ELSEVIER Mathematical Social Sciences 29 (1995) 195–212

mathematical social sciences

On determining the importance of attributes with a stopping problem

Zvika Neeman

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

Received September 1993
Revised March 1994

Abstract

In this paper we 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-attribute products. From each product, he can observe only one attribute. At each period the agent has to decide whether he wants to stop and take the best product he has observed so far, or whether he prefers to continue the observation process and observe an attribute of the next product in the sequence. We find the optimal observation policy and the conditions under which it observes only one attribute, rendering it the most 'informative'. When the sequence of products is finite, second-order stochastic dominance characterizes the case in which an optimal strategy observes only one attribute in the sense that if it holds between any two random variables induced by the expected utility given an attribute, it is never optimal to observe the 'dominating' attribute. When the sequence of products is infinite, observing one attribute only is always optimal. The seeming discrepancy between finite and infinite horizon models vanishes for a sufficiently large horizon, making the infinite horizon optimal attribute the one chosen for a long period in finite horizon problems as well.

Keywords: Search; Product attributes; Consumer preferences; Stopping problem; Learning

1. Introduction and summary

One approach in consumer theory, first presented by Lancaster (1966), views products as bundles of attributes. Consumers' preferences are accordingly defined over the set of different combinations of attributes. Consumers' preferences over products, rather than assumed as a primitive of the theory, are determined by the different composition and magnitude of products' attributes.

SSDI 0165-4896(94)00764-Y

In this paper we try to develop a systematic way of answering the question: What makes a certain attribute important? In order to address this question we introduce the following decision problem: a rational decision-maker is facing a sequence of multi-attribute products. He can choose one product, preferably the best, from this sequence. However, he is restricted by the following information constraint, namely, he can observe only one attribute from each product in the sequence. Now, since observing different attributes may yield different information, the decision-maker has to decide in what kind of information he is most interested. As we shall see, it may well be the case that the optimal attribute to observe will vary across time and history. In the sequel, we shall identify necessary and sufficient conditions under which there will be only one such attribute.

An example illustrating this decision problem is the following. Suppose that one is interested in buying a used car. She visits several car dealers which in turn make her an offer. However, for various reasons (e.g. the dealer does not have time, there are more customers waiting to examine the car, these are the prevailing social norms), she can have only a limited amount of time in which she can test each car. After visiting each car dealership, she has to decide whether she wants to buy the suggested car or whether she is interested in observing yet another car. At any stage, she can come back and buy a car that she has already observed before. Her problem, then, is to decide which attributes of the car should she be most interested in examining. Is there one attribute that provides the best information about cars or should she examine a different attribute at each time, maybe as a function of the outcome of previous observations.

Another example is the following. 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 and interview new candidates. We can think of the candidates as having two attributes, their research ability and their teaching skills. Before reviewing each candidate, the department has to decide which attribute of the two it wishes to examine. In this example, the information restriction imposed by the model seems more natural. It is customary that the department flies the candidate and hosts him for a couple of days during which only limited information can be gathered on the candidate. Furthermore, it is not customary to interview a candidate twice. Therefore, the department's decision can only be based upon partial information. The question is what partial information is more useful.

To put it in the context of a stopping problem, we try 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, namely, what might appear at a particular moment as an interesting attribute to observe may indeed be an optimal one. In other words, when people are confronted with problems of this kind what may seem at first as indeterminate and inconsistent behavior may actually be the result of unconscious optimizing rather than sheer arbitrariness. Furthermore, this approach can be justified on an evolutionary basis. More generally, the notion of a decision-maker who unconsciously chooses optimally seems more plausible to us than that of a (boundedly) rational decision-maker who is fully aware of the model and calculates an optimal strategy.

Similar models of decision processes were considered in psychological literature. These models (see, for example, Coombs, 1964; Fishburn, 1968; and 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 which is either fixed beforehand or determined by a lottery. All the products (alternatives) that do not rank highly enough when this attribute is used as the criterion for ranking are disqualified. This process is repeated until only one product (alternative) is left, and this last product ends up being chosen. Yet, whether the particular way of choosing the attributes is deterministic or probabilistic, the way in which they are chosen is left implicit in these papers.

As mentioned above, bounded perception is reflected in our model in the assumption that from each product only one attribute may be observed. 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 virtually impossible (for instance, in the problem of interviewing job applicants). In addition, the model allows to lump several attributes together into one attribute. For example, 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. Some natural mechanisms force a limitation very similar to the one we have described. The human eye, for instance, can observe only a limited range in a certain time, although one can choose where one wants to look.

We characterize the optimal observation policy in the case when the sequence of products is finite (the finite horizon problem) and in the case where the sequence of products is infinite (the infinite horizon problem). In particular, we identify necessary and sufficient conditions under which only one attribute is observed by the optimal strategy, rendering it the most 'informative' and simplifying the optimal strategy. We present 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 example, one only one attribute is observed by the optimal strategy.

Example 1. Suppose that the decision-maker is facing two products, that is, it is a two-period problem. Each product has two attributes p and q. Each attribute may take the values 0 or 1 according to the following distributions which are assumed to be independent.

$$X^p = \begin{cases} 0\,, & \text{with probability } \frac{3}{4}\,, \\ 1\,, & \text{with probability } \frac{1}{4}\,, \end{cases}$$
$$X^q = \begin{cases} 0\,, & \text{with probability } \frac{1}{4}\,, \\ 1\,, & \text{with probability } \frac{3}{4}\,. \end{cases}$$

The decision-maker's utility function is $U(X^p, X^q) = X^p + X^q$ and it is discounted by $\beta < 1$. Thus, observing attribute p yields an expressed utility value of $\mathrm{E}[U(X^p, X^q) \mid X^p = x^p] = \frac{3}{4} + x^p$ and observing attribute q yields an expected utility of $\mathrm{E}[U(X^p, X^q) \mid X^q = x^q] = \frac{1}{4} + x^q$. The analysis of the problem is carried out by backward induction. Computation shows that the optimal strategy (for $\beta \geq \frac{2}{3}$) is to observe attribute p first. In the case $x^p = 1$, the optimal strategy is to stop and take the product. If, however, $x^p = 0$, the optimal strategy is to continue and observe attribute q of the second product. If $x^q = 1$, the optimal strategy is to take this product. However, in the case $x^q = 0$, the optimal strategy is to go back and take the previous product (the one with $x^p = 0$). As we shall see later, in the infinite horizon problem, for $\beta \geq \frac{16}{19}$ the optimal strategy always observes attribute p, regardless of previous realizations.

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 optimal strategy is to bear the risk and observe attribute p. If the high value is observed, the optimal strategy is to stop. If the low value is observed, the optimal strategy is to move over and to observe attribute q in the hope that it, at least, will guarantee a better than average conditional expected payoff. If it does, the optimal strategy takes this product. Otherwise, there is no choice but to take the first product. Since X^p is more likely to be zero then X^q, the q attribute functions 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. On the other hand, when the sequence of products is infinite the decision-maker can always choose to observe the more risky attribute p. In the case of 'failure', he can always continue and make another observation.

Example 2. Again, we present a two period problem. As before, each product has two attributes p and q with the following independent distributions,

$$X^p = \begin{cases} 0, & \text{with probability } \frac{1}{2}, \\ 3, & \text{with probability } \frac{1}{2}, \end{cases}$$

$$X^q = \begin{cases} 1, & \text{with probability } \frac{1}{2}, \\ 2, & \text{with probability } \frac{1}{2}. \end{cases}$$

As before, the utility function is $U(X^p, X^q) = X^p + X^q$ and is discounted by $\beta < 1$. Now, observing attribute $i \in \{p, q\}$ yields an expected utility value of $\mathrm{E}[U(X^p, X^q) \mid X^i = x^i] = \frac{3}{2} + x^i$. Again, analysis of the problem is carried out by backward induction. The optimal strategy (for the case where $\beta \geq \frac{1}{2}$) is as follows. First, observe attribute p. In the case when $x^p = 3$ is observed, stop and take the product. In the case when $x^p = 0$ is observed, the optimal strategy is to make another observation. It does not matter whether attribute p or q is the one observed, either one will give the same expected utility value. In this example an optimal strategy can be restricted to observing only attribute p. In this sense, attribute p is more interesting than attribute q.

Thus, we see a major difference between the two examples. In Example 1 an optimal strategy observes both attributes, while in Example 2 an optimal strategy can be restricted to observe only one attribute. In Section 3 we generalize this distinction and characterize it in terms of the relationship between the distributions of the attributes.

We show that in the finite case, for any time horizon N, a necessary and sufficient condition for an optimal strategy to observe only one attribute is that the random variable induced by the expected utility given this attribute is second-order stochastically dominated by the random variables 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, regardless of past realizations. We prove this result and characterize the infinite horizon optimal attribute in Section 4. Consequently, when the finite and infinite horizon optimal strategies differ, the latter fails 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. In Section 4 we show that the apparent discrepancy between the finite and infinite horizon optimal strategies 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 just 'being informative' in the sense of being stochastically dominated. Attractive attributes are, in the long run, those that have the possibility of pulling the whole product assessment sharply upwards, no matter how small is the probability of that occurring at any given stage.

The problem addressed in this paper relates to the search literature and to the secretary and multi-armed bandit problems. The similarity between the problem studied in this paper and an ordinary search problem (for a survey of search literature, see McMillan and Rothschild, 1992) is due to the fact that a search for the optimal product is taking place and once a good enough product is found, the search stops. The difference lies in the motivation for the search. In this paper we 'search' for the attribute that will convey the best information so that we can conduct our (real) search for the best product optimally. As a result, there is no 'optimal attribute' that we search for; rather, the most informative attribute may vary along time and history.

The main differences between the problem presented in this paper and the secretary problem (see, for instance, Chow et al. , 1971) is that here we impose an information constraint on the decision-maker. He can observe only one attribute of each product and, indeed, his problem is to decide which one. By contrast, in the secretary problem this problem does not arise since no information restriction is imposed. Another important difference is that the search strategies in the secretary problem are restricted to be without recall. Namely, one cannot call back a secretary that has been turned down.

Lastly, in the multi-armed bandit problem (see, for instance, Berry and Fristedt, 1985), generally, several distributions whose parameters are unknown are given. The problem is to maximize the sequence of payoffs obtained from these distributions. The problem, then, is to identify the 'best' distribution through experimentation. The difference between a multi-armed bandit problem and the problem presented here is that here there is no need for experimentation because all the parameters are known a priori. In this respect, the problem presented here is a degenerate bandit problem.

The general structure of the paper is as follows. In Section 2 we present the model. In Section 3 we deal with the finite horizon case. We present the optimal strategy and characterize the case in which it can be restricted to observe only one attribute in terms of second-order stochastic dominance. Finally, in Section 4 we handle the infinite horizon case and demonstrate that observing only one attribute is optimal. We explain the relationship between the finite and infinite horizon cases by showing the continuity of the finite horizon optimal strategy at infinity.

2. The model

Let $X_1, X_2, \ldots$ denote a sequence of i.i.d. random variables, $X_n : \Omega \to \mathbb{R}^k$ for $n \geq 1$, which represent multi-attribute products. Each product in the sequence is represented by a vector of k random variables $X_n = (X_n^1, \ldots, X_n^k)$, $X_n^i : \Omega \to \mathbb{R}$, which are interpreted as the products' attributes. The preferences of the decision-maker over the products are given by a non-negative, bounded and continuous von Neumann–Morgenstern utility function, $U : \mathbb{R}^k \to [0, M]$, which is discounted by $0 < \beta < 1$.

The information restriction imposed on the decision-maker is that he can observe only one attribute of each product he is facing. We denote the attribute of the nth product that the decision-maker observes by $a_n \in \{1, \ldots, k\}$. The objective of the decision-maker is to choose the product that maximizes his preferences subject to his information restriction. Thus, the decision-maker wishes to select, if possible, the product that maximizes $\mathrm{E}[U(X^1, \ldots, X^k) \mid X^a = x^a]$, where a denotes the ath attribute which the decision-maker is free to choose himself. Obviously, without the information restriction, the objective of the decision-maker would have been to choose the product that maximizes $U(X^1, \ldots, X^k)$ rather than the one that maximizes $\mathrm{E}[U(X^1, \ldots, X^k) \mid X^a = x^a]$. Thus, the k random variables X^i, $i \in \{1, \ldots, k\}$, yield k random variables $\mathrm{E}[U(X) \mid X^i] : \Omega \to [0, M]$ whose realizations are the actual objects among which the decision-maker chooses. For each $i \in \{1, \ldots, k\}$, we denote $Z^i \equiv \mathrm{E}[U(X) \mid X^i]$. In words, Z^i denotes the expected utility that the decision-maker derives from a product whose ith attribute he has observed. Since the value of the ith attribute is a random variable, so is Z^i. Let $G_i(\cdot)$, $i \in \{1, \ldots, k\}$, denote the cumulative distribution function of Z^i and let z^i denote a realization of Z^i. Note that while the X^i's did not relate to each other in any special way, for any $i, j \in \{1, \ldots, k\}$, $\mathrm{E}[Z^i] = \mathrm{E}[Z^j]$. This observation is an immediate consequence of the fact that for any two random variables A and B, $\mathrm{E}[\mathrm{E}[A \mid B]] = \mathrm{E}[A]$.

The decision-maker is going through the following observation procedure. He goes through the sequence of products $X_1, X_2, \ldots$. From each product X_n he observes one attribute a_n, until he decides to stop. When he decides to stop, he takes the best product he has seen so far in the sequence (subject to his information restriction). We denote the decision-maker's utility from the best product seen up to stage n by Y_n. Notice that Y_n is actually the reservation value of the decision-maker after observing the nth product in the sequence. Namely, the decision-maker can guarantee to himself a utility value of Y_n by stopping the observation procedure and taking the best product so far. Thus, the observation procedure yields a sequence of random variables $Y_0, Y_1, Y_2, \ldots$, where $Y_0 \equiv 0$, that is, the decision-maker has the choice of not observing any product whatsoever and getting a utility value 0, and where for $n \geq 1$, $Y_n = \beta^n \max_{t \in \{1,\dots,n\}} \{Z_t^{a_t}\}$.[1]

The introduction of the Y_n's allow us to describe this problem as a stopping problem. At any time n, the decision-maker has to decide whether to stop the observation procedure and get a utility value of y_n, which denotes the realization of Y_n, or to continue and observe the realization of the a_{n+1} attribute of the $n+1$th product.

In order to describe this problem as a dynamic programming problem, we need to specify a states space, an actions space, a transition function and a payment function. We denote the states space by $S \equiv \mathbb{R}_+ \cup \{\emptyset\}$ with the following interpretation. In period n, or after n products have been observed, a state $s_n \in S \backslash \{\emptyset\}$ stands for $s_n = \max_{t \in \{1,\dots,n\}} \{z_t^{a_t}\}$, or $s_n = y_n / \beta^n$. That is, it represents the utility value of the best among the n products that have been observed so far net of the discounting. $\emptyset$ is an absorbing state, denoting the end of the observation process. Sometimes, we shall denote the state succeeding s by s'. Notice that knowing the state does not imply knowing the history of the process nor the number of periods that have passed since the observation procedure began. However, it contains all the relevant information for designing an optimal strategy.

The actions space is denoted by $A \equiv \{1, \dots, k, \text{Stop}\}$. At each state $s \in S$, the decision-maker chooses an action which may either be to observe an attribute $a \in \{1, \dots, k\}$, or to Stop, that is, to terminate the observation procedure, and take the best product observed so far.

We denote the transition function by $q : S \times A \to F(S)$. The transition function, given a state and an action, describes the distribution of states which follows. $F(S)$ denotes the family of cumulative distribution functions over the states space S. $q(s, a) = \emptyset$ with probability 1 if $s = \emptyset$ or $a = \text{Stop}$. Otherwise, that is, when $s \neq \emptyset$ and $a \in \{1, \dots, k\}$,

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

Notice that as long as the observation process continues $s' \geq s$, that is, the state is always (weakly) 'improving' after each observation regardless of which attribute is the one observed.

We denote the immediate payment function by $r : S \times A \times S \to \mathbb{R}$. The payment function describes the immediate payoff to the decision-maker given that the observation process is in state s, action a is chosen, and the resulting state is s':

[1] Notice that this formulation allows us to describe a finite sequence of products as well. When the product sequence is of length N, $Z_t^{a_t} \equiv 0$ for all $t > N$.

$$r(s, a, s') = \begin{cases} s\,, & s \neq \emptyset, a = \text{Stop}, s' = \emptyset\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

The objective function of the decision-maker is then

$$\max_{a_1, a_2, \ldots} \mathrm{E}\left[\sum_{n=0}^{\infty} \beta^n r(s_n, a_{n+1}, s_{n+1})\right], \text{where } s_0 = 0\,.$$

We are interested in finding an optimal strategy (policy, or rule, interchangeably) for this decision 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 the random variables denoting the attributes observed until the stopping time t.
- At any time n, the decisions (namely, whether t stops and if not a_{n+1}) must be a function of what is known at that time.

We denote the value of the objective function of the decision-maker when he uses τ as his stopping rule by Y_τ. Sometimes it will be more convenient to use a different notation, namely we denote the expected value of the objective function of the decision-maker when he faces a sequence of N products, uses τ as his stopping rule, is currently in state s, and have already observed n products, by $V_{n,N}(\tau)(s)$. When the sequence of products is infinite, we write $V_{n,\infty}(\tau)(s)$. The decision-maker's problem then is to find a stopping rule τ that will maximize $\mathrm{E}[Y_\tau]$ or $V_{0,N}(\tau)(0)$. In the following two sections we describe the optimal observation policy for the cases where the decision-maker faces a finite sequence of products and for the case where the sequence of products is infinite.

3. The finite horizon problem

When the sequence of products is finite, the number of products that have not been observed yet is of crucial importance in determining the optimal strategy. The analysis is carried out by backward induction. First, the last period, or the period in which the last product is observed, 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 the results of the actions taken are known in terms of the distribution over states in the following period.

Suppose that the decision-maker is facing a finite sequence of products X_1,

$X_2, \ldots, X_N$. Before we compute the optimal strategy, we need to establish some more notation. Let

$$V_N \equiv Y_N = \beta^N s_N$$

be the random variable denoting the reservation value after observing the last product N.

For $0 \leqslant n \leqslant N-1$, let

$$U^i_{n+1}(s_n, Z^i) \equiv \beta^{n+1} \cdot \max\{s_n, Z^i\}$$

be the random variable denoting the value of the objective function after n products have already been observed, the current state is s_n, and the ith attribute of the $n+1$th product is observed.

$$U_{n+1}(s_n) \equiv \max_{i \in \{1,\ldots,k\}} \{\mathrm{E}[U^i_{n+1}(s_n, Z^i)]\}$$

denote the expected value of the objective function in the $n+1$th period when the decision maker observes the best attribute and when the current state is s_n. Finally, let

$$V_n \equiv \max\{Y_n, U_{n+1}(s_n)\}$$

denote the expected value of the objective function in the nth stage.

Notice that for all $0 \leqslant n \leqslant N-1$,

$$U_{n+1}(s_n) = \mathrm{E}[V_{n+1} \mid s_n] = \mathrm{E}[V_{n+1} \mid Y_n] ,$$

and

$$V_n \equiv \max\{Y_n, U_{n+1}(s_n)\} = \max\{\beta^n s_n, U_{n+1}(s_n)\} .$$

The sequence of V_n's incorporates an optimal stopping rule. For each n, V_n expresses the expected value of the objective function when the optimal attributes $a_1, \ldots, a_n$ have been observed and the decision whether to stop or not was made optimally. The following proposition formalizes this claim.

Proposition 1. *The optimal observation policy* $\tau^*_N = (t^*; a^*_1, \ldots, a^*_{t^*})$ *is defined as follows.* $t^* \equiv \min\{0 \leqslant n \leqslant N \mid U_{n+1}(s_n) \leqslant \beta^n s_n\}$ *and for* $n < t^*$, a^*_n *is the attribute that maximizes* $\mathrm{E}[U^i_{n+1}(s_n, Z^i)]$.

Proof. We show that τ^*_N is the optimal stopping rule by showing that for any other stopping rule $\tau = (t; a_1, \ldots, a_t)$, $\mathrm{E}[Y_\tau] \leqslant \mathrm{E}[Y_{\tau^*_N}] = V_0$. For any stopping rule τ define $\alpha_\tau(N) \equiv \mathrm{E}[Y_\tau]$, $\alpha_\tau(n) \equiv \mathrm{E}[Y_\tau \cdot 1_{\{t<n\}}] + \mathrm{E}[V_n \cdot 1_{\{t \geqslant n\}}]$ for $1 \leqslant n < N$, and $\alpha_\tau(0) = \mathrm{E}[V_0] = V_0$. Notice that $\alpha_\tau(n)$ is the expected utility associated with a stopping rule that coincides with τ until period n, and coincides with τ^*_N afterwards, yielding a utility of V_n. Recall that $Y_0 \equiv 0$; hence, $V_0 = \max\{Y_0, U_1(0)\}$ is a constant.

The proof consists of two steps. (1) We show that for any stopping rule τ, $\alpha_\tau(n)$

is a decreasing sequence in n, and thus V_0 is a bound for the value of the objective function. And (2) we show that $\alpha_{\tau_N^*}(n)$ is a constant sequence, and thus τ_N^* is optimal since $\mathrm{E}[Y_{\tau_N^*}] = \alpha_{\tau_N^*}(N) = \alpha_{\tau_N^*}(0) = V_0$.

To prove (1), observe that $\alpha_\tau(n-1) = \mathrm{E}[Y_\tau \cdot 1_{\{t<n-1\}}] + \mathrm{E}[V_{n-1} \cdot 1_{\{t \geq n-1\}}]$ and $\alpha_\tau(n) = \mathrm{E}[Y_\tau \cdot 1_{\{t<n-1\}}] + \mathrm{E}[Y_\tau \cdot 1_{\{t=n-1\}}] + \mathrm{E}[V_n \cdot 1_{\{t \geq n\}}]$. We show that $\alpha_\tau(n-1) \geq \alpha_\tau(n)$ by showing that $\mathrm{E}[Y_\tau \cdot 1_{\{t=n-1\}}] + \mathrm{E}[V_n \cdot 1_{\{t \geq n\}}] \leq \mathrm{E}[V_{n-1} \cdot 1_{\{t \geq n-1\}}]$. Notice that by the definition of V_{n-1}, $\mathrm{E}[Y_{n-1} \cdot 1_{\{t=n-1\}}] \leq \mathrm{E}[V_{n-1} \cdot 1_{\{t=n-1\}}]$. Since for any two random variables A and B, $\mathrm{E}[\mathrm{E}[A \mid B]] = \mathrm{E}[A]$, $\mathrm{E}[V_n \cdot 1_{\{t \geq n\}}] = \mathrm{E}[\mathrm{E}[V_n \cdot 1_{\{t \geq n\}}] \mid Y_{n-1}]$. In period $n-1$, Y_{n-1} is realized and the decision whether to stop or not is made. Thus, it becomes known whether the stopping time t is greater or equal to n or not. Formally, this means that the random variable $1_{\{t \geq n\}}$ is measurable with respect to Y_{n-1}, and therefore $\mathrm{E}[\mathrm{E}[V_n \cdot 1_{\{t \geq n\}}] \mid Y_{n-1}] = \mathrm{E}[\mathrm{E}[V_n \mid Y_{n-1}] \cdot 1_{\{t \geq n\}}]$. Now, since $U_n(s_{n-1}) = \mathrm{E}[V_n \mid s_{n-1}] = \mathrm{E}[V_n \mid Y_{n-1}]$, $\mathrm{E}[\mathrm{E}[V_n \mid Y_{n-1}] \cdot 1_{\{t \geq n\}}] = \mathrm{E}[U_n(s_{n-1}) \cdot 1_{\{t \geq n\}}] \leq \mathrm{E}[V_{n-1} \cdot 1_{\{t \geq n\}}]$ by the definition of V_{n-1}. Therefore, $\mathrm{E}[Y_\tau \cdot 1_{\{t=n-1\}}] + \mathrm{E}[V_n \cdot 1_{\{t \geq n\}}] \leq \mathrm{E}[V_{n-1} \cdot 1_{\{t=n-1\}}] + \mathrm{E}[V_{n-1} \cdot 1_{\{t \geq n\}}] = \mathrm{E}[V_{n-1} \cdot 1_{\{t \geq n-1\}}]$, and (1) is proved.

Next, we show that (2) holds. We do this by showing that the inequalities in (1) are satisfied as equalities for τ_N^*. By the definition of τ_N^*, t^* is the first time in which the maximum between Y_n and $U_{n+1}(s_n)$ is obtained on Y_n. Therefore, under τ_N^*, $t^* = n-1$ implies $V_{n-1} = Y_{n-1}$ and therefore $\mathrm{E}[Y_{n-1} \cdot 1_{\{t^*=n-1\}}] = \mathrm{E}[V_{n-1} \cdot 1_{\{t^*=n-1\}}]$. $t^* \geq n$ implies that $V_{n-1} = U_n(s_{n-1})$ and therefore $\mathrm{E}[V_n \cdot 1_{\{t^* \geq n\}}] = \mathrm{E}[\mathrm{E}[V_n \cdot 1_{\{t^* \geq n\}}] \mid Y_{n-1}] = \mathrm{E}[\mathrm{E}[V_n \mid Y_{n-1}] \cdot 1_{\{t^* \geq n\}}] = \mathrm{E}[U_n(s_{n-1}) \cdot 1_{\{t^* \geq n\}}] = \mathrm{E}[V_{n-1} \cdot 1_{\{t^* \geq n\}}]$. This completes the proof of Proposition 1. □

Notice that the stopping rule τ_N^* satisfies the following two properties. The first one is that it is stationary. At each point in time its decision is independent of the number of products that have been observed so far. The second property is that it is myopic. At any stage in the observation process the decision whether to stop or not is made by comparing the current reservation value with the expected utility from making another single observation. Furthermore, the attributes that τ_N^* chooses to observe are those that maximize the one observation ahead expected utility.

We now show that second-order stochastic dominance between the Z^i's characterizes the case in which an optimal strategy can be restricted to observe only one attribute. Namely, if there exists an attribute p such that Z^p is second-order stochastically dominated by all the other Z^i's, then there exists an optimal strategy that will choose to observe only attribute p. If, on the other hand, there is no such attribute, every optimal strategy will observe at least two attributes under some continuity assumptions. Before turning to the statement and proof of this result, we present the definition of second-order stochastic dominance as it appears in Rothschild and Stiglitz (1970).

Definition. Suppose that P and Q are two random variables with a bounded support $[0, M]$, and with the corresponding cumulative distribution functions, $F_P(\cdot)$ and $F_Q(\cdot)$, P is second-order stochastically dominated by Q if

$$\text{(i)} \quad \int_0^M (F_P(x) - F_Q(x))\,\mathrm{d}x = 0$$

and

$$\text{(ii)} \quad \int_0^m (F_P(x) - F_Q(x))\,\mathrm{d}x \geq 0 \quad \text{for all } 0 \leq m < M\,.$$

Notice that since the Z^i's are non-negative $\mathrm{E}[Z^i] = \int_0^M (1 - G_i(x))\,\mathrm{d}x$ for all $i \in \{1, \ldots, k\}$. This, together with the fact that $\mathrm{E}[Z^i] = \mathrm{E}[Z^j]$ for all $i, j \in \{1, \ldots, k\}$, implies that $\int_0^M G_i(x)\,\mathrm{d}x = \int_0^M G_j(x)\,\mathrm{d}x$ for any $i, j \in \{1, \ldots, k\}$, and thus (i) above is satisfied for any Z^i and Z^j. Regarding the examples given in the introduction, in Example 1 there is no second-order stochastic dominance between Z^p and Z^q and in Example 2, Z^p is stochastically dominated by Z^q.

Theorem 2. *When $N \geq 2$, there exists an optimal strategy that observes only attribute p if Z^p is second-order stochastically dominated by Z^i for all $i \in \{1, \ldots, k\}$. Conversely, if the cumulative distributions of the Z^i's $i \in \{1, \ldots, k\}$, are strictly monotone and differentiable, and an optimal strategy observes attribute p only, then Z^p is second-order stochastically dominated by all the other Z^i's.*

Proof. 'If': Recall that $U_n^i(s, z^i) = \beta^n \max\{s, z^i\}$ for all $0 \leq n \leq N$ and $i \in \{1, \ldots, k\}$. $U_n^i(s, z^i)$ is a convex function of z^i because maximum is a convex function. We use a theorem from Rothschild and Stiglitz (1970) that establishes the following fact. If a random variable A is second-order stochastically dominated by a random variable B, then for every convex function $f(\cdot)$, $\mathrm{E}[f(A)] \geq \mathrm{E}[f(B)]$. From applying this result to $U_n^i(s, z^i)$, it follows that $\mathrm{E}[U_n^p(s, Z^p)] \geq \mathrm{E}[U_n^i(s, Z^i)]$ for all $i \in \{1, \ldots, k\}$, $s \in [0, M]$ and $1 \leq n \leq N$. Hence, in every period n, observing attribute p is preferable to observing any other attribute i.

Conversely, we demonstrate that if there is no second-order stochastic dominance between any two attributes $i, j \in \{1, \ldots, k\}$, then observing attribute i in the last period in superior to observing attribute j with a positive probability and vice versa.

We show that $\int_0^m (G_i(x) - G_j(x))\,\mathrm{d}x > 0$ implies $\mathrm{E}[U_N^i(m, Z^i)] > \mathrm{E}[U_N^j(m, Z^j)]$. $\int_0^M G_i(x)\,\mathrm{d}x = \int_0^M G_j(x)\,\mathrm{d}x$ and therefore $\int_0^m (G_i(x) - G_j(x))\,\mathrm{d}x > 0$ implies $\int_m^M G_i(x)\,\mathrm{d}x < \int_m^M G_j(x)\,\mathrm{d}x$. Integration by parts implies that $m + \int_m^M (1 - G_i(x))\,\mathrm{d}x = G_i(m)m + \int_m^M x\,\mathrm{d}G(x) = \mathrm{E}[m \cdot 1_{\{Z^i \leq m\}}] + \mathrm{E}[Z^i \cdot 1_{\{Z^i > m\}}] = \mathrm{E}[\max\{m, Z^i\}]$, and therefore $\int_0^m (G_i(x) - G_j(x))\,\mathrm{d}x > 0$ implies $\beta^N \mathrm{E}[\max\{m, Z^i\}] > \beta^N \mathrm{E}[\max\{m, Z^j\}]$ or $\mathrm{E}[U_N^i(m, Z^i)] > \mathrm{E}[U_N^i(m, Z^j)]$.

If neither Z^i stochastically dominates Z^j nor Z^j stochastically dominates Z^i, then there exists an m_1 such that $\int_0^{m_1} (G_i(x) - G_j(x))\,\mathrm{d}x > 0$ and there exists an m_2 such that $\int_0^{m_2} (G_i(x) - G_j(x))\,\mathrm{d}x < 0$. Continuity of the integral implies that there exist two disjoint intervals of possible states $[s_i', s_i'']$ and $[s_j', s_j'']$ such that for all $m \in [s_i', s_i'']$, $\int_0^m (G_i(x) - G_j(x))\,\mathrm{d}x > 0$ and thus in the last period an optimal strategy will prefer observing attribute i to attribute j and for all $m \in [s_j', s_j'']$, $\int_0^m (G_i(x) - G_j(x))\,\mathrm{d}x < 0$ and in the last period an optimal strategy will prefer observing attribute j to attribute i. Since every Z^i has a strictly monotonic cumulative distribution, for any interval there is a positive probability that $s_{N-1} \in [s_i', s_i'']$ and that $s_{N-1} \in [s_j', s_j'']$. This completes the proof of Theorem 2. □

4. The infinite horizon problem

As opposed to the finite horizon problem, when the sequence of products is infinite the horizon faced by the decision-maker is the same at every stage. We show that in this case an optimal strategy always observes the same attribute. In order to prove this result, we use the fundamental theorem of discounted dynamic programming which is due to Blackwell (1965). According to this theorem, *excessivity* is a sufficient condition for the optimality of a policy τ. Formally, a policy τ satisfies the *excessivity* criterion if for all $s \in S$ and $n \geq 0$,

$$V_{n,\infty}(\tau)(s) \geq O(V_{n,\infty}(\tau))(s)\,,$$

where $O(\cdot)$ is an operator which is defined as follows. Let $F \equiv \{f : S \to \mathbb{R} \mid f$ is bounded and measurable$\}$, define $O : F \to F$ as

$$O(f)(s) \equiv \sup_{a \in A} \int (r(s, a, s') + \beta f(s'))\,\mathrm{d}q(s' \mid s, a)\,.$$

Verbally, a strategy τ satisfies the excessivity criterion if at any stage $n \geq 0$ and any state $s \in S$, delaying τ by one period, at period n taking the best possible action, and at period $n + 1$ reverting back to τ does not yield a higher expected payoff.

Before identifying the optimal strategy, we prove a preliminary result. For $i \in \{1, \ldots, k\}$, define a stopping time t_i:

$$t_i \equiv \min\{n \geq 0 \mid \beta \mathrm{E}[\max\{s_n, Z^i\}] \leq s_n\}\,.$$

Proposition 3. *For every attribute $i \in \{i, \ldots, k\}$, t_i determines a unique threshold value $b_i = \beta \mathrm{E}[\max\{b_i, Z^i\}]$ such that for states $s > b_i$, $\beta \mathrm{E}[\max\{s, Z^i\}] < s$ and for states $s < b_i$, $\beta \mathrm{E}[\max\{s, Z^i\}] > s$.*

Proof. The proof follows from the properties of $\beta \mathrm{E}[\max\{s, Z^i\}]$ as a function of s. Notice that,

(1) $\beta \mathrm{E}[\max\{s, Z^i\}]$ is a continuous, increasing, and convex function of s.

(2) Non-negativity of the decision-maker's utility function $U(\cdot)$ implies that $\beta \mathrm{E}[\max\{0, Z^i\}] = \beta \mathrm{E}[Z^i]$.

(3) Since $U(\cdot)$ is bounded by M, $\beta \mathrm{E}[\max\{s, Z^i\}] = \beta s$ for all $s \geqslant M$.

By (1), $f(s) \equiv \beta \mathrm{E}[\max\{s, Z^i\}] - s$ is a continuous function; since $f'(s) = \beta \mathrm{E}[\mathrm{d}(\max\{s, Z^i\})/\mathrm{d}s] - 1 \leqslant \beta \mathrm{E}[1] - 1 = \beta - 1 < 0$, $f(s)$ is decreasing; $f(0) = \mathrm{E}[Z^i] > 0$ by (2); and $f(M) = \beta M - M < 0$ by (3). Therefore, there exists a unique $0 < b_i < M$ satisfying $b_i = \beta \mathrm{E}[\max\{b_i, Z^i\}]$. For any $s > b_i$, there exists a $\lambda < 1$ such that $s = \lambda b_i + (1-\lambda)M$. Convexity of $\beta \mathrm{E}[\max\{s, Z^i\}]$ in s implies $\beta \mathrm{E}[\max\{s, Z^i\}] \leqslant \lambda \beta \mathrm{E}[\max\{b_i, Z^i\}] + (1-\lambda)\beta \mathrm{E}[\max\{M, Z^i\}]$ which, by the definition of b_i and (3) above, $= \lambda b_i + (1-\lambda)\beta M < \lambda b_i + (1-\lambda)M = s$. Therefore, $\beta \mathrm{E}[\max\{s, Z^i\}] < s$ for all $s > b_i$. For $0 \leqslant s < b_i$, observe that $\lambda = (b_i - s)/(M - s)$ satisfies $0 < \lambda < 1$ and $b_i = \lambda M + (1-\lambda)s$. Suppose that $\beta \mathrm{E}[\max\{s, Z^i\}] \leqslant s$. By convexity of $\beta \mathrm{E}[\max\{s, Z^i\}]$ in s, $b_i = \beta \mathrm{E}[\max\{b_i, Z^i\}] \leqslant \lambda \beta \mathrm{E}[\max\{M, Z^i\}] + (1-\lambda)\beta \mathrm{E}[\max\{s, Z^i\}] \leqslant \lambda \beta M + (1-\lambda)\beta s = \beta b_i < b_i$. A contradiction. Therefore, we conclude that for $0 \leqslant s < b$, $\beta \mathrm{E}[\max\{s, Z^i\}] > s$. □

We now turn to define the infinite horizon optimal strategy $\tau^*_\infty = (t^*; a^*_1, \ldots, a^*_{t^*})$. Let

$$t^* \equiv \min\{n \geqslant 0 \mid \beta \mathrm{E}[\max\{s_n, Z^i\}] \leqslant s_n \text{ for all } i \in \{1, \ldots, k\}\}\,.$$

Let $b \equiv \max_{i \in \{1,\ldots,k\}} \{b_i\}$ and let $j^* \in \arg\max_{i \in \{1,\ldots,k\}} \{b_i\}$. For $1 \leqslant n \leqslant t^*$ define $a^*_n = j^*$. In the case when there exists an $i^* \neq j^*$ such that $b_{i^*} = b_{j^*}$, the optimal strategy τ^*_∞ can alternate between observing i^* and j^*. Alternatively, as a function from the states space to the actions space the stopping rule τ^*_∞ can be defined as follows:

$$\tau^*_\infty(s) = \begin{cases} \text{Stop}\,, & s \geqslant b\,, \\ \text{observe attribute } j^*\,, & s < b\,. \end{cases}$$

Notice that by Proposition 3 the two alternative definitions of τ^*_∞ are equivalent, namely they induce exactly the same strategy or observation behavior. We now prove,

Theorem 4. *The stopping rule τ^*_∞ is optimal.*

Proof. We demonstrate that τ^*_∞ satisfies the excessivity property. We start by computing $V_{n,\infty}(\tau^*_\infty)(s)$. We claim that

$$V_{n,\infty}(\tau^*_\infty)(s) = \begin{cases} \beta^n s\,, & s \geqslant b\,, \\ \beta^n b\,, & s < b\,, \\ 0\,, & s = \emptyset\,. \end{cases}$$

$s = \emptyset$ implies that the observation process has already stopped and therefore

$V_{n,\infty}(\tau^*_\infty)(\emptyset) = 0$. For states $s \geqslant b$, τ^*_∞ stops and therefore $V_{n,\infty}(\tau^*_\infty)(s) = \beta^n s$. For states $s < b$, $V_{n,\infty}(\tau^*_\infty)(s) = \beta^n \cdot \mathrm{E}[\beta^m] \cdot \mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b]$, where m is distributed geometrically with probability of success $p \equiv \Pr(Z^{j^*} \geqslant b)$. Therefore, $\mathrm{E}[\beta^m] = \sum_{m=1}^{\infty} \beta^m (1-p)^{m-1} p = p\beta/[1-\beta(1-p)]$. By Proposition 3, $b = \beta\mathrm{E}[\max\{b, Z^{j^*}\}] = \beta((1-p)b + p\mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b])$. Rearranging terms yields $b = [p\beta/(1-\beta(1-p))]\mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b]$, or $\mathrm{E}[\beta^m] \cdot \mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b] = b$ and therefore $V_{n,\infty}(\tau^*_\infty)(s) = \beta^n b$.

We demonstrate excessivity. Suppose that the decision-maker has already observed $n \geqslant 0$ products. For states $s_n \geqslant b$, τ^*_∞ stops. Consider instead a different policy τ which continues by observing any attribute i and then continues according to τ^*_∞. Notice that since $s_{n+1} \geqslant s_n$, observing attribute i can only increase the state, and therefore τ, which now imitates τ^*_∞ stops immediately after the $n+1$th product is observed. Now, for states $s_n \geqslant b$, Proposition 3 established that $\beta\mathrm{E}[\max\{s, Z^i\}] < s$, and therefore τ yields an expected value of $\beta^{n+1}\mathrm{E}[\max\{s_n, Z^i\}]$, lower than $\beta^n s_n$ which is the value that τ^*_∞ would give by stopping at stage n.

For states $s_n < b$, τ^*_∞ continues by observing attribute j^*. A policy τ which stops immediately is doing worse since it gives a value of $\beta^n s_n$, while τ^*_∞ promises an expected value $\beta^n b > \beta^n s_n$. Consider a different policy τ which continues by observing attribute $i \neq j^*$ and then continues according to τ^*_∞. Distinguish between two cases. (1) After observing attribute i, the state is $s_{n+1} < b$ and τ continues as τ^*_∞. (Notice that after observing attribute i, the policies τ and τ^*_∞ coincide.) In this case, since $\mathrm{E}[V_{n,\infty}(\tau^*_\infty)(s_n) \mid s_{n+1} < b] = \beta^{n+1} b$, using τ yields an expected payoff of $\beta^{n+1} b$ which is not better than what τ^*_∞ gives. (2) After observing attribute i, the state is $s_{n+1} \geqslant b$ and τ continues as τ^*_∞, and therefore stops. We claim that strategy τ^*_∞ (that is, observing attribute j^*) would have given a higher expected payoff in this case as well. Conditional on $s_{n+1} \geqslant b$, observing attribute i yields an expected value of $\beta^{n+1}\mathrm{E}[Z^i \mid Z^i \geqslant b]$, while observing attribute j^* yields an expected value of $\beta^{n+1}\mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b]$. Therefore, it is sufficient to show that $\mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b] \geqslant \mathrm{E}[Z^i \mid Z^i \geqslant b]$. We claim that for any $i, j \in \{1, \ldots, k\}$, $b_j \geqslant b_i$ implies $\mathrm{E}[\max\{b_j, Z^j\}] \geqslant \mathrm{E}[\max\{b_j, Z^i\}]$. The reason is that by Proposition 3, $\beta\mathrm{E}[\max\{b_j, Z^j\}] = b_j$ and since $\mathrm{E}[\max\{s, Z^i\}]$ is increasing in s, $\beta\mathrm{E}[\max\{b_j, Z^i\}] \geqslant \beta\mathrm{E}[\max\{b_i, Z^i\}] = b_i$ by the proposition. Now we show that $\mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b] \geqslant \mathrm{E}[Z^i \mid Z^i \geqslant b]$. Suppose the opposite holds. It follows that $\Pr(Z^{j^*} < b)b + \Pr(Z^{j^*} \geqslant b)\mathrm{E}[Z^{j^*} \mid Z^{j^*} \geqslant b] < \Pr(Z^{j^*} < b)b + \Pr(Z^{j^*} \geqslant b)\mathrm{E}[Z^i \mid Z^i \geqslant b]$, and that $\mathrm{E}[\max\{b, Z^{j^*}\}] < \mathrm{E}[\max\{b, Z^i\}]$, which contradicts the previous fact. This completes the proof of Theorem 4. □

An important conclusion that we can draw from the optimality of the stopping rule τ^*_∞ is that for every decision-maker (who is characterized by a utility function $U(\cdot)$ and a discount factor β) there is only one interesting attribute, which is independent of the state or the realizations of his past observations. The attribute that τ^*_∞ chooses to observe depends on β, and it is possible that different β's will lead to observing different attributes. Thus, two decision-makers with the same utility function but different discount factors may choose to observe different attributes. More specifically, as β gets close to 1, the attribute that is chosen by τ^*_∞ is the following. Let $G_i^{-1}(p) \equiv \inf\{x \mid G_i(x) \geq p\}$ denote the inverse of the cumulative distribution function $G_i(\cdot)$. Note that the set $\arg\max_{i\in\{1,\ldots,k\}} \{G_i^{-1}(1)\}$ is non-empty. If it contains a single maximizer, this will be the infinite horizon optimal attribute when β is close to 1. If there is more than one maximizer, the infinite horizon optimal attribute will be the one that maximizes $\mathrm{E}[\max\{g-\varepsilon, Z^i\}]$ for all $\varepsilon < \varepsilon'$ for some $\varepsilon' > 0$ and where $g \equiv \max_{i\in\{1,\ldots,k\}} \{G_i^{-1}(1)\}$. Such a maximizer exists and it is the one with the highest associated b_i value for all $\beta > \beta'$ for some $\beta' < 1$. In words, it is the attribute which is capable of pulling the product to its highest possible value. When β approaches 1 and when the sequence of products is infinite, the decision-maker becomes 'patient' enough to risk observing only this attribute. The dependence of the optimal attribute on β in the infinite horizon problem contrasts with what happens in the finite horizon problem. When the sequence of products is finite, if it is the case that the optimal strategy continues to observe another product, then the choice of which attribute to observe is independent of β. β affects the finite horizon optimal strategy only through the decision of when to stop.

Another important parameter of the optimal strategy is b which, as shown in the proof of Theorem 4, is the expected utility of the decision-maker when starting the observation process. Not surprisingly, it is monotonically increasing in β.

Notice that the stopping rule τ^*_∞ is also stationary. However, it is only 'almost' myopic. It is myopic in the sense that the decision whether to stop or not depends only on observing the situation one step ahead. Yet, 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 ahead expected payment. It is therefore possible that in some states, if the observation process would have ended in the next period, an optimal strategy would have picked a different attribute than τ^*_∞, as demonstrated in Example 1. By contrast, τ^*_N is truly myopic.

This last observation suggests the existence of a discontinuity between the finite and infinite horizon problems. The two following observations show that this is not the case. For every $n \geq 0$, $V_{n,N}(\tau^*_N)(s)$ is increasing in N and is bounded by $V_{n,\infty}(\tau^*_\infty)(s)$. In fact,

Lemma 5. *Fix* $n \geq 0$, $V_{n,N}(\tau^*_N)(s) \xrightarrow[N\to\infty]{} V_{n,\infty}(\tau^*_\infty)(s)$ *for any* $s \in S$. *Moreover, the convergence is uniform over* $s \in S\backslash\{\emptyset\}$.

Proof. $U(\cdot)$ is bounded by M. Therefore for $s \in S\backslash\{\emptyset\}$, $V_{n,\infty}(\tau^*_\infty)(s) - V_{n,N}(\tau^*_N)(s) \leq \beta^{N-n}M \xrightarrow[N\to\infty]{} 0$. For $s = \emptyset$, $V_{n,N}(\tau^*_N)(\emptyset) = V_{n,\infty}(\tau^*_\infty)(\emptyset) = 0$. □

The previous lemma establishes the similarity between the finite and infinite horizon problems by showing that the payoffs associated with the optimal strategies coincide in the limit. The following theorem shows the similarity in the actions taken by the optimal finite and infinite horizon strategies.

Theorem 6. *There exists a finite N' such that for any finite horizon $N \geq N'$, an optimal strategy τ_N^* will observe an infinite horizon optimal attribute j^* of at least $N - N'$ products.*

Proof. In the proof of Theorem 4 it was established that

$$V_{n,\infty}(\tau_\infty^*)(s) = \begin{cases} \beta^n s\,, & s \geq b\,, \\ \beta^n b\,, & s < b\,, \\ 0\,, & s = \emptyset\,, \end{cases}$$

for all $n \geq 0$. By Lemma 5, for any $n \geq 0$ and $\varepsilon > 0$ there exists an $N'(\varepsilon)$ such that for all $N \geq N'$, $V_{n,N}(\tau_N^*)(s) > V_{n,\infty}(\tau_\infty^*)(s) - \varepsilon$ for all $s \in S$. Fix an $n \geq 0$, an

$$\varepsilon < \min_{\substack{i \notin \arg\max\{b_j\} \\ j \in \{1,\dots,k\}}} \{b - \beta \mathrm{E}[\max\{b, Z^i\}]\}$$

(notice that by Proposition 3, $\varepsilon > 0$) and $N \geq N'(\varepsilon)$. For $s_n \geq b$, both τ_N^* and τ_∞^* stop. Suppose that $s_n < b$. Observing j^* yields

$$\begin{aligned} &\beta \mathrm{E}[V_{n+1,N}(\tau_N^*)(\max\{\mathrm{s_n}, \mathrm{Z}^{j^*}\})] \\ &\quad > \beta \mathrm{E}[V_{n+1,\infty}(\tau_N^*)(\max\{s_n, Z^{j^*}\})] - \varepsilon \\ &\quad = \beta(\Pr(Z^{j^*} < b)b + \mathrm{E}[Z^{j^*} \cdot 1_{\{Z^{j^*} \geq b\}}]) - \varepsilon \end{aligned}$$

which, by Proposition 3, equals $b - \varepsilon$. Observing any other attribute $i \notin \arg\max_{j \in \{1,\dots,k\}} \{b_j\}$ yields no more than

$$\begin{aligned} &\beta \mathrm{E}[V_{n+1,N}(\tau_N^*)(\max\{s_n, Z^i\})] \\ &\quad \leq \beta \mathrm{E}[V_{n+1,\infty}(\tau_N^*)(\max\{s_n, Z^i\})] \\ &\quad = \beta(\Pr(Z^i < b)b + \mathrm{E}[Z^i \cdot 1_{\{Z^i \geq b\}}])\,. \end{aligned}$$

We want to show that the observing attribute j^* is better than observing attribute i. We have to show that $b - \varepsilon > \beta(\Pr(Z^i < b)b + E[Z^i \cdot 1_{\{Z^i \geq b\}}]) = \beta \mathrm{E}[\max\{b, Z^i\}]$. To complete the proof, note that this last inequality holds because $\varepsilon < \min_{\substack{i \notin \arg\max\{b_j\} \\ j \in \{1,\dots,k\}}} \{b - \beta \mathrm{E}[\max\{b, Z^i\}]\}$. □

Acknowledgements

I thank Itzhak Gilboa, Isaac Meilijson, Ariel Rubinstein, Asher Wolinsky and an anonymous referee for their helpful comments.

References

D.A. Berry and B. Fristedt, Bandit Problems – Sequential Allocation of Experiments (Chapman and Hall, London, 1985).

D. Blackwell, Discounted dynamic programming, Ann. Math. Stat. 36 (1965) 226–235.

Y.S. Chow, H. Robbins and D. Ziegmund, Great Expectations: The Theory of Optimal Stopping (Houghton Mifflin, Boston, MA, 1971).

C.H. Coombs, A Theory of Data (Wiley, New York, 1964).

P.C. Fishburn, Utility theory, Manage. Sci. 13 (1968) 435–453.

K.J. Lancaster, A new approach to consumer theory, J. Polit. Econ. 74 (1966) 132–157.

J. McMillan and M. Rothschild, Search, in: R.J. Aumann and S. Hart, eds., Handbook of Game Theory with Economic Applications (North-Holland, Amsterdam, 1992).

M. Rothschild and J.E. Stiglitz, Increasing risk: 1. A definition, J. Econom. Theory 2 (1970) 225–243.

A. Tversky, Elimination by aspects: A theory of choice, Psychol. Rev. 79 (1972) 281–299.