



Queen's Economics Department Working Paper No. 1119

## How to take an exam if you must: Decision under Deadline

Hsueh-Ling Huynh  
Boston University

Sumon Majumdar  
Queen's University

Department of Economics  
Queen's University  
94 University Avenue  
Kingston, Ontario, Canada  
K7L 3N6

2-2006

How to take an exam if you must:  
Decision under Deadline

Hsueh-Ling Huynh

Department of Economics, Boston University  
270 Bay State Road, Boston, MA 02215.

Sumon Majumdar<sup>1</sup>

Department of Economics, Queen's University  
Kingston, Ontario. Canada K7L 3N6.

<sup>1</sup>*Corresponding author*; e-mail: [sumon@qed.econ.queensu.ca](mailto:sumon@qed.econ.queensu.ca). We are grateful to Daniel Akerberg, Chirantan Ganguly, Kevin Lang, Dilip Mookherjee, Debraj Ray, Bob Rosenthal and seminar participants at Boston University, University of Toronto and the University of British Columbia for their very useful comments. The title of this paper is a pun on a book by Dubins and Savage (1965).

## **Abstract**

This paper uses the example of an exam to model multi-dimensional search under a deadline. When the dimension is two, an order-invariance property allows simple characterization of the optimal search policy. Behavior is shown to be highly sensitive to changes in the deadline, and a wide variety of policies can be rationalized (as being optimal) as the length of the deadline increases. This is contrasted with behavior under the traditional case of geometric discounting, in which a similar sensitivity to changes in the discount factor cannot hold. For dimensions higher than two, the invariance principle does not hold; this increases complexity of the problem of finding the optimal search policy.

# 1 Introduction

The pressure of a deadline is encountered in many situations. Often such deadlines are externally imposed, or are the result of a credible commitment that one has made. Graduate students aim to finish their thesis before financial support runs out; governments seek to maximize their ‘achievements’ by the end of their electoral term; firms’ research strategies are often influenced by the deadlines before patents expire. While the economic importance of deadlines is obvious, its exact impact has been relatively little studied.

In this paper we study the effect of deadlines on economic processes that involve multi-dimensional search. While searches with an infinite (or indefinite) horizon have inspired an impressive literature both within and beyond economics, it is perhaps easy to believe that finite horizon search problems offer little intellectual challenge or economic insight. We show that, to the contrary, the nature and length of the deadline can have strong and possibly counter-intuitive effects on the process and optimal strategy of the search.

A good paradigm for many decision situations under deadlines is one that most individuals have faced at some time or the other, namely, when taking an exam: credit is only awarded for completing correct answers within the stipulated time; after the time limit on the exam is past, the value from a correct answer diminishes to zero. When two or more questions are involved and one is unsure of one’s chance of success on the various questions, the existence of such a deadline will affect behavior with respect to the attempts that one makes on each question, and consequently on the outcome. In this paper, we seek to explore the behavioral effects of deadlines in decision problems involving multiple time-consuming projects, with uncertainty regarding the chance of success in each.

We model the situation as a multi-dimensional search problem: there are several questions, each with a possible set of answers. Although the candidate is aware of this set, she is unsure which one among them is the correct answer. This uncertainty could be either factual or logical; if it stems from an uncertainty in knowledge such as “What is the currency of Russia — the ruble or the peso?”, it may be impossible to resolve the uncertainty within the atmosphere of the exam. If however the uncertainty is logical, then time could be a useful resource in resolving it; for example, in a mathematics exam, knowledge of the axioms of mathematics should, in principle, allow one

to verify whether a given series of statements constitute a proof (or refutation) of the proposition in question. In this paper, we assume that the uncertainty the candidate has over the possible answers is logical, so that given enough time, she could arrive at the correct answer by verification and elimination. More precisely, we assume that in each period of time, it is possible for her to verify whether or not a particular answer is the right one for the corresponding question. Points are only awarded for arriving at the correct answer within the given deadline.

The candidate begins with some subjective beliefs about how likely it is for each of the possibilities to be the correct answer. This sort of probabilistic judgement may be a reflection of her preparation, expertise or experience. Given her subjective beliefs on the various sets of answers and the given deadline, the optimal search policy for the candidate will consist of deciding at each point of time, the question that she will attempt in the upcoming period, given her history of success or failure up to that stage.

We begin by establishing that when there are only two questions, changing the order of attempts on the two questions does not change the probability of finding the correct answer in either. Thus for example, the following two policies have the same value: (i) “Start by attempting question  $A$  once. If successful, use remaining time on  $B$ ; if unsuccessful in the first attempt on  $A$ , then try  $B$  until success.” (ii) “Attempt question  $B$  upto (and including) the last but one period. If successful at any intermediate period, use remaining time on  $A$ ; if not, then attempt question  $A$  (once) in the last period.” This order-invariance property allows a simple characterization of the optimal policy in the case of an exam with two questions. In most of the paper, we focus our analysis on this case.

In determining this optimal policy, the candidate trades off the immediate prospect of getting a reward against the informational value from eliminating some suspects (and thereby increasing the probability of future rewards). A simple example will illustrate the point: suppose there are two questions  $A$  and  $B$ . Suppose on question  $A$ , the candidate has 5 possible answers among which she is 99% sure that  $A_1$  is the correct one (with the remaining 1% equally distributed over the other four possibilities). For question  $B$ , she has two suspects, each of which she considers to be 50% likely. If she has only one period to use, she would obviously attempt question  $A$ , but facing an exam of length  $T = 2$ , it is optimal for her to start by attempting  $B$  (and sticking to it if she fails).

The above example also shows, albeit in a rather trivial way, the effect of the deadline in play in

determining the optimal policy. More generally, we may attempt to answer the following question:

- What behavior can be rationalized as optimal under a deadline, with respect to any system of subjective probabilities at all?

In this paper, we provide some partial answers. While it is not true that all sequences of policies are rationalizable, we can do so for several broad classes of policies. As we see later, even given the same system of beliefs, a wide variety of disparate behavior can be justified as being rational under various sets of deadlines.

While the appearance of deadlines is a ubiquitous fact of life, one may also ask whether deadlines could be interpreted as a particular form of discounting, namely, no discounting of payoffs till the deadline and then infinite discounting past it. The most widely used way of modelling impatience in economics is to assume that agents have a constant rate of inter-temporal substitution, thus giving rise to *geometric discounting*. Do deadlines and geometric discounting lead to similar sets of behavior? More specifically, we formulate our search model as an infinite-horizon geometric discounting model, and ask:

- Is every search policy that is rationalizable under an unbounded sequence of deadlines also rationalizable under a sequence of discount factors approaching unity?

In other words, is behavior under a long enough deadline similar to that without any deadline but with a high enough discount factor? We consider this question in section 4. As we shall see, the answer is negative. We find that while it is possible to justify behavior that is very sensitive to a change in the length of the deadline as being optimal, it is much more difficult to generate a similar sensitivity of optimal behavior to changes in the discount factor. Geometric discounting imparts the property of analyticity to the difference in value functions between two different policies; this makes infinite oscillation between them impossible. On the other hand, it is possible to rationalize behavior that involves switching infinitely often between such policies as “attempt question  $A$  until success” and “attempt question  $B$  until success” depending only on the length of the deadline.

As mentioned earlier, in the case of only two questions, an order-invariance property exists so that the value for any policy does not depend on the order of attempts on the two questions.

This property does not hold for three questions or more. Furthermore, as we show via a counter-example, the invariance property does not even hold for the optimal policy, thereby making its characterization more complex in such cases.

The rest of the paper is organized as follows. In Section 2, we define the basic search model under deadlines. Section 3 characterizes the optimal policy in the case of two questions. In section 4, we examine behavior that can be rationalized under deadlines versus that in an infinite-horizon geometric discounting framework. Section 5.1 presents a counter-example to show that the invariance property fails to hold in an exam involving more than two questions. In section 5.2, we consider an extension to the basic model for which the order invariance property holds for any number of questions, and we conclude by broadly discussing some possible applications.

## 2 A Search Model with Deadlines

### 2.1 Model of an Exam: the general problem

Consider an exam with  $n$  questions,  $A, B, C, \dots$ . For each question, the candidate has a list of possible answers  $\{A_1, A_2, \dots\}, \{B_1, B_2, \dots\}$ , etc. Each list contains exactly one correct answer to the question and the lists are mutually exclusive: thus for example,  $A_4$  cannot be the right answer to both questions  $A$  and  $B$ . These facts are known to the candidate.

The candidate's subjective probability on  $A_i$  being the correct answer to question  $A$  is denoted by  $a_i$ . Therefore  $a_i \geq 0$  and  $\sum_i a_i = 1$ . Likewise for the other questions, we adopt the convention of denoting subjective probabilities on the answers by the corresponding lower-case letter. The probability distributions  $\mathbf{a}, \mathbf{b}, \dots$  are assumed to be mutually independent; thus failure or success on one question does not provide any new information about the other questions. As discussed in the Introduction, this formulation is intended to capture uncertainty which is "logical" in nature and may be resolved over time by verification and elimination.

The duration of the exam consists of  $T$  periods of time. In each period, the candidate is able to pick exactly one question (say  $B$ ) and verify whether a particular possibility (say  $B_4$ ) is the correct answer for that question.  $\alpha$  marks are awarded to the candidate if she gets the right answer for question  $A$  within the stipulated time,  $\beta$  marks for  $B$ , and so forth. The candidate's objective is

to maximize the expected number of marks in the exam.<sup>1</sup>

Facing the deadline, upon deciding to attempt a particular question in the upcoming period, the candidate will clearly pick to verify the most likely possibility from the list of (remaining) answers to that question. Therefore, without loss of generality we may relabel the possible answers to a question such that the corresponding subjective probabilities are in non-increasing order i.e.  $a_i \geq a_{i+1}$  etc. If the candidate succeeds with her choice, she then shifts her focus to the remaining unanswered questions; if she does not succeed, she updates her belief on the question that she has just attempted. For example, if she has tried question  $B$  for the first  $k$  periods and not found the correct answer, then her subjective probability on  $B$  is updated, with support on  $\{B_{k+1}, B_{k+2}, \dots\}$  and the corresponding probabilities given by  $b_{k+j}/(1 - \sum_{i=1}^k b_i)$ ,  $j = 1, 2, \dots$

Although this model is reminiscent of ‘multi-armed bandit’ problems (see Gittins 1989), there are two important differences: (i) in contrast to the finite deadlines in our model, the time horizon in the bandit problems is typically infinite, and (ii) in the typical bandit problem, flow payoffs are realized at every period, as opposed to the one-shot nature of the payoffs in our model. As we shall see below, this makes the analysis in the two problems significantly different.

## 2.2 An Equivalent Formulation in terms of Hazard Rates

For future analysis, it is convenient to reformulate the model in terms of the hazard rates on the questions. We denote the *hazard rate* for the distribution  $\mathbf{a}$  by:

$$h_i^A = \frac{a_i}{(1 - a_1 - \dots - a_{i-1})} \quad i = 1, 2, \dots \quad (1)$$

This is the probability that  $A_i$  is the correct answer for question  $A$ , conditioned on the event that  $A_1, A_2, \dots, A_{i-1}$  are not. Thus, if the candidate has failed to find the right answer to question  $A$  among the first  $i - 1$  possibilities, then her subjective belief on  $A_i$  being the correct one is given by  $h_i^A$ . The hazard rates for the other questions are similarly defined. We adopt the convention that if  $a_1 + \dots + a_{i-1} = 1$ , then define  $h_i^A = 0$ .

For future reference, it may be also useful to note the reciprocal relation:

$$a_i = (1 - h_1^A)(1 - h_2^A)\dots(1 - h_{i-1}^A)h_i^A = g_1^A g_2^A \dots g_{i-1}^A h_i^A \quad (2)$$

---

<sup>1</sup>As we show in section 5.2, the analysis is isomorphic if the candidate is interested in correctly answering only one of two equally weighted questions.



where  $g_i^A$  is defined as  $g_i^A = 1 - h_i^A$ . Similarly,  $b_i = g_1^B g_2^B \dots g_{i-1}^B h_i^B$ .

### 3 The Case of Two Questions

In this section, we consider the problem when the candidate faces only two questions. In this case, a policy needs to only determine the sequence to be followed when none of the questions have been answered correctly. In the event of a success, the follow-up policy is trivial, namely, use the residual time to attempt the remaining (unanswered) question. As we show later, this makes analysis of the optimal policy in the case of two questions much simpler than that for a higher number of questions.

#### 3.1 An Invariance Principle

Let us denote the two questions by  $A$  and  $B$ . Here, a policy can be described as a sequence of the following form:  $(Q_i^{(t)})_{t=1}^T, Q_i^{(t)} \in \{A, B\}$ . This is understood to mean the following: if the candidate has failed to find the right answer to either question prior to period  $t$ , then she will attempt  $Q_i^{(t)}$  in that period. On failure, she continues with this policy (i.e. attempts  $Q_i^{(t+1)}$  in period  $t + 1$ ); if she succeeds, she only attempts the remaining question in the residual time.

We begin by showing that in the case of two questions, the particular order in which the questions are attempted does not matter: the value of a policy is determined only by the *number* of attempts on each question. For example, if the deadline is say  $T = 5$ , both policies  $BABAA$  and  $AAABB$  would yield the same value.

**Proposition 1** (*Invariance Principle*) *In an exam with two questions  $A$  and  $B$ , all policies that contain the same number of attempts on  $A$  have the same value.*

**Proof.** See Appendix. ■

The proof of the invariance principle can be seen from figure 1. In the figure, the x- and y-axes denote the number of attempts on questions  $A$  and  $B$  respectively, while the dark and dashed lines denote policies  $AABAABBA$  and  $AABAABAB$  respectively (with a deadline of  $T = 8$ ). Both have the same number of attempts on both questions (5 on question  $A$  and 3 on  $B$ ), but differ in their order of attempts. Now, if the pair of true answers  $(A_i, B_j)$  to the two questions lies in the

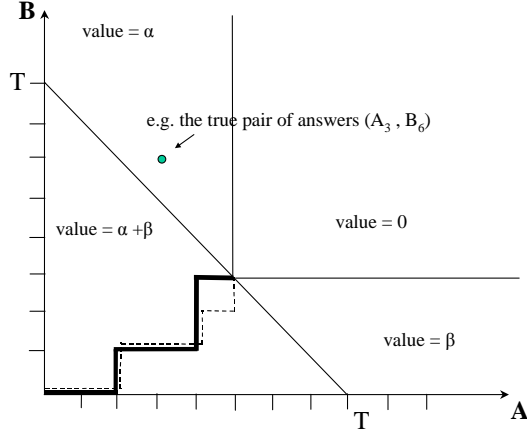


Figure 1: The Invariance Principle

region  $i + j \leq T$ , then either of the above two policies would lead to success in both questions and realize the value  $\alpha + \beta$ . Likewise, if the pair of true answers lie in the region  $i > 5$  and  $j > 3$ , then neither policy will lead to success in either question and so the realized value from both policies is 0. Finally, suppose  $i \leq 5$  while that for  $B$  is  $B_j$  with  $j > T - i$ ; now, pursuing either policy leads to success in  $A$  but failure in  $B$  and the realized value is  $\alpha$ . Similarly, both policies realize the value  $\beta$  when  $j \leq 3$  and  $i > T - j$ . This means that the policies  $AABAABBA$  and  $AABAABAB$ , which have the same number of attempts on questions  $A$  and  $B$ , produce the same outcome, and *a fortiori* the same value, in every possible state  $(A_i, B_j)$ .

This invariance principle greatly simplifies the problem of determining the optimal policy in the case of two questions: one requires to only determine the optimal number of attempts on one of the questions, not their order. So, from now on, we will generally refer to a policy only by the number of attempts on question  $A$ . Let us denote by  $V(m)$  the value from pursuing a policy involving  $m$  attempts on  $A$  (and  $T - m$  attempts on  $B$ );  $V(m)$  is given by:

$$V(m) = \sum_{i=1}^m a_i [\alpha + \beta \sum_{j=1}^{T-i} b_j] + (1 - \sum_{i=1}^m a_i) \sum_{j=1}^{T-m} b_j [\beta + \alpha \sum_{k=m+1}^{T-j} \frac{a_k}{(1 - \sum_{i=1}^m a_i)}] \quad (3)$$

The value function consists of two parts:  $[\alpha + \beta \sum_{j=1}^{T-i} b_j]$  is the expected value conditioned on having correctly answered  $A$  on the  $i$ -th attempt, and then using the remaining time to answer  $B$ ;

if the candidate does not succeed in answering  $A$  in the first  $m$  attempts, then she switches to  $B$ , where  $[\beta + \alpha \sum_{k=m+1}^{T-j} \frac{a_k}{(1 - \sum_{i=1}^m a_i)}]$  is the expected value conditioned on getting  $B$  right on the  $j$ -th attempt and then using the residual time to attempt  $A$  (with the updated probabilities).

It is useful to rewrite this expression in terms of the overall probabilities of correctly answering each question:

$$V(m) = \alpha \left[ \sum_{i=1}^m a_i + \sum_{i=m+1}^{T-1} a_i \left( \sum_{j=1}^{T-i} b_j \right) \right] + \beta \left[ \sum_{j=1}^{T-m} b_j + \sum_{j=T-m+1}^{T-1} b_j \left( \sum_{i=1}^{T-j} a_i \right) \right]$$

Thus this policy may uncover the correct answer to question  $A$  in two ways: first, if this answer lies within the first  $m$  possibilities; if not, then it will be uncovered only if the candidate finds the right answer to  $B$  and then has sufficient time left to get to it. The marginal effect of increasing the number of attempts on question  $A$  from  $m$  to  $m+1$  is now clear: the gain on  $A$  is the increased probability of uncovering  $A_{m+1}$ , while the loss on  $B$  is the decreased chance of uncovering  $B_{T-m}$  :

$$V(m+1) - V(m) = \alpha [a_{m+1} - a_{m+1} \sum_{j=1}^{T-(m+1)} b_j] - \beta [b_{T-m} - b_{T-m} \sum_{i=1}^m a_i] \quad (4)$$

The marginal effects can be added up to obtain the global effect of raising the number of attempts on question  $A$  from 0 to  $m$  :  $V(m) - V(0) =$

$$\begin{aligned} \sum_{i=0}^{m-1} \{V(i+1) - V(i)\} &= \sum_{i=0}^{m-1} \left\{ \alpha a_{i+1} [1 - F_B(T-m) - \sum_{j=T-(m-1)}^{T-(i+1)} b_j] - \beta b_{T-i} [1 - F_A(i)] \right\} \\ &= \alpha F_A(m) [1 - F_B(T-m)] - \beta [F_B(T) - F_B(T-m)] - (\alpha - \beta) \sum_{k=1}^{m-1} F_A(k) b_{T-k} \end{aligned}$$

where  $F_A(t) = \sum_{i=1}^t a_i$  and  $F_B(t) = \sum_{j=1}^t b_j$  are the cumulative distributions for  $A$  and  $B$  respectively. Therefore the optimal number of attempts on question  $A$  is the value of  $m$  that maximizes the above expression.

When the weights on the two questions,  $\alpha$  and  $\beta$ , are the same (and equal to 1), the expression simplifies to:

$$V(m) - V(0) = F_A(m) - [F_B(T) - F_B(T-m)] - F_A(m)F_B(T-m) \quad (5)$$

In the rest of the analysis, we will assume that the marks on both questions are equal i.e.  $\alpha = \beta = 1$ . We feel that putting unequal weights on the two questions complicates the problem without adding much extra insight.

### 3.2 Characterization of Optimal Policies

In this section, we try to characterize properties of the optimal policy under certain probability distributions. We begin with the case when the candidate is sure of one of the questions; the following proposition shows that it is always optimal to attempt the “sure-shot”.

**Proposition 2** *In an exam with two equally weighted questions, if the candidate is sure of correctly answering a question within the deadline, then attempting that question until she succeeds on it is an optimal policy.*

**Proof.** Suppose the candidate is sure of answering question  $A$  within the deadline i.e.  $F_A(T) = 1$ . Then from (5),  $V(T) - V(0) = 1 - F_B(T)$ .

Since  $F_A(m)$  and  $F_B(T - m)$  are both non-negative and no greater than 1, it follows that:

$$V(m) - V(0) + F_B(T) = F_A(m) + F_B(T - m) - F_A(m)F_B(T - m) \leq 1$$

Hence for every  $m \leq T$  we have,  $V(m) - V(0) \leq 1 - F_B(T) = V(T) - V(0)$ , i.e. attempting  $A$  until success is optimal. ■

The intuition for this result is simple. Suppose the true answer to question  $A$  is  $A_i$  and that for  $B$  is  $B_j$ . When one of the questions, say  $A$ , is a “sure-shot” (i.e. can be surely answered within the given deadline), then  $i \leq T$  and the policy of “attempt  $A$  until success” uncovers it with probability one. Any other policy  $\mathcal{P}$  of the form “attempt  $A$   $k$  times” does as well as the policy of focusing first on  $A$  if  $i + j \leq T$  (in this case, all policies will uncover answers to both the questions). However, if  $i + j$  exceeds  $T$ , then only one of the questions can be correctly answered. In this case, whereas, the policy of focusing first on  $A$  answers it correctly for sure, policy  $\mathcal{P}$  may not have a sure shot either at question  $A$  or question  $B$ .

This “sure-shot” principle can lead to situations that may seem non-intuitive at first glance. Consider the example in the Introduction: in an exam with two periods, the candidate has 5 possibilities for question  $A$ , among which she is 99% sure that  $A_1$  is the correct answer (with the remaining 1% equally distributed over the other four possibilities). For question  $B$ , she has two suspects, both of which she considers equally likely. Although there is the possibility  $A_1$  with a 99% chance of success, it is optimal for her to start by attempting  $B$  (and sticking to it if she fails),

since she is sure of getting it correct within the deadline. This example also shows the effect of the deadline in play (albeit in a trivial way) in determining the optimal policy: if the deadline was of one period only, she would attempt question  $A$ ; but with a deadline of two periods, she attempts  $B$  until success.

The following proposition characterizes optimal policies in the case of monotonic hazard rates.

**Proposition 3** *Assume that in an exam with two equally weighted questions, the candidate is not sure of being able to answer either question correctly within the deadline, i.e.  $F_A(T) < 1$  and  $F_B(T) < 1$ .*

(i) *Suppose the hazard rates on both questions are non-decreasing at all points i.e. for  $k \in \{A, B\}$ ,  $h_{m+1}^k \geq h_m^k$  for  $m = 1, \dots, T-1$ . Now, if  $F_A(T) \geq F_B(T)$ , then “attempting  $A$  until success” is an optimal policy; if not, then “attempting  $B$  until success” is an optimal policy.*

(ii) *Suppose the hazard rates on both questions are decreasing at all points. Then any optimal policy will consist of non-zero attempts on both questions unless  $h_1^A \leq h_T^B$  or  $h_T^A \geq h_1^B$ . If  $h_1^A \leq h_T^B$ , then zero attempts on  $A$  is an optimal policy; if  $h_T^A \geq h_1^B$ , then zero attempts on  $B$  is optimal.*

(iii) *In general, the optimal number of attempts,  $m$ , on question  $A$  minimize  $\prod_{i=1}^m (1-h_i^A) \prod_{i=1}^{T-m} (1-h_i^B)$*

**Proof.** Using (4), we can write the marginal effect of increasing the number of attempts on  $A$ :

$$V(m+1) - V(m) = [1 - F_A(m)][1 - F_B(T - m - 1)](h_{m+1}^A - h_{T-m}^B) \quad (6)$$

Thus the sign of  $V(m+1) - V(m)$  depends on whether  $h_{m+1}^A \geq h_{T-m}^B$ .

(i) If the hazard rates on both questions are non-decreasing,  $h_{m+1}^A - h_{T-m}^B$  is non-decreasing in  $m$ . Hence, either  $V(m+1) - V(m) > 0$  for all  $m$ , or it is negative for all  $m$ , or it is initially negative and then changes sign somewhere in the middle. Thus, in this case, the optimal policy will involve either 0 or  $T$  attempts on question  $A$ . From (5),  $V(T) - V(0) = F_A(T) - F_B(T)$ . Thus, according as  $F_A(T) \geq F_B(T)$ , “all  $A$ ” or “all  $B$ ” is optimal.

(ii) If the hazard rates on both questions are decreasing, the sign of  $h_{m+1}^A - h_{T-m}^B$  is non-increasing in  $m$ . Now, if  $h_1^A \leq h_T^B$ , then  $V(1) - V(0) \leq 0$ . This implies that  $V(m+1) - V(m) \leq 0$  for all  $m$ ; therefore, in that case, the optimal policy consists of 0 attempts on question  $A$ . Similarly, if  $h_T^A \geq h_1^B$ ,  $V(m+1) - V(m) \geq 0$  for all  $m$ , and the optimal policy consists of  $T$  attempts on  $A$ .

If neither of these two conditions holds, then  $V(m+1) - V(m)$  must change sign for some  $m$  in between 0 and  $T$ . Then the optimal policy will be given by one of the  $m$ 's for which it does so.

(iii) From the reciprocal relation (2),  $F_A(m) = \sum_{i=1}^m a_i = 1 - \prod_{i=1}^m g_i^A$ . Thus (5) can be written as:

$$V(m) - V(0) = \prod_{i=1}^T g_i^B - \prod_{i=1}^m g_i^A \prod_{i=1}^{T-m} g_i^B \quad (7)$$

Hence for a given  $T$ , the optimal policy requires minimization of  $\prod_{i=1}^m (1 - h_i^A) \prod_{i=1}^{T-m} (1 - h_i^B)$ . ■

The above proposition thus characterizes optimal policies in the case where hazard rates on both questions are either increasing or decreasing. In the case of increasing hazard rates, the intuition for a corner-solution is quite clear: if one starts with say question  $A$ , then a failure in the attempt only makes the distribution on  $A$  more attractive. Thus in this case, if one starts with  $A$ , one continues with it in the case of failure.<sup>2</sup>

Part (iii) of the above proposition illustrates an important issue in determination of the optimal policy. It shows that, in general, the entire distribution for each question matters. Suppose, there is a ‘gem’ hidden somewhere in the distribution i.e. a possibility with a very high hazard rate, then one may be willing to go through some ‘dirt’ (i.e. some low hazard rates) to reach it. Thus, without monotonicity in the hazard rates for the two questions, the objective function is not necessarily concave or convex and therefore the local maxima need not be global. Further both questions matter (see also footnote 2); in contrast to the standard “bandit problem”, the value of each question (or “arm”) cannot be computed independently of the other arm.

## 4 Behavior under Deadline versus Discounting

In this section, we begin by examining the type of policies that can be rationalized under deadlines. The specific question we ask is that given a sequence of deadlines and associated search policies, are these policies justifiable as being optimal under some set of beliefs on the two projects? Of course, if both sets of answers,  $A$  and  $B$ , have the same constant hazard-rate, then (from the previous

---

<sup>2</sup>This is not true if the hazard rate on  $B$  is not increasing at the same time. For example if  $T = 2$ ,  $h_1^A = 0.8$ ,  $h_2^A = 0.9$ , and  $h_1^B = 0.95$ ,  $h_2^B = 0.7$ , then the optimal policy is  $A_1B_1$ ; i.e. even if one starts with  $A$ , it is optimal to shift to  $B$  in the event of failure.

section) all policies give the same value and are trivially rationalizable. Hence, for rationalizability, we require the prescribed policies to be *strictly* optimal.

#### 4.1 Rationalizability of Behavior in the Deadline Model

Recall that in the case of two questions, a policy can be characterized solely by the number of attempts on either question. For a deadline  $T$ , let us denote by  $f_A(T)$  the number of attempts on question  $A$ . Given a function  $f_A(T)$ , we seek to determine if there exists a set of beliefs on questions  $A$  and  $B$ , such that the unique optimal policy under a deadline  $T$  consists of  $f_A(T)$  attempts on question  $A$ .

We provide a partial answer to the question by rationalizing a wide class of policies: (i) policies in which the candidate either attempts only one question or the other depending on the deadline i.e. functions of the form  $f_A(T) \in \{0, T\}$ , and (ii) policies in which the number of attempts on both questions are non-decreasing in the length of the deadline i.e. functions with the property that both  $f_A(T)$  and  $f_B(T) = T - f_A(T)$  are non-decreasing in  $T$ . The proof is constructive in both cases: for a given function  $f_A$ , we construct probability distributions on the two questions by constructing the corresponding  $g$ -functions ( $g^k(t) = 1 - h_t^k$ ,  $k = A, B$ ) to support  $f_A(T)$  as the unique optimal policy for a deadline of  $T$  periods.

**Proposition 4** *Suppose  $f_A(T) \in \{0, T\}$  for all  $T = 1, 2, \dots$ . Then  $f_A$  is rationalizable.*

**Proof.** See Appendix. ■

Although the proof explicitly constructs a system of beliefs to justify any behavioral function  $f_A(T)$  of the above form as being optimal, it may be instructive to study the constructed probabilities in the case of extreme oscillation: say, the optimal policy when the deadline is odd (i.e. for  $T = 1, 3, 5, \dots$ ) is to “attempt  $A$  until success”, and when the deadline is even is to “attempt  $B$  until success”. The following sequence of  $g$ -functions will rationalize this behavior as being optimal:

$t$	1	2	3	4	5	6	7
$g_t^A$	$s^2$	$(1 + \epsilon)s^3$	$s^6$	$(1 + \epsilon)s^7$	$s^{10}$	$(1 + \epsilon)s^{11}$	$s^{14}$
$g_t^B$	$s$	$s^4$	$(1 + \epsilon)s^5$	$s^8$	$(1 + \epsilon)s^9$	$s^{12}$	$(1 + \epsilon)s^{13}$

where it is assumed that  $s \in (0, 1/2)$ ,  $\epsilon > 0$  and  $(1 + \epsilon)s < 1/2$ .

Since the hazard rates on both questions are increasing, in this case the optimal policy (from proposition 3) for any deadline will consist of attempting one of the questions until success. Comparing  $\prod_{i=1}^T g_i^A$  to  $\prod_{i=1}^T g_i^B$ , we see that the ratio is  $s < 1$  when  $T$  is odd; hence with an odd deadline, it is optimal to try only question  $A$ . When  $T$  is even, this ratio is  $1 + \varepsilon$ , thus justifying the policy of attempting only  $B$  for such deadlines.

Next, we rationalize the class of policies in which the number of attempts on either question does not decrease with an increase in the deadline  $T$ .

**Proposition 5** *Suppose  $f_A(T)$  and  $T - f_A(T)$  are both non-decreasing in  $T$ . Then  $f_A$  is rationalizable.*

**Proof.** See Appendix. ■

The proof itself is constructive, and it may be useful to see the construction for a particular function  $f_A(T)$ . Suppose we wish to rationalize the policy “attempt  $A$  half the time (rounded to the nearest higher integer)”; thus, in this case, the  $f_A$  function (up to  $T = 8$ ) is as follows:

<b>T</b>	1	2	3	4	5	6	7	8
<b>f<sub>A</sub>(T)</b>	1	1	2	2	3	3	4	4

The following  $g$ -functions on  $A$  and  $B$  rationalize this behavior:

$$\begin{aligned} \{g_1^A &= x_1, g_2^A = x_3, g_3^A = x_5, g_4^A = x_7, \dots\} \\ \{g_1^B &= x_2, g_2^B = x_4, g_3^B = x_6, g_4^B = x_8, \dots\} \end{aligned}$$

where  $\{x_n\}_{n \in \mathbb{N}}$  is given by the increasing sequence  $x_n = s - \epsilon^n$  with  $0 < \epsilon < \frac{1}{2}$  and  $\epsilon + \frac{1}{2} < s < 1$ .

Recall that from proposition 3, the optimal number of attempts,  $m$ , on question  $A$  minimizes  $\prod_{i=1}^m g_i^A \prod_{i=1}^{T-m} g_i^B$ . Since  $x_n$  is an increasing sequence, when the deadline is  $T = 7$  for example, to minimize the product of the  $g$ -functions, we include  $g_1^A = x_1, g_1^B = x_2, g_2^A = x_3, \dots, g_3^B = x_6$ , and  $g_4^A = x_7$ , but not  $g_4^B = x_8$  which exceeds  $g_4^A$ . Thus for  $T = 7$ , it is optimal to attempt question  $A$  exactly four times. Essentially, the above program takes an increasing sequence  $\{x_n\}_{n \in \mathbb{N}}$  and distributes it sequentially between the two questions depending on whether at each stage an increase in the deadline calls for an increase (by one) in the number of attempts on  $A$  or  $B$ .



Although we have thus been able to rationalize a wide class of policies, not all sequences of behavior are justifiable as being optimal under any set of beliefs on the two questions. For example, it can be shown that if  $f_A(s) = s$  and  $f_A(t) = 0$ , then  $f_A(s+t) \neq t$ . In other words, if it is optimal to try question  $A$   $s$  times when the deadline is  $s$  and to try question  $B$   $t$  times when the deadline is  $t$ , then it cannot be optimal to attempt  $A$   $t$  times and  $B$   $s$  times when the deadline is  $s+t$ ; in such a situation, the policy of attempting  $A$   $s$  times (and  $B$   $t$  times) does better.<sup>3</sup> Hence the question of what is the precise set of policies that are rationalizable under deadlines remains open.

However, as we show below, even this limited set of rationalizable policies under deadlines cannot be justified as being optimal under traditional geometric discounting.

## 4.2 Rationalizability of Behavior in a Model of Geometric Discounting

In the previous section, we have shown that several forms of sensitivity of behavior to changes in deadlines can be justified as being optimal under suitable distributions of beliefs on the two questions. Now, one may also view deadlines as a special form of discounting: no discounting till the deadline and complete discounting after it. Suppose instead one were to consider the same problem of two-dimensional search, but with an infinite deadline and in which the present value of obtaining a reward in the  $(t+1)$ -th period is given by  $\delta^t$  ( $\delta < 1$ ). In this more standard infinite-horizon geometric discounting framework, can a similar sensitivity of behavior to changes in the discount factor  $\delta$  be also rationalized? Are changes in the discount factor behaviorally similar to changes in the deadline? Following in the spirit of the previous section, here we ask that given a set of prescribed policies, does there exist a sequence of discount factors  $\{\delta_k\}$  going to unity for which the given policies are optimal?

Extreme oscillatory behavior, such as attempting only question  $A$  if the deadline is odd and attempting only  $B$  if the deadline is even was shown to be rationalizable in the case of deadlines. Under geometric discounting, we examine whether it is possible for the optimal policy to display this extreme form of sensitivity to slight changes in the discount factor. As the next proposition shows, the answer is negative.

---

<sup>3</sup>From (7), the fact that  $f_A(s) = s$  implies  $\prod_{i=1}^s g_i^B > \prod_{i=1}^s g_i^A$ ; similarly,  $f_A(t) = 0$  implies  $\prod_{i=1}^t g_i^B < \prod_{i=1}^t g_i^A$ . Together, this means that when the deadline is  $s+t$ ,  $V(s) - V(0) = \prod_{i=1}^{s+t} g_i^B - \prod_{i=1}^s g_i^A \prod_{i=1}^t g_i^B > \prod_{i=1}^{s+t} g_i^B - \prod_{i=1}^t g_i^A \prod_{i=1}^s g_i^B = V(t) - V(0)$  i.e. the policy of attempting  $A$   $t$  times is inoptimal.

For brevity, we will refer to the policies “attempt  $A$  until success” and “attempt  $B$  until success” in the model of geometric discounting as the “all  $A$ ” and “all  $B$ ” policies respectively.

**Proposition 6** *In the model with geometric discounting, for any given distribution of beliefs on  $\mathbf{A}$  and  $\mathbf{B}$ , there do not exist two sequences of discount factors going to unity such that for one sequence, the strictly optimal policy is “all  $A$ ” while for the other sequence, the strictly optimal policy is “all  $B$ ”.*

**Proof.** Let us define  $A_i(\delta) = \frac{(a_{i+1} + \delta a_{i+2} + \delta^2 a_{i+3} + \dots)}{(a_{i+1} + a_{i+2} + a_{i+3} + \dots)}$ . In the event that question  $B$  has been successfully answered, and  $i$  unsuccessful attempts have been made on question  $A$ , this then gives the discounted value of the (updated) distribution  $\mathbf{A}$ . Similarly,  $B_i(\delta) = \frac{(b_{i+1} + \delta b_{i+2} + \delta^2 b_{i+3} + \dots)}{(b_{i+1} + b_{i+2} + b_{i+3} + \dots)}$ .

The overall value from following the “all  $A$ ” policy is given by:

$$\begin{aligned} V_A(\delta) &= a_1(1 + \delta B_0(\delta)) + \delta(1 - a_1) \left\{ \frac{a_2}{1 - a_1} (1 + \delta B_0(\delta)) + \dots \right\} \\ &= A_0(\delta)(1 + \delta B_0(\delta)) \end{aligned}$$

Similarly, the value from following “all  $B$ ” is given by:  $V_B(\delta) = B_0(\delta)(1 + \delta A_0(\delta))$ .

We do the proof by contradiction. Suppose there exist two sequences of discount factors going to unity such that “all  $A$ ” is strictly optimal for one sequence and “all  $B$ ” for the other.

For this conjecture to hold,  $V_A(\delta) - V_B(\delta)$  must have an infinite number of roots  $\delta$  in some open interval  $(\delta^*, 1]$ . By theorem (see for example Shilov (1973)), an analytic function cannot have an infinite number of roots, unless it is the constant zero function. We establish contradiction by showing that if “all  $A$ ” and “all  $B$ ” are optimal along two sequences of discount factors going to unity, then the function  $V_A(\delta) - V_B(\delta)$  must be analytic at 1 and hence must be the zero function; but then “all  $A$ ” (or “all  $B$ ”) cannot be strictly optimal.

The proof is in two steps. We first show that if  $\liminf_n h_n^A = 0$  ( $\liminf_n h_n^B = 0$  respectively), then “all  $A$ ” (“all  $B$ ” resp.) cannot be optimal for any sequence of discount factors going to unity, i.e. if the distribution on  $A$  becomes sufficiently pessimistic after some stage, it cannot be optimal to ignore  $B$  after that. We then show that if  $\liminf_n h_n^A > 0$  and  $\liminf_n h_n^B > 0$ , then the function  $V_A(\delta) - V_B(\delta)$  must be analytic at 1, and thus cannot oscillate infinitely often.

**Step 1:** Suppose  $\liminf_n h_n^A = 0$ . Then the policy “all  $A$ ” cannot be optimal along a sequence of discount factors going to 1.

Proof: We will use the following claim in our proof.

**Claim:** Consider a sequence of discount factors  $\{\delta_k\}$  going to unity ( $\delta_k < 1$ ) along which “all  $A$ ” is the optimum policy. Then  $b_1 E_{A_n}(T) \leq 1$  for all  $n$ , where  $E_{A_n}(T)$  is the expected time to success in question  $A$  using the “all  $A$ ” policy, after having failed at it  $n$  times.

Proof of claim: Let us denote the updated probabilities on (the remaining suspects in)  $A$  after  $n$  failures by  $a_1^n, a_2^n, \dots$  i.e.  $a_i^n = \frac{a_{n+i}}{1 - F_A(n)}$ . Now, the payoff from following an “all  $A$ ” policy from this stage onwards is  $V_A^n(\delta) = A_n(\delta)(1 + \delta B_0(\delta))$ . Note that, question  $B$  has not been attempted so far and so the probability distribution on it is unchanged.

Consider the alternative policy “once  $B$  and then all  $A$ ”. The value from this policy is:

$$\begin{aligned} W^n(\delta) &= b_1(1 + \delta A_n(\delta)) + \delta(1 - b_1)A_n(\delta)(1 + \delta B_1(\delta)) \\ &= b_1 + \delta(1 - b_1)A_n(\delta) + \delta A_n(\delta)B_0(\delta) \end{aligned}$$

Now, for “all  $A$ ” to be optimal, we require that  $V_A^n(\delta) - W^n(\delta) > 0$  for all  $n$ , for all  $\delta$  along the sequence  $\{\delta_k\}$ . Let us denote by  $f(\delta) = V_A^n(\delta) - W^n(\delta) = A_n(\delta)(1 - \delta(1 - b_1)) - b_1$ . Now,

$$\begin{aligned} \lim_{\delta \rightarrow 1^-} \frac{d}{d\delta} f(\delta) &= (a_2^n + 2a_3^n + \dots)b_1 - (1 - b_1) \\ &= b_1 E_{A_n}(T) - 1 \end{aligned}$$

Note that  $f(1) = 0$ , since at  $\delta = 1$  both policies give the same value. Thus, for “all  $A$ ” to be optimal along the sequence  $\{\delta_k\} \rightarrow 1$ , we require that  $\lim_{\delta \rightarrow 1^-} \frac{d}{d\delta} f(\delta) \leq 0$ , i.e. that  $b_1 E_{A_n}(T) \leq 1$ .  $\square$

Continuing with the proof of step 1, let us now compute the expected time of success in question  $A$  using the “all  $A$ ” policy, after having failed at it  $n$  times:

$$\begin{aligned} E_{A_n}(T) &= \frac{a_{n+1}}{1 - F_A(n)} + 2\frac{a_{n+2}}{1 - F_A(n)} + \dots + (m+1)\frac{a_{n+m+1}}{1 - F_A(n)} + \dots \\ &> \frac{a_{n+1}}{1 - F_A(n)} + 2\frac{a_{n+2}}{1 - F_A(n)} + \dots + m\frac{a_{n+m}}{1 - F_A(n)} \\ &\quad + (m+1)\left[1 - \frac{a_{n+1}}{1 - F_A(n)} + \dots + \frac{a_{n+m}}{1 - F_A(n)}\right] \\ &= m+1 - \left[m\frac{a_{n+1}}{1 - F_A(n)} + (m-1)\frac{a_{n+2}}{1 - F_A(n)} + \dots + \frac{a_{n+m}}{1 - F_A(n)}\right] \\ &\geq m+1 - \frac{m(m+1)}{2} h_{n+1}^A \end{aligned} \tag{8}$$

with the last inequality following from the fact that  $\frac{a_{n+i}}{1-F_A(n)} \leq \frac{a_{n+1}}{1-F_A(n)} = h_{n+1}^A$  for all  $i > 1$ .

Since the above inequality holds for all  $m$ , choose  $m$  large enough so that  $mb_1 > 1$ ; then choose  $\varepsilon > 0$  such that  $1 - \frac{m(m+1)}{2}\varepsilon > 0$ .

Since  $\liminf_n h_n^A = 0$ , there exists a  $p$  such that  $h_{p+1}^A < \varepsilon$ .

Now from (8),  $E_{A_p}(T) \geq m + 1 - \frac{m(m+1)}{2}h_{p+1}^A > m + 1 - \frac{m(m+1)}{2}\varepsilon > m$ . Hence,  $b_1 E_{A_p}(T) > b_1 m > 1$ . Thus if  $\liminf_n h_n^A = 0$ , then from the above claim, “all  $A$ ” cannot be optimal after  $p$  failures on  $A$ .

Similarly for “all  $B$ ” to be optimal, we require that  $\liminf_n h_n^B > 0$ .  $\square$

**Step 2:** If  $\liminf_n h_n^A > 0$  and  $\liminf_n h_n^B > 0$ , then the function  $V_A(\delta) - V_B(\delta) = \sum_{i=1}^{\infty} \delta^{i-1}(a_i - b_i)$  must be analytic at 1.

Proof: Suppose  $\liminf_n h_n^A = \lambda > 0$ . Take an  $\varepsilon > 0$  such that  $\eta = \lambda - \varepsilon > 0$ . Then, by definition, there exists  $M$  such that for all  $n > M$ ,  $h_n^A > \eta$ .

Now from (2),  $a_{n+1} = (1 - h_1^A)(1 - h_2^A)\dots(1 - h_n^A)h_{n+1}^A$ .

Therefore, for all  $n > M$ ,  $a_{n+1} < K(1 - \eta)^{n-M}$  where  $K$  is a positive constant.

Thus  $|a_{n+1}|^{1/n} < (1 - \eta)\left\{\frac{K}{(1-\eta)^M}\right\}^{1/n}$ . Hence  $\limsup_n |a_{n+1}|^{1/n} < 1$ , which is the condition for  $\sum_{i=1}^{\infty} \delta^{i-1}a_i$  to be analytic at 1.

Similarly, if  $\liminf_n h_n^B > 0$ , then  $\sum_{i=1}^{\infty} \delta^{i-1}b_i$  is analytic at 1.

This then establishes analyticity of  $V_A(\delta) - V_B(\delta)$ .  $\square$

From the discussion at the start, this then completes the proof of the proposition.  $\blacksquare$

It may be useful to see why the construction (in proposition 4) that was used for rationalizing extreme oscillation in the case of deadlines does not work for geometric discounting. The probabilities there for rationalizing “all  $A$ ” for odd deadlines and “all  $B$ ” for even deadlines can be shown to have the following general form for  $n > 2$ :

$$\begin{aligned} a_n &= (1 - s^{2n})s^{(n-1)(2n-1)/2}(1 + \varepsilon)^{(n-1)/2} && \text{if } n \text{ is odd} \\ &= (1 - (1 + \varepsilon)s^{2n-1})s^{(2n^2-3n+2)/2}(1 + \varepsilon)^{(n-2)/2} && \text{if } n \text{ is even} \\ b_n &= (1 - (1 + \varepsilon)s^{2n-1})s^{(2n^2-3n+1)/2}(1 + \varepsilon)^{(n-3)/2} && \text{if } n \text{ is odd} \\ &= (1 - s^{2n})s^{(2n^2-3n)/2}(1 + \varepsilon)^{(n-2)/2} && \text{if } n \text{ is even} \end{aligned}$$

Here,  $a_n - b_n$  contains terms in which the exponent on  $s$  is of a higher power than  $n$ . Hence, in this case  $a_n - b_n$  goes to zero at a rate faster than  $1/n$ . This makes the difference  $V_A(\delta) - V_B(\delta) = \sum_{i=1}^{\infty} \delta^{i-1}(a_i - b_i)$  from following policies “all  $A$ ” and “all  $B$ ” analytic at 1. Such a function cannot oscillate infinitely often near 1, thereby ruling out extreme oscillation in the optimal policy with geometric discounting.

Although the above proposition considers only the particular policies “all  $A$ ” and “all  $B$ ”, the logic of the proof can be used to rule out infinite oscillations between any finite set of policies provided that  $\liminf_n h_n^A > 0$  and  $\liminf_n h_n^B > 0$ . The difference in value between any two policies can be written as a polynomial (possibly of infinite degree) in the discount factor  $\delta$ , and step 2 of the above proof can then be used to show that this difference is analytic at 1 (under the condition that  $h_n^A$  and  $h_n^B$  are both bounded away from 0). This implies that for a range of  $\delta$  close to 1, either one policy always does better than the other or both give the same value. Comparing among a finite set of policies, this implies that for any set of beliefs with  $\liminf_n h_n^A > 0$  and  $\liminf_n h_n^B > 0$ , there exists a  $\delta^*$  close to unity such that for all  $\delta \in (\delta^*, 1]$  the set of best policies (among this finite set) is the same.

Thus, even comparing the case of a very long deadline with that of geometric discounting with discount factors close to unity, we see that there is much less variation in behavior due to changes in the discount factor than due to changes in the deadline.

## 5 Additional Considerations

### 5.1 Three or More Questions: Breakdown of the Invariance Principle

In the case of two questions, the order-invariance property allowed a simple characterization of the optimal policy only in terms of the number of attempts on each question. The invariance principle does not hold when the number of questions is greater than two. It is easy to show that in such cases not all policies are order-invariant. More importantly, even the optimal policy need not be order-invariant. We illustrate this by means of an example.

**Example 1** *Consider an exam with 3 questions  $A$ ,  $B$  and  $C$ , and a deadline of 5 periods. The subjective probability distribution for answers to the three questions is represented by the hazard-*

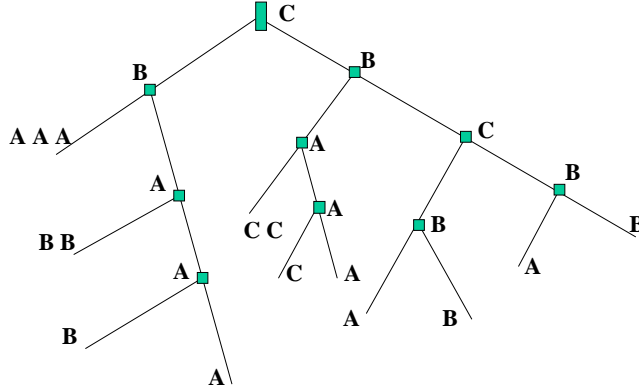


Figure 2: Optimal policy for Example 1. On success, go left; on failure, go right.

rates in the following table:

	$h_1$	$h_2$	$h_3$	$h_4$	$h_5$
A	0.15	0.17	0.2	0.01	0.01
B	0.18	0.18	0.18	0.01	0.01
C	0.225	0.18	0.01	0.01	0.01

The marks on all three questions are assumed to be equal to 1.

Here, a policy can be described as a sequence of the following form  $((Q_i^t, E^t))_{t=1}^5$ . This is understood to mean the following: if the candidate has failed in all her attempts prior to a given time  $t$ , then she will attempt  $Q_i^t$  in the current period. If she fails in this attempt too, she will move on to  $Q_i^{t+1}$  in the next period. But if she succeeds (in verifying that  $Q_i$  is the correct answer to  $Q$ ), she will then pursue the policy  $E^t$  that is optimal for the remaining unanswered question(s), with the updated probabilities and in the remaining time.  $E^t$  involves at most two questions and is defined as in the previous sections.

In this example it can be shown that the following is an optimal policy:

$$\begin{aligned} \mathbb{P}^{opt} &\equiv ((C_1, E^1), (B_1, E^2), (C_2, E^3), (B_2, E^4), (B_3, E^5)) \\ \text{with } E^1 &= (B_1, A_1, A_2, A_3) & E^2 &= (A_1, A_2, A_3) \\ E^3 &= (B_2, B_3) & E^4 &= (A_1) \end{aligned}$$

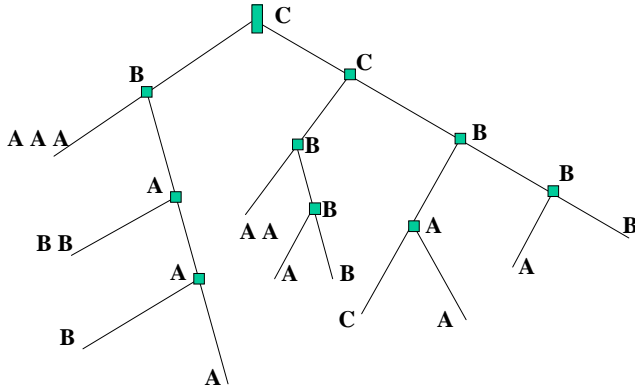


Figure 3: Alternative (with the order of attempts changed) to optimal policy in Example 1. On success, go left; on failure, go right.

This policy is shown in figure 2. The value from following this policy is 0.92495.

This is however not the unique optimal policy for this example. The following policies with associated exit policies at each stage are also optimal: (i)  $(B_1, C_1, B_2, C_2, B_3)$  (ii)  $(B_1, C_1, C_2, B_2, B_3)$  (iii)  $(C_1, B_1, B_2, B_3, C_2)$  (iv)  $(C_1, B_1, B_2, C_2, B_3)$  (v)  $(B_1, C_1, B_2, B_3, C_2)$ .

It may be interesting to note that none of these optimal policies are ‘lumpy’ i.e. it is not the case that the attempts on any question are bunched together.<sup>4</sup> To demonstrate a failure of the order-invariance property, let us consider the policy  $(C_1, C_2, B_1, B_2, B_3)$  (with optimal exit policies at each stage) which has the same number of attempts on each question as the optimal policy  $\mathbb{P}^{opt}$ , but differs in the order of the attempts. The value from this policy is 0.92298, which is clearly less than that from  $\mathbb{P}^{opt}$ .

To understand better the counter example, it may be useful to compute the optimal exit policies in this case. These are:  $E^1 = (B_1, A_1, A_2, A_3)$ ,  $E^2 = (B_1, B_2, B_3)$ ,  $E^3 = (A_1, A_2)$ ,  $E^4 = (A_1)$ ; this policy is shown in figure 3. The important difference with the optimal policy  $\mathbb{P}^{opt}$  occurs for the exit policy  $E^2$ . If the first success occurs in period 2, and one has only attempted question  $C$  so

<sup>4</sup>Hence, even to determine the best policy involving a given number of attempts on each question, it is not sufficient to compare only (the relatively small number of) policies with various permutations of ‘bunches of attempts’ on each question.

far, the distribution on  $A$  and  $B$  is unchanged. With only three periods left, the distribution on  $B$  is better, and the optimal course of action from this state onward is “attempt  $B$  until success”. In the optimal policy shown in figure 2, success in period 2 leaves questions  $A$  and  $C$  unanswered (having already attempted  $C_1$  and failed). Now the distribution on  $A$  contains a relative gem at  $h_3^A = 0.2$  which is only attainable in three periods; hence the optimal policy from this stage on is to attempt  $A$  in all the subsequent periods. If however, one did not find success in period 2, it would not be favorable to go for this gem, as the combined distributions on  $B$  and  $C$  are more favorable. Thus, the exit policy depends quite non-trivially on the question that one is exiting through. It is this feature that plays an important role in the failure of the invariance principle in the case with more than two questions.

If the exam consists only of two questions, the invariance principle (Proposition 1) applies. This means that, at worst, the candidate need only to compute the values  $V(0), V(1), \dots, V(T)$ , and then seek the maximum amongst them. Without this simplifying feature in the optimal policy, how complex is it when the number of questions exceeds two? With regard to our model, we can show that for the set of exams with a fixed number of questions  $n$ , the optimal search policy for a deadline  $T$  can be computed in a number of steps bounded by a polynomial in  $T$  of degree  $n + 1$ . But now as the number of questions  $n$  increases, the degree of the polynomial also grows. The algorithmic complexity of the general optimization problem is likely to be even higher; the computation time may grow exponentially with  $n$ , the number of exam questions.

## 5.2 An extension

In this section, we consider a variant to the model discussed so far. We shall show that for this variant, the order invariance principle holds for any number of questions. In the above model, the candidate’s objective was to maximize her expected total score from the exam. Suppose instead, the candidate is interested in correctly answering only *one* of the questions. Thus she seeks to maximize the probability of answering one question.

When rewards on all the questions are equal, the logic of the invariance principle applies here too. On the margin, a change in the order of attempts  $A_i B_j$  to  $B_j A_i$  will not alter the probability of being able to answer one question. For instance, suppose  $A_i$  is the right answer for question  $A$ ,



while  $B_j$  is not the correct one for  $B$ ; then either policy will uncover the right answer.

In fact, under this new objective function, the invariance principle not only holds for two questions, but for any number of questions as well. As we saw in the counter-example 1 in the previous section, it was the difference in value of different follow-up policies after success on one of the questions that caused breakdown of the invariance principle. However, in the case when the candidate seeks to answer only one of the questions correctly, this issue does not arise. Hence the validity of the invariance principle holds for any number of questions in this variant of the model. Consequently, finding the optimal policy here is much simpler and the analysis is analogous to that for the case of two questions in the original model.

In this modified version, the value from  $m$  attempts on question  $A$  is given by:

$$W(m) = \sum_{i=1}^m a_i + (1 - \sum_{i=1}^m a_i) \left( \sum_{j=1}^{T-m} b_j \right)$$

Now, the marginal value from increasing the attempts on  $A$  from  $m$  to  $m + 1$  is given by:

$$W(m + 1) - W(m) = a_{m+1} \left( 1 - \sum_{j=1}^{T-(m+1)} b_j \right) - b_{T-m} \left( 1 - \sum_{i=1}^m a_i \right)$$

which is identical to the expression for the marginal effect (4) in the earlier case. This similarity is intuitive: any change in value from increasing the attempts on question  $A$  can only occur from those states of nature in which only one of the two questions are answered correctly by either policy. Hence, the marginal effect must be the same whether one is interested in answering only one question or both.

Thus, the analysis in this variant is identical to that for the earlier model. Unlike the previous case, however, here the analysis is also simple when the number of questions is greater than two. In particular, an optimal policy is completely specified by the number of attempts it allocates to each question; the order in which they appear do not matter.

### 5.3 Discussion

Although our analysis in this paper has been primarily decision-theoretic, we can suggest some real-life situations in which the model may be applicable, and where its behavioral implications are economically relevant.

Consider for example, a government that is interested in leaving a legacy, establishing a reputation for being a “doer”, or simply in getting re-elected. It will typically have a finite horizon over which to achieve its goal, usually until the next election. Now, the government may choose to leave its mark on any of a possible number of areas. Suppose for example, health-care, solving an insurgency problem and developmental work are three important problems facing the country and in each year of governance, the government is able to focus its energy on any one of them. With subjective beliefs about its ability to solve any or all of these three problems, the decision problem facing the government is akin to the model in the paper. Changes in the deadline, as given by the length of an electoral cycle, may lead to drastic differences in the policies that get implemented. Similarly, the length of time before a firm’s patent on a product runs out maybe crucial in determining the research strategy that it will adopt.

Deadlines are also encountered within organizations. Projects in firms often have to be completed within given time limits. In recent times, one of the biggest deadlines that firms have had to face is the Year 2000 (Y2K) problem. This was a universal deadline, imposed externally, and quite infrangible. For a firm needing to eliminate the Y2K bug from a number of different applications, the optimal priority with which to proceed is an important decision and would depend on the values of these applications as well as on the probabilities of being able to fix the bug(s) in any or all of them within the deadline. It may also have to decide whether to employ separate employees on the various projects or to employ the same employee to search through the different applications, one at a time. This resource-allocation problem can be addressed by a variant of our basic model.

Consider the behavior of unemployed job seekers in labor markets. In this context, the duration of unemployment benefits has been a widely debated policy issue. It is believed to have important impacts on unemployment rates, on social welfare, and on labor movement across sectors and regions. Search models with deadlines, as the one we have studied in this paper, may be helpful in understanding some of these phenomena by explaining how deadlines may affect the decisions of workers to switch between sectors in searching for jobs. If workers’ beliefs about the chance of finding a job in their ‘parent’ sector differ from their belief about finding employment in other sectors, the deadline imposed by the expiry of unemployment benefits is likely to affect their sectoral search behavior, and consequently their overall probability of leaving unemployment.

Going beyond single-agent decision problems, one may also consider the effects of deadlines when agents strategically interact. On this front, several authors have studied the theory of bargaining when deadlines are introduced. See for example, Fershtman and Seidmann (1993), Ma and Manove (1993). There the main emphasis has been how deadlines can affect the delay in reaching agreement. While confirming the general point that behavior is sensitive to deadlines, the strategic effects exploited in these models are quite different from ours. As we have already noted in the Introduction, while deadlines are a ubiquitous fact of life, the economic reasons behind why and how they are set remain largely unexplored. One paper that attempts to explain how a deadline may arise endogenously in a principal-agent framework is O'Donoghue and Rabin (1999). If the agent has a propensity to procrastinate, and the principal suffers from delay, then they show that the solution may take the form of an incentive contract involving deadlines (in a generalized sense).

In this paper, we have characterized optimal behavior in a multi-dimensional search framework with deadlines and have shown that behavior can be extremely sensitive to changes in the deadline, even under very long deadlines. Furthermore, it can be sensitive in a way that cannot be explained merely as changes in impatience (or more precisely, impatience modelled as geometric discounting.) One may also ask: if a principal were to impose a deadline on an agent, how should the optimal deadline depend on the principal's objective? For example, for an examiner designing an exam that best elicits the candidate's knowledge, as against an employer wishing to maximize the expected return of the employee's efforts, we suspect that the deadlines involved in the two cases would be rather different. These and other issues remain in an understanding of the behavioral and welfare implications of deadlines, and promises to be a fertile area of future research.

## 6 References

- Dubins, L. E. and L. J. Savage. *How to gamble if you must: Inequalities for stochastic processes*. New York: McGraw-Hill, 1965.
- Fershtman, C. and D. J. Seidmann. "Deadline Effects and Inefficient Delay in Bargaining with Endogenous Commitment." *Journal of Economic Theory* 60(2) (1993): 306-321.
- Gittins, J.C. *Multi-Armed Bandit Allocation Indices*. New York: Wiley-Interscience, 1989.
- Ma, C. A. and M. Manove. "Bargaining with Deadlines and Imperfect Player Control." *Econo-*

*metrica*, 61(6) (1993): 1313-1339.

O'Donoghue, T. and M. Rabin. "Incentives for Procrastinators." *Quarterly Journal of Economics*, 114(3) (1999): 769-816.

Shilov, G. E.(English Translation: R. A. Silverman) *Mathematical Analysis, vol. 1*. Cambridge, MA: MIT Press, 1973.

## 7 Appendix: Proofs of Propositions

**Proof of Proposition 1.** If the deadline is  $T = 1$ , there is nothing to prove; thus we assume  $T \geq 2$ . Consider all policies  $\mathbf{P}$  which involve  $n$  attempts on question  $A$ . Let us denote the pair of true answers to the two questions as  $(A_i, B_j)$ . There are four possibilities: (i) if  $(A_i, B_j)$  lie in the region  $i + j \leq T$ , then all policies would lead to success in both questions and realize the value  $\alpha + \beta$ . (ii) if the pair of true answers lie in the region  $i > n$  and  $j > T - n$ , then all policies in  $\mathbf{P}$  will see  $n$  (unsuccessful) attempts on  $A$  and  $T - n$  (unsuccessful) attempts on  $B$ ; none will lead to success in either question and so the realized value is 0. (iii) suppose  $i \leq n$  while  $j > T - i$ ; now, all policies in  $\mathbf{P}$  lead to success in  $A$  but failure in  $B$  and the realized value is  $\alpha$ . (iv) similarly if  $j \leq T - n$  and  $i > T - j$ , all  $\mathbf{P}$  policies realize the value  $\beta$ . We have thus shown that all policies involving  $n$  attempts on  $A$  (irrespective of the order of the attempts) produce the same outcome, and *a fortiori* the same value, in every possible state. ■

**Proof of Proposition 4.** Without loss of generality let us assume that  $f_A(1) = 1$ . We further assume that there does not exist a time-period  $T^*$  such that  $f_A(T) = 0$  for all  $T > T^*$ . As will become clear, the construction below can be easily modified to rationalize such a function.

We will now use the following program to construct recursively the  $g$ -functions for  $A$  and  $B$ .

Take  $s \in (0, 1/2)$  and  $\varepsilon > 0$  such that  $2(1 + \varepsilon) < 1/s$ .

Start with  $g^A(1) = s^2$ ,  $g^B(1) = s$ , and initialize two artificial variables  $j = 2$ , "regime  $A$ " = 1.

For  $t = 2$  to  $\infty$ , do the following recursively:

if  $f_A(t) = t$ ,

(i) if regime  $A = 1$ , increment  $j$  by 1; define  $g^A(t) = g^B(t) = s^j$ ; leave "regime  $A$ " unchanged.

(ii) if regime  $A \neq 1$ , then set regime  $A = 1$ , define  $g^A(t) = s^{j+2}$ ,  $g^B(t) = (1 + \varepsilon)s^{j+1}$ , and increment  $j$  by 2.

else, if  $f_A(t) = 0$ ,

(i) if regime  $A = 0$ , increment  $j$  by 1; define  $g^A(t) = g^B(t) = s^j$ ; leave “regime  $A$ ” unchanged.

(ii) if regime  $A \neq 0$ , then set regime  $A = 0$ , define  $g^A(t) = (1 + \varepsilon)s^{j+1}$ ,  $g^B(t) = s^{j+2}$ , and increment  $j$  by 2.

First, we need to verify that what we have constructed are valid probability distributions and secondly that the corresponding probabilities i.e.  $\{a_i\}_{i=1}^\infty$  and  $\{b_i\}_{i=1}^\infty$  are non-decreasing. Using the reciprocal relation (2), we can reformulate the condition of non-increasing probabilities i.e.  $a_i \geq a_{i+1}$  in terms of  $g_i^A$  as:  $g_{i+1}^A \geq 2 - \frac{1}{g_i^A}$ .

By our choice of  $s$  and  $\varepsilon$ , we have  $2 - \frac{1}{(1+\varepsilon)s^n} < 0$  for all  $n \geq 1$ ; since  $g^A(t)$  is of the form  $Ks^m$  for some  $m$ , with  $K \in \{1, 1 + \varepsilon\}$ , the above condition is clearly satisfied. Secondly,  $\prod_{i=1}^\infty g_i^A = 1 - \sum_{i=1}^\infty a_i$ . Since in our construction,  $\prod_{i=1}^n g_i^A \leq (1 + \varepsilon)^n s^n$ , and  $(1 + \varepsilon)s < 1$ , this implies that  $\prod_{i=1}^\infty g_i^A = 0$ ; consequently  $\sum_{i=1}^\infty a_i = 1$ , and therefore  $\{a_i\}_{i=1}^\infty$  is a valid probability distribution. A similar logic also applies to  $\{b_i\}_{i=1}^\infty$ .

Now, by construction, the hazard rates on both questions are increasing [here either  $g_{i+1}^A = s^2 g_i^A$  or  $g_{i+1}^A = (1 + \varepsilon)s g_i^A$ ; since  $(1 + \varepsilon)s < 1$ ,  $g_{i+1}^A < g_i^A$  i.e.  $h_{i+1}^A = 1 - g_{i+1}^A > 1 - g_i^A = h_i^A$ ; similarly for  $B$ ] hence by proposition 3, the optimal policy will generally involve attempts on only  $A$  or only on  $B$  depending on whether  $\prod_{i=1}^T g_i^A$  is smaller (bigger) than  $\prod_{i=1}^T g_i^B$ . The construction is such that at points of regime switch i.e. say  $f_A(T-1) = T-1$  and  $f_A(T) = 0$ , the exponent on  $s$  in the two products  $\prod_{i=1}^T g_i^A$  and  $\prod_{i=1}^T g_i^B$  are equal. An extra  $(1 + \varepsilon)$  term is introduced into  $g^A(T)$  to make the former product bigger and thus get “all  $B$ ” to be optimal under a deadline of  $T$ .

In general, when regime  $A = 1$  (this occurs when  $f_A(t) = t$ ), the ratio of  $\prod_{i=1}^T g_i^A$  to  $\prod_{i=1}^T g_i^B$  is  $s < 1$ , and thus “all  $A$ ” is optimal as required; when regime  $A = 0$  (this is when  $f_A(t) = 0$ ) the same ratio is  $1 + \varepsilon$ , and then “all  $B$ ” is optimal. ■

**Proof of Proposition 5.** Without loss of generality let us assume that  $f_A(1) = 1$ .

Define the increasing sequence  $x_n = s - \varepsilon^n$  with  $0 < \varepsilon < \frac{1}{2}$  and  $\varepsilon + \frac{1}{2} < s < 1$ .

We will now use the following program to construct recursively the  $g$ -functions for  $A$  and  $B$ .

Start with  $g^A(t) = g^B(t) = s$  for all  $t$ , and initialize four artificial variables  $j = 1$ ,  $n_A = 1$ ,  $n_B = 0$ , “regime  $A$ ” = 1. Redefine  $g^A(n_A) = x_1$ .

For  $t = 2$  to  $\infty$ , do the following recursively:

if  $f_A(t) = f_A(t - 1)$ ,

(i) increment  $n_B$  by 1

(ii) if regime  $A = 1$ , then set regime  $A = 0$  and increment  $j$  by 1

(iii) redefine  $g^B(n_B) = x_j$ .

else, if  $f_A(t) \neq f_A(t - 1)$ ,

(i) increment  $n_A$  by 1

(ii) if regime  $A \neq 1$ , then set regime  $A = 1$  and increment  $j$  by 1

(iii) redefine  $g^A(n_A) = x_j$ .

To verify that the constructed  $g$ -functions imply valid probability distributions, note that  $\prod_{i=1}^{\infty} g_i^A = 1 - \sum_{i=1}^{\infty} a_i$ . Since in our construction  $\prod_{i=1}^n g_i^A$  is the product of  $x_i$ 's (which are each less than 1), this implies that  $\prod_{i=1}^{\infty} g_i^A = 0$ , i.e.  $\sum_{i=1}^{\infty} a_i = 1$ . Therefore  $\{a_i\}_{i=1}^{\infty}$  is a valid probability distribution. A similar logic also applies to  $\{b_i\}_{i=1}^{\infty}$ . To check that the constructed probabilities are non-increasing, the  $g$ -functions must satisfy the condition  $g_{i+1}^A \geq 2 - \frac{1}{g_i^A}$ . Now, according to the construction  $g^A(n) = x_k$  and  $g^A(n + 1) = x_m$  for some  $k$  and  $m$  with  $m \geq k$ . Since  $x_j$  is an increasing sequence,  $x_m \geq x_k > \frac{2x_k - 1}{x_k}$  and thus the above condition is satisfied.

Recall from proposition 3 that the optimal policy of  $m$  attempts on question  $A$  in an exam with deadline  $T$  must minimize  $\prod_{i=1}^m g_i^A \prod_{i=1}^{T-m} g_i^B$ . Now, by construction, the hazard rates on both questions are decreasing [here either  $g_{i+1}^A = g_i^A$ , or  $g_i^A = x_j$  and  $g_{i+1}^A = x_{j+1}$ ; since  $x_j$  is an increasing sequence,  $g_{i+1}^A \geq g_i^A$  i.e.  $h_{i+1}^A = 1 - g_{i+1}^A \leq 1 - g_i^A = h_i^A$ ; similarly for  $B$ ]; hence the optimal policy will generally involve non-zero attempts on both questions. From (6), the optimal number of attempts is given by  $m$  such that  $h_m^A - h_{T-m+1}^B > 0$  and  $h_{m+1}^A - h_{T-m}^B < 0$ . By construction, this occurs at  $m = f_A(T)$ . ■