CARL J. NORSTRØM

STUDIES IN CAPITAL BUDGETING

THE NORWEGIAN SCHOOL OF ECONOMICS AND
BUSINESS ADMINISTRATION
BERGEN 1975

PREFACE

In the preparation of the articles in this dissertation I have had valuable comments from several friends and colleagues. Some of my gratitude is expressed in the individual articles. However, special thanks must be given to Karl Borch, Jan Mossin and Agnar Sandmo. I will also use this opportunity to thank Leif Johansen for encouragement and advice especially in connection with the last article in the collection.

I am indebted to the Norway-America Association and to the Norwegian Research Council of Science and the Humanities for grants which made it possible to spend a year at the University of Michigan.

Finally I would like to thank the many secretaries who typed the different versions of these articles, especially Mrs. A. Hoem, Mrs. E. Stiegler and Mrs. T. Hageselle.

CONTENTS

A MODIFICATION OF THE INTERNAL RATE OF RETURN METHOD*

By CARL J. NORSTRØM

1. *Introduction.*

The internal rate of return method is one of the main methods used by economists for evaluating investments. Due to the pionering work of Joel Dean [5] and others, it has also gained a certain acceptance in practice. During the last twenty years several distinguished economists have attacked the internal rate of return and claimed that it, in contrast to the present value method, will not always lead to correct decisions. The aim of this article is to clarify the reasons for the weaknesses of the internal rate of return method, and to suggest a modification of the method. The most serious limitation in our analysis will be that the discount rate is the same in each period.

The most important characteristics of the internal rate of return compared with the present value are that

1. The internal rate of return depends only on the cash flows of the projects and not on the discount rate.
2. The internal rate of return is independent of the size of the investment in the sense that it is unchanged if each element of the cash flow is multiplied by the same number.

The present value has none of these properties. It is a function of the discount rate and proportional to the scale of the investment. It is a consequence of this difference in character of the two measures that each is better than the other for some purposes. Thus the internal rate of return is very suitable if we want to know for which values of the discount rate a project is profitable, while the present value is much simpler to use in the choice between mutually exclusive projects. Moreover property 2 above indicates that the internal rate of return is a better measure of the quality of an investment while the present value also takes scale into consideration. We mention these points to indicate that establishing a correct version of the internal rate of return method not only is a matter of academic interest, but has some practical relevance as well.

* The author is grateful to professor Jan Mossin for encouragement and advice.

The main objections against the internal rate of return are the following:

1. The internal rate of return is not unique. There exist cash flows with none as well as cash flows with more than one internal rate of return.
2. The internal rate of return method may lead to incorrect decisions in the choice between mutually exclusive investments.

The readers interested in a more detailed discussion of these weaknesses are referred to Lorie and Savage [13], Solomon [15], Hirshleifer [7], Bailey [1], and Bernhard [2].[1] These works have contributed to the understanding ot both the present value and internal rate of return method.

2. *Assumptions and Definitions.*

The article is based on the following fundamental assumptions:

1. Each project is fully described by a net cash flow, e. g. $(a_0, a_1, \ldots, a_n)$, where a_t is the net in- or outflow in period t. Each flow takes place at the end of the corresponding period. For convenience we shall assume $a_0 \neq 0$.
2. In any period the firm may borrow or lend an arbitrary large amount at a rate of interest, ϱ. This rate will be referred to as the cost of capital. Unless otherwise stated it will be assumed that $\varrho \geqq 0$, but as will become clear this is not an essential assumption.

We shall make use of the following definitions and notation:

3. Some mathematical notation will be used in the usual way. (a,b) denotes the open interval between a and b, and $[a,b)$ the half-open interval. Note that if $a=b$, then $(a,b)=[a,b)=\Theta$, the empty set. The union of two sets is denoted by $\cup$ and the intersection by $\cap$. $r \varepsilon \mathfrak{D}$ denotes that r is an element in $\mathfrak{D}$.
4. Since a project is fully described by the associated cash flow we shall not distinguish between the project and the cash flow but denote both by a capital roman letter without subscript, e.g. A, B, C. Thus we have $A=(a_0, a_1, \ldots, a_n)$.

[1] Hirshleifer seems later to have modified his view. See Hirshleifer [9] for a more recent statement of his opinion.

5. The cash flow $A+B$ is defined by $A+B=(a_0+b_0,\ a_1+b_1,\ \ldots,\ a_n+b_n)$. $A-B$ is defined in a similar manner. Note that the cash flow resulting from accepting A and B is not necessarily $A+B$, since the cash flows may be dependent.
6. The present value function (of project A) is defined as

$$P(i)=a_0+a_1/(1+i)+\ \ldots\ +a_n/(1+i)^n\,,$$

where i is a rate of interest.
7. The present value, $P(\varrho)$, is the value of the present value function when $i=\varrho$.
8. A number r is a root in the equation $P(i)=0$ if the equation holds for $i=r$. r is a simple root if $P(i)=0$ and $P(i)/(i-r)\neq 0$ when $i=r$. r is a repeated root if $P(i)=0$ and $P(i)/(i-r)=0$ when $i=r$.
9. An internal rate of return r of the cash flow $(a_0,\ a_1,\ \ldots,\ a_n)$ is a root of the equation

$$P(i)=a_0+a_1/(1+i)+\ \ldots\ +a_n/(1+i)^n=0\,.$$

It is customary also to restrict r to some set of numbers. We shall denote this set $\mathfrak{D}$. It is a natural choice to let $\mathfrak{D}=[0,\ \infty)$ and we shall do so unless otherwise stated.
10. A cash flow $(a_0,\ a_1,\ \ldots,\ a_n)$ is said to have a unique internal rate of return if there exists one and only one r in $\mathfrak{D}$ such that r is a root in $P(i)=0$ and that moreover it is a simple root.[2]
11. In the discussion of one project, $r_1,\ r_2,\ \ldots$ will denote the different rates of return of this project. When discussing ditferent projects, $r_A,\ r_B$ etc. will denote rates of return of projects $A,\ B$ etc. In the same way $P_A(\varrho),\ P_B(\varrho)$ etc. will denote the present values of the different projects.

[2] The restriction of r to be a simple root is made to make the internal rate of return method more easily applicable. With this definition an investment project with a unique internal rate of return will have a positive present value if and only if the internal rate of return is greater than the cost of capital. This is not necessarily the case, when r is a repeated root. It is therefor unpractical to define uniqueness in such a way that cash flows with one set of repeated roots are counted as cash flows with a unique internal rate of return. See Bernhard [3] who criticizes the internal rate of return for this reason.

3. *Independent Projects with Unique Internal Rates of Return.*

It is well known that the present value method and the internal rate of return method will always lead to the same results in accept-reject decisions on independent projects with a unique internal rate of return. This will be shown in this section in essentially the same way as by Lorie and Savage [13].

Multiplication with $(1+i)^n$ and substitution of $x=1+i$ transform the equation $P(i)=0$ into the more suitable form

$$f(x)=a_0x^n+a_1x^{n-1}+\ \dots\ +a_n=0\,.$$

Then $i=x-1$ and the relevant range of values for x is $[1,\infty)$. The following theorem is well known[3] from the theory of algebraic equations:

Theorem 1. An odd number or an even number of real roots of an equation $f(x)=0$ lie between the two values $x=a$ and $x=b$ according as $f(a), f(b)$ differ in sign or have the same sign.

We shall classify the cash flows with a unique internal rate of return into two classes, according to the sign of a_0.[4] The fact that there is an odd number of roots in $[1,\infty)$ implies that a_0 and $\sum_{t=0}^{n} a_t$ never will have the same sign, since $f(1)=\sum_{t=1}^{n} a_t$ and $f(x)$ for large values of x have the same sign as a_0. We define:

1. Projects with a unique internal rate of return,

   $a_0<0$ and $\sum_{t=0}^{n} a_t \geqq 0$, are called simple investment projects.

2. Projects with a unique internal rate of return,

   $a_0>0$ and $\sum_{t=0}^{n} a_t \leqq 0$, are called simple financial projects.

It follows from the theorem that a simple investment project has a positive present value if the cost of capital ϱ is less than the internal rate of return r and a negative present value if ϱ is greater than r. Hence there is no conflict between the present value method and the internal rate of return method in this case.

[3] See e.g. Turnbull [17], page 95–96.

[4] We have for convenience assumed that $a_0 \neq 0$. If $a_0 = 0$ for some cash flow, it may be classified according to its first non-zero element.

The same line of argument holds for simple financial projects, but now the present value is negative for ϱ less than r and positive for ϱ grater than r. These projects are accepted when the internal rate of return is lower than the cost of capital.

Many writers have presented examples of cash flows with either none or multiple internal rates of return. However, the following sufficient condition shows that uniqueness holds for a large and important class of cash flows. Let $A_t = a_0 + a_1 + \ldots + a_t$, i.e. the undiscounted accumulation of the cash flow from 0 to t.

Theorem 2. A cash flow $(a_0, a_1, \ldots, a_n)$ with accumulated cash flow $(A_0, A_1, \ldots, A_n)$ will have a unique nonnegative internal rate of return if the accumulated cash flow changes sign once and $A_n \neq 0$.

A proof of this theorem is given in Norstrøm [14].

The above theorem gives sufficient but not necessary conditions for a unique internal rate of return. A general procedure for finding the number of roots of $f(x) = 0$ in any interval was found by the French mathematician Sturm in 1829 and is known as Sturm's Theorem.[5]

4. *Dependent Projects.*

Two projects A and B are dependent if the cash flow resulting from accepting them both is different from the sum of the cash flows of each separate project. The type of dependence mostly discussed in capital budgeting is the case where the projects are mutually exclusive, i.e. when only one of the projects in question may be chosen. This case is important since any decision concerning acceptance of projects — dependent or independent — may be seen as the choice between mutually exclusive projects — or sets of projects. If e.g. the projects A and B are dependent, but not mutually exclusive, the decision may be seen as the choise between: A and B; A; B; neither A nor B. Furthermore the choice between many mutually exclusive projects may be seen as a sequence of choice between pairs of mutually exclusive projects. It follows that it will be sufficient to give a treatment

[5] For a statement and proof of Sturm's theorem see e.g. Turnbull [17]. The application of Sturm's theorem to determine whether a cash flow has a unique internal rate of return has been done by Kaplan [12]. If $f(x) = 0$ has repeated roots, they are counted as one in Sturm's theorem. See Bernhard [3] and Turnbull [17].

of the choice between two mutually exclusive projects, since then in principle all cases are considered.

Formally the different investment methods may be said to consist of

1. A function which transforms the cash flow (and eventual parameters like the discount rate) into a measure of merit.
2. Rules for using this function and the measure of merit. These may include guidelines on which cash flows should enter into the function and how the measure of merit should be used to reach a decision.

There has been some disagreement on whether the internal rate of return method will yield the same decisions as the present value method in the choice between two mutually exclusive investments. This disagreement is due to different definitions of what the internal rate of return method is in this case, or more precisely — a different opinion concerning the guidelines in 2. above. A specific example will clarify the point.

A company with cost of capital $\varrho=0.10$ has the choice between the two mutually exclusive A and B with cash flows $A=(-200, 264)$, $B=(-100, 143)$.

With the above cost of capital we have

$$P_A(\varrho)=40\ ;\quad r_A=0.32$$
$$P_B(\varrho)=30\ ;\quad r_B=0.43$$

The most obvious way to use the internal rates of return is to choose the project with the highest internal rate of return. If this is the guideline for using the internal rate of return method for this case, then clearly we have presented an example showing that the present value method and the internal rate of return method may lead to different decisions.[6]

[6] This version of the internal rate of return method has been criticized by e.g. Bernhard [2] and Hirshleifer [7]. It is easy to demonstrate that it leads to inconsistent results. Let C be another project, independent of A and B, and with cash flow $C = (-100, +115)$. Since $r = 0.15$ is greater than the cost of capital, C is accepted. But B and C together has a cash flow $B + C = (-200, 258)$ which clearly is inferior to A since $a_0 = b_0 + c_0$ and $a_1 > b_1 + c_1$. To accept B and C can not be optimal since there exists an alternative which is better. (Whether there exists still better alternatives is irrelevant.) Hence it has been demonstrated that this version of the internal rate of return method may lead to non-optimal decisions.

The advocates of the internal rate of return method do not, however, follow this procedure when choosing between two mutually exclusive projects.[7] They regard the marginal cash flow A–B to be the relevant cash flow in this situation. Cash flow A is preferred to cash flow B if and only if the marginal cash flow is preferred to nothing. Given that the decision on the marginal cash flow is in accordance with the present value method, this will lead to a correct choice between A and B, because $P_{A-B}(\varrho) > 0$ if and only if $P_A(\varrho) > P_B(\varrho)$. In the above example the marginal project is $A-B=(-100, +121)$. The internal rate of return $r_{A-B}=0.21$ is greater than the cost of capital ϱ and hence project A is chosen, which is in accordance with the present value method. Note that the decision is not influenced of whether A–B or B–A is regarded as the marginal cash flow, since B–A will be rejected if and only if A–B is accepted.

Essentially the same argument holds when the cash flow in question may be continuously varied through the choice of some input variables. We shall as an illustration consider the case that the cash flow at time t is a continuously differentiable function of a single variable λ; $a_t(\lambda)$. The optimization problem consists in choosing an optimal value for λ. Let $a_t'(\lambda)$ denote the derivative of $a_t(\lambda)$ with respect to λ. The marginal cash flow is then $(a_0'(\lambda), a_1'(\lambda), \ldots, a_n'(\lambda))$, and the first order condition for optimum according to the marginal version of the internal rate of return method is that an internal rate of return of this cash flow should be equal to the cost of capital ϱ. But this is exactly the same first order condition as obtained when using the present value method, since

$$\frac{dP}{d\lambda} = \frac{d}{d\lambda}\left\{\sum_{t=0}^{n} \frac{a_t(\lambda)}{(1+\varrho)^t}\right\} = \sum_{t=0}^{n} \frac{a_t'(\lambda)}{(1+\varrho)^t} = 0,$$

which implies that ϱ is an internal rate of return of the marginal cash flow.

The marginal use of the internal rate of return in principie reduces the choice between two mutually exclusive projects to an accept-reject decision of an independent project. It has previousiy been shown that the present value method and the internal rate of return method

[7] See e. g. Grant and Ireson [6] which is a standard textbook in the field.

will give the same decision in such cases if the internal rate of return is unique. It remains to consider the case of non-uniqueness. When we in the following discuss a project, it is either an independent one or the marginal cash flow between two mutually exclusive projects.

5. *Multiple Internal Rates of Return.*

In section 3 a theorem was stated which shows that it is only in exceptional cases that a project will have none or more than one internal rate of return. It can not be denied, however, that such cases may occur, especially as marginal cash flows, and the issue of non-uniqueness is undoubtly the most serious theoretical objection against the internal rate of return method.

The perhaps most famous example of a project with multiple rates of return is due to Lorie and Savage [13] and Solomon [15]. The project considered is the installation of a larger oil pump that would get a fixed amount of oil out of the ground more rapidly than the existing pump. The marginal cash flow resulting from installation of the larger pump is $(-1{,}600, +10{,}000, -10{,}000)$. This cash flow has two rates of return; $r_1=0.25$ and $r_2=4.00$.

Examples of cash flows with more than two internal rates of return are also easily constructed. The cash flow $(-1, +6, -11, +6)$ has three rates of return; $r_1=0.00$, $r_2=1.00$ and $r_3=2.00$. Furthermore, some cash flows have no internal rate of return, e.g., $(-1, 3, -2.5)$. (The two last examples are due to Hirshleifer [7].)

The natural question to ask is now, which of the internal rates of return is the relevant one. If in the pump example the cost of capital $\varrho=0.50$, should the larger pump be rejected because $r_1<\varrho$ or accepted because $\varrho<r_2$? It has been argued by many writers that the internal rate of return method breaks down in this case.[8] One of the best known critics is Solomon [15] who answers the above question by saying: "The answer is that neither of these rates of return is a measure of investment worth, neither has relevance to the profitability of the project under consideration, and neither, therefore, is correct."[9]

[8] See e.g. Lorie and Savage [13], Solomon [15], Hirshleifer [7], and Bernhard [2].

[9] Solomon [15], page 128.

The Solomon argument against the internal rate of return is, beside the existence of multiple rates, based on the fact that an increase in the cost of the larger pump leads to an increase in one of the internal rates of return of the project. See Table 1. Solomon concludes that any definition of "profitability" that leads to such absurd results must be in error.[10]

Table 1

Cost of Pump	r_1
0	0.00
827	0.10
1600	0.25
2500	1.00

We shall not pursue this discussion of the weaknesses of the internal rate of return due to non-uniqueness any further. Interested readers are referred to the articles mentioned at the end of the introduction.

6. *An Interpretation of the Internal Rate of Return.*

The economic justification for the use of the internal rate of return has been that it represents a rate of growth. Although this is not wholly untrue, it may be misleading. As pointed out by Hirshleifer [7], a project may then have more than one rate of growth. The purpose of this section is to give another interpretation of the internal rate of return. The interpretation will be used to explain the economic reality behind an internal rate of return both in projects with a unique and in projects with multiple rates of return.

The interpretation is based on the following result.

Theorem 3. The number $r_j \varepsilon \mathfrak{D}$ is an internal rate of return of the cash flow $A=(a_0, a_1, \ldots, a_n)$ if and only if A can be decomposed into two cash flows, $A^0=(a_0^0, a_1^0, \ldots, a_n^0)$ and $A^1=(a_0^1, a_1^1, \ldots, a_n^1)$, such that

$$a_t^0+a_t^1=a_t, \qquad t=0, 1, \ldots, n,$$

$$a_{t+1}^1=-(1+r_j)a_t^0, \qquad t=0, 1, \ldots, n-1,$$

$$a_n^0=0, \; a_0^1=0.$$

[10] Solomon [15], page 128.

The proof of this theorem follows directly from the fact that a polynomial $f(x)$ is divisible by $(x-x_j)$ if and only if x_j is a root in $f(x)=0$.

For example, the cash flow (−10, −15, +25, +30) has a unique non-negative internal rate of return, $r_1=0.50$. Using this rate of retur this cash flow is decomposed into the two cash flows, (−10, −30, −20, 0) and (0, +15, +45, +30).

The important characteristic of the decomposition mentioned in the theorem is that the cash flows A^0 and A^1 except for a shift of one period have a similar development over time, so that it is meaningful to compare the scale of the two cash flows. It is exactly the relative scale of the two cash flows which is measured through the internal rate of return. Nothing can be concluded from r_j about the profitability of the project before something is known about the cash flow A^0 and hence about A^1. However, if A is regarded as an investment, a_t^0 may be interpreted to be the unrecovered investment at time t.[11] One would then expect that like in the example $a_t^0 \leqq 0$ for all t. A large positive internal rate of return will then mean high profitability as the cash flow A^1 will consist of positive elements and be large relative to A^0. Since for each t a_{t+1}^1 is $(1+r_j)$ times the absolute value of a_t^0, but occurs one period later, the benefits of A^1 will dominate over the sacrifices of A^0 if and only if $r_j > \varrho$. This is the well known rule from section 3 that an investment should be accepted if and only if its internal rate of return is greater than the cost of capital.

A similar line of argument holds for the case that $a_t^0 \geqq 0$ and $a_t^1 \leqq 0$ for all t, conditions which usually hold for loans. A^1 will as above dominate over A^0 if and only if $r_j > \varrho$, but since A^0 now represents the benefits and A^1 the sacrifices, A should be accepted if and only if $r_j < \varrho$.

The logic behind the above analysis is simple. The original cash flow is decomposed by use of the internal rate of return into two cash flows, in such a way that these are comparable, i.e. their relative size may be expressed as a number. A condition for this decomposition to be helpful is of course that it is easier to evaluate one of these cash flows than the original one.

In both cases considered above it was true that A^0 dominated over A^1 when $r_1 < \varrho$ and A^1 over A^0 when $r_1 > \varrho$. It is easy to see that this

[11] See Bernhard [2].

holds generally since nothing in the argument depend of the nature of A^0 or A^1, but only on the relation between them. A consequence of this is that from a decisionmaking point of view the only interesting internal rates of return are those who lie in the range of possible values of the cost of capital ϱ. For example, if it is known that the cost of capital always will be non-negative, it will be of little interest to decompose a cash flow using a negative internal rate of return, since A^0 will dominate over A^1 with any possible value of ϱ. The above argument is to some extent an economic justification for the practice to disregard by definition negative internal rates of return.

As an example consider again the cash flow from the example above which in addition to the internal rate of return $r_1=0.50$ has two negative ones, $r_2=-2.00$ and $r_3=-3.00$. Using r_3 the cash flow is decomposed into the two cash flows, (-10, +5, +15, 0) and (0, -20, +10, +30). The decomposition achieves virtually nothing since the critical information about the profitability of the project is hidden in the cash flows A^0 and A^1 and not contained in the internal rate of return. Further calculations are necessary to find for which values of ϱ the project is profitable.

As will be clear from the above discussion the explanatory power of the decomposition is best when it is obvious which of the cash flows A^0 and A^1 that represents the advantage and which the disadvantage. Two important classes of cash flows have this property when decomposed.

The first of these consists of all cash flows which change sign once, from negative to positive, and with total receipts exceeding total outlays. Such projects have been called conventional investments by some writers.[12] It is well known that these cash flows have a unique non-negative internal rate of return, and it is easy to prove that when they are decomposed we obtain cash flows A^0 and A^1 with $a_t^0 \leqq 0$ and $a_t^1 \geqq 0$ for all t. The cash flow in the example above belongs to this class.

The other class to be mentioned consists of projects with disposal costs, such that there are net outlays in the beginning and the end of the project's life and receipts in the middle. It is still assumed that total receipts exceed total outlays. Such projects will have one internal rate

[12] Bierman and Smidt [4].

of return in $(-1,0)$ and one in $(0,\infty)$.[13] When the cash flows are decomposed, A^0 will contain positive as well as negative elements, but the accumulated cash flow consisting of the elements $A_t^0 = a_0^0 + a_1^0 + \ldots + a_t^0$ will be negative or zero for all t. It is easily seen from this that A^0 will be a clear disadvantage for all $\varrho \geqq 0$. An example of such a cash flow is (-100, +110, +72, -18), with internal rates of return r_1=-0.80 and r_2=0.50. Using r_2 it is decomposed into (-100, -40, +12, 0) and (0, +150, +60, -18). A^0 contains one positive element, but is a clear disadvantage. The project should be accepted if and only if $r_2 > \varrho$.

We shall now use the decomposition on projects with multiple internal rates of return. Let A be a project with exactly two rates, r_1 and r_2. A can be decomposed using either of these rates. The rate not used in the decomposition will be a unique internal rate of return in A^0 and in A^1. Two cases are now possible depending on the sign of a_0: either A^0 is a simple investment project and A^1 a simple financial project, or vice versa. We may conclude that any project with two internal rates of return is a mixture of a simple investment and a simple financial project.

An example of a project with two internal rates of return is the oil pump project described in the previous section with the cash flow (-1,600, +10,000, -10,000) and the two rates r_1=0.25 and r_2=4.00. Using r_1 this cash flow decomposed into (-1,600, +8,000, 0) and (0, 2,000, -10.000).

Note that not only does the cash flow A^0 contain both negative and positive elements, there is also a change in the accumulated cash flow. It is easy to prove that this always will be the case when the project A has more than one positive internal rate of return. Hence it is no longer obvious whether a high internal rate of return indicates high or low profitability. The decomposition is still helpful, however, because it makes it possible to discuss projects with two internal rates of return in terms of the more familiar projects with a unique one.

The case $a_0 < 0$ will be considered first. Let r_1 and r_2 be numbered

[13] The fact that such cash flows have a unique internal rate of return has been proved by Jean [10]. It does also follow as a corollary to Theorem 2 in section 3. It is necessary that total receipts exceed total outlays. See the discussion between Hirshleifer [8] and Jean [11].

such that $r_1 \leqq r_2$ and let e.g. A be decomposed by use of r_1 such that $a_{t+1}^1 = -(1+r_1)a_t^0$. Since $a_0^0 = a_0$, A^0 is a simple investment project and A^1 a simple financial project, both with internal rate of return r_2. Hence A^0 is profitable and A^1 unprofitable when $r_2 > \varrho$; A^0 is unprofitable and A^1 profitable when $r_2 < \varrho$. As known from the previous discussion A^1 will dominate over A^0 when $r_1 > \varrho$, and A^0 over A^1 when $r_1 < \varrho$. These results are combined in Table 2.

Table 2.

$\varrho < r_1$	A^1 dominates	A^1 unprofitable	A unprofitable
$r_1 < \varrho < r_2$	A^0 dominates	A^0 profitable	A profitable
$r_2 < \varrho$	A^0 dominates	A^0 unprofitable	A unprofitable

The project A is unprofitable when $\varrho < r_1$, profitable when $r_1 < \varrho < r_2$ and unprofitable when $r_2 < \varrho$. The reader may easily verify that this conclusion is independent of which of the rates r_1 or r_2 is used in the decomposition.

When $a_0 > 0$, A^0 is a simple financial project which is profitable when $r_2 < \varrho$ and A^1 a simple investment project. By the same line of argument as above the following table is obtained.

Table 3.

$\varrho < r_1$	A^1 dominates	A^1 profitable	A profitable
$r_1 < \varrho < r_2$	A^0 dominates	A^0 unprofitable	A unprofitable
$r_2 < \varrho$	A^0 dominates	A^0 profitable	A profitable

Project A is now profitable when $\varrho < r_1$, unprofitable when $r_1 < \varrho < r_2$ and profitable when $r_2 < \varrho$.

The analysis holds also for the case $r_1 = r_2$, i.e. repeated roots. Then no value of ϱ will satisfy $r_1 < \varrho < r_2$ and a project with e.g. $a_0 < 0$ will not be profitable for any value of ϱ.

The above analysis has shown that the occurrence of two internal rates of return may be explained by the fact that the project in question is a mixture between an investment and a financial project. Both the internal rates of return are relevant when making a decision on such a project.[14] Moreover, just as the unique internal rate of return does in a simple investment or financial project, these rates determine the range of values of the cost of capital which makes the project profitable.

The analysis could be extended to a discussion of projects with three internal rates of return in terms of projects with two rates and so on. In this way it is possible by economic arguments to obtain decision rules for projects with any number of internal rates of return. With these decision rules it is possible to use the internal rate of return method correctly. A different approach will be taken here. In the next section a measure of investment worth will be suggested, which may be said to be another version of the internal rate of return, since it contains essentially the same information.

7. *A Modification of the Internal Rate of Return Method.*

Before turning to our suggestion for a modification of the internal rate of return method, two other approaches will be mentioned briefly. To solve the problem of dual rates of return in the oil pump example Solomon [15] introduces the reinvestment rate k. He defines a new rate of return p as the root of the equation

$$-1{,}600(1+i)^2+10{,}000(1+k)-10{,}000=0\,.$$

With the assumptions taken in this article the reinvestment rate will be equal to the cost of capital. The project is accepted if and only if $p>\varrho$. It is easy to see that p is unique for a given value of k and that the approach leads to a decision in accordance with the present value method.

A related approach has been taken by Teichroew, Robichek and Montalbano [16]. In cases of non-uniqueness they define a measure q, which in the pump example is the root in the equation

$$[-1{,}600(1+i)+10{,}000](1+k)-10.000=0$$

[14] The fact that an internal rate of return also have some significance when it is not unique, has been pointed out by Wright [18, 19].

Teichroew, Robichek and Montalbano's approach contains too many ideas to be given an adequate treatment here.[15] The important points are again that q is unique for a given value of k, and that the approach leads to the same decisions as the present value method.

Both the approaches mentioned above represent modifications of the internal rate of return method, which eliminates the problem of non-uniqueness of the internal rate of return. However, both have also the weakness that the rate p or q, which is used when a project has multiple rates of return is a function not of the cash flow alone, but also of the reinvestment rate — or with our assumptions — the cost of capital. This violates one of the fundamental characteristics of the internal rate of return mentioned in the introduction, and makes the method less useful when we want to know for which values of the cost of capital a project is profitable.

It was mentioned in section 4 that the different investment methods may be said to consist of a function which defines a measure of merit and certain rules for reaching a decision from this measure. The most obvious is to choose a real-valued function such that the measure of merit is a real number. As pointed out in last section, however, the information given by the internal rate of return method is the values of the cost of capital which makes the project profitable. This suggests that we should let the measure of merit be this set, and make the function used a set-valued one.

It is assumed that $\mathfrak{D}$ is an interval which includes all possible values of the cost of capital ϱ. We define

$$\begin{aligned} \mathfrak{A}(A) &= \{i \mid i \varepsilon \mathfrak{D} \text{ and } P(i) > 0\}, \\ \mathfrak{J}(A) &= \{i \mid i \varepsilon \mathfrak{D} \text{ and } P(i) = 0\}, \\ \mathfrak{R}(A) &= \{i \mid i \varepsilon \mathfrak{D} \text{ and } P(i) < 0\}. \end{aligned}$$

We shall call $\mathfrak{A}(A)$ the acceptance range, $\mathfrak{J}(A)$ the indifference range and $\mathfrak{R}(A)$ the rejection range. $\mathfrak{A}(A)$ is the proposed measure of investment worth.

[15] One of the important contributions in Teichroew, Robichek and Montalbano [16] is a clear demonstration of the fact that a project has to be a mixture between an investment and a financial project in order to have more than one internal rate of return.

An independent project is accepted if and only if $\varrho \varepsilon \mathfrak{A}(A)$. The project is rejected if $\varrho \varepsilon \mathfrak{R}(A)$ and one is indifferent with respect to the project if $\varrho \varepsilon \mathfrak{J}(A)$.

It is easy to see that the choice of the interval $\mathfrak{D}$ has no influence on the decision as long as $\mathfrak{D}$ contains all possible values of the cost of capital. The choice of $\mathfrak{D}$ is therefore primarily one of convenience.

Certain properties of the above sets are immediate.

$$\mathfrak{A}(A) \cup \mathfrak{J}(A) \cup \mathfrak{R}(A) = \mathfrak{D}, \tag{1}$$

$$\mathfrak{A}(A) \cap \mathfrak{J}(A) = \mathfrak{A}(A) \cap \mathfrak{R}(A) = \mathfrak{J}(A) \cap \mathfrak{R}(A) = \Theta, \tag{2}$$

$$\mathfrak{A}(A) = \mathfrak{R}(-A), \mathfrak{J}(A) = \mathfrak{J}(-A) \,. \tag{3}$$

(1) and (2) show that ϱ is a number of one and only one of the three sets $\mathfrak{A}(A)$, $\mathfrak{J}(\mathfrak{A})$ and $\mathfrak{R}(A)$, such that the decision rule lead to a unique decision.

In the choice between two mutually exclusive projects B and C, B is accepted if $\varrho \varepsilon \mathfrak{A}(B-C)$, C is accepted if $\varrho \varepsilon \mathfrak{R}(B-C)$, and one is indifferent between B and C if $\varrho \varepsilon \mathfrak{J}(B-C)$. (3) assures that the decision is not influenced by whether $(B-C)$ or $(C-B)$ is taken to be the marginal project.

The approach we here have suggested will obviously lead to the same decisions as the present value method. Its advantage over the internal rate of return method is that the acceptance range is unique. The information given by the set $\mathfrak{A}(A)$ is, however, essentially identical with that given by the internal rates of return. This will be clear from the following theorem.

Theorem 4. Let A be a cash flow with $m \geqq 1$ internal rates of return in the interval $\mathfrak{D} = [0, \infty)$. Let the rates of return be numbered such that $r_1 \leqq r_2 \leqq \dots \leqq r_m$. Then the acceptance range of the project, $\mathfrak{A}(A)$, is given by

1) If $a_0 < 0$ and m odd: $\mathfrak{A}(A) = [0, r_1) \cup (r_2, r_3) \cup \dots \cup (r_{m-1}, r_m)$.
2) If $a_0 < 0$ and m even: $\mathfrak{A}(A) = (r_1, r_2) \cup (r_3, r_4) \cup \dots \cup (r_{m-1}, r_m)$.
3) If $a_0 > 0$ and m odd: $\mathfrak{A}(A) = (r_1, r_2) \cup (r_3, r_4) \cup \dots \cup (r_m, \infty)$.
4) If $a_0 > 0$ and m even: $\mathfrak{A}(A) = [0, r_1) \cup (r_2, r_3) \cup \dots \cup (r_m, \infty)$.

Proof: The theorem is an immediate consequence of theorem 1 in section 3. It may also be proved by induction using the decomposition principle used in the previous section. We shall do this for the case that m is odd and $a_0 < 0$, i.e. case 1. The proofs of the other cases are similar.

Decompose A into A^0 and A^1 using the internal rate of return r_m. Since m is odd and $a_0 < 0$, $(m-1)$ will be even and $a_0^0 < 0$. Hence A^0 represents case 2 and A^1 case 4 in the theorem. By induction hypothesis $\mathfrak{A}(A^0) = (r_1,r_2) \cup (r_3,r_4) \cup \ldots \cup (r_{m-2},r_{m-1})$ and $\mathfrak{A}(A^1) = [0,r_1) \cup (r_2,r_3) \cup \ldots \cup (r_{m-1},\infty)$. A^1 dominates in the interval $[0,r_m)$ and hence A will be profitable in $\{\mathfrak{A}(A^1) \cap [0,r_m)\} = [0,r_1) \cup (r_2,r_3) \cup \ldots \cup (r_{m-1},r_m)$. A^0 dominates in the interval (r_m,∞), but since $r_{m-1} \leqq r_m$ A^0 and hence A will be unprofitable in all of this interval. Combining the results prove case 1.

We have in the theorem assumed that $\mathfrak{D} = [0,\infty)$. It is easily extended to other choices of $\mathfrak{D}$. Note that $\mathfrak{A}(A)$ is defined also when there is no internal rate of return. When $\mathfrak{D} = [0,\infty)$, $\mathfrak{A}(A) = \Theta$ if $a_0 < 0$ and $\mathfrak{A}(A) = \mathfrak{D}$ if $a_0 > 0$.

As a consequence of the uniqueness of $\mathfrak{A}(A)$ certain apparent weaknesses in the usual internal rate of return method are eliminated. It was mentioned in section 5 that one of Solomon's arguments against the internal rate of return method was that an increase in the cost of the large oil pump resulted in an increase in one of the rates of return. (See Table 1, section 5.) This paradox is eliminated by the introduction of the acceptance range. We see from Table 4 that although r_1 increases with an increase in the cost of the pump, the acceptance range $\mathfrak{A}(A)$ decreases.

Table 4.

Cost of Pump	r_1	r_2	$\mathfrak{A}(A)$
0	0.00	∞	$(0,\infty)$
827	0.10	10.00	(0.10,10.00)
1600	0.25	4.00	(0.25, 4.00)
2500	1.00	1.00	Θ
Cost > 2500	No rate of return		Θ

Norwegian School of Economics and Business Administration,

Bergen.

REFERENCES

[1] BAILEY, MARTIN J., "Formal Criteria for Investment Decisions." *Journal of Political Economy*, LXVII, No. 6, October 1959.

[2] BERNHARD, RICHARD H., "Discount Methods for Expenditure Evaluation — A Clarification of Their Assumptions." *The Journal of Industrial Engineering*, Volume XIII, No. 1, January—February 1962.

[3] Bernhard, Richard H., "On the Consistency of the Soper and Sturm-Kaplan Conditions for Uniqueness of the Rate of Return." *The Journal of Industrial Engineering*, Volume XVIII, No. 8, August 1967.

[4] BIERMAN, HAROLD, JR., and SEYMOR SMIDT, *The Capital Budgeting Decision*. 2nd Ed., Macmillan, 1966.

[5] DEAN, JOEL, *Capital Budgeting*. Columbia University Press, New York, 1951.

[6] GRANT, EUGENE L., and W. GRANT IRESON, *Principles of Engineering Economy*. 4th Ed., The Ronald Press, New York, 1960.

[7] HIRSHLEIFER, J., "On the Theory of Optimal Investment Decisions." *Journal of Political Economy*, LXVI, No. 5, August 1958.

[8] HIRSHLEIFER, J., "On Multiple Rates of Return: Comment." *Journal of Finance*, Vol. XXIV, No. 1, March 1969.

[9] HIRSHLEIFER, J., *Investment, Interest and Capital*. Prentice Hall Inc., Englewood Cliffs, N.J., 1970.

[10] JEAN, WILLIAM H., "On Multiple Rates of Return." *Journal of Finance*, Vol. XXIII, No. 1, March 1968.

[11] JEAN, WILLIAM H., "Reply". *Journal of Finance*, Vol. XXIV, No. 1, March 1969.

[12] KAPLAN, SEYMOR, "A Note on a Method for Precisely Determining the Uniqueness or Nonuniqueness of the Internal Rate of Return for a Proposed Investment." *The Journal of Industrial Engineering*, Volume 16, No. 1, January—February 1965.

[13] LORIE, J. and SAVAGE, L. J., "Three Problems in Capital Rationing." *Journal of Business*, Vol. XXVIII, No. 4, October 1955.

[14] NORSTRØM, CARL J., "A Sufficient Condition for a Unique Internal Rate of Return." Forthcoming in *The Journal of Financial and Quantitative Analysis*.

[15] SOLOMON, EZRA, "The Arithmetic of Capital Budgeting Decisions." *Journal of Business*, XXIX, No. 2 April 1956.

[16] TEICROEW, D., ROBICHEK, A. A., and MONTALBANO, M., "An Analysis of Criteria for Investment and Financing Decisions under Certainty." *Management Science*, Vol. 12, No. 3, November 1965.

[17] TURNBULL, H. W., *Theory of Equations*. 5th. Ed., Oliver and Boyd, Edinburgh and London, 1952.

[18] WRIGHT, J. F., "Notes on the Marginal Efficiency of Capital." *Oxford Economic Papers*, Vol. XV, July 1963.

[19] WRIGHT, J. F., "Some Further Comments on the Ambiguity and Usefulness of Marginal Efficiency as an Investment Criterion." *Oxford Economic Papers*, Vol. XVII, March 1965.

A mathematical connection between the present value, the rate of return and the scale of an investment

by Carl J. Norström
Assistant Professor, Norwegian School of Economics and Business Administration, Bergen

Abstract: The purpose of this note is to show how discrepancies between the present values and rate of returns of investments are due to differences in scale. A measure of the scale of a project is introduced and a mathematical connection between the present value, rate of return and scale established.

THE two most important measures for evaluating investments are the present value and the (internal) rate of return. It is well known that these measures are not in complete correspondence: an investment A may have a higher present value than investment B, while B has a higher rate of return than A.[1] A bit imprecisely the reason for this may be said to be that the rate of return measures only the quality of the investment, while the present value takes into consideration both the quality and the scale.[2] The purpose of this note is to make this idea more precise and show that it indeed is a correct one.

ONE-PERIOD INVESTMENTS

It is convenient first to discuss the simple case of one period investments, where an amount is invested at time 0 and the benefits from the investment are received at time 1, one period later. The cash flow from a project is of the form $A=(a_0, a_1)$, where for a usual investment $a_0<0$ and $a_1>0$. Let ρ denote the market interest rate, and r the rate of return of the cash flow. The present value of the project is

$$V(A) = a_0+\frac{a_1}{(1+\rho)}, \tag{1}$$

and the rate of return is given by

$$0 = a_0+\frac{a_1}{(1+r)} \tag{2}$$

In the one-period case the investment outlay $(-a_0)$ is an obvious measure of the scale of the investment.

[1] It is not an issue here whether the two measures lead to different decisions. A discussion of that question may be found in eg, Bailey [1], Bernhard [2], Hirshleifer [4, 5], Lorie and Savage [6], Solomon [7] and Teichroew, Robichek and Montalbano [8].

[2] See Bierman and Smidt [3], pp. 40–41.

Using (2) a_1 may be eliminated from (1) by regular substitution. This gives

$$V(A) = \frac{(r-\rho)}{(1+\rho)}(-a_0) \tag{3}$$

Equation (3) states that the present value of a project is equal to the difference between the rate of return and the market interest rate, discounted one period, times the scale of the project measured by the investment outlay. (3) makes, for the one-period case, explicit and precise how the present value takes into account both the quality and the scale of a project, and explains why one project may have a larger rate of return than another and yet a smaller present value.

MULTI-PERIOD INVESTMENTS

The ideas in the previous section will now be generalized to multiperiod investments. The following theorem will play a central part in the discussion:

Theorem. The number r is a rate of return of the cash flow $A=(a_0, a_1,\ldots, a_n)$ if and only if A can be decomposed into two cash flows $A^0=(a_0^0, a_1^0,\ldots, a_{n-1}^0)$ and $A^1=(a_1^1, a_2^1,\ldots, a_n^1)$ such that

$$a_0^0 = a_0 \tag{4a}$$

$$a_t^0+a_t^1 = a_t \qquad t = 1, 2,\ldots, n-1 \tag{4b}$$

$$a_n^1 = a_n \tag{4c}$$

$$a_{t+1}^1 = (1+r)(-a_t^0) \qquad t = 0, 1,\ldots, n-1 \tag{5}$$

The proof of the theorem follows directly from the fact that a polynomial $f(x)$ is divisible by $(x-\alpha)$ if and only if α is a root in the equation $f(x)=0$.

We note that the first element of the cash flow A^0

refers to the present, while the first element in A^1 takes place at time 1. The cash flow A^0 plays a role similar to a_0 in the one-period investments, and A^1 to a_1. (When $n=1$, $A^0=(a_0)$ and $A^1=(a_1)$).

The important characteristic of the decomposition is that A^0 and A^1 have a similar development over time. Mathematically they belong to a one-dimensional vectorspace, and it is therefore meaningful to compare their relative size. It is exactly this relative size which is reflected in the rate of return r.

An Example.

It is easy to verify that the cash flow

$$A = (-1000, -1500, +2500, +3000)$$

has a rate of return $r=0{\cdot}50$. Decomposing A using this rate gives

$$A^0 = (-1000, -3000, -2000)$$
$$A^1 = (+1500, +4500, +3000)$$

The reader may verify that the decomposition satisfies (4)–(5).

We want to find a measure of the scale of the investment which is an extension of the investment outlay $(-a_0)$, the measure in the one-period case. It is easily seen that the cash flow A of the actual investment project is identical with the total cash flow resulting from n hypothetical, one-period investments where the tth project takes place at time t and has cash flow (a_t^0, a_{t+1}^1). The natural economic interpretation of $A^0=(a_0^0, a_1^0, \dots, a_{n-1}^0)$ is therefore that $(-a_t^0)$ is the unrecovered investment at time t of the actual project.[3] We shall measure the scale of the actual project by the present value of the unrecovered investments

$$V(-A^0) = \sum_{t=0}^{n-1} \frac{(-a_t^0)}{(1+\rho)^t} \qquad (6)$$

It is now possible to derive a result for multiperiod investments, which is similar to (3) and include this equation as a special case. It follows easily from (4)–(5) that the present value of the actual project is

$$V(A) = \frac{(r-\rho)}{(1+\rho)} \cdot V(-A^0) \qquad (7)$$

[3] This interpretation is used in eg, Bernhard [2].

The economic interpretation of (7) is similar to that of (3). The present value of a project is the product of the quality measured by $(r-\rho)$ and the scale measured by $(V(-A^0))$, discounted one period. Eventual discrepancies between the present values and rates of return of two projects are the result of differences in scale.

The economic interpretation of A^0 suggests that a_t^0 is negative and hence $-a_t^0$ positive, for all t. It may be proved that this will be the case for all cash flows $A=(a_0, a_1, \dots, a_n)$ with one change of signs. In this case it is obvious that $V(-A^0)$ is positive for all non-negative values of ρ, and that the information of whether the project has a positive present value or not is contained in r. Although not so obvious from an economic point of view, the same is true for cash flows where some a_t^0 are positive as long as A has a unique rate of return and a_0 is negative.

The economic interpretation of a positive a_t^0 is that the unrecovered investment is negative, or that the project at this point of time is a loan. It is well known that there exist projects with more than one rate of return. If such a cash flow is decomposed using one of these rates, the remaining rates will also be rates of return in the cash flow A^0. $V(-A^0)$ will then be positive or negative according to the value of ρ, reflecting the well known fact that projects with multiple rates of return will act predominantly as an investment for some interest rates and as a loan for others.

REFERENCES

[1] Bailey, Martin J., "Formal Criteria for Investment Decisions", *Journal of Political Economy*, LXVII, No. 6, October 1959.

[2] Bernhard, Richard H., "Discount Methods for Expenditure Evaluation—A Clarification of Their Assumptions", *The Journal of Industrial Engineering*, Volume XII, No. 1, January-February 1962.

[3] Bierman, Harold Jr. and Smidt, Seymor, *The Capital Budgeting Decision*, 3rd ed., Macmillan, 1971.

[4] Hirshleifer, Jack, *Investment, Interest and Capital*, Prentice Hall, Inc., 1970.

[5] Hirshleifer, Jack, "On the Theory of Optimal Investment Decisions", *Journal of Political Economy*, LXVI, No. 5, August 1958.

[6] Lorie, James H., and Savage, Leonard J., "Three Problems in Capital Rationing", *Journal of Business*, Vol. XXVIII, No. 4, October 1955.

[7] Solomon, Ezra, "The Arithmetic of Capital Budgeting Decisions", *Journal of Business*, Vol. XXIX, No. 2, April 1956.

[8] Teichroew, Daniel, Robichek, Alexander A., and Montalbano, Michal, "An Analysis of Criteria for Investment and Financing Decisions under Certainty", *Management Science*, Vol. 12, No. 3, November 1965.

22.

UNIQUENESS OF THE INTERNAL RATE OF RETURN WITH VARIABLE LIFE OF INVESTMENT: A COMMENT [1]

IN the article " Uniqueness of the Internal Rate of Return with Variable Life of Investment " published in the September 1969 issue of the ECONOMIC JOURNAL, K. J. Arrow and D. Levhari consider the uniqueness of the internal rate of return when it is possible to truncate the investment project at any moment of time. It has earlier been proved by Soper [3] and Karmel [2] that if the investor chooses the truncation period so as to maximise the internal rate of return, then the truncated project has a unique internal rate of return. Arrow and Levhari point out that with a perfect capital market the truncation period should be chosen so as to maximise the present value of the truncated project and not its internal rate of return. They then go on to prove that if, with a given constant rate of discount, the truncation period is chosen so as to maximise the present value of the project, then the internal rate of return is unique.

The purpose of this comment is to demonstrate that the internal rate of return Arrow and Levhari find, in reality, is Soper's maximal internal rate of return, *i.e.*, the one that is obtained when the truncation period is chosen to maximise the internal rate of return.

In correspondence with Arrow and Levhari's article let $x(t)$ be a given continuous stream of net income and define

$$(1) \qquad \phi(r,T) = \int_0^T x(t)e^{-rt}dt$$

Thus $\phi(r, T)$ is the present value of the stream $x(t)$ when the discount rate is r and the truncation period T. Further, let

$$(2) \qquad \psi(r) = \underset{T}{\mathrm{Max}}\, \phi(r, T)$$

$$(3) \qquad T(r) = \{T|\phi(r, T) \text{ Max}\}$$

$\psi(r)$ is the maximum present value, for a given r, with the appropriate choice of the truncation period, T; and $T(r)$ is the set of truncation periods leading to the highest present value when the discount rate is r.

On the other hand, let T^* be one of the truncation periods obtained when the internal rate of return of the truncated project is maximised, and let r^* be the corresponding unique maximal internal rate of return.

In their article, Arrow and Levhari show that the maximum present value $\psi(r)$ is a decreasing monotonic function of the discount rate r, and hence that there is at most one solution of $\psi(r) = 0$. This result is not

[1] The author wishes to acknowledge valuable comments by Professor K. J. Arrow on the first version of this note.

disputed.[1] What is claimed here is that the solution of $\psi(r) = 0$ (if such a solution exists) is exactly Soper's maximal internal rate of return r^*, *i.e.*, that $\psi(r^*) = 0$.

In the following it will be assumed that the net stream $x(t)$ is negative in some interval before it turns positive for the first time, because otherwise there will be no finite solution of the equation $\psi(r) = 0$. This assumption assures that for any truncation period T, $(\phi(r, T)$ is negative when r is large enough.[2]

It will now be shown that $\psi(r^*) = 0$. From the definition of r^* and T^* it follows that $\phi(r^*, T^*) = 0$. Let us show that $\phi(r^*, T^*) = \underset{T}{\text{Max}}\, \phi(r^*, T)$. Suppose, to the contrary, that there exists a truncation period T_1 such that $\phi(r^*, T_1) > \phi(r^*, T^*) = 0$. Since $\phi(r, T_1)$ by the assumption above will be negative when r is large enough, there exists an $r_1 > r^*$ such that $\phi(r_1, T_1) = 0$. But this is impossible, since r^* is the largest possible internal rate of return. Hence $\phi(r^*, T^*) = \underset{T}{\text{Max}}\, \phi(r^*, T)$. It follows that

$$\psi(r^*) = \underset{T}{\text{Max}}\, \phi(r^*, T) = \phi(r^*, T^*) = 0. \tag{4}$$

This concludes the proof.

CARL J. NORSTRØM

The Norwegian School of Economics and Business Administration.

REFERENCES

1. K. J. Arrow and D. Levhari, " Uniqueness of the Internal Rate of Return with Variable Life of Investment," ECONOMIC JOURNAL, September 1969, pp. 560–6.
2. P. H. Karmel, " The Marginal Efficiency of Capital," with Appendix by B. C. Rennie, *Economic Record*, Vol. XXXV, No. 72 (1959), 423–34.
3. C. S. Soper, " The Marginal Efficiency of Capital: A Further Note," ECONOMIC JOURNAL, March 1959, pp. 174–7.

[1] Arrow and Levhari thus define the internal rate of return in the following way. They compute for any given rate of interest the optimal truncation period, and define the internal rate of return to be the rate of interest which, when the truncation period has been optimised with respect to it, yields a present value of zero.

[2] If $x(t) \geqslant 0$, $0 \leqslant t \leqslant \tau$, and $\int_0^\tau x(t)dt > 0$ for some $\tau > 0$, then for any finite r

$$\psi(r) = \underset{T}{\text{Max}}\, \phi(r, T) \geqslant \phi(r, \tau) > 0.$$

If, on the other hand, there exists a $\tau > 0$ such that $x(t) \leqslant 0$, $0 \leqslant t \leqslant \tau$, and $\int_0^\tau x(t)dt < 0$, then for any finite T and bounded function $x(t)$ there will be an r such that $\int_0^T x(t)e^{-rt}dt < 0$, since $\int_\tau^T x(t)e^{-rt}dt$ approaches zero faster than $\int_0^\tau x(t)e^{-rt}dt$ when r approaches infinity.

JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS
June 1972

A SUFFICIENT CONDITION FOR A UNIQUE NONNEGATIVE INTERNAL RATE OF RETURN

*Carl J. Norstrøm**

I. Introduction

A proposition is proved which shows that each member of an important class of investment and financing projects has a unique nonnegative internal rate of return. Nonuniqueness of the internal rate of return is thus shown to occur less frequently than formerly believed. The correspondence between the proposition and previous results on the uniqueness of the internal rate of return is briefly indicated.

II. Method

One of the problems in applying the internal rate of return in the evaluation of investment projects is that it is not necessarily unique. This fact has been noted by a number of writers, who have presented examples of cash flows with either no or more than one rate of return.[1] We shall here prove a sufficient condition for uniqueness, which shows that the members of a large and important class of cash flows do have a unique rate of return.

Consider an investment or financing project with cash flow $\{a_o, a_1, \ldots, a_n\}$. An internal rate of return is usually defined as a rate of interest r having the property that

$$a_o + a_1(1+r)^{-1} + \ldots + a_n(1+r)^{-n} = 0. \tag{1}$$

Since an n'^{th} order equation always will have n roots, real or complex, the range of r has to be restricted if the question of uniqueness is to be meaningful. We shall restrict r to being a nonnegative real number. The cash flow $\{a_o, a_1, \ldots, a_n\}$ is said to have a unique internal rate of return if there is

**University of Michigan.*

[1]See, e.g., Wright [12], Samuelson [8], Lorie and Savage [6], Solomon [9], and Hirshleifer [3].

one and only one $r \geqq 0$, such that (1) holds and moreover that this r is a simple root in (1).[2]

Multiplication with $(1+r)^n$ and substitution of $x = 1+r$ transforms (1) into the more suitable form

$$(2) \qquad f(x) = a_o x^n + a_1 x^{n-1} + \ldots + a_n = 0 .$$

Here $r = x-1$ and the relevant range of values for x is $(1, \infty)$. We define

$$(3) \qquad A_t = \sum_{\tau=0}^{t} a_\tau .$$

A_t is the undiscounted accumulation of the cash flow from $\tau = 0$ to $\tau = t$. The main result in this note is:

Proposition 1. A cash flow $\{a_o, a_1, \ldots, a_n\}$ with accumulated cash flow $\{A_o, A_1, \ldots, A_n\}$ will have a unique nonnegative internal rate of return if the accumulated cash flow changes signs once and $A_n \neq 0$.

Proof. Since $f(x) = 0$ is equivalent to $-f(x) = 0$, it may be assumed without loss of generality that $A_o = a_o < 0$ and hence $A_n > 0$. There is at least one root of f(x) in $(1, \infty)$, since $f(1) = A_n$ and f(x) will have the same sign as A_o when x is large enough. $x = 1$ is not a root since $A_n \neq 0$.

It remains to show that there is at most one root. Suppose that there is more than one root and that these roots are numbered $x_1 \leqq x_2 \leqq \ldots$. It will be shown that $x_k > 1$ and $f(x_k) = 0$ implies that the derivative $f'(x_k) < 0$. This implies that x_1 is unique. $f'(x_1) < 0$ implies in the first place that x_1 is a simple root and moreover that $f(x) < 0$ for $x_1 < x < x_2$, but, then must $f'(x_2) \geqq 0$, which is a contradiction.

We have

$$(4) \qquad f(x) = (x-1)\ (A_o x^{n-1} + A_1 x^{n-2} + \ldots + A_{n-1}) + A_n.$$

Define

[2]The restriction of r to being a simple root in (1) is made to make the internal rate more easily applicable. With this definition an investment project with a unique internal rate of return will have a positive present value if and only if the internal rate of return is greater than the calculation rate. This is not necessarily the case when r is a repeated root in (1). It is therefore impractical to define uniqueness in such a way that cash flows with one set of repeated roots are counted as cash flows with a unique internal rate of return. See Bernhard, [2], who criticizes the internal rate of return for this reason.

(5) $$g(x) = A_o x^{n-1} + A_1 x^{n-2} + \ldots + A_{n-1}$$

such that

(6) $$f(x) = (x-1)g(x) + A_n, \text{ and}$$

(7) $$f'(x) = (x-1)g'(x) + g(x).$$

Let $x_k > 1$ and $f(x_k) = 0$. We shall show that $g(x_k) < 0$ and $g'(x_k) < 0$ such that from (7) $f'(x_k) < 0$. $g(x_k) < 0$ follows immediately from (6) and $A_n > 0$. As there is only one change of signs in the accumulated cash flow, there must be an integer m such that

$$A_t \leq 0 \quad \text{for } t = 0, 1, \ldots, m, \text{ and}$$
$$A_t \geq 0 \quad \text{for } t = m+1, \ldots, n.$$

Hence,

(8) $$x_k g'(x_k) - (n-m-1)g(x_k) = mA_o x_k^{n-1} + \ldots + A_{m-1}x_k^{n-m} - A_{m+1}x_k^{n-m-2} - \ldots - (n-m-1)A_{n-1} < 0,$$

since all the coefficients are negative. Since $n-m-1 \geq 0$ and $g(x_k) < 0$, it follows from (8) that $g'(x_k) < 0$. Q.E.D.

The proposition gives sufficient but not necessary conditions for a unique internal rate of return.[3]

We shall now briefly consider the correspondence between Proposition 1 and some other results on the uniqueness of the internal rate of return.

Descartes Rule of Signs[4] states that the number of positive roots in (2) cannot exceed the number of signs of the coefficients. Hence, there will be at most one internal rate of return if the cash flow $\{a_o, a_1, \ldots, a_n\}$ changes signs once. The advantage of this result over Proposition 1 lies in that it is applicable to interest rates $r > -1$ (not only to $r \geq 0$). When r is restricted to being nonnegative, Descartes's Theorem obviously is less general.

[3]The cash flow {-1, 2, -2, 4} has a unique internal rate of return, r = 1.0, although the accumulated cash flow {-1, 1, -1, 3} changes sign three times.

[4]See Turnbull [11], pages 99-102.

It follows as a corollary to Proposition 1 that an investment will have a unique nonnegative internal rate of return when there are two changes of signs in the original casle flow $\{a_o, a_1, \ldots, a_n\}$ and moreover $A_o = a_o$ and $A_n = a_o + a_1 + \ldots + a_n$ have different signs. This result is immediate when one notices that the above conditions imply that the accumulated cash flow $\{A_o, A_1, \ldots, A_n\}$ will change signs once and that $A_n \neq 0$. The corollary is of some interest since it covers the case where there is an outlay at the end of an investment's life, due, e.g., to disposal costs. A slightly less general result than the corollary has previously been proved by Jean [3].

The French mathematician Sturm found in 1829 a method for finding the exact number of distinct roots of (2) within any given range of values. Thus, Sturm's theorem may be used to find for any given cash flow whether it has none, one, or many rates of return.[5] The weakness of Sturm's theorem is that it is relatively cumbersome to apply and that it has no immediate economic interpretation.

Proposition 1 provides a convenient way of checking uniqueness, which may be incorporated easily in computer programs calculating the internal rate of return. Its economic significance lies in showing that the class of projects with a unique internal rate of return is much larger than the class of projects with one change of signs in the original cash flow. A project with nonunique internal rates of return must have an accumulated cash flow which changes signs more than once or not at all. This result is in accordance with the arguments made by Teichroew, Robichek, and Montalbano [10] and Bernhard [1] that a project has to be a mixture of an investment and a loan in order not to have a unique internal rate of return. Teichroew, Robichek, and Montalbano [10] have given an extended version of the internal rate of return method which yields the same decisions as the present value method also for such projects.

REFERENCES

[1] Bernhard, R. H. "Discount Methods for Expenditure Evaluation - A Clarification of Their Assumptions." *The Journal of Industrial Engineering*, Volume XII, No. 1, January-February 1962.

[2] __________. "On the Consistency of the Soper and Sturm-Kaplan Conditions for Uniqueness of the Rate of Return." *The Journal of Industrial Engineering*, Volume XVIII, No. 8, August 1967.

[3] Hirshleifer, J. "On the Theory of Optimal Investment Decision." *The Journal of Political Economy*, LXVI, No. 5, August 1958.

[4] Jean, W. H. "On Multiple Rates of Return." *The Journal of Finance*, Volume XXIII, No. 1, March 1968.

[5] For a statement and proof of Sturm's theorem see, e.g., Turnbull [11]. The application of Sturm's theorem to determine whether a cash flow has a unique internal rate of return has been done by Kaplan [5]. If f(x) has repeated roots, they are counted as one in Sturm's theorem. See Bernhard [2] and Turnbull [11].

[5] Kaplan, S. "A Note on a Method for Precisely Determining the Uniqueness of Nonuniqueness of the Internal Rate of Return for a Proposed Investment." *The Journal of Industrial Engineering*, Volume 16, No. 1, January-February 1965.

[6] Lorie, J. and L. J. Savage. "Three Problems in Capital Rationing." *The Journal of Business*, Volume XXVIII, No. 4, October 1955.

[7] Norstrøm, C. J. "Noen betraktninger om beregning og bruk av internrenten." Unpublished discussion paper, NHH, 1966.

[8] Samuelson, P. A. "Some Aspects of a Pure Theory of Capital." *The Quarterly Journal of Economics*, Volume LI, May 1937.

[9] Solomon, E. "The Arithmetic of Capital Budgeting Decisions." *Journal of Business*, Volume XXIX, No. 2, April 1956.

[10] Teichroew, D.; A. A. Robichek; and M. Montalbano. "An Analysis of Criteria for Investment and Financing Decisions under Certainty." *Management Science*, Volume 12, No. 3, November 1965.

[11] Turnbull, H. W. "Theory of Equations," 5th ed. Edinburgh: Oliver and Boyd, 1952.

[12] Wright, C. A. "A Note on 'Time and Investment'." *Economica*, New Series, Volume III, 1936.

Artikkelen finnes her

A COMMENT ON TWO SIMPLE DECISION RULES IN CAPITAL RATIONING*

uction

l rationing problem, optimal allocation of scarce
available investment projects, was introduced by
Savage in their famous article [5]. Weingartner [7]
an elegant solution to the problem through the use
or integer programming. His contribution and later
y Baumol and Quandt [1], Myers [6] and others, have
panded the understanding of the capital rationing
l related problems in finance. The application of
ch in real life situations is, however, severly limited by the strong assumptions about the available information; that the cash flows of all future projects up to the planning horizon of the firm is known. This requirement will not be satisfied in most real situations, and the firm is left with the choice between more conventional procedures, like the present value per dollar outlay or the internal rate of return method.[1] The purpose of this article is to give conditions under which these simpler methods lead to essentially the same solution as Weingartner's linear programming approach. The conditions will be stated in terms of the shadow discount rates from the linear programming solution.

---

[1] A similar view on this point has been expressed e.g. by Hughes and Lewellen [4].

* Forthcoming in the Journal of Business Finance and Accounting. The comments of Karl Borch, Steinar Ekern, Jan Mossin and Cary L. Sundem are gratefully acknowledged.

2. Formulation of the Problem

In its typical linear programming form the general capital rationing problem is

(1) $$\text{Max} \sum_{j=1}^{N} V_j(k_1,\ldots,k_T)\cdot x_j$$

subject to

(2) $$-\sum_{j=1}^{T} a_{jt}x_j \leqq b_t \quad , \quad t = 0,1,\ldots,T$$

(3) $$0 \leqq x_j \leqq 1 \quad , \quad j = 1,2,\ldots,N$$

Here N is the number of projects, T the planning horizon and

k_t - a discount rate for period t,

a_{jt} - the net cash flow obtained from a unit of project j during period t,

x_j - the accepted proportion of project j,

b_t - amount of cash made available from sources external to the project during period t.

(4) $$V_j(k_1,\ldots,k_T) = a_{jo} + \frac{a_{j1}}{(1+k_1)} + \ldots + \frac{a_{jT}}{(1+k_1)\ldots(1+k_T)}$$

is the present value of project j at the discount rates $k_1,\ldots,k_T$.

The choice of discount rates in the objective function has been discussed by Baumol and Quandt [1], Weingartner [8], Carleton [2], Elton [3] and Myers [6].

In the following we shall distinguish between current ($a_{jo} < 0$) and future investment projects ($a_{jo} = 0$).[2] It is assumed that the cash flows of current projects are known, but that the future projects, and a forteriori the corresponding cash flows, are unknown. It follows that only the current financial constraint ($t = 0$ in (2)) is known. To distinguish this problem from the linear programming problem (1)-(3), we shall call it Problem A.

We shall for convenience make some simplifying assumptions. It is assumed that the current financial constraint is binding in the optimal solution of the primal problem (1)-(3) as well as in the solutions obtained by the simpler methods to be considered below. Otherwise there would not be capital rationing in the usual sense. Moreover, we shall assume that the optimal solution of the problem (1)-(3) is unique and non-degenerate.

3. Analysis

The following result is an immediate consequence of the duality theorems.

Proposition 1. Assume that the primal problem (1)-(3) has a unique, non-degenerate optimal solution. Then there exist shadow interest rates $\rho_1,\ldots,\rho_T$ such that:

---

[2] Deferment of projects is not considered.

(a) Any project which is fully accepted in the optimal solution has a positive present value evaluated at these interest rates.

(b) Any project which is partially accepted has present value zero.

(c) Any project which is rejected has a negative present value.

A firm which in addition to the cash flows of its current projects knows the shadow interest rates $\rho_1,\ldots,\rho_T$ may hence classify these projects into three classes.

Class 1. Projects which in the optimal solution of the primal problem (1)-(3) is fully accepted ($x_j = 1$).

Class 2. Projects which are partially accepted ($0 < x_j < 1$).

Class 3. Projects which are rejected ($x_j = 0$).

In general it is not possible on the basis of $\rho_1,\ldots,\rho_T$ and the current financial restriction to determine the proportion of each partially accepted project. Weingartner [7] has shown that the optimal solution of the primal problem (1)-(3) will include at most T partially accepted projects. With a reasonable number of accepted projects in each period, the relative number of partially accepted projects will be small, and the classification above will be a good approximation to the optimal solution of the primal.

Definition 1. A correct solution of Problem A is defined as a classification of the current investment projects into the three classes above, which is in accordance with the optimal solution of the primal problem (1)-(3).

Lorie and Savage [5] considered two methods for solving the one-period capital rationing problem:

Method 1. The projects are ranked according to the present value per dollar of outlay required. To be more precise, let

(5) $$v_j(i_1,\ldots,i_T) = \frac{V_j(i_1,\ldots,i_T)}{(-a_{jo})} ,$$

where $i_1,\ldots,i_T$ are the interest rates used in evaluating the investments. The interest rates in the objective function, $k_1,\ldots,k_T$ is a natural, but not the only possible choice for $i_1,\ldots,i_T$. The projects are ranked according to v_j.

Method 2. The projects are ranked according to the internal rate of return, r_j. We shall assume that this rate exists and is unique for each project.

Lorie and Savage [5] regarded Method 1 to be correct and Method 2 incorrect in case of conflict. We shall now give sufficient conditions for each method to result in a correct solution of Problem A, which may be viewed as a more realistic version of the one-period rationing problem.

Proposition 2. A sufficient[3] condition for Method 1, always to give a correct solution of Problem A is that

$i_2 = \rho_2, \ldots, i_T = \rho_T$.

Proposition 3. A sufficient[3] condition for Method 2. always to give a correct solution of Problem A is that

$\rho_1 = \rho_2 = \ldots = \rho_T$.

Proofs of these propositions are given in the Appendix.

The sufficient conditions of the two problems are seen to be of a different nature. Method 1. will give a correct solution to any problem as long as the firm is capable of estimating the shadow prices of future periods. This estimation is likely to be difficult, however, unless the capital rationing is of temporary character. In particular Lorie and Savage's conclusion is correct in the case they considered, capital rationing in only the current period. The conditions of Method 2. put restrictions on the problem itself. The method is likely to give a good approximation to the optimal solution if the investment and financing opportunities in the future are approximately as in the current period. If the scarcity of capital is temporary, Method 2. will have a bias in favor of projects with a short pay-back period.

---

[3] The conditions are also necessary in the following sense. If the conditions are not satisfied, there will exist problems where Method 1. (Method 2.) lead to an incorrect solution. However, in a particular problem either method may happen to result in a correct solution.

Some corollaries follow from the proofs of Proposition 2. and 3.

Corollary 1. The ranking by Method 1. is independent of the interest rate i_1.

The corollary implies that this interest rate may be chosen arbitrarily, e.g. $i_1 = 0$.

The necessary and sufficient conditions in Proposition 2. and 3. put (T-1) restrictions on the shadow discount rates. The values of $\rho_1,\ldots,\rho_T$ may consequently be determined by adding the current financial constraint.

Corollary 2. Suppose that the condition in Proposition 2. is satisfied, i.e. $i_2 = \rho_2,\ldots,i_T = \rho_T$. Then

(6) $$\rho_1 = v_m(0,i_2,\ldots,i_T)$$

where project m is a partially accepted project.[4]

Corollary 3. Suppose that the condition in Proposition 3. is satisfied, i.e. $\rho_1 = \ldots = \rho_T = \rho$.
Then

(7) $$\rho = r_m$$

where project m is a partially accepted project.[4]

---

[4] The existence of a partially accepted project follows from the assumption of non-degeneracy in the optimal solution of the linear programming problem.

Corollary 2. provides in the application of Method 1. a possibility to check the realism of the discount rates $i_2,\ldots,i_T$. If the discount rate obtained from (6) e.g. is much larger than $i_2,\ldots,i_T$, this may indicate that these rates are too low.

As seen above, Method 1. and Method 2. both make implicit assumptions on future shadow prices, which restrict these rates to a one-dimensional curve in the T-dimensional space. This order may be reversed, i.e. one may postulate a set of conditions which restrict the shadow discount rates to a one-dimensional curve, and then apply an iterative procedure to select the investment projects. Whatever approach is taken a record of past estimates of the shadow discount rates will be valuable information in the firm's investment and financial planning.

Appendix

Proof of Proposition 2.

Sufficiency. Suppose that current project m belongs to Class 1. according to the optimal solution of the linear programming problem (1)-(3) and project n to Class 2. Then

(8) $$V_m(\rho_1,\ldots,\rho_T) > V_n(\rho_1,\ldots,\rho_T) = 0,$$

and since for all current investment projects $a_{jo} < 0$,

(9) $$v_m(\rho_1,\ldots,\rho_T) > v_n(\rho_1,\ldots,\rho_T).$$

It is easily derived that for any current project j

(10) $$v_j(\rho_1,\ldots,\rho_T) = -1 + \frac{1}{(1+\rho_1)} w_j(\rho_2,\ldots,\rho_T)$$

where

(11) $$w_j(\rho_2,\ldots,\rho_T) = \left(\frac{1}{-a_{jo}}\right)\left[a_{j1} + \frac{a_{j2}}{(1+\rho_2)} + \ldots + \frac{a_{jT}}{(1+\rho_2)\ldots(1+\rho_T)}\right]$$

From (9) and (10) it follows that

(12) $$w_m(\rho_2,\ldots,\rho_T) > w_n(\rho_2,\ldots,\rho_T)$$

and from (10) and (12) that

(13) $$v_m(i_1,\rho_2,\ldots,\rho_T) > v_n(i_1,\rho_2,\ldots,\rho_T)$$

for all $i_1 > -1$. Hence, if $i_2 = \rho_2,\ldots,i_T = \rho_T$, project m is ranked before project n by Method 1. In a similar manner it can be shown that project m is ranked before project n by Method 1. if the former belongs to Class 1. or 2. and the latter to Class 3. Moreover, all projects in Class 2. are tied in the ranking by Method 1. under the assumption in the proposition, since

(14) $$v_m(\rho_1,\ldots,\rho_T) = v_n(\rho_1,\ldots,\rho_T) \quad (= 0)$$

implies by (10) that

(15) $$v_m(i_1,\rho_2,\ldots,\rho_T) = v_n(i_1,\rho_2,\ldots,\rho_T)$$

for all $i_1 \neq -1$. Sufficiency now follows from the assumption of tightness in the current financial constraint.

To prove necessity[5] we must show that it is possible to construct a counterexample if $i_t \neq \rho_t$ for some $t \geq 2$. Construct a primal problem (1)-(3) with a unique non-degenerate solution with shadow discount rates $\rho_1,\ldots,\rho_T$. If the solution obtained by Method 1. is the same as the linear programming solution, add a project n with cash flow given by

(16a) $$a_{no} = -\varepsilon$$

(16b) $$a_{n\tau} = \varepsilon(1+\rho_1)\ldots(1+\rho_{\tau-1})(2+\rho_\tau+i_\tau)/2$$

(16c) $$a_{nt} = 0 \qquad , \quad t \neq 0\ ,\ \tau,$$

---

[5] See footnote 3.

where τ is smallest integer such that $i_\tau \neq \rho_\tau$ and $\tau \geq 2$, and $\varepsilon > 0$ is sufficiently small to preserve the basis in the primal problem if project n is accepted. Then project n is classified in Class 1. by the linear programming solution and in Class 3. by Method 1. if $i_\tau > \rho_\tau$, and vice versa if $i_\tau < \rho_\tau$. Q.E.D.

Proof of Proposition 3.

It has previously been assumed that each project has a unique internal rate of return. Suppose that project m belongs to Class 1. and project n to Class 2. according to the linear programming solution and that $\rho_1 = \ldots = \rho_T = \rho$. Then

(17) $$V_m(\rho,\ldots,\rho) > V_n(\rho,\ldots,\rho) = 0$$

(17) implies that

(18) $$r_m > r_n = \rho.$$

The remaining part of the proof is similar the proof of Proposition 2.

Proof of Corollary 2.

Suppose that project m is a partially accepted project. From (10)

(19) $$v_m(0,\rho_2,\ldots,\rho_T) = -1 + w_m(\rho_2,\ldots,\rho_T)$$

(20) $$v_m(\rho_1,\ldots,\rho_T) = -1 + \frac{1}{1+\rho_1} w_m(\rho_2,\ldots,\rho_T) = 0$$

Substitution of the last equation in (20) into (19) gives (6).

REFERENCES

1. W.J. Baumol and R.E. Quandt, "Investment and Discount Rates Under Capital Rationing - A Programming Approach", The Economic Journal, LXXV (June 1965), pp. 317-329.

2. W.T. Carleton, "Linear Programming and Capital Budgeting Models", Journal of Finance, XXIV (December 1969), pp. 825-833.

3. E.J. Elton, "Capital Rationing and External Discount Rates", Journal of Finance, XXV (June 1970), pp. 573-584.

4. J.S. Hughes and W.G. Lewellen, "Programming Solutions to Capital Rationing Problems," Journal of Business Finance and Accounting, Vol. I (Spring 1974), pp. 55-74.

5. J.H. Lorie and L.J. Savage, "Three Problems in Rationing Capital", Journal of Business, Vol. XXVIII (October 1955), pp. 229-239.

6. S.C. Myers, "A Note on Linear Programming and Capital Budgeting", Journal of Finance, Vol. XXVII (March 1972), pp. 89-92.

7. H.M. Weingartner, "Mathematical Programming and the Analysis of Capital Budgeting Problems", Englewood Cliffs, N.J. Prentice Hall Inc. 1970.

8. H.M. Weingartner, "Criteria for Programming Investment Project Selection", Journal of Industrial Economics, Vol. XV (November 1966), pp. 65-76.

THE ABANDONMENT DECISION UNDER UNCERTAINTY

Carl J. Norstrøm

The Norwegian School of Economics and Business Administration, Bergen, Norway

Summary

Two approaches to the abandonment problem under uncertainty are considered; the ECF-approach based on expected cash flows and the DP-approach using dynamic programming. The ECF-approach is demonstrated to sometimes give incorrect solutions. The DP-approach is correct, but may be unfeasible due to the amount of numerical calculations. Under certain conditions Markov programming may be used to obtain the correct solution. A general result on the connection between the ECF-approach and the DP-approach is established.

1. One of the decisions in capital budgeting which has received special interest, is that concerning optimal replacement of equipment. A special case of this problem arises when it is assumed that the equipment will not be replaced, but that the production will come to an end when the equipment is sold. The choice of optimal time for abandoning production is easily solved under certainty. In this article we shall consider some of the aspects of the abandonment problem under uncertainty.

2. It is assumed that the equipment in each period results in a net operating inflow Z_τ which is realised at time τ, the end of the period. The company considers in the end of each period whether production should halt and the equipment be sold for its abandonment value S_τ. If a decision is made not to sell the equipment, production continues and the same question is considered again at the end of the next period. When the decision is made at time τ, the firm is assumed to know the net operating inflows and the abandonment values up to that time, i.e. Z_t and S_t for $t \leqq \tau$; as well as the conditional probability distribution of Z_{t+1} and S_{t+1} for each $t \geqq \tau$,

$$F_t(Z_{t+1}, S_{t+1} | Z_t, Z_{t-1} \ldots, S_t, S_{t-1} \ldots) \qquad (1)$$

The objective of the firm is to maximize, at a given rate of discount r, the expected present value of the total cash flow resulting from the equipment, and our problem is to find the abandonment time which achieves this.

3. One approach to finding the optimal abandonment time is to base the decision on the expected values for net operating inflows and abandonment values. We shall call this the Expected Cash Flow approach, or for short the

ECF-approach.[1] Let $Z_{t\cdot\tau}$ be the expected value of the net operating inflow from the equipment at time t as estimated at time τ (with knowledge of Z_τ, $Z_{\tau-1}$, ..., S_τ, $S_{\tau-1}$...) and $\bar{S}_{t\cdot\tau}$ the expected abandonment value at time t as estimated at time τ. Furthermore, let $R=(1/1+r)$, the discount factor. Define

$$\bar{G}_\tau^T = \sum_{t=\tau+1}^{T} R^{(t-\tau)}\bar{Z}_{t\cdot\tau} + R^{(T-\tau)}\bar{S}_{T\cdot\tau} \tag{2}$$

In economic terms $\bar{G}_\tau^T$ is the expected present value of keeping the equipment from τ to T and then selling it.

According to the ECF-approach the firms should at time τ maximise $\bar{G}_\tau^T$ with respect to T. It is easy to show that the following conditions are necessary for T to be optimal in this sense.

$$\bar{Z}_{T\cdot\tau} - (\bar{S}_{T-1\cdot\tau} - \bar{S}_{T\cdot\tau}) - r\bar{S}_{T-1\cdot\tau} \geq 0 \tag{3}$$

$$\bar{Z}_{T+1\cdot\tau} - (\bar{S}_{T\cdot\tau} - \bar{S}_{T+1\cdot\tau}) - r\bar{S}_{T\cdot\tau} \leq 0 \tag{4}$$

If the decision made at time τ is to postpone the sale of the equipment at least another period, it is necessary to calculate the optimal selling time anew at the end of this period, since normally $\bar{Z}_{t\cdot\tau+1}$ and $\bar{S}_{t\cdot\tau+1}$ will be different from $\bar{Z}_{t\cdot\tau}$ and $\bar{S}_{t\cdot\tau}$. Expectations have to be revalued as the firm obtains knowledge of $Z_{\tau+1}$ and $S_{\tau+1}$.

4. The ECF-approach is in principle equivalent to the approach taken under certainty, except that expected values are used instead of certain values. As may be suspected, this is too simple to be correct. It will in this section be shown by a counterexample that the ECF-approach is not in general correct.

In the example the abandonment value is constant, $S_t=25$, and the net operating inflow Z_t can take on only the values 8, 6 and 4, that is $Z_t=Z(i)=10-2i$, $i=1, 2, 3$. We let p_{ij} denote the conditional probability that $Z_{t+1}=Z(j)$ given that $Z_t=Z(i)$ where[2]

$$P = [p_{ij}] = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

The discount rate $r=0.25$.

We shall first find the true optimal abandonment policy for the example and then show that the decision given by the ECF-approach is not optimal. In the example the economic situation at time t is completely described by the value of Z_t. Hence we shall call Z_t the state variable and say that the process is in state i when $Z_t=Z(i)$. It is easy to see that the optimal policy[3] for the

[1] This is the approach taken by Robichek & Van Horne [6], [7]; and Dyl & Long [3].

[2] It follows from the definition of p_{ij} that $p_{ij} \geq 0$ and $\sum_{j=1}^{3} p_{ij} = 1$.

[3] By a policy we mean a rule which, dependent on the situation, determines a decision.

Table 1

i	$H(i\|i^*)$ $i^*=0$	$i^*=1$	$i^*=2$	$i^*=3$
1	25	27.5	27.8	27
2	25	25	25,6	24
3	25	25	25	21

firm must be either never to abandon the equipment or to abandon it the first time Z_t falls below some critical limit; that is, the policy must be of the form

$$\begin{array}{ll} \text{Abandon when } Z_t = Z(i),\ i > i^* & \\ & i^* = 0, 1, 2, 3 \\ \text{Keep} \quad \text{when } Z_t = Z(i),\ i \leq i^* & \end{array}$$

It remains to find the optimal value of i^*. Let $H(i|i^*)$ be the expected present value of the equipment when the process is in state i and a given i^* is used. Since the expected present value of the equipment if it is not abandoned, is the discounted sum of the expected net inflow in the following period and the expected present value of the equipment at the end of the period, we have

$$H(i|i^*) = \begin{cases} S & i > i^* \\ R\sum_{j=1}^{3} p_{ij} Z(j) + R\sum_{j=1}^{3} p_{ij} H(j|i^*) & i \leqslant i^* \end{cases} \tag{5}$$

$i = 1, 2, 3.$

(5) represents three equations in three unknowns and $H(i|i^*)$ can be found for any given i^*. The optimal i^* can now be obtained by straightforward enumeration. This is done in Table 1, and the optimal i^* is seen to be $i^* = 2$.[1] The firm should keep the equipment at time τ if $Z_\tau = 8$ or $Z_\tau = 6$, and sell it only if $Z_\tau = 4$.

The problem will now be solved using the ECF-approach. It is easy to show[2] that the expected net operating inflow in period t as estimated in period τ, $Z_{t \cdot \tau}$, in this example is

$$\bar{Z}_{t \cdot \tau} = \begin{cases} 6 + 2(\frac{3}{4})^{t-\tau} & \text{when } Z_\tau = 8 \\ 6 & \text{when } Z_\tau = 6 \\ 6 - 2(\frac{3}{4})^{t-\tau} & \text{when } Z_\tau = 4 \end{cases}$$

Since the abandonment value S_t is a constant, the necessary conditions for T to be the optimal abandonment time ((3) and (4) in section 3) are simplified to

$$\bar{Z}_{T.\tau} \geqslant r S_t$$
$$\bar{Z}_{T+1.\tau} \leqslant r S_t$$

[1] Note that $H(i/2)$ is a maximum for all values of i.
[2] See e.g. Howard [5], chapter 1.

Since $rS_t = 6.25$, the equipment should be sold at τ if $Z_\tau = 4$ or $Z_\tau = 6$, and kept for at least another period only if $Z_\tau = 8$. But this is not the true optimal policy obtained in Table 1. On the contrary, it corresponds to the nonoptimal $i^* = 1$. Hence it has been demonstrated that the ECF-approach may lead to wrong conclusions and that it is not generally correct.

5. The essential characteristic of the abandonment problem is that the decisions are made sequentially and under uncertainty. Through the development of dynamic programming, Richard Bellman has created a mathematical approach which is particularly suited to cope with such problems. We shall now outline the Dynamic Programming approach, or for short the DP-approach, to the abandonment problem.

Let $\bar{V}_t$ be the expected present value at time t given that the firm from t onward uses an optimal abandonment policy. According to the Principle of Optimality[1] it follows that

$$\bar{V}_\tau = \text{Max}\begin{cases} S_\tau \\ R\bar{Z}_{\tau+1\cdot\tau} + R \cdot E[\bar{V}_{\tau+1}] \end{cases} \tag{6}$$

Here $E[\bar{V}_{\tau+1}]$ is the expected value of $\bar{V}_{\tau+1}$ taken over the values of $Z_{\tau+1}$ and $S_{\tau+1}$. The economic interpretation of (6) is as follows. In the case in which it is optimal to sell the equipment at time τ, the expected value of the equipment $\bar{V}_\tau$ is equal to the abandonment value S_τ. If it is not optimal to sell at time τ, the expected value of the equipment will be equal to the discounted expected inflow generated by the equipment in the following period, plus the discounted expected value of the equipment at time $\tau + 1$, given that an optimal policy is used from that time.

(6) does not in general represent an explicit solution of our problem, since the function $\bar{V}_{\tau+1}$ is not known. In order to obtain a solution, further analytic or numerical analysis is required. The weakness of the DP-approach is that this derivation of the explicit solution from (6) may be difficult or impossible (with present computers). The difficulty involved is strongly dependent on the generality of the underlying stochastic process.

6. In this section we shall consider the interesting subclass of problems which is obtained when the stochastic process for Z_t and S_t is a Markovian process with stationary transition probabilities,[2] and Z_t and S_t only can take on a finite number of values.[3] From the first part of these assumptions it follows that the problem at any time t is fully described by the combination of values of Z_t and S_t. In dynamic programming terminology each such combination of

[1] Bellman [1], page 83.
[2] When the stochastic process is a Markovian process with stationary transition probabilities, the probability distribution for Z_{t+1} and S_{t+1} is only dependent on the values of Z_t and S_t. The function F_t in (1) reduces to $F(Z_{t+1}, S_{t+1} | Z_t, S_t)$.
[3] The example in section 4 is taken from this subclass of problems.

values is called a state, and Z_t and S_t are called the state variables. Since each of the variables Z_t and S_t only can take on a finite number of values, the number of combinations of values of Z_t and S_t is also finite. Hence it is possible to number the states of the process, i.e. to assign to each combination of values of Z_t and S_t an integer i, $i=1, 2, ..., N$. Let $Z(i)$ and $S(i)$ denote the values of the variables when the process is in state i. The probability distribution (1) may now be replaced by a matrix

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \cdot & \cdot & \cdot & \cdot \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}$$

where the element p_{ij} denotes the conditional probability that the process will be in state j at time $t+1$ given that it is in state i at time t.

With these assumptions, equation (6) takes the form

$$\bar{V}(i) = \text{Max}\begin{cases} S(i) \\ R\sum_{j=1}^{N} p_{ij} Z(j) + R\sum_{j=1}^{N} p_{ij} \bar{V}(j) \end{cases} \tag{7}$$

Here $\bar{V}(i)$ is the expected present value of the equipment when the process is in state i and an optimal policy is used. Again (7) is no explicit solution to the problem, but the explicit solution may now be found by Markov Programming—a branch of dynamic programming developed by Ronald Howard [5]. The optimal solution is found by successive approximations alternately to the values $\bar{V}(i)$ and to the optimal policy. This iterative procedure has been shown to converge to the optimal solution in a finite number of steps.[1]

7. Since the ECF-solution in general is simpler to apply than the DP-solution, it is of interest to obtain possible connections between the two. The following theorem holds for the general stochastic process described by (1).

THEOREM. *If the decision resulting from the Dynamic Programming approach consists in selling at present, the decision given by the Expected Cash Flow approach will also consist in selling at present.*

Proof. Suppose the DP-approach yields the decision to sell at present. Then $\bar{V}_\tau = S_\tau$. But $\bar{V}_\tau \geqslant \text{Max}_T\, \bar{G}_\tau^T$ so $S_\tau \geqslant \text{Max}_T\, \bar{G}_\tau^T$. On the other hand $\text{Max}_T \bar{G}_\tau^T \geqslant \bar{G}_\tau^0 = S_\tau$. Hence $\text{Max}_T\, \bar{G}_\tau^T = \bar{G}_\tau^0$.

Corollary. If the decision resulting from the Expected Cash Flow approach consists in keeping the equipment for at least another period, then the deci-

[1] Readers who are not familiar with Markov-programming, are referred to Howard [5], which beside being the pioneer work still is the best introduction to the topic. A proof of the convergence of the iteration procedure will be found there. Although the ECF-approach may lead to wrong decisions, it will no doubt in most cases give a good starting point for the successive approximations.

sion given by the Dynamic Programming approach will also consist in keeping the equipment for at least another period.[1]

Due to the corollary it is not necessary to find the correct optimal policy when the optimal decision according to the ECF-approach is to keep the equipment.

8. As will be clear from the previous sections, we have not obtained a feasible procedure which generally solves the abandonment problem under uncertainty.[2] Dynamic programming will lead to correct solutions, but may be unfeasible due to the amount of numerical calculations required. If the stochastic process is a Markov process with stationary transition probabilities and finite state space, then Markov programming may be used. Other approaches which have not been considered here, are applicable under other assumptions.[3] The ECF-approach is demonstrated to give incorrect results, but will no doubt usually be a good approximation. As the theorem in section 7 shows, the ECF-approach leads to a tendency too sell too soon. When the optimal decision according to this approach is to keep the equipment, this decision is known to be correct even when the optimal policy is not known.

References

1. Bellman, R.: *Dynamic Programming.* Princeton University Press, Princeton, 1957.
2. Breiman, L.: Stopping rule problems, in Beckenbach (ed.) *Applied Combinatorial Mathematics.* J. Wiley & Son, 1964.
3. Dyl, E. A. & Long, H. W.: Abandonment value and capital budgeting: Comment. *Journal of Finance XXIV*, no. 1, 1969.
4. Hart, A. G.: Anticipations, uncertainty and dynamic planning. *Studies in Business Administration XI* (1940). University of Chicago Press, 1940. Also available as a Reprint of Economic Classics, Augustus M. Kelley, 1951.
5. Howard, R.: *Dynamic Programming and Markov Processes.* The Technology Press of MIT and J. Wiley & Son, 1960.
6. Robichek, A. A. & Van Horne, J. C.: Abandonment Value and Capital Budgeting. *Journal of Finance XXII*, no. 4, 1967.
7. Robichek, A. A. & Van Horne, J C.: Abandonment value and capital budgeting: Reply. *Journal of Finance XXIV*, no. 1, 1969.

[1] The theorem and corollary do of course hold for the example in section 4 as the reader may readily verify.

[2] Even when we disregard problems like risk aversion and estimation of probability distributions.

[3] In particular the reader is referred to the theory of Optimal Stopping Rules. For an exposition of this theory, see Breiman [2].

WHEN TO DROP A PRODUCT:

THE ABANDONMENT DECISION UNDER ATOMISTIC COMPETITION*

1. Introduction

The decision to drop unprofitable products forms an important part of the overall marketing strategy. Excellent discussions of the many aspects of this decision are given by Alexander [1] and Kotler [6]. This paper is more narrow. The abandonment of a product is in principle a capital budgeting problem, and this will be the approach taken. More specifically we shall develop a capital budgeting model for the abandonment decision of an atomistic firm when the future prices of the product form a stochastic process. Allthough many firms are not atomistic, this may be a fair approximation of the situation at the end of a product's life, when many close substitutes are in the market.[1]

The paper is organized as follows. The model is outlined in section 2. In section 3 the optimal production policy and the resulting net operating income is derived. The abandonment decision is discussed in section 4. In section 5 some consequences of the abandonment decision are discussed. Finally, in section 6 some examples are given which illustrates the material in the previous sections.

* Published in Markedskommunikasjon, No. 2. 1975.

2. The Model

We consider a product, which is independent of the other activities of the firm. To avoid som mathematical difficulties, time is regarded to be discrete and is partioned into time periods of equal length. The following events take place at the outset of each period in the given order.

(i) The firm decides whether to decrease the capacity to produce the product.

(ii) The firm decides how much to produce in the period.

(iii) A chance event occurs which determines the price of the product during the period.

A more detailed description of each event is given below.

(i) Capacity. Let.

k_t - capacity at the outset of period t
y_t - decrease in capacity at the outset of period t.

The capacity prevailing during period t is consequently $k_t - y_t$. This also forms the beginning capacity of next period, i.e.[2]

(1) $$k_{t+1} = k_t - y_t \quad , \quad 0 \leqq y_t \leqq k_t \; .$$

From the sale of the capacity the firm receives an amount sy_t, where s is a constant. The cash inflow (or outflow if s is negative) takes place at the beginning of the period.

(ii) Production. Further assumptions on the cost function will be given in the next section.

(iii) Price. Let

p_t - the price at the beginning of the period, the beginning price.

q_t - the price prevailing during the period, the prevailing price.

We shall let the continuous function $F(q|p)$ be the probability distribution of the prevailing price q for a given beginning price p. Prices are assumed to be bounded above. The prevailing price in period t forms the beginning price in period $t+1$, $p_{t+1} = q_t$. Hence the prices $q_t, q_{t+1}, \ldots$ form a stationary Markov process. There are two natural cases for the domain of q_t.

(a) Continuous case. The domain is an interval $[\alpha,\beta]$ (α may be $-\infty$, but β is finite), and

(2) $$F(q|p) = \int_{\alpha}^{q} f(\xi|p)d\xi$$

where $f(q|p)$ is the probability density of q given p.

(b) Discrete case. The domain consists of a finite (or infinite sequence of points $\{p(1), p(2), \ldots, p(N)\}$ and the matrix

$$[a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \cdot & \cdot & \cdots & \cdot \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}$$

gives the probability of a transition $p_t = p(i)$ to $p_{t+1} = p(j)$.

We shall in the development of the model only consider the continuous case, but some of the illustrations in section 6 will be discrete. Mathematically oriented readers will observe that both cases may have been treated simultaneously by the use of the Stieltjes integral.

We shall assume that a high prevailing price is more likely when the beginning price is high, and similarily for low prices. More precisely, it is assumed that for any two beginning prices, p^1 and p^2 : if $p^1 < p^2$, than

(3) $$\int_{\alpha}^{q} f(\xi|p^1)d\xi \geqq \int_{\alpha}^{q} f(\xi|p^2)d\xi$$

for all q, with strict inequality holding for at least one value of q.

The expactation of the prevailing price q based on the information of the beginning price p is

(4) $$\bar{q}(p) = \int q \; f(q|p)dq$$

The expectation is a continuous function of the beginning price, and from assumption (3) it is seen also to be strictly increasing.

The objective of the firm is to maximize the expected present value of all future cash flows. The cost of capital is assumed to be given and equal to r in each period. The discount factor R is defined as $1/(1+r)$.

3. The Production Decision

We note as a preliminary observation that the quantity produced has no influence on future periods. It may be classified as a static or short run decision.

In cost accounting a usual assumption is that total short run costs consist of fixed costs and variable costs which increase linearly with output. If in addition it is assumed that there are constant returns to scale in the long run, the cost function may be written in the form

(5) $$C(x,\ k-y) = (c-b)x + b(k-y) \quad ; \quad 0 \leqq x \leqq k-y.$$

Here x is output, $(k-y)$ is the capacity, and the constants $(c-b)$, b and c denote respectively average variable costs, average fixed costs at full capacity and average total costs at full capacity.

The expected net operating income at the outset of a period before the prevailing price is known, but after an eventual decrease in capacity, is

(6) $$\Pi(p, k-y) = \bar{q}(p)x - (c-b)x - b(k-y).$$

The production policy takes the simple form of producing nothing if the expected price is less than average variable costs and to produce up to capacity when the expected price exeeds this cost. In terms of the beginning price it is optimal to produce nothing when this price is less or equal to a certain level w, and to produce at full capacity when it exeeds this level.[3] Defining

(7) $$\pi(p) = \begin{cases} -b & \text{when} \quad p \leqq w \\ \bar{q}(p) - c & \text{when} \quad p > w \end{cases}$$

the optimal expected net operating income may be written

(8) $$\Pi^*(p, k-y) = \pi(p)(k-y).$$

$\pi(p)$ is the expected net operating income pr. unit of capacity and is a continuous and monotonically increasing function of the beginning price p.

4. The Abandonment Decision

The decision to decrease the capacity has an influence on the state in future periods. The determination of the optimal policy is a sequential decision problem. Such problems are well suited for the use of dynamic programming, which will be the mathematical approach taken.

If the net operating income takes place at the end of the period, it follows from the description of the model in section 2 and (8) that the expected present value of the cash flow in the first period is

(9) $$sy + R\pi(p)(k-y) ,$$

where the first term represents the cash flow from sale of capacity. We define the $V(p,k)$ to be the expected present value of all future cash flows from the product with an optimal abandonment policy. Proceeding formally from the Principle of Optimality[4] it follows from (1) and (9) that

(10) $$V(p,k) = \max_{0 \leq y \leq k} \{sy + R\pi(p)(k-y) + R \int V(q,k-y)f(q|p)dq\}$$

The first two terms on the right hand side represent the cash flow (9) of the first period, and the last term all later cash flows. The optimization is achieved by balancing immediate and future gains, the latter being a function of the beginning price and the beginning capacity of next period. Some results may now be derived from (10).

Proposition 1. There exists a unique bounded and continuous function $V(p,k)$ satisfying (10).

This proposition may be derived from a general result in Denardo [4]. For a proof of a similar result, see Norstrøm [9].

Proposition 2. The expected present value of all future cash flows is a linear homogeneous function of the beginning capacity,

$$V(p,k) = A(p)k, \tag{11}$$

where $A(p)$ is a continuous and monotonically increasing function.

This proposition may be derived by the method of successive approximations. See e.g. Norstrøm [9] for a similar result. An alternative proof is given in Norstrøm [8]. The result is reasonable intuitively, since both the expected net operating income and the cash flow from sale of capacity are linear. $A(p)$ is the average expected present value pr. unit of capacity.

Proposition 3. There exists an optimal abandonment policy $y(p,k)$ of the form

$$y(p,k) = \begin{cases} k & \text{for } p \leq v \\ 0 & \text{for } p > v, \end{cases} \tag{12}$$

where v is a real number.

Proof. Substitute (11) into the right hand side of (10) and differentiate with respect to y.

The form of the optimal abandonment policy is a consequence of the assumptions of atomistic competition and the cost structure.[5]

In general the constant in (12) must be found by some iterative procedure, e.g. by successive approximations, or, in the discrete case, by Markov programming. However, an explicit solution may be found for the special case that the prices never increase.

First, we note that there are two possibilities with respect to the relation between the abandonment and the production decision. If v is less than w, it will be optimal to keep the capacity forever without production when p falls below w. If v is greater than w, it is optimal to produce up to full capacity until the product is abandoned. It may be shown that (under non-increasing prices) $w > v$ if and only if the scrap value of capacity is negative with an absolute value greater than fixed cost divided by the interest rate, $s < -b/r$.

Proposition 4. Assume that $s \geq -b/r$ and that prices are non-increasing. Then it is optimal to abandon the product the first time the expected prevailing price falls below the sum $c+rs$.

This proposition follows from a well known result in Stopping-rule problems.[6] It is also easy to see intuitively. The difference between the expected present value of abandoning the product after one period and abandoning it now is

(13) $R[\bar{q}(p) - (c+rs)]$

When (13) is positive, it pays to wait at least one period. When (13) is negative, the expected value of waiting one period is negative, and since prices are non-increasing the same will be true for any future period.

In the general case when prices are not necessarily non-increasing, Proposition 4 does not hold. Instead we have

Proposition 5. It is never optimal to abandon the product as long as the expected prevailing price is above c+rs, but it may be optimal to keep the capacity if the expectation is lower than this sum.

A proof of this proposition may be found in [7]. The reason for this somewhat paradoxical result is an asymmetry in the effect of price increases and decreases created by the opportunity to abandon the product. A favorable prevailing price will have an effect not only on the net operating income in the present period, but also on the value of the production capacity at the end of the period. The corresponding negative effect on the value of capacity of an unfavorable prevailing price is eliminated or reduced by the abandonment of the product.

5. Some Consequences of the Abandonment Decision

The Time of Abandonment

In this section some of the implications of the abandonment decision. We shall first consider the time of abandonment, which will be denoted T and obviously is a stochastic variable. Not only the expected time of abandonment, but the entire probability distribution of T, ought to be a valuable input in the firm's planning. We define for $p > v$ the sequence of functions $\{g_n(q|p)\}$ by

(14a) $$g_1(q|p) = f(q|p)$$

(14b) $$g_n(q|p) = \int_v^\beta f(q|\xi)g_{n-1}(\xi|p)d\xi$$

The function $g_n(q|p)$ denotes for a given beginning price p the joint event that the prevailing prices are above v in first $(n-1)$ periods and that the prevailing price in the n'th period has the value q. The probability that the product will be abandoned at time n (the end of n'th or beginning $(n+1)$'st period) is consequently

(15) $$G_n(v|p) = \int_\alpha^v g_n(q|p)dq$$

It is not always the case that

(16) $$\sum_{n=1}^{\infty} G_n(v|p) = 1$$

since there may be a positive probability of never abandoning the product. If (16) holds, the expected time to avandonment is

(17) $$\bar{T}(p) = \sum_{n=1}^{\infty} nG_n(v|p)$$

However, as pointed out above, the entire distribution of the time of abandonment ought to be of interest to the firm.

The Expected Present Value of Capacity

In the previous section it was found that the expected present value of all future cash flows from the product was $A(p)k$, where k is the present capacity. $A(p)$ is consequently the average and marginal expected present value pr. unit of capacity. Clearly $A(p) = s$ for $p \leqq v$.

If we for convenience assume that $w \leqq v$ (that it is never optimal to keep the capacity idle) it may be shown[7] that for $p > v$

(18) $$A(p) = \sum_{n=1}^{\infty} R^n \{\int (q-c)g_n(q|p)dq + sG_n(q|p)\}$$

Keeping in mind (14) and (15) it is seen that the n'th term in (18) represents the sum of the expected net operating income pr. unit in the n'th period and the expected cash from selling the unit at time n.

Implications for Cost Reduction

The function $A(p)$ is of particular interest for the decision to increase the capacity, a decision which is not discussed in this paper.[8] However, we shall derive an interesting and perhaps somewhat surprising result from (18). Differentiation with respect to c gives[9]

$$\frac{\delta A(p)}{\delta c} = - \sum_{n=1}^{\infty} R^n G_n(\beta|p). \tag{19}$$

$G_n(\beta|p)$ denotes the probability that the beginning prices of the n first periods will exeed v. Due to assumption (3) this probability is an increasing function of the beginning price at the outset of the first period. It follows that a cost increase or decrease has less impact on the value of the product when the price is low. In particular, a one shot cost - reducing investment is more profitable when the price is high than when it is low. This result contradicts statements like "it is necessary to rationalize because prices are low", if these are made out from profit considerations. The economic argument behind the result is that the life of the cost reducing investment is longer when the beginning price is high.

The Effect of Uncertainty

It has often been an implicit assumtion in much of the economic literature that uncertainty is undesirable. It will be demonstrated that this is not necessarily true in the present case. We assume first that the prices $q_t, q_{t+1}, \dots$ is a stationary process with independent increments. To be consistent with the assumptions for existence and uniqueness in Proposition 1, it is further assumed that prices are nonincreasing. Let the price decrease be $\eta_t = p_t - q_t$. The assumptions may then be summarized into the statements that $\eta \geq 0$ and that the probability density of η, $h(\eta)$ is independent of p. Consider two probability densities $h_1(\eta)$ and $h_2(\eta)$ having the same mean, but where the dispersion of $h_1(\eta)$ is greater.[10] More precisely, for some number η_o

(20)
$$\int_o^\eta h_1(\xi)d\xi \geq \int_o^\eta h_2(\xi)d\xi \qquad \text{for all } \eta \leq \eta_o$$
$$\int_o^\eta h_1(\xi)d\xi \leq \int_o^\eta h_2(\xi)d\xi \qquad \text{for all } \eta > \eta_o$$

We define $V_j(p,k)$ to be the value of the product when the distribution is $h_j(\eta)$, and $A_j(p)$ in a similar manner.

Proposition 6. If the prices q_1, q_2, ... form a stationary process with independent increments, the value of the product will increase with an increase in uncertainty in the sense of (20).

Proof. By the method of successive approximations it may be shown that $A_j(p)$ is convex, and using convexity that $A_1(p) \geq A_2(p)$. The result follows from (11).

The conclusion in Proposition 6 does not necessarily hold when the assumptions of independent increments is lifted. We shall prove a somewhat weaker result for the general case. The expected prevailing price after n periods as for a given p

(21) $\bar{q}_n = \int q \, f_n(q|p)$,

where $f_n(q|p)$ is the n'th convolution of $f(q|p)$. Suppose that the firm in the planning uses the expected prevailing price $\bar{q}_t$ instead of the stochastic q_t. We shall call this the "expected price approach".

Proposition 7. The value of the product calculated by the "expected price approach" is always less or equal to the correct expected present value of the product defined by (10).

Proof. In the "expected price approach" the time of abandonment would be deterministic time T. The policy to sell at time T irrespective of the values of $p_1, \ldots, p_{T+1}$ is possible (but usually non-optimal) for the stochastic case. The expected value of the product with this policy is equal to the value of the product by the "expected price approach".

6. Some Illustrations

We shall in this section give some examples, which will serve as illustrations of some of the points made in the earlier sections.

Example 1

It is assumed that the domain of the prices is {18, 16, 14} i.e. $p(1) = 18$, etc. The transition matrix is

$$[a_{ij}] = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

Let further $r = 0.25$, $c = 10$, $s = 25$ and $k = 1$. It is easily calculated that

$$\bar{q}_n = \begin{cases} 16 + 2(\frac{3}{4})^n \\ 16 \\ 16 - 2(\frac{3}{4})^n . \end{cases}$$

$c + rs = 16.25$

Since the expected prevailing price is below $c+rs$ in all future periods if the present beginning price is 16 or 14, it is tempting to abandon the product if either of these cases occur. Actually this would be the policy used if the firm's planning was based on the "expected price approach". However, this policy is nonoptimal. The expected present values with this policy are respectively 27.5 , 25 and 25 when p is 18, 16 and 14. The optimal policy, which may be found by Markov programming[11] (or simply by trial and error), is to abandon the product only if p is 14. The optimal policy results in the present values 27.8 , 25.6 and 25 when p is respectively 18, 16 and 14.[12]

Example 2

It is assumed that prices are discrete and that

$$q = \begin{cases} p & \text{with probability } 1 - \gamma \\ p - \omega & \text{with probability } \gamma \end{cases}$$

This is a stationary process with independent increments. The expected prevailing price is

$$\bar{q}(p) = p - \gamma\omega .$$

Since prices are non-increasing it follows from Proposition 4 that

$$v = c + rs + \gamma\omega .$$

We define the integer m by

$$(p-v) + \omega > m\omega \geqq p-v.$$

m denotes the number of times the price will drop before the product is abandoned. The probability that the abandonment takes place at time n is for $p > v$.

$$G_n(v|p) = \begin{cases} 0 & \text{for } n < m \\ \binom{n-1}{m-1} \gamma^m (1-\gamma)^{n-m} & \text{for } n \geqq m. \end{cases}$$

This is the negative binomial density.

In the derivation of the function $A(p)$ it is convenient to restrict p to the values $m\omega + v$, $m=0, 1,...$ Intermediate values of $A(p)$ may be obtained by linear interpolation.

$$A(p) = \frac{p-c}{r} - \frac{(1+r)\gamma m}{r^2} + \frac{\gamma m}{r^2}\left(\frac{\gamma}{r+\gamma}\right)^m$$

In Table 1 we have calculated the values for $A(p)$ under the condition that $c = 19$, $s = 0$, $r = 0.10$.

Table 1.

p	Value of A(p), γω=1			Value of cost reduct
	γ=1	γ=0,5	γ=0,25	
20	0	0	0	0,91
21	0,91	1,67	2,98	1,73
22	2,64	3,33	5,97	2,49
23	5,13	6,38	8,95	3,17
24	8,30	9,44	11,94	3,79
25	12,09	13,65	16,89	4,36
26	16,45	17,87	21,85	4,87
27	21,32	23,05	26,80	5,33
28	26,65	28,23	31,76	5,76
29	32,41	34,21	38,13	6,14
30	38,55	40,19	44,50	6,50
31	45,05	46,83	50,87	6,81
32	51,86	53,48	57,24	7,11

Three set of values for γ and ω are used: $\gamma = 1$ and $\omega = 1$, $\gamma = 0.5$ and $\omega = 2$, and $\gamma = 0.25$ and $\omega = 4$. Since $\gamma\omega = 1$ in all the three cases and $q_t, q_{t+1}, \ldots$ is a process with independent increments, it is easily seen that the expected prices in any future period are identical. However, the degree of uncertainty is different, the first case representing certainty and the third ($\gamma = 0.25$) greater uncertainty than the second in the sense of (20) as well as in terms of variances.

(The variances are respectively 0, 1, and 3). According to Proposition 6 the values of $A(p)$ will be smallest for $\gamma = 1$ and largest for $\gamma = 0.25$ for all values of p. This result is confirmed in the table.

It may be shown that a cost reduction will have the same effect as a price increase in the case of a process with independent increments. A reduction ψ in average total costs will hence increase the expected present value of the product by

$$[A(p+\psi) - A(p)]k.$$

In the last column of Table 1 we have calculated the values of $[A(p+\psi) - A(p)]$ for $\psi = 1$, $\gamma = 1$, $\omega = 1$ and values of c, s and r as specified above. The expected present value of the cost reduction is seen to be an increasing function of p.

Example 3

As in Example 2 it is assumed that the prices form a stationary process with independent, but the domain is now continuous with the density function

$$h(\eta) = \lambda e^{-\lambda\eta}$$

It follows that

$$\bar{q}(p) = p - \frac{1}{\lambda}$$

and from Proposition 4

$$v = c + rs + \frac{1}{\lambda}$$

The probability that the product will be abandoned at time n is for $p > v$

$$G_n(v|p) = \frac{[\lambda(p-v)]^{n-1}}{(n-1)!} e^{-\lambda(p-v)}$$

which is the density of the Poisson-distribution.

The function $A(p)$ is for $p > v$

$$A(p) = \frac{p-c}{r} + \frac{(1+r)}{\lambda r^2} + \frac{1}{\lambda r^2} e^{-\frac{\lambda r(p-v)}{1+r}}$$

Numerical values for $A(p)$ are easily calculated, and the sensitivity with respect to the parameters may be investigated. It is easily verified that $A(p)$ is continuous, increasing and convex.

Footnotes

1) This point was brought to my attention by Johan Arndt.

2) The capacity is constant or decreasing. An extension to investment in new capacity is given in Norstrøm [9].

3) The existence of w follows from the assumption that $\bar{q}(p)$ is continuous and strictly increasing. w is defined by $\bar{q}(w) = (c-b)$.

4) Bellman [2], page 83.

5) Proposition 3 may be shown to hold under the following, less restrictive assumptions with respect to the cost function.
(a) Average total costs are not lower at any lower production or capacity.
(b) Average fixed costs pr. unit of capacity are not lower at any lower capacity.
(c) Average variable costs (including eventual escapable fixed costs) are for any capacity lowest at full capacity.
A proof is given in Norstrøm [8].

6) See Breiman [3], Theorem 10.1.

7) It follows from (10) and (12) that $A(p)$ for $p > v$ satisfies the integral equation

$$A(p) = R\pi(p) + R \int A(q)f(q|p)dq$$

Expansion into the Neumann series gives (18)

8) See footnote 2.

9) The change in c will have an indirect effect through the abandonment, but this will not influence the direction of the change in A. To see this let $A(v,c)$ be any function of the variable v and the parameter c, where v is chosen so that $A(v,c)$ is maximized. Assume $c_1 \leq c_2$ implies that $A(v,c_1) \geq A(v,c_2)$ for any value of v. Let v_1 and v_2 denote the optimal values of v corresponding to c_1 and c_2 respectively.

Then

$$A(v_1,c_1) \geqq A(v_2,c_1) \geqq A(v_2,c_2)$$

Hence the secondary effect working through the decision variable may be ignored. See also Bellman [2], pp. 157-8.

10) See Rothschild and Stiglitz [10].

11) The best introduction to Markov programming is still the original work, Howard [5].

12) A more detailed discussion of this example is given in Norstrøm [7].

References

1. Alexander, R.S., The Death and Burial of 'Sick' Products, Journal of Marketing, April 1964.

2. Bellman, R., Dynamic Programming, Princeton University Press, N.J. 1957.

3. Breiman, L., Stopping Rule Problems, in Applied Combinatorial Mathematics, E.F. Beckenback (ed), John Wiley and Son, Inc., N.Y. 1964.

4. Denardo, E.V., Contraction Mappings in the Theory underlying Dynamic Programming, SIAM Review 9 (1967), 165-177.

5. Howard, R., Dynamic Programming and Markov Processes, The MIT Press and John Wiley & Son, Inc., N.Y. 1960.

6. Kotler, P., Phasing Out Weak Products, Harvard Business Review, March-April 1965.

7. Norstrøm, C.J., The Abandonment Decision under Uncertainty, Swedish Journal of Economics 72 (1970), 124-129.

8. Norstrøm, C.J., The Abandonment Decision under Atomistic Competition, Manuscript, 1974.

9. Norstrøm, C.J., Optimal Capital Adjustment under Uncertainty, Journal of Economic Theory 8 (1974), 139-148.

10. Rothschild, M. and J.E. Stiglitz, Increasing Risk: I A definition, Journal of Economic Theory 2 (1970), 66-84.

A STOCHASTIC MODEL FOR THE GROWTH PERIOD DECISION IN FORESTRY*

1. Introduction

The growth period decision is a familiar problem in economic theory, known from the work of Jevons, Wicksell and Irving Fisher. The aim of this paper is to strengthen the realism of the model when applied to forestry, by explicit recognition of uncertainty in the future values of the trees.[1] In order to investigate whether this extension may be important, a comparison is made between a policy which takes account of uncertainty and one where uncertainty is ignored.

2. The Model

In the deterministic growth period model the objective is to maximize the present value

$$W(T) = - C + (1+r)^{-T}[g(T) + G] \tag{1}$$

where

g(t) - the net value of the forest stand when cut t years old.
C - the cultivation cost, e.g. cost of seedling.
G - the value of land.[2]
r - the market interest rate.
T - the time the forest stand is cut.

The model is well known, and will not be elaborated here.[3]

* Forthcoming in the Swedish Journal of Economics.

The first task in the construction of a stochastic model is to identify the main sources of uncertainty. The most critical factor in the determination of the optimal time to cut a forest stand is the value increase from waiting one period,

(2) $g(t+1) - g(t)$.

This value is normally decomposed into three components: the increase in total volume, the quality increase due to the larger dimension of each tree, and the increase or decrease in the basic price of timber. Due to research done by forest economists, reasonably good estimates exist for the volume and quality increase, dependent on such factors as age, fertility of the soil, intensity of thinning. There are, however, very substantial fluctuations in the price of timber from season to season.[4] These fluctuations are here taken to be the only source of uncertainty.

The basic assumption of the model is that the value of the forest stand at a future time t is a product

(3) $p_t f(t)$

where $f(t)$ is a deterministic growth function and p_t a stationary stochastic process. The two elements in the product may be given various economic interpretations. We shall let $f(t)$ represent the expected value of the asset and p_t an index reflecting the variation in the basic price of timber relative to the long run expectation.

In addition to the above basic assumption, we shall in this and the next three paragraphes make some others, which are not essential for the approach. - It is assumed that the sequence $p_1, p_2, \ldots,$ is a stationary Markow process with a finite state space $P = \{p(1), p(2), \ldots, p(N)\}$, where $p(1) > p(2) > \ldots > p(N)$. The transition probabilities will be written in the form of a transition matrix

$$(4) \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ . & . & \cdots & . \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}$$

where the element a_{ij} denotes the probability that $p_{t+1} = p(j)$ given that $p_t = p(i)$.

The cutting policy will depend on the information available to the decisionmaker. Several alternative assumptions are possible, and the choice must depend on the real situation at hand. It is assumed here that p_t is fully known, when the decision whether to sell at time t is taken.[5]

The objective is taken to be maximization of the expected present value of future cash flows. This is the obvious generalization of the objective in the deterministic model, but contains the weakness that the possibility of risk aversion is ignored.

Two cases will be consideres with respect to the value of land. In the first the value of land is taken to be a given function of the state of the price index p_t;

$$(5) \qquad G_t = G(i) \quad \text{when} \quad p_t = p(i).$$

The second approach consists in determining the value of land endogenously from the use in forestry.

The principal characteristic of the optimization problem resulting from the specifications above is that the decisions are made sequentially and without perfect foresight. The importance of such problems for economics was first pointed out by A.G.Hart.[6] The mathematical approach to be used, dynamic programming, is due to the mathematician R. Bellman.[7]

Case 1. G exogenous

To facilitate the solution of the optimization problem, we introduce the artificial constraint

(6) $$T \leq \tau$$

where τ is some integer. For the given τ we define the function $V(i,t)$ to be the expected present value of the forest stand and the land at time t when $p_t = p(i)$ and an optimal cutting policy is used. Proceeding formally from the Principle of Optimality[8] we obtain the following sequence of recursive equations:

(7a) $$V(i,\tau) = p(i)f(\tau) + G(i)$$

(7b) $$V(i,t) = \text{Max}\begin{cases} p(i)f(t) + G(i) \\ (1+r)^{-1} \sum_j a_{ij} V(j,t+1) \end{cases} ; \quad t = \tau-1, \tau-2, \ldots, 1.$$

(7c) $$V(i,o) = \text{Max}\begin{cases} G(i) \\ -C + (1+r)^{-1} \sum_j a_{ij} V(j,1) \end{cases}$$

The values of $V(i,t)$ are given by (7) for all i and t. To formulate an optimal policy define the sets

(8) $$S_t = \{p(i) \varepsilon P | V(i,t) = p(i)f(t) + G(i)\}$$

$t = 1,2,\ldots, \tau$, and similarily for S_0. In presence of the artificial constraint (6), the following policy is optimal:

Cut the forest stand the first period the price index p_t belongs to the set S_t.

In general the values of $V(i,t)$ and the sets S_t will depend on the number τ in (6). Ideally τ should be chosen large enough to make it uneconomical to keep the forest stand beyond this date regardless of the value of p_τ. This is not always possible, but it may be shown under fairly broad conditions that $V(i,t)$ for any fixed t can be made arbitrarily close to the true values by choosing τ sufficiently large.

Case 2. G endogenous

We shall now consider the case that the value of land is not given, but must be derived. For simplicily forestry is assumed to be the only economic alternative.

The solution is obtained by solving a sequence of optimization problems with the value of land taken from the previous problem in the sequence. For a given value of τ

(9a) $$G(i,0) = 0$$

(9b) $$V(i,\tau,n) = p(i)f(\tau) + G(i,n)$$

$$(9c) \quad V(i,t,n) = \text{Max}\begin{cases} p(i)f(t) + G(i,n) \\ (1+r)^{-1}\sum_j a_{ij} V(j,t+1,n) \end{cases}$$

$$(9d) \quad V(i,o,n) = -C + (1+r)^{-1} \sum_j a_{ij} V(j,1,n)$$

$$(9e) \quad G(i,n+1) = V(i,o,n).$$

From (9) we obtain a sequence of functions $\{V(i,t,n)\}$ which under fairly broad conditions will converge to a function $V(i,t)$ representing the desired value. Clearly $G(i) = V(i,o)$ when determined endogenously.

In the previous presentation it has been assumed that the prices of timber and the value of land are stationary over time. A more general assumption is that there is a positive (or negative) trend due to inflation or to an increase in the value of timber relative to other goods. When this increase (or decrease) takes place at a constant rate, (3) may be replaced by

$$(10) \quad (1+\gamma)^t \; p_t f(t),$$

where γ is the growth rate. Clearly (10) includes (3) as a special case. However, the greater generality does not lead to important new results. Provided the value of land and the cultivation cost increase at the same rate, it may be shown that the problem (10) is essentially identical with the problem (3), when the interest rate used in the latter is $(r-\gamma)/(1+\gamma)$. For this reason we shall in the following use the simpler model.

3. The Effect of Ignoring Uncertainty

We shall in this section make a comparison between the economic result in the stochastic model of the previous section and the same model when uncertainty is ignored in the decision process. For convenience the latter will be called the corresponding ignorance model.

The optimization problem in the ignorance model may be formulated as the maximization of

(11) $$Z(T) = E\{-C + (1+r)^{-T} [p_T f(T) + G_T]\}$$

where E denotes the expectation operator p_T and G_T are random variables as defined in the previous section. Given that the Markov chain (4) is ergodic, the probability that $p_t = p(i)$ will converge to a limiting probability q_i independent of the initial state. Since the actual value of p_t is ignored in the ignorance model, it is reasonable to use for the expectations of p_t and G_t the constants

(12a) $$\bar{p} = \sum_i q_i p(i)$$

(12b) $$\bar{G} = \sum_i q_i G(i)$$

The optimization problem in the ignorance model is then reduced to the maximization of

(13) $$Z(T) = -C + (1+r)^{-T} [\bar{p} f(T) + \bar{G}]$$

In the above we have assumed that the value of land is the same random variable as in the stochastic model. This is inconsistent when this value is determined endogenously. The best use of land is obtained by applying the optimal decision policy of the stochastic model, which is unfeasible since the decisionmaker is assumed to ignore uncertainty. Instead the value of land will be defined from the best use when uncertainty is ignored, i.e. as the limit $\bar{G}$ of the sequence $\{\bar{G}(n)\}$ defined by

14a) $\bar{G}(o) = 0$

$$Z(T,n) = -C + (1+r)^{-T}\,[\bar{p}f(T) + \bar{G}(n)]$$

$$\bar{G}(n+1) = \max_{T} Z(T,n)$$

We shall let T^* denote the best cutting time in the ignorance model.

Although the expected present value of a forest stand when the optimal stochastic policy is used obviously will be at least as large as the corresponding value with the best ignorance policy, the absolute difference and proportion between these numbers will depend on t and p_t. In the following we shall assume that $t = 0$ and that $p_o = p(i)$ with probability q_i. This seems to be a fairly neutral choice of situation [9]. The expected present value when the optimal stochastic model is used is then

(15) $\bar{V} = \sum_{i} q_i\, V(i,o)$

and correspondingly with the best ignorance policy

(16) $\bar{Z} = Z(T^*)$.

Before proceeding to an actual comparison between the stochastic and the ignorance models, some results will be derived for the deterministic model. Comparison of (1) and (13) shows that these two optimization problems are mathematically identical when

(17a) $g(t) = \bar{p}f(t)$

(17b) $G = \bar{G}$.

The deterministic model satisfying (17) will be said to correspond to the stochastic model. The following results are now immediate.

Proposition 1. The deterministic and ignorance models corresponding to the same stochastic model are equivalent in the sense that the decisions and expected present values are identical.

Proposition 2. The expected present value in a stochastic model is at least as great as the present value in the corresponding deterministic model.

Proposition 2 shows that the individual forest owner on the average is better off with fluctuations in prices and the resulting uncertainty, than he would have been with a constant price equal to the long run average of actual prices. This result may be somewhat unexpected, but is easily understood if a proper distinction is made between the fluctuations themselves and the uncertainty following from the fluctuations. The effect of the latter will always be negative or zero and represents the loss due to lack of perfect information. The effect of the fluctuations themselves may well be positive as earlier has been pointed out by Waugh [11], and does in the present case always outweigh the negative effect of the lack of perfect information.[10]

An Illustration from Norwegian Forestry.

In order to get some idea of the magnitude of the gain by using the stochastic model we shall make a comparison between this model and the corresponding ignorance model based on data from Norwegian forestry. It should be emphasized that the presentation has more the character of an illustration than a systematic study.

Estimation of the transition matrix (4) is based on calculated net price of timber deflated by the wholesale price index in the periode 1918-62. After the elimination of a growth trend ($\gamma = .011$) the probabilities of transition are found by the maximum likelihood procedure.[11]

$$(18)\qquad A = \begin{bmatrix} 1/6 & 1/2 & 0 & 0 & 1/3 \\ 4/13 & 4/13 & 4/13 & 1/13 & 0 \\ 0 & 4/11 & 4/11 & 2/11 & 1/11 \\ 1/10 & 1/10 & 3/10 & 3/10 & 2/10 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

where

$$(19)\qquad p(1) = 1.4,\ p(2) = 1.2,\ p(3) = 1.0,\ p(4) = .8,\ p(5) = .6\ .$$

The transition matrix is ergodic, and the limiting probabilities $(q_1, q_2, \ldots, q_5)$ are

$$(20)\qquad (.126,\ .256,\ .245,\ .257,\ .116)$$

and

(21) $\bar{p} = 1.004$.

The growth function $f(t)$ will depend on such factors as the quality of the soil, intensity of thinnings etc.. However, since growth functions have the same general shape, the choice will be of small importance in the comparison. We shall use an example taken from Risvand [7] . See Table 1.

Table 1

1. Age	2. Thinning value	3. Value of remaining stand
34	0	38.46
38	0	146.68
42	0	301.87
46	-1.03	444.06
50	0	647.29
54	133.81	686.25
58	0	909.63
62	0	1,146.41
66	0	1,397.85
70	0	1,637.97
74	0	1,881.58

Source: Risvand [7].

There are two thinnings, one in the 46th and one in the 54th year.We shall restrict the final cut to the years after the last thinning. The growth in the value of the remaining stand is in this period nearly linear, and the values of the intermediate years are estimated by linear interpolation. The value of τ used in the artificial constraint (6) is taken to be 74 years.

In the comparison the value of land will be determined endogenously. We let the interest rate $r = .05$ and the cultivation cost $C = 0$.[12]

The results of the comparison are shown in Tables 2. and 3.. Table 2. shows the values of $\bar{V}(n)$ and $\bar{Z}(n)$.[13] The convergence to $\bar{V}$ and $\bar{Z}$ is seen to be fairly rapid. Table 3. shows the optimal policy of the stochastic model represented by the sets S_t. To facilitate comparison the best cutting time T^* in the ignorance is also expressed in terms of these sets by letting $S_t = \emptyset$ when $t < T^*$ and $S_t = P$ when $t \geq T^*$. ($T^* = 63$).

Table 2.

n	$\bar{V}(n)$	$\bar{Z}(n)$
1	73.29	56.25
2	76.77	58.74
3	76.94	58.86
4	76.95	58.86
5	76.95	58.86

Table 3.

1. Age	2. Stochastic Model	3. Ignorance Model
55	$\emptyset$	$\emptyset$
56 - 62	$\{p(1)\}$	$\emptyset$
63 - 68	$\{p(1)\}$	P
69 - 72	$\{p(1), p(2)\}$	P
73	$\{p(1), p(2), p(3)\}$	P
74	P	P

As is seen from Table 2. the expected present value of the optimal stochastic policy is 76.95 , compared with 58.86

when uncertainty is ignored. The increase in present value is 17.04 , or 30.3 per cent.

4. Conclusion.

In the illustration above we have neglected several factors of real life, and the result must hence be interpreted with some care. None the less the comparison is a clear indication of the importance of taking price fluctuations into account in the determination of the final cut.

The stochastic model presented here may be further refined in several directions, e.g. by adding som constraints on the decisionmaker's actions. Another possibility is to add other state variables in order to improve the decisionmaker's information about future prices, and in this way obtain a still better policy. Either addition will, however, lead to great increases in computational requirements.

Footnotes:

*) The author would like to express his indebtness to several menmbers of the Norwegian Forest Research Institute, in particular John Eid, Jens Risvand, Asbjørn Svendsrud and Rolf Sæther, and to Agnar Sandmo. Thanks also go to Helge Gundersen, who carried out the computer programming.

Footnotes continued

1) Although formally valid, the model is probably not a realistic stochastic representation for other growth processes, like wine-aging.

2) A more general approach would be to let the value of land be a function of time.

3) Interested readers are referred to Lutz [4], Lutz and Lutz [5] and Bierman [2].

4) See the study by Sæther [10], pp. 391-414.

5) This seems to be in reasonable agreement with the conditions in Norwegian forestry. See Sæther [10], pp. 415-417 for details.

6) Hart [3], especially chapter 4.

7) Bellman [1].

8) Bellman [1], p. 83.

9) The assumption may be greatly weakened with negligible consequences for the actual comparison below.

10) For discussion of Waugh's result, see Samuelson [8,9], Oi [6] and Waugh [12].

11) The data are taken from the study by Sæther [10]. The matrix is based on observation of 6 transitions from state 1, 13 from state 2, 11 from state 3, 10 from state 4 and 4 from state 5.

12) The growth function in Table 1. is derived under the condition that $r = .03$ in which case $T^* = 74$. The choice of $r = .05$ is made in order to have T^* somewhat smaller than τ.

13) The definitions of $\bar{V}(n)$ and $\bar{Z}(n)$ are obvious extensions of (15) and (16).

References:

1. Bellman, R.: Dynamic Programming, Princeton University Press, N.J. 1957.

2. Bierman, H. Jr.: The Growth Period Decision , Management Science, Vol. 14, No. 6, February 1968.

3. Hart, A.G.: Anticipations, Uncertainty and Dynamic Planning, Studies in Business Administration XI (1940), University of Chicago Press, 1940. Also available as a Reprint of Economic Classics, Augustus M. Kelley, 1951.

4. Lutz, F.A.: The Criterion on Maximum Profits in the Theory of Investment, Quarterly Journal of Economics, Vol. LX, No. 1, November 1945.

5. Lutz, F.A. and V.Lutz: The Theory of Investment of the Firm, Princeton University Press, Princeton, N.J., 1951.

6. Oi, W.Y.: The Consumer Does Benefit from Feasible Price Stability: A Comment, Quarterly Journal of Economics, Vol. LXXXVI, No. 3, 1972.

7. Risvand, J.: Economic Analysis of Cutting Programs Applying Dynamic Programming, in Readings in Forest Economics, (A.Svendrud, Ed.), 73-80, Universitetsforlaget, Oslo, 1969.

8. Samuelson, P.A.: The Consumer Does Benefit from Feasible Price Stability, Quarterly Journal of Economics, Vol. LXXXVI, No. 3, August 1972.

9. Samuelson, P.A.: Rejoinder, Quarterly Journal of Economics, Vol. LXXXVI, No. 3, August 1972.

10. Sæther, R.: Supply of Industrial Softwood in Norway. A Statistical Inquiry into Annual Cuts and Factors Affecting Quantities Cut during the Period 1918-1960, Norwegian Forest Research Institute, Vollebekk, Norway. Written in Norwegian with English Summary.

11. Waugh, F.V.: Does the Consumer Benefit from Price Instability? Quarterly Journal of Economics, Vol. LVIII, 602-14, August 1944.

12. Waugh, F.V.: A Comment. Quarterly Journal of Economics, Vol. LXXXVI, No. 3, August 1972.

Reprinted from JOURNAL OF ECONOMIC THEORY
Vol. 8, No. 2, June 1974

Optimal Capital Adjustment under Uncertainty*

CARL J. NORSTRØM

Norwegian School of Economics and Business Administration, 5000 Bergen, Norway

Received March 6, 1973

1. INTRODUCTION

The purpose of this article is to analyze the effect of existing capital stock on the optimal investment policy of the firm under conditions of uncertainty.

Pioneering work on the optimal capital adjustment of the firm under certainty has been done by Arrow, Beckmann, and Karlin [3], and Eisner and Strotz [6]. Their work has later been extended by other authors like Arrow [1, 2], Gould [7], Lucas [9, 10], Rothschild [12], and Treadway [14]. More recently Hartman [8] introduced uncertainty in a model with constant returns to scale and strictly convex cost of adjustment.

The article is organized as follows. After the presentation of the model in Section 2, existence of an optimal investment policy is proved in Section 3. The connection between the static returns to scale and the value of the firm is established in Section 4. The implications for the optimal investment policy are discussed in Section 5. It is assumed throughout that the firm's objective is to maximize the expected present value of the future cash flows, or in other words, that it is risk neutral.

2. THE MODEL

To avoid mathematical difficulties, time will be taken to be discrete. The situation of the firm at the beginning of period t is assumed to be completely described by the capital stock K_t and a finite-dimensional vector P_t. This vector may consist of elements like the price of the product produced by the firm, the wage rate, parameters for the cost of investment, etc., in the current and a finite number of previous periods.

* This paper is an extension of a part of the author's licentiat thesis [10]. The present version owes much to a stay at the University of Michigan, made possible by a scholarship from the Norway-American Association. The author is indebted to Karl Borch, Jan Mossin, Agnar Sandmo, Sidney G. Winter, Jr., and a referee for valuable comments.

The vector P_t is restricted to lie in a bounded set $\mathscr{P}$. We let the function $F(P_{t+1} \mid P_t)$ defined on $\mathscr{P} \times \mathscr{P}$ be the conditional probability distribution of P_{t+1} given P_t. Hence the vectors P_t, P_{t+1},... form a stationary Markov process.

The dynamic behavior of the capital stock is governed by

$$K_{t+1} = (1 - \delta) K_t + I_t, \tag{1}$$

where δ is the depreciation rate and I_t is the gross investment in period t measured in physical units. The cost of the gross investment is given by a cost function $C(I_t, P_t)$. It is assumed that $C(I, P)$ is continuous in P and twice differentiable in I, that the first derivative $C_1(I, P) \geqslant 0$ and that $C(0, P) = 0$. Moreover the gross investment is assumed to be non-negative and bounded, i.e.,

$$0 \leqq I_t \leqq \bar{I}, \tag{2}$$

where $\bar{I}$ is a given constant. The net operating profit of the firm in any period is a given function $g(K_t, P_t)$ of the capital stock and the vector P_t. Any optimization necessary to obtain this function is assumed to be of myopic character. $g(K, P)$ is taken to be continuous in P, twice differentiable in K, and bounded for finite K and $P \in \mathscr{P}$.

The objective of the firm is to maximize the expected present value of the future net cash flow, that is, the expectation of

$$\sum_{t=1}^{\infty} R^t[g(K_t, P_t) - C(I_t, P_t)], \tag{3}$$

where the initial capital stock K_1 is given and $R = (1 + r)^{-1}$ is a discount factor $(0 < R < 1)$.

3. Existence of an Optimal Investment Policy

It is known from the formulation of the model that the situation of the firm at the beginning of period t is completely described by the capital stock K_t and the vector P_t. We define $V(K_t, P_t)$ to be the expected present value of all future cash flows of the firm, when an optimal investment policy is used. In the absence of risk aversion the function will also represent the value of the firm. Proceeding formally from the *principle of optimality*,[1] it is seen that

$$V(K_t, P_t) = \sup_{I_t} R \Big[g(K_t, P_t) - C(I_t, P_t) + \int_{\mathscr{P}} V(K_{t+1}, P_{t+1})\, dF(P_{t+1} \mid P_t) \Big]. \tag{4}$$

[1] Bellman [4], p. 83.

Here the two first terms on the right side represent the net cash inflow in the present period, and the last term the expected present value of the total cash flow from next period onward. Using (1) and the fact that P_{t+1} enters only as a vector of integration, (4) may be reformulated as

$$V(K, P) = \sup_I R \left\{ g(K, P) - C(I, P) + \int_{\mathscr{P}} V[(1 - \delta)K + I, Y]\, dF(Y \mid P) \right\}, \tag{5}$$

where for convenience subscripts have been dropped.

It must be shown that the function $V(K, P)$ exists and is unique. In the proof we shall use the fact that the capital stock is bounded.

LEMMA 1. *The stock of capital in any period lies in the interval* $[0, \bar{K}]$, *where* $\bar{K} = \max[K_1, \delta^{-1}\bar{I}]$.

Proof. The proof is by induction.

The existence and uniqueness of the value of the firm now seems plausible from an economic point of view, since the net cash flow in any period is bounded, and the discount factor is less than unity.

PROPOSITION 1. *There exists a unique bounded and continuous function* $V(K, P)$ *satisfying* (5).

Proof.[2] Let Γ be the space of all bounded, continuous functions defined on $[0, \bar{K}] \times \mathscr{P}$ with the metric

$$\rho(V, U) = \sup_{K,P} | V(K, P) - U(K, P)|. \tag{6}$$

It is known from real analysis that Γ is a complete metric space. We shall show that the transformation

$$T[V(K, P)] = \sup_I R \left\{ g(K, P) - C(I, P) + \int_{\mathscr{P}} V[(1 - \delta)K + I, Y]\, dF(Y \mid P) \right\} \tag{7}$$

is a contraction mapping, i.e., there exists a positive number $b < 1$ such that

$$\rho[T(V), T(U)] \leq b\rho(V, U) \tag{8}$$

[2] A more general approach to existence and uniqueness in dynamic programming may be found in Denardo [5].

for any pair of functions V and U in Γ. The proposition then follows, since it is known that every contraction mapping defined on a complete metric space has a fixed point, or that there is a unique function in Γ satisfying (7). From (7) and Lemma 1

$$\begin{aligned} &\sup_{K,P} \{T[V(K, P)] - T[U(K, P)]\} \\ &\quad \leqq \sup_{I,K,P} R \int_{\mathscr{P}} \{V[(1-\delta)K + I, Y] - U[(1-\delta)K + I, Y]\}\, dF(Y \mid P) \\ &\quad \leqq R \int_{\mathscr{P}} \sup_{K,Y} [V(K, Y) - U(K, Y)]\, dF(Y \mid P) \leqq R\rho(V, U). \end{aligned} \tag{9}$$

Repeating (9) with V and U interchanged gives (8) with $b = R$. Q.E.D.

We shall next establish the existence of an optimal investment policy. Let $\Phi(K, P)$ be the set of optimal investments for a given pair (K, P). Since the functions

$$W(K_{t+1}, P_t) = \int_{\mathscr{P}} V(K_{t+1}, P_{t+1})\, dF(P_{t+1} \mid P_t) \tag{10}$$

and $C(I, P)$ are continuous and the domain of I compact, it follows from (5) that $\Phi(K, P)$ will be nonempty. We define an optimal investment policy as a function $I(K, P)$ with value belonging to the set $\Phi(K, P)$ for each pair (K, P). The existence of an optimal investment now follows from the fact that $\Phi(K, P)$ is nonempty.

4. Returns to Scale and the Value of the Firm

We shall in this section investigate the effect on the value of the firm of three different assumptions on the returns to scale in the net operating profit. It should be noted that "returns to scale in $g(K, P)$" is essentially a static and short run concept describing how the net operating profit in a given period varies with the capital stock. In contrast, the value of the firm $V(K, P)$ incorporates the opportunity to adjust the capital stock to a more desired level, and in this way takes into account the dynamics of the model. The three cases to be considered are

(1) Globally decreasing returns to scale; $g_{11}(K, P) \leqq 0$ for all K and P.

(2) Constant returns to scale; $g(K, P) = \gamma(P)\, K$.

(3) Globally increasing returns to scale; $g_{11}(K, P) \geqq 0$ for all K and P.

The effect on the value of the firm is given in the following three propositions.

PROPOSITION 2. *If for all values of P the net operating income is a concave function in K and the cost of investment a convex function in I, then the value of the firm V(K, P) is a concave function in K.*

Proof. Consider the sequence of functions $\{V_n(K, P)\}$ defined by

$$V_0(K, P) = 0, \tag{11}$$

$$V_n(K, P) = T[V_{n-1}(K, P)]. \tag{12}$$

It follows from Proposition 1 that the sequence converges uniformly to $V(K, P)$, which hence will be concave in K if $V_n(K, P)$ is concave in K for any n. We shall proceed by induction. $V_0(K, P)$ is concave in K. Assume $V_{n-1}(K, P)$ is the same. Let K^1 and K^2 be arbitrary values of K, and

$$K^\lambda = \lambda K^1 + (1 - \lambda) K^2, \tag{13}$$

where $0 \leqq \lambda \leqq 1$. Let I^1 and I^2 be optimal investments for K^1 and K^2, respectively, with the given n, and define

$$I^\lambda = \lambda I^1 + (1 - \lambda) I^2. \tag{14}$$

It follows immediately from (13) and (14) that

$$(1 - \delta) K^\lambda + I^\lambda = \lambda[(1 - \delta) K^1 + I^1] + (1 - \lambda)[(1 - \delta) K^2 + I^2]. \tag{15}$$

Since the optimal investment for K^λ (and n) will give at least as good a result as I^λ,

$$V_n(K^\lambda, P) \geqq R \Big\{ g(K^\lambda, P) - C(I^\lambda, P) + \int_{\mathscr{P}} V_{n-1}[(1 - \delta) K^\lambda + I^\lambda, Y]\, dF(Y \mid P) \Big\}. \tag{16}$$

The fact that

$$V_n(K^\lambda, P) \geqq \lambda V_n(K^1, P) + (1 - \lambda) V_n(K^2, P)$$

follows easily from the assumptions in the proposition, the induction hypothesis, and Eqs. (15) and (16). Q.E.D.

PROPOSITION 3. *If for all values of P the net operating income g(K, P) is a linear homogeneous function of K, say $\gamma(P) K$, than the value of the firm is of the form*

$$V(K, P) = \alpha(P) K + \beta(P). \tag{17}$$

Proof. It is sufficient to show that the function $V_n(K, P)$ defined by (11) and (12) is of the required form for any n. This is easily done by induction.

PROPOSITION 4. *If for all values of P the net operating income g(K, P) is a convex function in K, then the value of the firm V(K, P) is convex in K.*

The proof is similar to the proof of Proposition 2 and is omitted.

5. Implications for the Optimal Investment Policy

The results derived in the previous sections will now be used to derive properties of the investment policy under different assumptions on the net operating income and cost of investment. In the analysis we shall disregard the possibility that there may be more than one optimal investment for certain values of the pair (K, P), and assume that the optimal investment policy $I^*(K, P)$ is unique.[3]

Using Eq. (10), (5) may be reformulated:

$$V(K, P) = \sup_I R\{g(K, P) - C(I, P) + W[(1 - \delta)K + I, P]\}. \quad (18)$$

Economically interpreted $W(K_{t+1}, P_t)$ is the expectation at time t of the firm's value at time $t + 1$ with the capital stock K_{t+1}. We shall assume at this point that the function is twice differentiable in K_{t+1}. Differentiation of the expression in the parentheses on the right hand side twice with respect to I, gives the sufficient conditions for an interior maximum:

$$-C_1(I, P) + W_1[(1 - \delta) K + I, P] = 0, \quad (19)$$

$$-C_{11}(I, P) + W_{11}[(1 - \delta) K + I, P] < 0. \quad (20)$$

From the equilibrium condition we obtain for a fixed value of P and an interior solution of I^*, that

$$\partial I^*/\partial K = (1 - \delta) W_{11}/(C_{11} - W_{11}). \quad (21)$$

It is seen from (20) and (21) that the effect on investment depends only on the sign of W_{11}. From Definition (10) it is easily established that W is, respectively, concave, linear, or convex in K_{t+1} when V is concave, linear, and convex in K. Hence we obtain from Propositions 2–4:

PROPOSITION 5. *Let $I^*(K, P)$ be the optimal investment policy. Then*

$$\partial I^*/\partial K \leqslant 0$$

[3] A more general approach would be to consider the correspondence $\Phi(K, P)$, which may be shown to have properties similar to the properties given in Proposition 5 for the investment policy. For example, $\max\Phi(K, P)$ and $\min\Phi(K, P)$ are both monotonically increasing in K if $g(K, P)$ is convex, constant if $g(K, P)$ is linear homogeneous, and monotonically decreasing if $g(K, P)$ is concave and $C(I, P)$ convex.

if $g(K, P)$ is concave in K and $C(I, P)$ is convex in I for all P;

$$\partial I^*/\partial K = 0$$

if $g(K, P)$ is linear homogeneous in K for all P;

$$\partial I^*/\partial K \geqq 0$$

if $g(K, P)$ is convex in K for all P.

Comment. It is possible to show by construction of counterexamples, that "decreasing returns to scale in the net operating income" is not alone a sufficient condition for the value of the firm to be concave in K or the optimal investment policy to be a decreasing function of K.

With the possible exception of the case of decreasing returns to scale, it is seen from Proposition 5 that the effect of existing capital stock on the investment policy is as anticipated by economic intuition. *Ceteris paribus* gross investment will increase with the size of the firm under increasing returns to scale, not be influenced under constant returns, and (normally) decrease with size under decreasing returns. Some care must be taken, however, in comparison between different firms. There is no *a priori* reason for all firms to have an identical cost structure, and the fact that a firm is big may well be a consequence of a uniformly high net operating income function, in which case the *ceteris paribus* assumption does not hold.

While Proposition 5 gives some information on the form of the investment policy under fairly general conditions, stronger results may be obtained by a strengthening of the assumptions. The remainder of the section will be used to discuss such cases, of which some will correspond to known results in deterministic theory.

A. *Constant Returns to Scale*[4]

It is seen from Proposition 5 that the investment decision is not influenced by the level of existing capital stock. As has been pointed out by Hartman [8], this implies that the current investment decision has no effect on the profitability of future investments and may be taken independently of these.

With constant returns to scale (19) and (20) are reduced to

$$-C_1(I, P) + \eta(P) = 0,$$
$$-C_{11}(I, P) < 0,$$

[4] The results in this subsection have been proved by Hartman [8] for the case of a strictly convex cost of investment.

where the function $\eta(P)$ is given by the integral equation

$$\eta(P) = R \int_{\mathscr{P}} \gamma(Y)\, dF(Y \mid P) + R(1 - \delta) \int_{\mathscr{P}} \eta(Y)\, dF(Y \mid P).$$

Expansion into the Neumann series gives

$$\eta(P) = \sum_{t=1}^{\infty} \left[R^t(1 - \delta)^t \int_{\mathscr{P}} \gamma(Y)\, dF^{(t)}(Y \mid P)\right],$$

where $F^{(t)}(Y \mid P)$ is the tth convolution of $F(Y \mid P)$. Economically interpreted $\eta(P)$ is the expectation as evaluated at time t of the marginal (and average) value of capital stock at time $t + 1$. The optimal investment is found where the marginal value of capital equals the marginal cost of acquiring it.

B. *Cost of Investment Is Linear in I*

Arrow assumes in Ref. [1] that the cost of investment is linear homogeneous. The consequences of this assumption for the stochastic case will now be considered. Let

$$C(I, P) = \psi(P)\, I, \qquad 0 \leqq I \leqq \bar{I}.$$

Equations (19) and (20) are reduced to

$$-\psi(P_t) + W_1(K_{t+1}, P_t) = 0,$$
$$W_{11}(K_{t+1}, P_t) < 0.$$

Constant or increasing "returns to scale" leads to bang–bang solutions where the optimal investment is either zero or $\bar{I}$. With decreasing returns to scale there is for each P an optimal capital stock $K^*(P)$, independent of the initial capital stock. The following result is easily proved:

PROPOSITION 6. *Suppose that $g(K, P)$ is concave in K and $C(I, P)$ is linear homogeneous in I. Then for each P there exists an optimal long run capital stock $K^*(P)$, and there is an optimal investment policy of the form*

$$\begin{aligned}
I^*(K, P) &= 0, & K^*(P) &\leqq (1 - \delta) K, \\
&= K^*(P) - (1 - \delta) K, & (1 - \delta) K < K^*(P) &\leqq (1 - \delta) K + \bar{I}, \\
&= \bar{I}, & (1 - \delta) K + \bar{I} &< K^*(P).
\end{aligned}$$

Proposition 6 is a stochastic counterpart to a result in Arrow [1]. Arrow's net operating income is, however, more general than a concave function.

C. *Minimum Size of Investment*

While many authors assume that the cost of investment is globally convex, Rothshild [12] has given convincing arguments for a more general cost of adjustment, e.g., cost functions which are initially concave and ultimately convex. It will be shown that under constant or increasing returns to scale there is a lower bound for a positive optimal investment, viz., the highest point where the average cost of investment is at a minimum.

Let $J(P)$ be defined for each P as follows:

$$J = 0 \quad \text{if} \quad \lim_{J \to 0} [C(J, P)/J] < C(I, P)/I, \quad 0 < I \leqq \bar{I},$$

and otherwise the largest number in $(0, \bar{I}]$ satisfying

$$C(J, P)/J \leqslant C(I, P)/I, \qquad 0 < I \leqq \bar{I}.$$

Clearly $J(P) > 0$ if $C(I, P)$ is initially concave.

PROPOSITION 7. *Suppose that $g(K, P)$ is convex in K for all P. Then for each value of P such that $I = 0$ is not an optimal policy, $I^*(P) \geqq J(P)$.*

Proof. Since I^* is optimal and $I = 0$ is not,

$$\lambda W[(1 - \delta) K, P] + (1 - \lambda) W[(1 - \delta) K + J, P] - (1 - \lambda) C(J, P)$$
$$< W[(1 - \delta) K + I^*, P] - C(I^*, P)$$

for $0 < \lambda \leqq 1$. If $0 < I^* < J$, a contradiction is obtained by setting $\lambda = (J - I^*)/J$. Q.E.D.

D. *Independence*

The assumption of constant prices in deterministic theory does not correspond in stochastic models to stationarity, but to stationarity and independence. Significant dynamic aspects of the model disappear when the probability distribution $F(P_{t+1} \mid P_t)$ is assumed to be independent of P_t.

From (10) it follows that W under independence is a function only of the capital stock. Uncertainty of the future investment levels comes only from the uncertainty in the cost of investment. If in addition the function C is independent of P, the first order condition (19) reduces to

$$-C'(I) + W'[(1 - \delta) K + I] = 0,$$

and the path of the capital stock is deterministic. The degree of uncertainty may still influence the level of investment. As shown by Hartman [8], a higher degree of uncertainty in the sense of Rothschild and Stiglitz [13] will in the case of constant returns to scale lead to an increase in the expected benefit from the investment and hence to a larger investment. This follows from the first order condition, which now takes the form

$$-C'(I) + \int_{\mathscr{P}} \gamma(Y)\, dF(Y)/(r + \delta) = 0, \tag{22}$$

and the fact that $\gamma(P)$ is convex in P. Equation (22) is a stochastic counterpart to a first order condition under constant prices given by Gould [7].

References

1. K. J. Arrow, Optimal capital adjustment, *in* "Studies in the Mathematical Theory of Investment and Production" (K. J. Arrow, S. Karlin, and H. Scarf, Eds.), Stanford U. P., Stanford, CA, 1962, pp. 1–17.
2. K. J. Arrow, Optimal capital policy with irreversible investment, *in* "Value, Capital and Growth, Papers in Honour of Sir John Hicks" (J. Wolfe, Ed.), Aldine, Chicago, IL, 1968, pp. 1–19.
3. K. J. Arrow, M. Beckmann, and S. Karlin, The Optimal expansion of the capacity of a firm, *in* "Studies in the Mathematical Theory of Investment and Production" (K. J. Arrow, S. Karlin, and H. Scarf, Eds.), Stanford U. P., Stanford, CA, 1958, pp. 92–105.
4. R. Bellman, "Dynamic Programming," Princeton U. P., Princeton, NJ, 1957.
5. E. V. Denardo, Contraction mappings in the theory underlying Dynamic Programming, *SIAM Review* **9** (1967), 165–177.
6. R. Eisner and R. Strotz, Determinants of business investment, research study two, in "Impact of Money Policy," prepared for the Comission on Money and Credit, Prentice Hall, Englewood Cliffs, NJ, 1963.
7. J. D. Gould, Adjustment costs in the theory of investment of the firm, *Rev. Econ. Stud.* **35** (1968), 47–55.
8. R. Hartman, The effect of price and cost uncertainty on investment, *J. Econ. Theory* **5** (1972), 258–266.
9. R. Lucas, Adjustment costs and the theory of supply, *J. Pol. Econ.* **75** (1967), 321–334.
10. R. Lucas, Optimal investment policy and the flexible accelerator, *Int. Econ. Rev.* **8** (1967), 78–85.
11. C. J. Norstrøm, An essay in capital budgeting, Licentiat thesis, The Norwegian School of Economics and Business Administration, 1967 (mimeographed).
12. M. Rothschild, On the cost of adjustment, *Quart. J. Econ.* **85** (1971), 605–622.
13. M. Rothschild and J. E. Stiglitz, Increasing risk: I. A definition, *J. Econ. Theory* **2** (1970), 66–84.
14. A. B. Treadway, On rational entrepreneurial behaviour and the demand for investment, *Rev. Econ. Stud.* **36** (1969), 227–239.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.