Discussion Paper No. 1002

ON DETERMINING THE IMPORTANCE OF ATTRIBUTES WITH A STOPPING PROBLEM*

by

Zvika Neeman**

June 1992

* I would like to thank Ariel Rubinstein, Isaac Meilijson, Asher Wolinsky and Itzhak Gilboa for helpful discussions.

** Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208.

ABSTRACT

One of the approaches in consumer theory considers each product as a collection of attributes. Consequently, consumer preferences are defined over attributes. As opposed to the traditional approach, according to which consumer preferences for products are the underlying feature of economic modeling, they are now derived from the composition and strength of products' attributes.

In this paper I try to answer the question of how one can determine the relative importance of the different attributes of a product. In order to answer this question a stopping problem model is constructed. An agent faces a sequence of i.i.d. multi-dimensional products of which he can observe only one attribute. At each stage the agent has to decide whether he wants to stop, taking the best product so far, or whether he prefers to continue by observing a specific attribute of the next product. The model is solved for an optimal observing policy.

In the finite case, second order stochastic dominance characterizes the optimal strategy in the sense that if it holds between the two random variables induced by the expected utility given an attribute, it is never optimal to observe the "dominating" one.

In the infinite horizon case, observing one attribute only is always optimal. However, the infinite horizon optimal strategy may not be myopically optimal. The seeming discrepancy between finite and infinite horizon models vanishes for a sufficiently large horizon, thus making the infinite case optimal attribute the one chosen for a long period in finite horizon cases also. Identifying the infinite case optimal attribute allows us to determine the performance of the model in the long run even when second order stochastic dominance does not hold.

1. INTRODUCTION AND SUMMARY

One approach in consumer theory, first presented by Lancaster 1966, states that the economic agent or the consumer regards each product as a bundle of attributes. Therefore, the consumer's preferences are defined over the set of different combinations of attributes which are identified as products, rather than assuming the preferences over products as a primitive. Different composition and magnitude of attributes among products determines the consumer's preferences over the different products.

This paper tries to develop a systematic way of answering the question, "What makes a certain attribute important ?". In order to address this problem we introduce the following decision problem: A rational decision maker is facing a sequence of multi-dimensional products. He has to choose one product, preferably the best, from this sequence. However, the decision maker is restricted by the following information constraint, he can observe only one attribute from each product in the sequence[1].

An example illustrating this process might be the following; One is interested in buying a used car. He visits several car dealers which in turn, each make him an offer. However, for various reasons (e.g. the dealer doesn't have time, there are more costumers waiting to examine the car, these are the prevailing social norms), he can have only a limited amount of time in which he can test each car. Then he can decide if he wants to buy the car or is he interested in observing another car. At any stage he can come back and buy the car that he already observed before. His problem then, is, which attributes of the car should he be most interested in examining. Is there one attribute that provides the best information about the car or should he examine each time a different attribute, maybe as a

[1] Given a full description of the products he is facing, the decision maker can choose optimally - He is facing a usual stopping problem.

function of previous results.

Another example is the following: Suppose that a university department is interested in hiring a young assistant professor. It summons up candidates and interviews them one at a time. After reviewing each candidate the department has to decide whether to make an offer to the strongest candidate so far or to continue to interview new candidates. In this example, the restrictions of the model seem more natural, since each interview is costly (the department has to fly the candidate and host him for a couple of days.) Furthermore, it is not customary to interview a candidate twice and during these couple of days only limited information can be gathered on the candidate. The department therefore has to decide in what kind of information it is more interested.

In this context, this paper tries to understand the dynamic nature of what makes a certain attribute interesting, where "interest" is defined implicitly by the optimal observation policy. However, this analysis can be thought of as having descriptive implications as well, suggesting that what might appear at a particular moment as interesting is indeed an optimal observation. In other words, when people are confronted with problems of this kind their indeterminacy and inconsistency may be the result of unconscious optimizing rather than sheer arbitrariness. The notion of a decision maker that unconsciously chooses optimally seems more plausible than that of the (boundedly) rational decision maker that is fully aware of the model and calculates his optimal strategy. Furthermore, the first approach can be justified on an evolutionary basis.

Similar models of decision processes were considered in the psychological literature. These models, (see Coombs 1964; Fishburn 1968; Tversky 1972), concern themselves with the following problem: A decision maker faces a set of different products (alternatives) from which he has to choose only one. They propose the following decision rule: The decision maker chooses one attribute (either deterministically or probabilistically). All the products (alternatives) that fail to have this attribute are disqualified. This process is repeated until only one product (alternative) is left, and this last product ends up being chosen. However, whether the particular way of choosing the attributes is deterministic or probabilistic, the way in which this attribute is being chosen is left implicit.

Two other related areas of research are: Search literature and the multi-armed bandit problem. In search literature (for a survey, see McMillan and Rothschild 1989), generally, a search problem is constructed and various search strategies are discussed. However, in this paper the motivation is quite different. The basic question we ask is, which attribute observation policy will better facilitate our search. Consequently, there is no "optimal attribute" that we search for, rather, the most informative attribute may vary along time and history. However, our problem is embedded in a search problem in the following sense: using the attribute observation policy we search for a good enough product and when we find one, we stop.

In the multi-armed bandit problem (see Berry and Fristedt 1985) one chooses among several distributions with unknown parameters. Each distribution yields a certain payoff and the objective function is to maximize the expected cumulative payoff. Since the parameters are unknown, each draw from a distribution provides, apart from the immediate payoff, some information about the distribution. Therefore, when devising a strategy one must consider incurring a certain loss at an early stage in order to gain information that will be valuable later. However, in the present model there is no reason for experimentation since all the parameters are known a-priori. While one can think of it as a degenerate bandit problem it is focused on the information acquisition aspect of the strategy.

As mentioned above, bounded perception is reflected in our model in the assumption that only one attribute may be observed at a time. While this extreme assumption may be too restrictive (and is made in our model mainly for tractability reasons) we find that it does not make the model much less realistic. Indeed, very often products are way too complex for the consumer to observe all their attributes, whether a used car, or a job applicant is concerned. Furthermore, the cost of having an observation can be very high considering time and other resources, rendering more than one observation per product practically impossible (e.g. the problem of interviewing job applicants). In addition, the model allows to lump several attributes together into one attribute. For instance, in the recruiting-new-faculty example, applicants are questioned about their former studies and publications and are required to give a research seminar, all of which can be redefined as a "research" attribute, but usually are not tested for their performance in front of an ordinary class (the "teaching" attribute).

Yet another justification for the bounded perception model stems from biological evidence: there are natural mechanisms that force a limitation very simi liar to the one described. The human eye, for instance, can observe only a limited range in a certain time, although one can choose the place one wants to look at.

In this paper the optimal observation problem is solved for the finite and infinite horizon cases. In the finite case, we construct two examples that help to clarify and motivate the main results of the paper. The first example shows a problem where an optimal strategy observes two attributes, whereas in the second one only one attribute is observed by the optimal solution.

Example 1: Consider a two period problem, N=2. The decision maker has to choose among products which are each defined by a pair of values, one for attribute **p** and one for attribute q. Each attribute may take the values 0 or 1 according to the following distributions which we assume to be independent.

x^p	0	1
	3/4	1/4

x^q	0	1
	1/3	2/3

The utility function is $V(x^p,x^q)=x^p+x^q$ and is discounted with $\beta<1$. Full analysis of the problem is carried out by backward induction.

Computation, presented in the appendix, shows that the optimal strategy (for $12/13<\beta$) is: First, observe attribute p. If you see $x^p=1$, stop and get a utility value of $5\beta/3$. If, however, you observe $x^p=0$, continue to observe attribute q in the next round. If you see $x^q=1$, take it and get a utility value of $5\beta^2/4$. If you see $x^q=0$ go back and take the previous observation ($x^p=0$), in this case you get a utility value of $2\beta^2/3$.[2]

As we shall see later, in the infinite horizon case, for $3/4<\beta$ an optimal strategy will always prefer observing the p attribute.

The economic interpretation of this example can be presented as follows. The q attribute has a large probability of success, much larger than that of p. However, a "success" in the p-dimension guarantees a higher conditional expected payoff. In the first period the decision maker is willing to bear a risk and he observes attribute p. If he wins, he stops. If he fails he moves over to observe the q attribute in the hope that it,

[2] However, for lower β values ($3/11<\beta<12/13$) in the first period an optimal strategy will prefer observing characteristic q because observing $x^p=0$ will not allow an additional observation while observing $x^q=0$ will.

at least, will guarantee a better than average conditional expected payoff. If it does he takes it. If not, he has no choice and he takes the first product. Since X^p is more likely to be zero than X^q, the q attribute functions here as some sort of "insurance". Loosely, the fact that X^q is "safer" than X^p allows the decision maker to bear risk in the first period, knowing that unless a "disaster" occurs (i.e. with a small probability) he is covered.

Example 2: Again we present a two period problem, N=2. The decision maker is faced with two attributes p,q, with the following distributions,

x^p	0	3
	1/2	1/2

x^q	0	1
	1/2	1/2

The utility function is as before $V(x^p,x^q)=x^p+x^q$. Again, analysis of the problem will be carried out by backward induction, in the same manner as in example 1. The computation is done in the appendix.

The optimal strategy is: First observe attribute p. If you see $x^p=3$, stop and get a value of $9\beta/2$. If you see $x^p=0$ observe again, it does not matter whether you choose to observe attribute p or attribute q, either one will give an expected utility value of $3\beta^2$. We see that an optimal strategy here can be restricted to observing attribute p alone. In this sense, attribute p is more interesting than attribute q.

We see a major difference between the two examples, in the first one an optimal strategy observes both attributes while in the second example an optimal strategy can be restricted to observe only one attribute. We wish to generalize this distinction by referring to the differences in the distributions of the attributes only.

In the finite case, a necessary and sufficient condition for an optimal strategy to observe only one attribute for any time horizon N, is that the random variable induced by the expected utility given this attribute is second order stochastically dominated by those corresponding to all other attributes. At first sight it might appear unintuitive that the dominated variable is chosen by the optimal solution. Recall, however, that for a random variable to be second order stochastically dominated implies that it reveals more information and that the decision here is what attributes to observe rather than what random variables to consume.

By contrast, in the infinite horizon case, observing one attribute only is always optimal. Yet, the infinite horizon optimal strategy may fail to be myopically optimal. Namely, there are cases in which, if we consider only a one-period future, the optimal infinite horizon strategy will be strictly sub-optimal. The apparent discrepancy between the finite and infinite horizon cases vanishes asymptotically. Although in the finite horizon case one can always construct examples where every optimal strategy may decide to observe two attributes (or more), the probability that this event actually occurs approaches zero as the horizon tends to infinity.

Therefore, it turns out that what makes an attribute attractive involves more than "being informative" in the sense of being stochastically dominated. Attractive attributes are, in the long run, those which have the possibility of pulling the whole product assessment sharply upwards, no matter how small is the probability of that occurring in any given stage.

The general structure of the paper is as follows: The second section presents the model. The third section deals with the finite horizon case, presenting the characterization of an optimal strategy in terms of second order stochastic dominance. Finally, the fourth section deals with the infinite horizon case, demonstrates that observing one attribute only is optimal and explains the connection between the finite and infinite horizon cases by showing continuity of the finite horizon analysis at infinity.

2. THE MODEL

Denote the attributes of a generic product by X^i, $i=1,..,k$. The X^i's are independent random variables $X^i:\Omega \longrightarrow R$ and denote the cumulative distribution function of X^i by $G_i(\cdot)$.

The sequence of products is denoted by $X_1, X_2, \ldots$ where $X_t=(X_t^1, X_t^2, \ldots, X_t^k)$, $t\in\mathbf{N}$. The X_t's are i.i.d. random variables $X_t:\Omega \longrightarrow R^k$ where $t\in\mathbf{N}$. By independence of the attributes, each X_t has $G(\cdot)=\prod_{i=1}^{k} G_i(\cdot)$ as its cumulative distribution function.

We will assume that the decision maker has preferences which are given by a bounded and continuous von Neumann-Morgenstern utility function, $V(\cdot)$, $V:R^k \longrightarrow R^+$, which will be discounted by $0<\beta<1$ and where R^k is the products' space.

The information restriction can be described by the sequence of attributes the decision maker wishes to observe at each stage $a_1, a_2, \ldots$ where $a_t\in\{1,\ldots k\}$. Thus, the decision maker would seek to maximize $E[V(X^1,\ldots,X^k)|x^a]$, where x^a is the realization of the a-th attribute that the decision maker chose to observe. This is as close as he could get to maximizing $V(x^1,\ldots,x^k)$ (which he would do in the case of no restriction). Here and throughout the paper capital letters X,Y, will denote random variables and lower case letters x,y, will denote realizations (numbers) of the corresponding random variables.

The observation procedure is as follows: The decision maker goes through the sequence $X_1, X_2, \ldots$. From each product X_t he observes one attribute a_t, until he decides to stop. When a decision to stop is reached, the decision maker takes the best product he has seen so far.

This procedure yields a sequence of random variables $Y_1, Y_2, \ldots$ which are defined as:

$$Y_t = \beta^t \max_{j=1..t} \left\{ E[V(X_j)|x_j^{a_j}] \right\}$$

Describing this problem as a stopping problem, at time t the decision maker has to decide - whether to take y_t and thus stop the procedure , or continue the observation procedure by choosing a_{t+1}, the attribute of the t+1-th product he wishes to observe[3].

Describing this setup as a dynamic programming problem one has to characterize the states set, the actions set, a transition function and a payment function.

States set: $S = R_+ \cup \{\phi\}$

Where a state $s \in S$ has the following interpretation: Suppose we are in period n, then $s = \max_{j=1..n} \{E[V(X_j)|x_j^{a_j}]\}$. Notice that knowing the state does not imply knowing the history of the process nor the number of periods that have passed since the beginning of the problem. Still, at time n, $s = y_n/\beta^n$ where y_n is the realization of Y_n, which means that the state description contains all the relevant information for the problem. ϕ is an absorbing state, denoting the end of the process.

Actions set: $A = \{1,\ldots,k,\text{"STOP"}\}$

At each state the decision maker chooses an action which may be to observe an attribute a, $1 \le a \le k$, or to "stop" in which case he terminates the observation procedure, and chooses the best product observed so far.

Transition function: $q: S \times A \longrightarrow F(S)$

The transition function, given a state and an action, describes the distribution of states which follows. (F(S) is the family of cumulative distribution functions over the states set S.)

$q(s,a) = \phi$ with probability one if $s=\phi$ or a="STOP".

3 Notice that if we want to describe the case where the decision maker is facing a finite product sequence with length N, we can do that by fixing $Y_i = 0$ for $i > N$.

Otherwise, $(s\neq\phi, a=j, 1\leq j\leq k)$

$$q(s,a) = \begin{cases} 0 & s'\leq s \\ \dfrac{G_j(s')}{1-G_j(s)} & s<s' \end{cases}$$

Immediate payment function: $r: S\times A\times S \longrightarrow R$

The payment function r describes the immediate payoff given a state s and an action a which transforms the process into state s'.

$$r(s,a,s') = \begin{cases} s & s\neq\varphi, a=\text{"stop"}, s'=\phi \\ 0 & \text{otherwise} \end{cases}$$

The target function: $$\underset{a_1,a_2,\ldots}{\text{Max}} \quad E\left[\sum_{i=0}^{\infty} \beta^i r(s_i,a_i,s_{i+1})\right]$$

We are interested in finding an optimal strategy for this problem. A strategy for this decision problem should tell us what we should do in every period: stopping and taking the best product so far or continuing by observing an attribute of the next product. Any such strategy induces a random variable which is referred to as a *stopping rule*.

Formally, a *stopping rule* is a random variable $\tau=\{t;a_1,\ldots,a_t\}$ such that:

- t is a random variable denoting the stopping time.
- $a_1,\ldots,a_t$ are random variables denoting the attributes observed until stopping time.
- In time n, the decisions (a_n and whether t stops) must be a function of what is known until time n.

Denote the random variable representing the utility for the decision maker if he uses τ as his stopping rule by Y_τ.

Denote $V^* \equiv \underset{\tau}{\text{Sup}}\ EY_\tau$. The decision maker wishes to guarantee himself a utility as close as possible to V^{*}[4]. In this problem he can actually get V^{*}, i.e., there is a stopping rule τ^{*} that guarantees a value of V^{*}. This is because the problem defined here is a *monotone problem*. Extending Chow, Robbins and Ziegmund's definition to this stopping problem, *Monotonicity* for the stochastic sequence $\{Y_n, a_n\}_{n=1}^{\infty}$ is defined as follows. First denote by A_n the event that there is no expected gain from continuing one step past stage n, i.e.,

$$A_n = \left\{ \omega : E[Y_{n+1} | y_n, a_{n+1}] \leq y_n \ , \text{ for all } a_{n+1} \right\}$$

This is a *monotone case* if: $A_1 \subseteq A_2 \subseteq \cdots$

and, $\bigcup_{n=1}^{\infty} A_n = \Omega$ (up to zero probability events)

Verbally, this definition considers the choice between stopping at a certain point and continuing just one more period. A stochastic sequence is monotone if this "myopic" decision to stop will never be regretted, i.e., if a decision to stop at time t implies the same at time t+1 and if such a decision will always occur for some t.

Observation: *This problem*, $\underset{\tau}{\text{Max}}\ EY_{\tau}$, *is a monotone problem.*

Proof: Since the proof follows naturally after some more results are obtained, we present the proof later, in section 4.

[4] A necessary and sufficient condition for the existence of EY_{τ} is that $E[V(X)|x^{a}]$ is bounded for all n. Since we assumed that the utility function V( ) is bounded, EY_{τ} is always defined.

3. THE FINITE CASE

In the finite case, the time left until the end of the process is of crucial importance in devising an optimal strategy. The analysis is carried out by backward induction. First, the last period is analyzed, then, the next to last period situation is analyzed given the results of the previous analysis. In much the same way, in each period, the analysis is carried out when one knows the result of his actions in terms of distribution over states in the following period. The solution for the general finite horizon problem appears in the appendix.

The k random variables X^i yield k random variables $E[V(X)|X^i]:\Omega \longrightarrow R$ which are non negative and bounded.

It is shown that second order stochastic dominance between the $E[V(X)|X^i]$'s characterizes the optimal strategy in the finite horizon case in the following sense; If there exists an attribute p such that $E[V(X)|X^p]$ is second order stochastically dominated by the $E[V(X)|X^q]$'s of all of the other attributes, then there exists an optimal strategy that will choose to observe attribute p alone. If, on the other hand, there is no such attribute, every optimal strategy will observe at least two attributes. Before we turn to the proof, we present the definition of second order stochastic dominance as it appears in Rothschild and Stiglitz, 1970.

Definition: *For two bounded random variables P and Q, with bounded support [0,M] and with the corresponding cumulative distribution functions, $F_P(\cdot)$, $F_Q(\cdot)$, we say that P is second order stochastically dominated by Q if:*

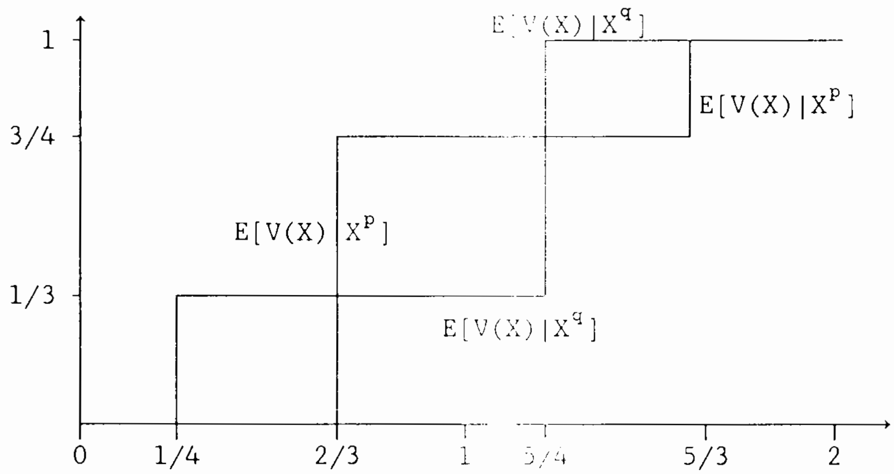
(i) $$\int_0^M (F_P(x) - F_Q(x))dx = 0$$

(ii) $\int_0^m (F_P(x) - F_Q(x))dx \geq 0$ for all $0 \leq m < M$

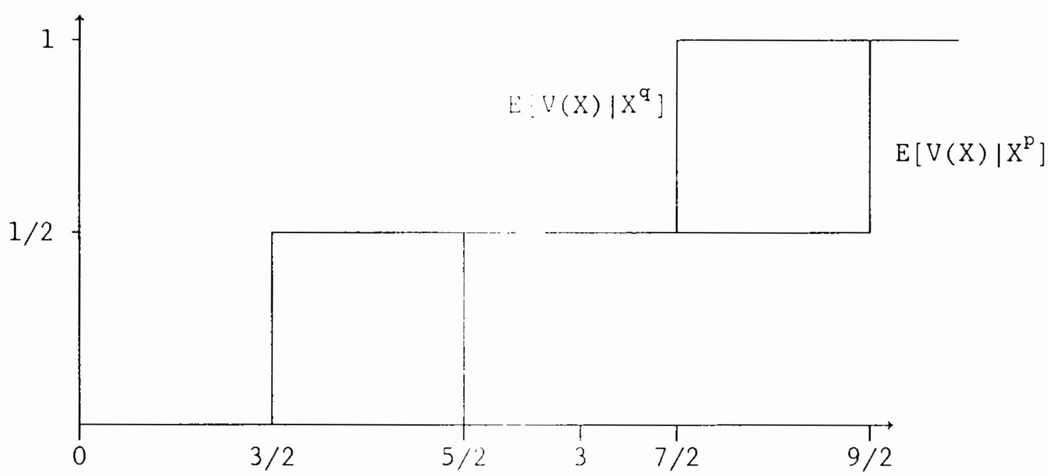
We will also use this term for attributes, namely:

Definition: *For two attributes p and q, p is said to second order stochastically dominate q if* $E[V(X)|X^p]$ *second order stochastically dominates* $E[V(X)|X^q]$.

In the examples given in the introduction, example 1 does not have second order stochastic dominance and in example 2, attribute p is stochastically dominated by attribute q.



Example 1



Example 2

Theorem 1: *In an $N\geq 2$ stage problem, there exists an optimal strategy that observes attribute p only with probability 1 if attribute p is second order stochastically dominated by all other attributes. Conversely, if the cumulative distributions of $E[V(X)|X^i]$, $1\leq i\leq k$, are strictly monotone and differentiable, and an optimal strategy observes attribute p only with probability 1, then attribute p is second order stochastically dominated by all other attributes.*

Proof: "If": First, we need some definitions. Define the random variable $U_i^1(v)$ to be,

$$U_i^1(v) = \mathrm{Max}\{v, E[V(X)|X^i]\}$$

Now we recursively define $U^n(v)$ and $U_i^{n+1}(v)$. For $n\geq 1$ let $U^n(v)$ be,

$$U^n(v) = \mathrm{Max}\left\{ v \ , \ \beta\cdot \mathrm{Max}_{1\leq i\leq k}\left\{EU_i^n(v)\right\} \right\}$$

and for $n\geq 2$, let $U_i^{n+1}(v)$ be,

$$U_i^{n+1}(v) = U^n\Big(\mathrm{Max}\{v, E[V(X)|X^i]\}\Big)$$

with the following interpretation: $U_i^1(v)$, which is a function of v and X^i, is the random variable that the decision maker is choosing by observing attribute i at the last period, with a reservation value v. $U^n(v)$, which is a function of v only, is interpreted as the value of the problem, n stages from the end, with reservation value v. $U^n(v)$ is indeed the value of the problem, it is the most the decision maker can hope to get. Furthermore, it encapsulates an optimal policy and $U^N(v)=V^*$ (as implied by the proof in the appendix). $U_i^{n+1}(v)$, again, a function of v and X^i, is the random variable that the decision maker is choosing by observing attribute i, n+1 periods before the end of the problem, with a reservation value v.

Lemma: $U^1_i(v)$, $U^n(v)$ *and* $U^{n+1}_i(v)$ *are convex functions of* v *for* $i=1,..,k$, $1\le n\le N$. $U^n_i(v)$ *is also convex in* $E[V(X)|X^i]$ *for all* $i=1,..,k$, $1\le n\le N$.

Proof: Since we always consider $Max\{v,E[V(X)|X^i]\}$ which is convex in $E[V(X)|X^i]$, showing that the above functions are convex in v suffices for $U^n_i(v)$ to be convex in $E[V(X)|X^i]$ for all $i=1,..,k$, $1\le n\le N$.

We prove the convexity of $U^n_i(v)$, $U^n(v)$ by induction on n:

n=1: We prove that $U^1_i(v)$ is convex in v.

Pick any $\lambda\in[0,1]$, $v_1,v_2\ge 0$. Denote $v_3=\lambda v_1+(1-\lambda)v_2$

$$\lambda\cdot Max\{v_1,E[V(X)|X^i]\} + (1-\lambda)\cdot Max\{v_2,E[V(X)|X^i]\} - Max\{v_3,E[V(X)|X^i]\}$$
$$= Max\{\lambda v_1,\lambda E[V(X)|X^i]\} + Max\{(1-\lambda)v_2,(1-\lambda)E[V(X)|X^i]\} - Max\{v_3,E[V(X)|X^i]\}$$
$$\ge Max\{\lambda v_1+(1-\lambda)v_2,\lambda E[V(X)|X^i]+(1-\lambda)E[V(X)|X^i]\} - Max\{\lambda v_1+(1-\lambda)v_2,E[V(X)|X^i]\}$$
$$= 0$$

n=k: We assume $U^k_i(v)$ is convex in v for $i=1,..,k$, and we prove that $U^k(v)$ and $U^{k+1}_i(v)$ are convex functions of v for $i=1,..,k$. First we show that convexity of $U^k_i(v)$ implies convexity of $EU^k_i(v)$.

Pick any $\lambda\in[0,1]$, $v_1,v_2\ge 0$. Denote $v_3=\lambda v_1+(1-\lambda)v_2$

$$\lambda EU^k_i(v_1) + (1-\lambda)EU^k_i(v_2) - EU^k_i(v_3)$$
$$= E\left[\ \lambda U^k_i(v_1) + (1-\lambda)U^k_i(v_2) - U^k_i(v_3)\ \right]$$
$$\ge 0$$

Because convexity of $U^k_i(v)$ for all values of X^i implies

$$\lambda U^k_i(v_1) + (1-\lambda)U^k_i(v_2) - U^k_i(v_3)\ge 0.$$

Now, $U^n(v)$ is convex since the maximum operation is convex. Similarly, $U^{n+1}_i(v)$ is convex since the maximum operation is convex and $U^n(v)$ is convex. ■

Using a theorem from Rothschild and Stiglitz 1970 we know that if a random variable X is second order stochastically dominated by another random variable Y, then for every convex function $U(\cdot)$, $EU(X)\ge EU(Y)$. Therefore, using the above lemma, we conclude $EU_p^n(v) \geq EU_q^n(v)$ for all $q=1,..,k$, $v \in [0,M]$ and $1 \leq n \leq N$. Which means that in every period observing attribute p is preferable to observing any other attribute q. Thus completing the proof of the "if" direction.

Conversely, We will demonstrate that in a two attribute case, if there is no second order stochastic dominance between the attributes then an optimal strategy will observe both attributes with positive probability. Without loss of generality , this implies that in a k attribute setting the absence of second order stochastic dominance implies that an optimal strategy will observe at least two attributes.

First we show that $\int_0^m (F_i(x)-F_j(x))dx > 0$ implies $EU_i^1(m) > EU_j^1(m)$.

Assume $\int_0^m (F_i(x)-F_j(x))dx > 0$. This implies $\int_m^M F_i(x)dx < \int_m^M F_j(x)dx$, since $\int_0^M F_i(x)dx = \int_0^M F_j(x)dx$ for all i and j. (Observe that this property holds in general since for any three random variables X,X',Y, $E[E[Y|X]]=E[E[Y|X']]$ and $E[X]=\int 1-F_X$ for bounded non negative X.)

Therefore,

$$m + \int_m^M (1-F_i(x))dx > m + \int_m^M (1-F_j(x))dx$$

However,

$$m + \int_m^M (1-F_i(x))dx \underset{\substack{\text{integ.}\\ \text{by parts}}}{=} m + (1-F_i(x))x\Big|_m^M + \int_m^M xf_i(x)dx$$

$$= m + (-m(1-F_i(m)) + \int_m^M xf_i(x)dx$$

$$= F_i(m)m + \int_m^M x f_i(x)\,dx$$

$$= E\{m \cdot 1_{\{E[V(X)|X^i] \le m\}}\} + E\{E[V(X)|X^i] \cdot 1_{\{E[V(X)|X^i] > m\}}\}$$

$$= EU_i^1(m)/\beta$$

Which implies, $EU_i^1(m) > EU_j^1(m)$.

If there is no stochastic dominance between i and j then there exists m_1 such that $\int_0^{m_1}(F_i(x)-F_j(x))dx > 0$ and there exists m_2 such that $\int_0^{m_2}(F_j(x)-F_i(x))dx > 0$. Continuity of the integral implies that there exist two intervals $[v_1', v_1'']$ and $[v_2', v_2'']$ such that for all $m \in [v_1', v_1'']$, $\int_0^m (F_i(x)-F_j(x))dx > 0$ and for all $m \in [v_2', v_2'']$, $\int_0^m (F_j(x)-F_i(x))dx > 0$.

Requiring that every $E[V(X)|X^i]$ will have strictly monotonic cumulative distribution insures that in the last period an optimal strategy has a positive probability of getting into any interval. In particular it insures a positive probability of getting into $[v_1', v_1'']$ and $[v_2', v_2'']$ thereby insuring a positive probability of an optimal strategy to observe both attributes i and j. ∎

4. THE INFINITE CASE

As opposed to the finite case, in the infinite case the horizon faced by the decision maker is the same at every stage. I will show that in the infinite case, an optimal strategy will always observe the same attribute. Using the fundamental theorem of discounted dynamic programming, Blackwell 1965, we know that excessivity is a sufficient condition for the optimality of a policy τ. Therefore, if we denote by $U^{\infty}(\tau)(s)$ the payoff function for the decision maker, which represents the expectation of payment for the decision maker in state s using policy τ, it is sufficient to show that, for all s,

$$U^{\infty}(\tau)(s) \geq O(U^{\infty}(\tau))(s)$$

where $O(\cdot)$ is an operator defined as follows:

Let $F \equiv \{f:S \longrightarrow R \mid f \text{ is bounded and measurable}\}$, define $O:F \longrightarrow F$ as,

$$O(f)(s) = \operatorname*{Sup}_{a \in A} \int \left[r(s,a,s') + \beta f(s')\right] dq(s' \mid s,a) \qquad \text{for } f \in F.$$

Verbally, a strategy satisfies the excessivity criterion if, when we delay its exercising by one period and in the meantime do the best thing we can, we do not improve our payment.

Define a stopping time t_i as follows,

$$t_i = \min\,\{n \geq 1 : y_n \geq E[Y_{n+1} \mid y_n, a_{n+1}=i]\}$$

Lemma: *For each attribute $i=1,\ldots,k$, t_i determines a threshold value b_i such that for states s which are greater or equal to b_i, observing attribute i worsens the state in expectation and for states s which are smaller than b_i, observing attribute i improves the state in expectation.*

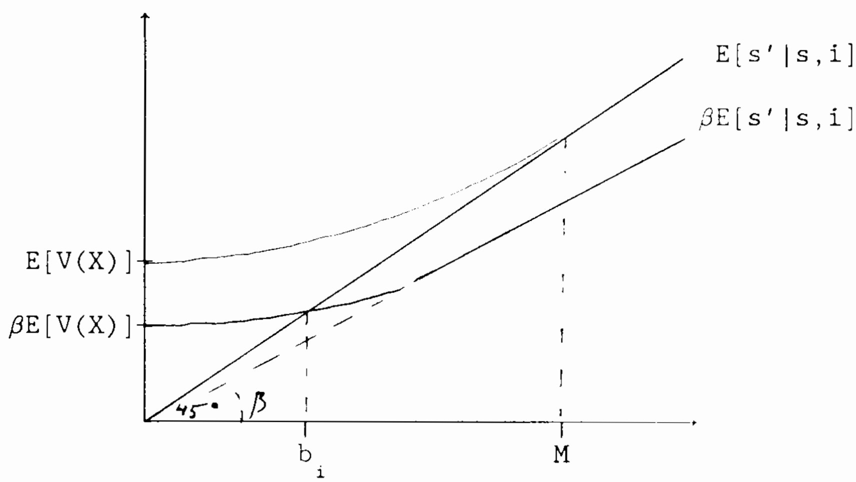
Proof: The existence of b_i can be shown by looking at the graph of the expectation of states in the next period when we are in state s and are about to observe attribute i. I.e., the graph of $E[s'|s,i]$ as a function of s and i. Observe:

(i) Since $E[s'|s,i]$ is equal to $E[\mathrm{Max}\{s,EV(X)\}|\text{observing attribute } i]$ it is continuous and non decreasing with respect to s. Furthermore, $E[s'|s,i]$ is a convex function of s for all $i=1,..,k$. Since $E[s'|s,i] = E[\mathrm{Max}\{s,EV(X)\}|\text{observing attribute } i] = E[\{s,E[V(X)|X^i]\}]$ which was proven to be a convex function of s for all $i=1,..,k$, in the proof of the lemma in theorem 1.

(ii) Non negativity of $V(\cdot)$ implies $E[s'|0,i]=E[E[V(X)|X^i]]=E[V(X)]$.

(iii) Since $V(\cdot)$ is bounded (say by M) $E[s'|s,i]=s$, for all $M\leq s$.

Since $E[Y_{n+1}|y_n,i]\leq y_n$ if and only if $\beta E[s'|s,i]\leq s$ (we use the notation $s=y_n/\beta^n$) the graph shows the value b_i corresponding to the choice of each of the attributes $i=1,...,k$. The monotonicity and continuity guarantee a unique value for b_i[5].



■

We can use the lemma to prove the following observation.

Observation: *This problem, $\mathrm{Max}_\tau\, EY_\tau$, is a monotone problem.*

5 In case there is $i\neq j$ such that $b_h<b_i=b_j$ for all $h\neq i,j$, then we are indifferent between observing characteristic i and characteristic j.

Proof: First we show $A_n \subseteq A_{n+1}$ for $n \geq 0$ by induction on n.

For n=0: $y_0=0$ (the reservation value is 0) then non negativity of $V(\cdot)$ implies $A_0=\phi$.

For general n, consider two cases: The first is that in period n+1 $E[V(X_{n+1})|x^{an+1}] \leq \underset{j=1..n}{Max} \left\{E[V(X_j)|x_j^{aj}]\right\}$ is observed, therefore, $s_{n+1}=s_n$ or $y_{n+1}=\beta y_n$.

$$E[Y_{n+2}|Y_{n+1}=y_{n+1}, a_{n+2}=i]$$

$$= \beta E[Y_{n+1}|Y_n=\beta y_n, a_{n+1}=i] \qquad (s_{n+1}=s_n \Rightarrow \beta Y_{n+1}= Y_{n+2} \text{ and } y_{n+1}=\beta y_n)$$

$$\leq \beta E[Y_{n+1}|Y_n=y_n, a_{n+1}=i] \qquad (\text{because of monotonicity of } E[Y_{n+1}|Y_n=y_n, a_{n+1}=i] \text{ in } y_n)$$

$$= \beta y_n \qquad (\text{by the assumption } s_{n+1}=s_n)$$

$$= y_{n+1} \qquad (\text{by the assumption } y_{n+1}=\beta y_n)$$

Otherwise, $E[V(X_{n+1})|x^{an+1}] > \underset{j=1..t}{Max} \left\{E[V(X_j)|x_j^{aj}]\right\}$ which means $s_{n+1}>s_n$ or $y_{n+1}>\beta y_n$.

Given that $\omega \in A_n$,

$$\beta^n s_n \underset{\text{def.}}{=} y_n \underset{An}{\geq} E[Y_{n+1}|Y_n=y_n, a_{n+1}=i] \qquad (*)$$

$$\underset{\text{def.}}{=} \beta^{n+1}E[s_n 1_{\{Sn+1=sn\}}] + \beta^{n+1}E[S_{n+1} 1_{\{Sn+1>sn\}}]$$

We want to show $\omega \in A_{n+1}$, namely,

$$\beta^{n+1} s_{n+1} \underset{\text{def.}}{=} y_{n+1} \underset{An+1}{\geq} E[Y_{n+2}|Y_{n+1}=y_{n+1}, a_{n+2}=i]$$

$$\underset{\text{def.}}{=} \beta^{n+2}E[s_{n+1} 1_{\{Sn+2=sn+1\}}] + \beta^{n+2}E[S_{n+2} 1_{\{Sn+2>sn+1\}}]$$

Multiplying by $s_n/\beta^{n+1}s_{n+1}$ this is equivalent to showing,

$$s_n \geq \beta E[s_n 1_{\{Sn+2=sn+1\}}] + \beta E[S_{n+2} 1_{\{Sn+2>sn+1\}}] s_n/s_{n+1}$$

which holds since we know,

$$\beta E[s_n 1_{\{Sn+2=sn+1\}}] + \beta E[S_{n+2} 1_{\{Sn+2>sn+1\}}] s_n/s_{n+1}$$

$$= \beta E[s_n 1_{\{S_{n+1} \le s_{n+1}\}}] + \beta E[S_{n+1} 1_{\{S_{n+1} > s_{n+1}\}}] s_n / s_{n+1}$$

Because knowing the state does not imply knowing the time and assume we're in time n with $s_n = s_{n+1}$. This in turn is smaller or equal than,

$$\le \beta E[s_n 1_{\{S_{n+1} = s_n\}}] + \beta E[S_{n+1} 1_{\{S_{n+1} > s_{n+1}\}}] s_n / s_{n+1}$$

Since, observe that $\Pr\{S_{n+1} \le s_{n+1}\} \ge \Pr\{S_{n+1} = s_n\}$ and $\Pr\{S_{n+1} > s_{n+1}\} \le \Pr\{S_{n+1} > s_n\}$ because $s_n < s_{n+1}$. So now the average is taken with a larger weight on the higher value. And this is smaller or equal to,

$$\le s_n \qquad \{\text{ by } (*)\ \}$$

Second, we show $\bigcup_{n=1}^{\infty} A_n = \Omega$.

Since $A_n \subseteq A_{n+1}$ we have,

$$\Pr\{\bigcup_{n=1}^{\infty} A_n\}^c = \Pr\{\bigcap_{n=1}^{\infty} A_n^c\} = \lim_{n\to\infty} \Pr\{A_n^c\}$$

Using $A_n \subseteq A_{n+1}$ again, we get, $\Pr\{A_n^c\} = \Pr\{A_n^c | A_{n-1}^c\} \cdot \Pr\{A_{n-1}^c\}$

$$= \Pr\{A_1^c\} \cdot \prod_{k=2}^{n} \Pr\{A_k^c | A_{k-1}^c\}$$

Using the lemma we know that for values greater or equal than b another observation worsens the state in expectation. Therefore there exists a $0<q<1$, $q \equiv \Pr\{E[V(X)|X^j] \ge b\}$ such that $\Pr\{A_k^c | A_{k-1}^c\} \le 1-q$ for all $k>1$, and,

$$\Pr\{A_n^c\} = \Pr\{A_1^c\} \cdot \prod_{k=2}^{n} \Pr\{A_k^c | A_{k-1}^c\}$$

$$\le \Pr\{A_1^c\} \cdot (1-q)^n \xrightarrow[n\to\infty]{} 0 \quad \blacksquare$$

We now turn to define τ^*, define t^* as the first time $y_n \ge E[Y_{n+1} | y_n, a_{n+1} = i]$ for all i, or $t^* = \mathrm{Max}_{i=1..k} \{t_i\}$. Denote $b = \mathrm{Max}_{i=1..k} \{b_i\}$ and $j = \mathrm{argMax}_{i=1..k} \{b_i\}$. Therefore, the stopping rule τ^* is defined as follows:

$$\tau^*(s) = \left\{ \begin{array}{ll} \text{"stop"} & b \le s \\ \\ \text{observe characteristic } j & s < b \end{array} \right\}$$

Theorem 2: *the stopping rule τ^* is optimal.*

Proof: First, with the previous lemma, observe that τ^* is well defined. Second, we demonstrate the excessivity property. Before that we need to find $EY_{\tau^*}(s)$ (which is actually $U^\infty(\tau^*)(s)$, the expectation of future payment when we are in state s and acting according to policy τ^*).

Claim: *$EY_{\tau^*}(s)$ is as follows:*

$$EY_{\tau^*}(s) = \begin{cases} s & b \le s \\ b & s < b \end{cases}$$

Proof: For $b \le s$ it is obvious. (since the process stops).
For $s<b$ the expected payment in state s, according to policy τ^* is $E[\beta^t] \cdot E\left[E[V(X)|X^j] \,|\, b \le E[V(X)|X^j]\right]$ where t is distributed geometrically with probability of success denoted by p. Computing $E[\beta^t]$:

$$E[\beta^t] = \sum_{n=1}^{\infty} \beta^n (1-p)^{n-1} p$$

$$= p\beta \sum_{n=1}^{\infty} \beta^{n-1} (1-p)^{n-1} = \frac{p\beta}{1-\beta(1-p)}$$

Using the lemma, b is such that for states s which are greater or equal to b, observing attribute j worsens the state in expectation and for states s which are smaller than b, observing attribute j improves the state in expectation. Therefore b is the solution to the equation:

$$b = \beta\left[(1-p)b + pE\left[E[V(X)|X^j] \,|\, b \le E[V(X)|X^j]\right]\right]$$

Rearranging terms yields:

$$\frac{p\beta}{1-\beta(1-p)} \cdot E\left[E[V(X)|X^j] \,|\, b \le E[V(X)|X^j]\right] = b$$

Which means,

$$E[\beta^t] \cdot E\left[E[V(X)|X^j] \,|\, b \le E[V(X)|X^j]\right] = b$$

This ends the proof of the claim. ■

Now we demonstrate excessivity: for $b \leq s$ the policy stops. If instead, it continues by observing any attribute i and then continuing according to τ^*, it will give a lower expected value since we are in the region $b \leq s$ where any additional observation worsens the state in expectation.

For $s<b$ the policy continues by observing attribute j, if instead, it continues by observing attribute i distinguish between two cases: If it stops immediately after that, which means it is in a state $s'>b$, it could not have given an expected utility value of more than b_i, which is smaller than b - the expected utility value if it does not deviate from τ^*. If it does not stop, then $s'<b$ and continuing according to τ^* will give no more than b, which is not an improvement. If, instead of observing a different attribute, it stops, it gives a utility value of s which is smaller than b. ■

The most important conclusion from the optimality of the stopping rule τ^* is that for every decision maker (who is characterized by a utility function $V(\cdot)$ and discounting factor β) there is only one interesting attribute, not depending on states. The value b is monotonically increasing in β. It is possible that different β values will lead to different chosen attributes in the strategy τ^*. Thus, the interesting attributes for two decision makers with the same utility function but different discounting factor may not be the same.

Notice that the stopping rule τ^* has two simple properties. First, it is stationary. That is to say, in each state the decision is not dependent upon time. Second, the stopping rule is "almost" myopic. It is myopic in the sense that the decision whether to stop or not depends on observing the situation one step ahead only. However it is not entirely myopic: once it decides to continue, the stopping rule might pick up an attribute which does not give the highest one-step expected payment. It is therefore possible that in some states, if the process would have ended in the next period, an optimal strategy would have picked a different attribute than τ^*, as demonstrated in example 1.

In order to better understand the relation between the finite and infinite cases, notice that $U^n(v)$, the expected payoff n periods from the end with a reservation value v, is an increasing sequence bounded by $U^\infty(\tau^*)(v)$. In fact,

Lemma: $U^n(v) \xrightarrow[n\to\infty]{} U^\infty(\tau^*)(v)$, *moreover the convergence is uniform.*

Proof: Since we assumed that $V(\cdot)$ is bounded (say by M), we have,

$$U^\infty(\tau^*)(v) - U^n(v) \le \beta^n M \xrightarrow[n\to\infty]{} 0 \quad \blacksquare$$

The following theorem demonstrates a stronger form of similarity, the similarity in the actions taken by optimal finite and infinite case strategies.

Theorem 3: *There exists n′ such that if the decision maker is faced by any horizon, n≥n′, then an optimal strategy will observe the infinite case dominant attribute for at least n-n′ periods.*

Proof: $U^n(v) \xrightarrow[n\to\infty]{} U^\infty(\tau^*)(v)$, implies that for all $\epsilon>0$ there exists an n′ such that for all $n\ge n'$, $U^n(v) > U^\infty(\tau^*)(v)-\epsilon = b-\epsilon$.

Suppose, without loss of generality, that p is the infinite case optimal attribute. Denote $P\equiv E[V(X)|X^p]$ and $Q\equiv E[V(X)|X^q]$ for attributes p and q. Observing attribute p gives,

$$\beta\left\{ E(P\cdot 1_{\{P\geq b\}}) + E(U^{n-1}(\mathrm{Max}\{v,P\})\cdot 1_{\{P<b\}}) \right\} > b-\epsilon$$

Observing any other attribute q gives,

$$\beta\left\{ E(Q\cdot 1_{\{Q\geq b\}}) + E(U^{n-1}(\mathrm{Max}\{v,Q\})\cdot 1_{\{Q<b\}}) \right\}$$

Observe that,

$$\beta\left\{ E(Q\cdot 1_{\{Q\geq b\}}) + E(U^{n-1}(\mathrm{Max}\{v,Q\})\cdot 1_{\{Q<b\}}) \right\}$$

$$\leq \beta\left\{ E(Q\cdot 1_{\{Q\geq b\}}) + E(b\cdot 1_{\{Q<b\}}) \right\}$$

Since for states at which the strategy does not stop , an optimal infinite strategy yields only an expected payoff of b.

This, in turn, is smaller than b, because for $b > b_q$,

$$b > \beta\left[(1-p)b + pE\left[E[V(X)|X^q]\,|\,b\leq E[V(X)|X^q]\right]\right]$$

by the definition of b_q and the lemma. For b_q it holds as an equality, for $b>b_q$ taking another observation worsens the state in expectation. Therefore, choosing n' large enough guarantees that observing attribute p will give higher expected payoff. ■

References:

Blackwell, D. (1965) Discounted dynamic programming. *The Annals of Mathematical statistics*, 36, 226-235.

Berry, D.A and Fristedt, B. (1985) Bandit Problems - Sequential allocation of experiments. Chapman and Hall.

Chow, Y.S., Robbins, H. and Ziegmund, D. (1971) Great expectations: The theory of optimal stopping. Houghton Mifflin.

Coombs, C.H. (1964) A theory of data. New York: Wiley.

Fishburn, P.C. (1968) Utility theory. *Management Science*, 13, 435-453.

Lancaster, K.J. (1966) A new approach to consumer theory. *Journal of Political Economy*, 74, 132-157.

McMillan, J. and Rothschild, M. (1989) Search. Forthcoming in the Handbook of Game Theory with Economic Applications, edited by R.J. Aumann and S. Hart.

Rothschild, M. and Stiglitz, J.E. (1970) Increasing risk: 1. A definition. *Journal of Economic Theory*, 2, 225-243.

Tversky, A. (1972) Elimination by aspects: A theory of choice. *Psychological Review*, 79, 281-299.

Appendix:

First, we present a detailed analysis of the examples,

Example 1: Consider a two period problem, N=2. The decision maker is faced by two attribute p,q, products of binary distributions,

x^p	0	1
	3/4	1/4

x^q	0	1
	1/3	2/3

The utility function is $V(x^p,x^q)=x^p+x^q$. Full analysis of the problem is carried out by reviewing the problem in backward induction.

after the last period, t=2:

$x^p=1$: $E[V(x)|x^p=1]=\beta^2 5/3$

$x^q=1$: $E[V(x)|x^q=1]=\beta^2 5/4$

$x^p=0$: $E[V(x)|x^p=0]=\beta^2 2/3$

$x^q=0$: $E[V(x)|x^q=0]=\beta^2/4$

In the second (last) period, t=1:

	observing x^p	observing x^q
$x^p=1$:	$(5\beta/3)$ stop	$(5\beta/3)$ stop
$x^q=1$:	$\beta^2\{3/4(5/4)+1/4(5/3)\}=65\beta^2/48$	$(5\beta/4)$ stop
$x^p=0$:	$\beta^2\{3/4(2/3)+1/4(5/3)\}=11\beta^2/12$	$\beta^2\{1/3(2/3)+2/3(5/4)\}=19\beta^2/18$
$x^q=0$:	$\beta^2\{3/4(2/3)+1/4(5/3)\}=11\beta^2/12$	$\beta^2\{1/3(1/4)+2/3(5/4)\}=11\beta^2/12$

In the first period, t=0:

observing x^p	observing x^q
$\{3/4(19\beta^2/18)+1/4(5\beta/3)\}=\beta(30+57\beta)/48$	$\{1/3(11\beta^2/12)+2/3(65\beta^2/48)\}=87\beta^2/72$

Example 2: Again we present a two period problem, N=2. The decision maker is faced by two attribute p,q, products of binary distributions,

x^p	0	3
	1/2	1/2

x^q	1	2
	1/2	1/2

The utility function is $V(x^p,x^q)=x^p+x^q$. Again, analysis of the problem will be carried out by reviewing the problem in backward induction, in the same manner as in example 1.

after the last period, t=2:

$x^p=3$: $E[V(x)|x^p=3]=\beta^2 9/2$

$x^q=2$: $E[V(x)|x^q=2]=\beta^2 7/2$

$x^q=1$: $E[V(x)|x^q=1]=\beta^2 5/2$

$x^p=0$: $E[V(x)|x^p=0]=\beta^2 3/2$

In the second (last) period, t=1:

	observing x^p	observing x^q
$x^p=3$:	$(9\beta/2)$ stop	$(9\beta/2)$ stop
$x^q=2$:	$\beta^2\{1/2(7/2)+1/2(9/2)\}=4\beta^2$	$(7\beta/2)$ stop
$x^q=1$:	$\beta^2\{1/2(5/2)+1/2(9/2)\}=7\beta^2/2$	$\beta^2\{1/2(5/2)+1/2(7/2)\}=3\beta^2$
$x^p=0$:	$\beta^2\{1/2(3/2)+1/2(9/2)\}=3\beta^2$	$\beta^2\{1/2(5/2)+1/2(7/2)\}=3\beta^2$

In the first period, t=0:

observing x^p	observing x^q
$\{1/2(3\beta^2)+1/2(9\beta/2)\}=\beta(9+6\beta)/4$	$\{1/2(7\beta^2/2)+1/2(4\beta^2)\}=15\beta^2/4$

Next, we solve for the optimal strategy for the general finite problem. Some notation first:

$e_N = y_N(a_1,\ldots,a_N)$ = The value of the target function after observing the last product, N (y_N is a random variable as defined before, i.e. a function of the decisions $a_1,\ldots,a_N$)

$f_{N-1}^{(i)} = E[e_N|H_1^{N-1},a_N=i]$ = The expectation of the target function value in stage N, conditioned on the history until stage

$N-1$, H_i^{N-1} (a sufficient statistic for H_1^{N-1} is $y_{N-1}(a_1,\ldots,a_{N-1})$) and when we choose to observe attribute i.

$f_{N-1} = \underset{i=1..k}{\mathrm{Max}} \{f_{N-1}^{(i)}\}$ = The value of the best observation in stage $N-1$.

$e_{N-1} = \mathrm{Max}\{f_{N-1}, y_{N-1}(a_1,\ldots,a_{N-1})\}$ = The value of the target function in stage $N-1$.

Similarly, this notation is extended to all periods ($0 \le t \le N$).

Proposition: *the optimal rule* τ^* *is defined as follows:* $\tau^* = Min\{0 \le n < N \mid f_n \le y_n\}$ *when for* $t<n$ *we choose to observe the attribute on which* f_t *is obtained. which means, for any other stopping rule* τ, $Ey_\tau \le Ey_{\tau^*} = e_0$[6].

Proof: for τ any stopping rule define:

$$\alpha(N)=E[y_\tau]$$

$$k<N \qquad \alpha(k)=E[y_\tau I_{\{\tau<k\}}]+E[e_k I_{\{k\le\tau\}}]$$

$$\alpha(0)=E[e_0]=e_0$$

Suppose $y_0=0$, if the decision maker prefers not to enter the process his utility is 0. Hence $e_0=\mathrm{Max}\{y_0,f_0\}$ is a constant equal to $E[e_0]$.

I will show: (i) $\alpha(k)$ is a decreasing sequence in k (e_0 is a bound for payment)

(ii) For $\tau=\tau^*$, $\alpha(k)$ is a constant sequence (τ^* is optimal since it gives e_0)

Observe that,

$$\alpha(k-1) = E[y_\tau I_{\{\tau<k-1\}}] + E[e_{k-1} I_{\{k-1\le\tau\}}]$$

$$\alpha(k) = E[y_\tau I_{\{\tau<k-1\}}] + E[y_\tau I_{\{\tau=k-1\}}] + E[e_k I_{\{k\le\tau\}}]$$

6 Notice that e_0 is a constant because we will define y_0 as a reservation value constant and $f_0 = \underset{i=1..k}{\mathrm{Max}} \{E[e_1|y_0,a_1=i]\}$ is also a constant.

Consider: $$E[e_k I_{\{k\le\tau\}}] = E[E[e_k I_{\{k\le\tau\}}|H_0^{k-1}]]$$
$$= E[I_{\{k\le\tau\}} E[e_k|H_0^{k-1}]]$$
$$\le E[I_{\{k\le\tau\}} f_{k-1}]$$
$$\le E[I_{\{k\le\tau\}} e_{k-1}]$$

And also: $$E[y_{k-1} I_{\{\tau=k-1\}}]$$
$$\le E[e_{k-1} I_{\{\tau=k-1\}}]$$

So we have: $$E[e_k I_{\{k\le\tau\}}] + E[y_{k-1} I_{\{\tau=k-1\}}] \le E[I_{\{k\le\tau\}} e_{k-1}] + E[e_{k-1} I_{\{\tau=k-1\}}]$$
$$= E[e_{k-1} I_{\{k-1\le\tau\}}]$$

which proves (i).

(ii) To show that for τ^* $\alpha(k)$ is a constant sequence, I will show that for τ^* the inequalities in (i) are equalities.

From the definition of τ^* we have that τ is the first time the maximum between f and y is obtained on y. Therefore under τ^*, $k\le\tau^*$ implies $e_{k-1}=f_{k-1}$.

$$E[I_{\{k\le\tau^*\}} E[e_k|H_0^{k-1}]] = E[I_{\{k\le\tau^*\}} f_{k-1}]$$
$$= E[I_{\{k\le\tau^*\}} e_{k-1}]$$

And from the same reason:

$$E[y_{k-1} I_{\{\tau^*=k-1\}}] = E[e_{k-1} I_{\{\tau^*=k-1\}}]$$

Because $\tau^*=k-1$ implies $e_{k-1}=y_{k-1}$ which completes the proof of (ii). ■