 - t_1, t_3 - t_2, \ldots \) being independently and identically distributed. The common distribution \( M \) satisfies \( M(t) \to 0 \), as \( t \to 0 \), and we suppose \( \int_0^\infty t \, dM = 1 \).

An opportunity is characterized by a family of distributions which depend on the resources when the opportunity arrives. The individual picks a member of the family. Then a number \( z \) is selected randomly according to this distribution and the resources are changed instantaneously by adding that amount; i.e., if an opportunity arrives at \( t \), \( x(t') = x(t) + z \). Let \( \mathcal{F}_x \) be the family of distributions available if the opportunity arrives when the resources are at \( x \), and suppose that \( \mathcal{F}_x \) contains all \( F \) such that \( \int z \, dF = 0 \) and \( F(z) = 0 \) for \( z < -x \). In other words, the decision maker has his choice at each opportunity of any fair gamble which does not involve loss of more resources than he has at the time; the problem is choice of risks.

Between opportunities, the resources \( x(t) \) decrease continuously at a constant rate, which without loss of generality may be taken to be unity. As soon as \( x(t) \) is zero, it remains at zero thereafter.

Suppose the \( n^{\text{th}} \) opportunity arrives at \( t_n \) and let \( z_n \) be the change at that time. A policy \( R \) is a rule which determines an element of \( \mathcal{F}_x(t_n) \) at each \( t_n, n = 1, 2, \ldots \), in such a manner that \( t_1, z_1, t_2, z_2, \ldots \) is a well-defined sequence of random variables. Then starting with \( x(0) = x \), the time \( T = T(x, R) = \inf \{ t : x(t) = 0 \} \) at which the resources run out is itself a well-defined random variable for each \( x \) and \( R \).
3. The Optimal Policies

Let \( v(x) \) be the l.u.b. of \( ET(x, R) \) over all policies \( R \). Let \( m(x) = E \min(x, t_1) \), and for any non-negative function \( f \), continuous on \([0, \infty)\), let \( Lf \) be the function defined by

\[
Lf(x) = \begin{cases} 
  m(x) + \int_0^x \sup_{t \in \mathbb{R}} \int f(x - t + z) \, dF(z) \, dM(t), & \text{for } x > 0, \\
  0, & \text{for } x = 0. 
\end{cases}
\]

Theorem 1. If \( f \geq Lf \), then \( v \leq f \).

The proof given here, which is based on Blackwell’s proof in the paper cited above, requires only that for \( F \in \mathfrak{F} \), \( \int zdF \) is bounded. With obvious modifications, the theorem and proof also extend to the case \( x \) is a point in \( n \)-space and the interval \([0, \infty)\), especially pertinent here, can be replaced by an arbitrary region, with \( T \) becoming the time until \( x(t) \) leaves this region. A number of very general inequalities, similar to Theorem 1, are given by Dubins and Savage [3].

Proof. Let \( L^n \) be the \( n \)th iterate of \( L \) and suppose that for every initial level of resources \( x \), and some \( n, \)

\[
E \min(T, t_n) \leq L^n f. 
\]

Let \( z \) be the resource change at \( t_1 = t \), and let \( I_A \) be the characteristic function of the set \( A \). Since \( T = x \) if \( t \geq x \), we have

\[
E[\min(T, t_{n+1}) | t, z] = \min(t, x) + E[\min(T - \min(t, x), t_{n+1} - t) | t, z]. 
\]

The resources given \( t, z \) are \( I_{\{t < x\}}(x - t + z) \) and given \( t, z \), \( T - \min(t, x) \) is the time until the resources run out starting from this level, that is, at \( I_{\{t < x\}}(x - t + z) \). Considering that \( t_1, t_2 - t_1, \text{ etc.} \), are independent and identically distributed, so that \( t_{n+1} - t \) has the same distribution as \( t_n \), and that any policy available at \( t = t_1 \) for maximizing \( T - \min(t, x) \) is available at \( t = 0 \) for maximizing \( T \), the hypothesis (2) can be applied, with initial level \( I_{\{t < x\}}(x - t + z) \), to the expectation in the right hand side of (3). Thus

\[
E[\min(T, t_{n+1}) | t, z] \leq \min(t, x) + L^n f(I_{\{t < x\}}(x - t + z)). 
\]

Since \( L^n f(0) = 0, L^n f(Ix) = IL^n f(x) \). Hence

\[
E[\min(T, t_{n+1}) | t, z] \leq \min(t, x) + I_{\{t < x\}} \sup_{t \in \mathbb{R}} \int L^n f(x - t + z) \, dF, 
\]

and

\[
E(\min(T, t_{n+1})) \leq m(x) + \int_0^x \sup_{t \in \mathbb{R}} \int L^n f(x - t + z) \, dF \, dM
\]

\[= L^{n+1} f. \]
Since $E \min(T, t_n) = E \min(T, t) = m(x) \leq L^f$, (2) holds for $n = 1$. By induction then, $E \min(T, t_n) \leq L^n f$ for all $n$. Inspection of (1) shows that if $f \geq g, Lf \geq Lg$. By hypothesis, $f \geq Lf$, and iterating $L$ gives $f \geq Lf \geq L^2 f \geq \cdots$ and $\sup_n E \min(T, t_n) \leq \sup_n L^n f \leq f$. Finally, $E \min(T, t_n) \to ET$ ($T$ a.s. finite or not) by the monotone convergence theorem. This completes the proof of Theorem 1.

To apply Theorem 1 to the problem of maximizing $ET$, consider the function $f = x$. The right hand side of (1) becomes

$$\int_0^x t \, dM + x \int_0^x dM + x \int_0^x dM - \int_0^x t \, dM = x$$

so that $x \geq Lx$. Thus, $v(x) \leq x$. However, $x$ can be achieved by the simple expedient of not taking any risk, that is, by always choosing $F$ degenerate at 0. Since $v(x) \geq x$ by definition, $v(x) = x$, and this conservative policy maximizes $ET$.

Turning now to the problem of maximizing $Eu(T)$ with respect to all policies $R$, suppose $u$ is any continuous, bounded, increasing function, and consider the related problem of maximizing $\int u(T) \, dP$ with respect to the class of all probability distributions $P$ on $[t, \infty)$ satisfying $\int T \, dP \leq a$. As illustrated in Figure 1, let $c$ be the concave l.u.b. of $u$ on $[t, \infty)$; that is, let $c(T) = \inf_{\gamma \in \Psi} \gamma(T)$ where $\Psi$ is the class of all lines $\gamma$ satisfying $\gamma(T) \geq u(T)$ for $T \geq t$.

Suppose that $u(a) < c(a)$ as shown in Figure 1. Because $u$ is continuous and bounded, there are two values of $T$, say $a_1$ and $a_2$, satisfying $t \leq a_1 < a < a_2 < \infty$, and for a certain line $\gamma \in \Psi$, $c(a_1) = u(a_1) = \gamma(a_1)$ and $c(a_2) = u(a_2) = \gamma(a_2)$. If now a distribution $P'$ is determined so as to assign probability $p = (a - a_1)/\; (a_2 - a_1)$ to $a_2$ and $1 - p$ to $a_1$, $\int u(T) \, dP' = \int c(T) \, dP'$. Since $c$ is concave and increasing $c(a) \geq \int c(T) \, dP$ for all $P$ satisfying $\int T \, dP \leq a$ (Jensen's inequality.) But for $P'$, $c(a) = \int c(T) \, dP'$, $c$ being equal to the line $\gamma$ over $[a_1, a_2]$. Since $\int u(T) \, dP \leq \int c(T) \, dP$, $P'$ maximizes $\int u(T) \, dP$.

![Figure 1](image-url)
If \( u(a) = c(a) \), essentially the same argument with \( a_1 = a_2 \), shows that the distribution \( P' \) which assigns probability \( p = 1 \) to the point \( a \) maximizes \( \int u(T) dP \).

We can now describe the optimal policies. Suppose that starting at time \( t = 0 \) with \( x(0) = x \), the first opportunity arrives at \( t_1 = t ≤ x \) (See Figure 1.). Before the outcome \( z \) is known, the maximum expected (remaining) time until \( x(t) = 0 \) is \( x - t \) as was shown above. At this point \( u(t) \) has already been achieved and the conditional distribution of \( T \) is restricted under any policy to \([t, \infty)\). With \( a = x \), the above argument can be applied to maximizing \( \int u(T) dP \) for distributions \( P \) on \([t, \infty)\) subject to \( \int T dP \leq x \), since while not every such \( P \) can be generated by a policy, the maximizing one \( P' \) described above can be, and this, of course, is sufficient. At \( t \), the distribution \( F \) for the change \( z \) is chosen to have a salutus equal to \( p \) at \( z = a_2 - a = a_2 - x \) and another salutus equal to \( 1 - p \) at \( -a - a_1 = a_1 - x \); the mean of such a distribution is easily seen to be zero, so that in fact it belongs to \( F_{-t} \), thereafter, the conservative policy of choosing \( F \) degenerate at \( 0 \) is used. The outcome \( z = a_2 - a \) changes the resources to \( a_1 - t \), and the outcome \( z = -(a - a_1) \) changes the resources to \( a_2 - t \). Thus, the resources run out either at \( a_1 \) or \( a_2 \) with just the probabilities required by \( P' \).

The optimal policy may be described conveniently as follows. There are two functions of time, say \( T_0(t) \) and \( T_1(t) \), which correspond respectively to a lower and upper additional amount of time which it is desirable to have the resources last when the first opportunity arrives at time \( t \). The optimal policy is to risk lasting only to \( T_0(t) \) in order to have a chance of lasting until \( T_1(t) \), and thereafter take no risk at all.

Suppose the amount of resources \( X_0(t) \) which will last until \( T_0(t) \) is called the security level and the amount of resources \( X_1(t) \) which will last until \( T_1(t) \), is called the goal level. In these terms, the optimal policy is to attempt to achieve the goal level \( X_1(t) \), but to avoid falling below the security level, \( X_0(t) \). After the first opportunity, the resources are used up steadily until they are exhausted. The goal level is a non-increasing function of time and the security level is non-decreasing. The optimal behavior becomes more conservative with time.

The functions \( X_0 \) and \( X_1 \) depend only on the initial level of resources \( x \) and the utility function \( u \). With the time \( t \) of the first opportunity fixed, both \( X_0 \) and \( X_1 \) are non-decreasing in \( x \). To illustrate, suppose \( u \) is strictly convex increasing up to a point \( T^* \) and then strictly concave increasing; i.e., the curve is "lazy S" shaped. Regarding the time \( t \) of the first opportunity as fixed, consider the line which passes through the point \( t, u(t) \), is on or above \( u \) for \( T > t \), and is just tangent to \( u \) at certain point \( T^{**} \). If the initial resources are less than \( T^{**} \), then \( X_0(t) = 0 \) and \( X_1(t) = T^{**} \). Thus, all the resources are risked at \( t \) in order to have a chance of lasting until \( T^{**} \). If the initial resources \( x \) exceed, or equal, \( T^{**} \), then \( X_0(t) = X_1(t) = x \) and no risk is taken. As \( x \) increases, the probability assigned to running out immediately at an opportunity decreases.
4. Interpretation of \( u(T) \)

To illustrate the various ways a non-decreasing utility function of the type considered above might arise, consider the related problem of “survival”. The individual is supposed to perish if he runs out of resources. Being alive through the interval \((T, T + dT)\) has value \( \alpha(T) \, dT \), \( \alpha(T) \) being positive. The total value of surviving to time \( T \) is thus \( \int_0^T \alpha(t) \, dt \). The expected total value will then be maximized by maximizing the expectation of \( u(T) = \int_0^T \alpha(t) \, dt \). For \( \alpha(T) = 1 \), being alive one time is as good as another and \( u(T) = T \).

Suppose that survival is still the primary concern, but now, completely independent of anything the individual may do, his life will be terminated at a random time \( T^* \) with distribution \( V \). His actual survival time is then \( \min(T, T^*) \). Now suppose he wishes to maximize \( E u_0(\min(T, T^*)) \), where \( u_0 \) is non-decreasing. Let \( u(T) = E(u_0(\min(T, T^*)) \mid T) = \int_T^\infty u_0(t) \, dV + u_0(T)(1 - V(t)) \).

Then \( u \) is non-decreasing and maximizing \( E u(T) \) maximizes \( E u_0(\min(T, T^*)) \).

In particular, if the time \( T^* \) has distribution \( 1 - e^{-\lambda T} \), corresponding to the situation where the individual’s life may be terminated at each instant with a constant probability, and \( u_0(T) = T \), then \( u(T) = \lambda^{-1}(1 - e^{-\lambda T}) \).

Suppose that the individual does not want to run out before perishing. In other words, not being able to do anything about \( T^* \) he wishes to at least avoid the indignity of actually running out. Thus, if \( T \geq T^* \), he will have succeeded and if \( T < T^* \) he will have failed. His object then is to maximize the probability of success, \( P(T \geq T^*) \). But \( P(T \geq T^*) = E(P(T \geq T^*) \mid T) = V(T) \).

Suppose that there are several favorable events \( A_1, A_2, \ldots A_n \) which occur respectively at the random times \( T_1^*, T_2^*, \ldots T_n^* \). These times are independent of one another and have distributions respectively \( V_1, V_2, \ldots V_n \). If any one of these events occur, the individual is able to achieve his goal. If none of the events occurs before \( T \), all further chances are lost. Thus, given \( T \), the probability of successfully achieving his goal is

\[
u(T) = P[\min(T_1^*, T_2^*, \ldots T_n^*) \leq T] = 1 - \prod_{i=1}^n (1 - V_i(T)),\]

and maximizing \( E u(T) \) maximizes the overall probability of success.

Consider the case where \( u \) is decreasing from some point on. In survival terms, this corresponds to the case where perishing from that point on becomes desirable. This may be seen in terms of the function \( \alpha \) introduced above, which represents the relative worth of being alive at time \( T \). If \( u(T) = \int_0^T \alpha(t) \, dt \) is decreasing above a certain value of \( T \), \( \alpha \) can only be negative there.

* This interpretation is due to Tom Ferguson.
The analysis of the optimal behavior in Section 3 applies to the situation where over some interval, \( u \) is decreasing (but bounded from below, of course.) For example, if \( u \) is increasing up to a maximum at a point \( T' \), and thereafter decreasing, the problem becomes that of making the resources run out in the vicinity of \( T' \) while avoiding having them last beyond. If \( x(0) = x \) is less than \( T' \), this presents no problems. The decreasing region need never be entered and the policies previously described are optimal. If \( x \) exceeds \( T' \), it is possible to achieve within \( \varepsilon \) of the maximum value of \( u \), that is, \( u(T') \), provided the first opportunity arrives before \( T' \). Suppose \( B \) is the lower bond for \( u \). At the first opportunity, the distribution \( F \) is chosen to assign almost probability 1 to the change \( z \) such that thereafter, if no more risk is taken the resources will run out at \( T' \). Thus, with probability almost 1, \( u(T') \) is achieved and with probability almost zero, a utility of at least \( B \) is the result. Thus, if the first opportunity arrives before \( T' \), \( E u(T) \) can be made to exceed \( u(T') - \varepsilon \) for every \( \varepsilon > 0 \). If the first opportunity arrives at \( t \geq T' \), high probability is placed on running out immediately. Thus, in this case, it is possible to achieve an expected utility within \( \varepsilon \) of \( u(t) \). Integrating with respect to the distribution \( M \) of the time until the first opportunity shows that if \( x \geq T' \), it is possible to get within \( \varepsilon \) of \( u(T') \) \( M(T') + \int_{T'}^{x} u(t) \, dM \).

References