Probability in the Engineering and Informational Sciences, 7, 1993, 85–119. Printed in the U.S.A.

ON SEQUENCING TWO TYPES OF TASKS ON A SINGLE PROCESSOR UNDER INCOMPLETE INFORMATION

APOSTOLOS N. BURNETAS
Graduate School of Management
Rutgers University
92 New Street
Newark, New Jersey 07102-1895

MICHAEL N. KATEHAKIS
Graduate School of Management and RUTCOR
Rutgers University
92 New Street
Newark, New Jersey 07102-1895

Two types of tasks are to be scheduled on a single processor under incomplete information about the task lengths. We derive the structure of optimal scheduling rules w.r.t. flowtime, as well as asymptotic approximations for a large number of tasks, when the length distributions belong to a one-parameter exponential family.

1. INTRODUCTION

We study the following version of a scheduling problem under incomplete information. There are two types of tasks, denoted by E_j ($j = 1,2$). Associated with a task of type E_j are i.i.d. random variables, which model the requirements of the tasks in terms of processing time. There is a single station avail-

Research partially supported by a research grant from Graduate School of Management, Rutgers University.

able for the processing of tasks from both types. Only nonpreemptive sequencing is allowed. The objective is to determine a sequence for performing the tasks, so that the total expected flowtime is minimized. The solution to this problem is known when the probability distributions of the task lengths are completely known (see, e.g., Baker [2]). In this article we assume that the characteristics of tasks from E_1 are known in advance, while those of E_2 are not; i.e., type E_1 corresponds to a process presently in use, while E_2 corresponds to a new process that is to be evaluated. In Section 2 we postulate a prior on the unknown parameter of the second type and formulate the problem of minimizing the expected flowtime for a finite number of available tasks. We point out that this is equivalent to minimizing a suitably defined regret (expected loss function).

In Section 3 we derive dynamic programming optimality equations for the determination of optimal policies in several equivalent forms that are convenient for our analysis. We also show that the optimal policy does not depend on the number of the tasks from the known type that have to be processed, a property that simplifies the optimality equations. The results of Section 3 make possible the transformation of the optimality equations into those of a stopping problem (Berry and Fristedt [3], Bradt, Johnson, and Karlin [5]).

In Sections 4 and 5 the case where the task lengths follow a distribution from the one parameter exponential family is considered. We obtain characterizations concerning the structure and properties of optimal sequential allocation policies in Section 4. In Section 5 we derive a simple explicit approximation to the optimal policy in the case that the number of remaining jobs of the unknown type is large.

For related work in this area see Gittins and Glazebrook [9], Lai and Robbins [13], Lai [12], and Agrawal, Hedge, and Teneketzis [1] as well as references given there.

2. PROBLEM FORMULATION

A given number of jobs has to be scheduled on a single processor with the objective of minimizing the total expected flowtime. The jobs belong to two types, E_1 and E_2, with respect to their duration. With type E_i, $i = 1,2$, there are associated

(i) an integer $N_i \geq 1$, which denotes the number of type E_i jobs,

(ii) a parameter θ_i belonging to some parameter set Θ, and

(iii) a sequence of random variables $X_i, Y_{i1}, Y_{i2}, \ldots$, such that Y_{ij} represents the duration (length) of the jth job of type E_i, while X_i is a generic random variable used to denote the length of a job from E_i. Given the value of $\theta_i = \theta$, the random variables $X_i, Y_{i1}, Y_{i2}, \ldots$ are i.i.d., with a probability density function (p.d.f.) $f(x|\theta)$ with respect to a nondegenerate measure ν. Also, let $\mu(\theta)$ denote the expected value of a random variable X distributed according to $f(x|\theta)$.

Parameter θ_1 is known in advance, while θ_2 is unknown. Following the Bayesian approach, we consider θ_2 as a random variable with prior distribution denoted by $H_0(\theta)$, $\theta \in \Theta$.

We define the optimization problem as follows. Let t denote the number of jobs of both types that have already been scheduled. In the sequel t will be referred to as time or stage of the problem. At $t = 0$, $X_i \sim f(x|\theta_i)$, $i = 1,2$ with respect to $\nu(dx)$, with θ_1 known and $\theta_2 \sim H_0(\theta)$.

A set of size k_i of performed type E_i jobs will be denoted by $d_i(k_i) = (y_{i1}, \ldots, y_{i,k_i})$, $k_i \leq N_i$, $i = 1,2$, $k_1 + k_2 = t$. Let $\underline{k} = (k_1, k_2)$, $\underline{d}(\underline{k}) = (d_1(k_1), d_2(k_2))$.

Because θ_1 is known, the lengths of unfinished jobs from $E_1, X_1, Y_{1,k_1+1}, Y_{1,k_1+2}, \ldots$, given $d_1(k_1)$, are i.i.d. random variables with p.d.f. $f(x|\theta_1)$, with respect to $\nu(dx)$.

Because θ_2 is unknown, the lengths of unfinished jobs from $E_2, X_2, Y_{2,k_2+1}, Y_{2,k_2+2}, \ldots$, given $\{d_2(k_2)$ and $\theta_2 = \theta\}$, are i.i.d. random variables with p.d.f. $f(x|\theta)$, with respect to $\nu(dx)$. Given $d_2(k_2)$, θ_2 is a random variable with (posterior) distribution $H(\theta|d_2(k_2))$, defined as follows:

$$dH(\theta|d_2(k_2)) = \frac{\bar{f}(d_2(k_2)|\theta)\, dH_0(\theta)}{\bar{f}(d_2(k_2)|H_0)} = \frac{f(y_{2,k_2}|\theta)\, dH(\theta|d_2(k_2-1))}{\int_\Theta f(y_{2,k_2}|\theta)\, dH(\theta|d_2(k_2-1))}, \tag{2.1}$$

where $d_i(k_i) = (d_i(k_i - 1), y_{i,k_i})$, $H(\theta|d_2(0)) = H_0(\theta)$, and $\bar{f}(d_2(k_2)|\theta)$, $\bar{f}(d_2(k_2)|H_0)$ denote the joint p.d.f. of the sample $d_2(k_2)$, given $\theta_2 = \theta$ and H_0, respectively.

Given $d_2(k_2)$, unconditional on the value of θ_2, the future observations from $E_2, X_2, Y_{2,k_2+1}, Y_{2,k_2+2}, \ldots$, are i.i.d. random variables with distribution determined by the marginal p.d.f. (with respect to $\nu(dx)$)

$$f(x|d_2(k_2)) = f(x|H) = \int_\Theta f(x|\theta)\, dH(\theta|d_2(k_2)). \tag{2.2}$$

Also, in this Bayes framework

$$\hat{\mu}_2(d_2(k_2)) = E_{H(\cdot|d_2(k_2))}[\mu(\theta_2)] = E_{f(\cdot|d_2(k_2))}[X_2] \tag{2.3}$$

denotes the Bayes estimate of $\mu(\theta_2)$.

For notational convenience we use the same symbol f to denote the p.d.f. of an outcome given a specific parameter value as well as the marginal p.d.f. of an outcome from E_2 given the history of observations $d_2(k_2)$. Although they are different quantities, there is no danger of confusion.

Let $\mathcal{H}$ denote the set of prior distributions of θ_2. A sequential scheduling policy is defined as a function $\pi: \{0,1,\ldots,N_1\} \times \{0,1,\ldots,N_2\} \times \mathcal{H} \to \{1,2\}$, where $\pi(n_1, n_2, H)$ is equal to 1 or 2, according to whether we schedule a job from E_1 or E_2, respectively, at a stage when $n_1(n_2)$ jobs are unfinished from type E_1 (E_2) and the current information about θ_2 is summarized by the posterior d.f. $H(\cdot)$.

To characterize the performance of a policy, we define the total flowtime corresponding to a policy π

$$F(N_1, N_2, \pi) = \sum_{t=1}^{N_1+N_2} \sum_{j=1}^{t} Y(j, \pi), \tag{2.4}$$

where $Y(j, \pi)$ denotes the length of the job scheduled at stage j, under π. Then,

$$\begin{aligned} E_\theta F(N_1, N_2, \pi) &= E[F(N_1, N_2, \pi) \mid \theta_2 = \theta] \\ &= \sum_{t=1}^{N_1+N_2} (N_1 + N_2 + 1 - t)\mu(\theta_{i(t,\pi)}), \end{aligned} \tag{2.5}$$

where $i(t, \pi)$ is the type of the job scheduled at stage t, under π. In Eq. (2.5) the flowtime is expressed as the expected accumulated number of unfinished jobs over time. This expression follows by taking the expectations in Eq. (2.4) and rearranging the terms.

Define the expected flowtime achieved by a policy π with respect to the prior distribution of θ_2:

$$M(N_1, N_2, H_0, \pi) = E_{H_0}[E_\theta F(N_1, N_2, \pi)] = E_{f(\cdot|H_0)}[F(N_1, N_2, \pi)]. \tag{2.6}$$

We say that a policy π^* is optimal for the problem of minimizing the expected flowtime under incomplete information and initial prior $H_0(\theta)$ on θ_2, if and only if

$$M(N_1, N_2, H_0, \pi^*) = \min_\pi M(N_1, N_2, H_0, \pi), \tag{2.7}$$

where the minimum is taken over all sequential policies defined earlier.

We now turn to an alternative way to describe the problem, which will be useful in the following development. First we observe that, if θ_2 were known, the optimal policy for minimizing the expected flowtime would be the Shortest Expected Processing Time (SEPT) rule, according to which all the jobs of the type with the least expected length would be scheduled first. Given $\theta_2 = \theta$, define the priority and nonpriority indices $a(\theta)$ and $b(\theta)$ such that $\mu(\theta_a) \equiv \mu^*(\theta) \equiv \min\{\mu(\theta_1), \mu(\theta)\}$ and $\mu(\theta_b) \equiv \max\{\mu(\theta_1), \mu(\theta)\}$, respectively. Using these definitions as well as Eq. (2.5), we can express the minimum expected flowtime under complete information for θ_2, as a function of θ, for any value of $\theta_2 = \theta$:

$$\begin{aligned} F^*(\theta; N_1, N_2) &= \sum_{t=1}^{N_{a(\theta)}} (N_{a(\theta)} + N_{b(\theta)} + 1 - t)\mu(\theta_{a(\theta)}) \\ &\quad + \sum_{t=1}^{N_{b(\theta)}} (N_{b(\theta)} + 1 - t)\mu(\theta_{b(\theta)}) \\ &= \frac{N_1(N_1+1)}{2}\mu(\theta_1) + \frac{N_2(N_2+1)}{2}\mu(\theta) + N_1 N_2 \mu^*(\theta). \end{aligned} \tag{2.8}$$

When θ_2 is thought of as a random variable, the expected value of the preceding expression with respect to the prior distribution $H_0(\theta)$ provides a lower bound for the achievable flowtime under uncertainty:

$$M(N_1, N_2, H_0, \pi) \geq E_{H_0} F^*(\theta; N_1, N_2), \quad \text{for all } \pi. \tag{2.9}$$

We finally define the regret due to uncertainty corresponding to a policy π as

$$R(N_1, N_2, H_0, \pi) = M(N_1, N_2, H_0, \pi) - E_{H_0} F^*(\theta; N_1, N_2). \tag{2.10}$$

Since the quantity $E_{H_0} F^*(\theta; N_1, N_2)$ in Eq. (2.10) is independent of π, minimization of M is equivalent to minimization of R. Thus, an optimal policy π^* can also be defined as

$$R(N_1, N_2, H_0, \pi^*) = \min_{\pi} R(N_1, N_2, H_0, \pi). \tag{2.11}$$

Before we proceed to the next section, we will state a martingale-type property of the information-updating mechanism that will be used in what follows. The proof is immediate.

Let $H(\theta)$ denote the distribution of θ_2 given the current information and $\tilde{H}(\theta|x)$ the same distribution after one additional job from type E_2 has been performed and its length has been observed to be equal to x.

LEMMA 2.1: *For any function $g: \Theta \to \mathbb{R}$ such that $E_H[|g(\theta_2)|] < \infty$,*

$$E_{f(\cdot|H)}[E_{\tilde{H}(\cdot|X_2)} g(\theta_2)] = E_H[g(\theta_2)]. \tag{2.12}$$

3. PRELIMINARY RESULTS

In this section we obtain dynamic programming optimality equations for the optimization problem defined in Section 2. The equations are given in terms of minimization of the expected total flowtime as well as of the regret. The main result in this section is Theorem 3.1, according to which the optimal policy does not depend on the number of unfinished jobs of type E_1. The assumption of processing time distributions from the one-parameter exponential family is introduced at the end of the section.

Let V, U denote the optimal value functions for the flowtime and the regret, respectively:

$$V(n_1, n_2, H) = \inf_{\pi} M(n_1, n_2, H, \pi), \tag{3.1}$$

$$U(n_1, n_2, H) = \inf_{\pi} R(n_1, n_2, H, \pi). \tag{3.2}$$

In the following two propositions the dynamic programming equations corresponding to Eqs. (3.1) and (3.2) are obtained. The proof of Proposition 3.1 is omitted, because it is based on well-known results on Markovian decision processes with incomplete information (Bertsekas [4], Dynkin and Yushkevich [7]). The proof of Proposition 3.2, given in the Appendix, is based on the fact that

$$U(n_1, n_2, H) = V(n_1, n_2, H) - E_H F^*(\theta; n_1, n_2), \tag{3.3}$$

which is immediate from Eq. (2.10).

Proposition 3.1: The functions $V(n, k, y)$ defined in Eq. (3.1) are the unique solutions of Eqs. (3.4)–(3.6):

$$V(n_1, n_2, H) = \min\{(n_1 + n_2) r(H; \alpha = 1) + V(n_1 - 1, n_2, H),$$
$$(n_1 + n_2) r(H; \alpha = 2) + E_{f(\cdot|H)} V(n_1, n_2 - 1, \tilde{H}(\cdot | X_2))\},$$
$$n_i = 1, 2, \ldots, N_i,\ i = 1, 2, \tag{3.4}$$

$$V(n_1, 0, H) = \frac{n_1(n_1 + 1)}{2} r(H; \alpha = 1), \qquad n_1 = 0, 1, \ldots, N_1, \tag{3.5}$$

$$V(0, n_2, H) = \frac{n_2(n_2 + 1)}{2} r(H; \alpha = 2), \qquad n_2 = 0, 1, \ldots, N_2, \tag{3.6}$$

where

$$r(H; \alpha = 1) = E_{\theta_1}\{X_1\} = \mu(\theta_1), \tag{3.7}$$

$$r(H; \alpha = 2) = E_{f(\cdot|H)} X_2 = E_H[E_\theta X_2] = \int_\Theta \mu(\theta)\, dH(\theta). \tag{3.8}$$

Proposition 3.2: The functions $U(n, k, y)$ defined in Eq. (3.2) are the unique solutions of Eqs. (3.9) and (3.10):

$$U(n_1, n_2, H) = \min\{n_2 c(H; \alpha = 1) + U(n_1 - 1, n_2, H),$$
$$n_1 c(H; \alpha = 2) + E_{f(\cdot|H)} U(n_1, n_2 - 1, \tilde{H}(\cdot | X_2))\},$$
$$n_i = 1, 2, \ldots, N_i,\ i = 1, 2, \tag{3.9}$$

$$U(n_1, 0, H) = U(0, n_2, H) = 0, \qquad n_i = 0, 1, \ldots, N_i, \tag{3.10}$$

where

$$c(H; \alpha = 1) = \int_{\mu(\theta) < \mu(\theta_1)} (\mu(\theta_1) - \mu(\theta))\, dH(\theta) \tag{3.11}$$

and

$$c(H; \alpha = 2) = \int_{\mu(\theta) > \mu(\theta_1)} (\mu(\theta) - \mu(\theta_1))\, dH(\theta) \tag{3.12}$$

are the one-step regret or loss functions.

Remark 3.3: Since the state and action spaces satisfy the conditions of Section 8.2 in Dynkin and Yushkevich [7], the infimum in Eqs. (3.1) and (3.2) is attained by an optimal policy π^* and can be replaced by minimum.

In the following theorem we will establish a property of the optimal policy that simplifies the study of the optimality equations and makes apparent the usefulness of the regret approach to the problem.

THEOREM 3.4:

(a) *The minimum regret function* $U(n_1, n_2, H)$ *has the following property*:

$$U(n_1, n_2, H) = n_1 U(1, n_2, H), \quad \forall n_1, n_2, H. \tag{3.13}$$

(b) *The optimal policy* π^* *satisfies*

$$\pi^*(n_1, n_2, H) = \pi^*(1, n_2, H), \quad \forall n_1, n_2, H. \tag{3.14}$$

PROOF: We prove (a) and (b) simultaneously by induction on n_1 and n_2. Let

$$U_1(n_1, n_2, H) = n_2 c(H; \alpha = 1) + U(n_1 - 1, n_2, H) \tag{3.15}$$

and

$$U_2(n_1, n_2, H) = n_1 c(H; \alpha = 2) + E_{f(\cdot|H)} U(n_1, n_2 - 1, \bar{H}(\cdot | X_2)). \tag{3.16}$$

For $n_1 = 0,1$ and all n_2, as well as for $n_2 = 0$ and all n_1, Eq. (3.13) is obvious. Assume that it is true for some n_1 and all n_2 as well as for $n_1 + 1$ and $n_2 - 1$. Then for $n_1 + 1$ and n_2 we have

$$\begin{aligned} U_1(n_1 + 1, n_2, H) &= n_2 c(H; \alpha = 1) + n_1 U(1, n_2, H) \\ &= U_1(1, n_2, H) + n_1 U(1, n_2, H), \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} U_2(n_1 + 1, n_2, H) &= (n_1 + 1) c(H; \alpha = 2) \\ &\quad + (n_1 + 1) E_{f(\cdot|H)} U(1, n_2 - 1, \bar{H}(\cdot | X_2)) \\ &= (n_1 + 1) U_2(1, n_2, H). \end{aligned} \tag{3.18}$$

Thus,

$$\begin{aligned} &U_1(n_1 + 1, n_2, H) - U_2(n_1 + 1, n_2, H) \\ &\quad = U_1(1, n_2, H) - U_2(1, n_2, H) + n_1 (U(1, n_2, H) - U_2(1, n_2, H)). \end{aligned} \tag{3.19}$$

Now we consider two cases.

Case 1: $U_1(1, n_2, H) \le U_2(1, n_2, H)$. Then $U(1, n_2, H) = U_1(1, n_2, H)$ and

$$\begin{aligned} &U_1(n_1 + 1, n_2, H) - U_2(n_1 + 1, n_2, H) \\ &\quad = (n_1 + 1)(U_1(1, n_2, H) - U_2(1, n_2, H)) \le 0, \end{aligned} \tag{3.20}$$

thus

$$U(n_1+1,n_2,H) = U_1(n_1+1,n_2,H) = (n_1+1)U_1(1,n_2,H) = (n_1+1)U(1,n_2,H) \quad (3.21)$$

and also

$$\pi^*(n_1+1,n_2,H) = 1 = \pi^*(1,n_2,H). \quad (3.22)$$

Case 2: $U_1(1,n_2,H) > U_2(1,n_2,H)$. Then $U(1,n_2,H) = U_2(1,n_2,H)$ and

$$U_1(n_1+1,n_2,H) - U_2(n_1+1,n_2,H) = U_1(1,n_2,H) - U_2(1,n_2,H) > 0, \quad (3.23)$$

thus

$$U(n_1+1,n_2,H) = U_2(n_1+1,n_2,H) = (n_1+1)U_2(1,n_2,H) = (n_1+1)U(1,n_2,H) \quad (3.24)$$

and

$$\pi^*(n_1+1,n_2,H) = 2 = \pi^*(1,n_2,H). \quad (3.25)$$

Since the required properties are true for both cases, the induction step is complete. ■

Remark 3.5:

(a) Theorem 3.4 implies the following. If at some state (n_1,n_2,H), $\pi^*(n_1,n_2,H) = 1$; then, because the next state will be (n_1-1,n_2,H) and π^* is independent of n_1, $\pi^*(k,n_2,H) = 1$, for $k = n_1-1,\ldots,1$.

(b) Theorem 3.4 and the preceding observation lead to the next simplification of the dynamic programming equations for U:

$$U(n,H) = \min\big[nc(H;\alpha=1), c(H;\alpha=2) + E_{f(\cdot|H)}U(n-1,\tilde{H}(\cdot|X_2))\big], \quad n = 1,2,\ldots,N, \quad (3.26)$$

$$U(0,H) = 0, \quad (3.27)$$

where $U(n,H)$ corresponds to $U(1,n,H)$.

(c) With Eq. (3.26) the original problem has been reduced to a problem of optimal stopping, where the stopping cost at state (n,H) is the first term of the right-hand side.

The problem of optimal stopping is well known in the literature (Bradt et al. [5], Berry and Fristedt [3], Kumar and Varayia [11]). In the remainder of this article we study the solution of Eqs. (3.26) and (3.27) under the additional assumption that the underlying distribution of the job lengths belongs to a one-parameter exponential family. More precisely we adopt the following.

Assumption 3.6:

(a) The p.d.f. $f(x|\theta)$ belongs to an exponential family with a single natural parameter θ; i.e.,

$$f(x|\theta) = e^{\theta x - \psi(\theta) + s(x)}. \tag{3.28}$$

(b) The parameter set is an interval of the form $\Theta = (\underline{\theta}, \bar{\theta})$, with endpoints that can be infinite, and satisfies the following conditions:

$$E_\theta |X_i| = \int_{-\infty}^{+\infty} |x| f(x|\theta)\nu(dx) < \infty, \quad \forall \theta \in \Theta, \, i = 1,2, \tag{3.29}$$

$$\zeta_1 = \inf_{\theta \in \Theta} \psi''(\theta) > 0, \qquad \zeta_2 = \sup_{\theta \in \Theta} \psi''(\theta) < \infty. \tag{3.30}$$

Remark 3.7:

(a) The exponential family assumption of Eq. (3.28) is quite general, because several interesting distributions have this form, among them the normal with known variance and parametric mean, Poisson, negative exponential, Bernoulli, etc.

(b) Since the term $e^{s(x)}$ in Eq. (3.28) is independent of θ, it can be included in the measure ν.

(c) In a one-parameter exponential family of the type described in Eq. (3.28) the distributions are ordered both in mean and in likelihood ratio (cf. Cox and Hinkley [6]). More specifically $\mu(\theta) = \psi'(\theta)$, and $\text{var}(X|\theta) = \psi''(\theta)$. Thus, $\mu(\theta)$ is strictly increasing in $\theta \in \Theta$, and the set $\{\mu(\theta) : \theta \in \Theta\}$ is an interval of the form $(\mu(\underline{\theta}), \mu(\bar{\theta}))$. The likelihood ratio ordering refers to the fact that for any $\theta_1, \theta_2 \in \Theta$, such that $\theta_1 < \theta_2$, the likelihood ratio $f(x|\theta_1)/f(x|\theta_2)$ is decreasing in x. We will use both of these ordering properties in the rest of the paper.

(d) The results of Section 5 require that θ_1 is an interior point of the parameter set Θ. This does not affect the generality of the model described thus far. Indeed, if $\theta_1 = \underline{\theta}$ ($\theta_1 = \bar{\theta}$) then the problem is trivial, because then one should always choose E_1 (E_2). Thus, from now on we shall assume that $\underline{\theta} < \theta_1 < \bar{\theta}$.

The one-parameter exponential family assumption has the following important implication. The posterior distribution $H(\theta|d_2(k_2))$ and the marginal density $f(x|d_2(k_2))$ defined in Eq. (2.1) and Eq. (2.2), respectively, are uniquely determined by $(k_2, \bar{y}_{2,k_2})$, where

$$\bar{y}_{2,k} = \frac{1}{k} \sum_{j=0}^{k} y_{2,j} \tag{3.31}$$

is the sample mean; i.e., the pair $(k, \bar{y}_k)$ is a sufficient statistic for the unknown parameter, for this family of distributions (cf. Cox and Hinkley [6]).

Thus, we can assume that in Eqs. (2.1)–(2.3) $d_2(k_2)$ is simply the two-dimensional vector

$$d_2(k_2) = (k_2, \bar{y}_{2,k_2}). \tag{3.32}$$

Note that given $d_2(k-1) = (k-1, y)$ and $Y_{2,k} = y_{2,k}$, $d_2(k)$ is defined by the following updating scheme:

$$d_2(k \,|\, d_2(k-1), y_{2,k}) = (k, m(k-1, y, y_{2,k})), \tag{3.33}$$

where

$$m(k, y, x) = \frac{ky + x}{k+1}, \tag{3.34}$$

and thus optimality Eqs. (3.26) and (3.27) can be simplified as follows, where the current information is represented by the sufficient statistic (k, y), as already defined:

$$\begin{aligned} U(n,k,y) = \min\big[& nc(k,y;\alpha=1), c(k,y;\alpha=2) \\ & + E_{f(\cdot|H)} U(n-1, k+1, m(k,y,X_2)) \big], \\ & n = 1,2,\ldots,N,\ k = 1,2,\ldots,N \end{aligned} \tag{3.35}$$

$$U(0,k,y) = 0. \tag{3.36}$$

Definition 3.8 allows us to use a change of measure transformation to obtain a further simplification of the optimality equations.

Definition 3.8: *Let*

$$l(\theta, \theta_1 \,|\, y) = \log \frac{f(y \,|\, \theta)}{f(y \,|\, \theta_1)} \tag{3.37}$$

$$\Lambda(k, y) = \int_\Theta e^{kl(\theta, \theta_1|y)}\, dH_0(\theta) \tag{3.38}$$

$$d(\theta) = \theta - \theta_1, \tag{3.39}$$

$$\delta(\theta) = \mu(\theta) - \mu(\theta_1), \tag{3.40}$$

$$\omega(\theta) = \psi(\theta) - \psi(\theta_1). \tag{3.41}$$

Remark 3.9:

(a) From the sufficiency of (k, y) it follows that the quantity $e^{kl(\theta,\theta_1|y)}$ is equal to the likelihood ratio of a sample d_2 with size k and sample average y:

$$e^{kl(\theta,\theta_1|y)} = \frac{f(d_2;\theta)}{f(d_2;\theta_1)}. \tag{3.42}$$

Therefore, relation Eq. (2.1) for the posterior distribution of θ_2 given the sample $d_2 = (k, y)$ becomes

$$dH(\theta \mid (k,y)) = \frac{e^{k(\theta y - \psi(\theta))}\, dH_0(\theta)}{\int_\Theta e^{k(\theta y - \psi(\theta))}\, dH_0(\theta)} = \frac{e^{kl(\theta,\theta_1|y)}\, dH_0(\theta)}{\Lambda(k,y)}. \quad (3.43)$$

(b) From Eq. (3.28)

$$l(\theta, \theta_1 | y) = d(\theta) y - \omega(\theta), \quad (3.44)$$

and thus

$$kl(\theta,\theta_1|y) + l(\theta,\theta_1|x) = (k+1) l(\theta,\theta_1 | m(k,y,x)). \quad (3.45)$$

The main idea in the next proposition is to transform the optimal regret function U to $u = U\Lambda$, with Λ as defined in Eq. (3.38). Then u satisfies a set of optimality equations equivalent to Eqs. (3.35) and (3.36), but they are easier to study. The one-step costs for the transformed system have a simpler form and also the expectation of the future cost is taken with respect to $f(\cdot|\theta_1)$ instead of the marginal $f(\cdot|H)$. Because we are more interested in the optimal policy than in the value of the regret function, this transformation is beneficial. The proof of the proposition is given in the Appendix together with a necessary auxiliary lemma.

Proposition 3.10: Optimality Eqs. (3.35) and (3.36) are equivalent to the following:

$$u(n,k,y) = \min\{ n\bar{c}(k,y;\alpha=1), \bar{c}(k,y;\alpha=2) + E_{f(\cdot|\theta_1)} u(n-1, k+1, m(k,y,X_2))\},$$
$$n = 1,2,\ldots,N,\ k = 0,1,\ldots,N-n,\ y \in \mathbb{R}, \quad (3.46)$$

$$u(0,k,y) = 0, \quad (3.47)$$

where

$$u(n,k,y) = U(n,k,y)\Lambda(k,y), \quad (3.48)$$

$$\bar{c}(k,y;\alpha=1) = -\int_{\theta<\theta_1} \delta(\theta) e^{kl(\theta,\theta_1|y)}\, dH_0(\theta), \quad (3.49)$$

$$\bar{c}(k,y;\alpha=2) = \int_{\theta\geq\theta_1} \delta(\theta) e^{kl(\theta,\theta_1|y)}\, dH_0(\theta). \quad (3.50)$$

We conclude this section with an auxiliary lemma, which is proved in the Appendix. Let

$$q(k,y) = \bar{c}(k,y;\alpha=2) - \bar{c}(k,y;\alpha=1) = \int_\Theta \delta(\theta) e^{kl(\theta,\theta_1|y)}\, dH_0(\theta). \quad (3.51)$$

LEMMA 3.11: *The quantities* $\bar{c}(k,y;\alpha), q(k,y)$ *satisfy*

(a) $\bar{c}(k,y;\alpha) > 0,\ \forall k, y,\ \alpha = 1,2.$ (3.52)

(b) $\forall k, \bar{c}(k,y;\alpha = 1)$ *is decreasing in* y.
$\forall k, \bar{c}(k,y;\alpha = 2)$ *is increasing in* y.
$\forall k, q(k,y)$ *is increasing in* y.

(c) $E_{\theta_1}[\bar{c}(k+1, m(k,y,X);\alpha)] = \bar{c}(k,y;\alpha).$ (3.53)

Remark 3.12: Equation (3.53) is an intuitive martingale property of the information update mechanism. It can be regarded as the transformation, according to the change of measure introduced in Proposition 3.10, of the following equation about the one-step regret functions:

$$E_{f(\cdot|H)}[c(\bar{H}(\cdot|X);\alpha)] = c(H;\alpha). \tag{3.54}$$

Equation (3.54) expresses the fact that, before an additional sample from E_2 is observed, the expected value of the one-step cost after the observation is equal to the current value. This is a well-known property of Bayes sequential estimation methods.

4. FINITE HORIZON

In this section we shall prove two results that give the structure of the optimal policy for the problem stated in Sections 2 and 3. Theorem 4.1 describes the optimal policy with respect to stopping and continuation intervals for $y = \sum_{j=1}^{k} y_{2,j}$, whereas Corollary 4.2 gives a more intuitive characterization in terms of factors subtracted from the Bayes estimate of $\mu(\theta_2)$.

THEOREM 4.1:

(a) *For each* n, k *there exists a number* $y_n(k)$ *with the property*

$$\pi^*(n,k,y) = \begin{cases} 1, & \text{if } y > y_n(k) \\ 2, & \text{if } y \le y_n(k) \end{cases}, \tag{4.1$_n$}$$

where $\pi^*(n,k,y)$ *is the action indicated by the optimal policy in state* (n,k,y).

(b) *The sequence* $y_n(k)$ *is nondecreasing in* n.

PROOF:

(a) Define

$$u(n,k,y) = \min\{n\bar{c}(k,y;\alpha = 1), \bar{c}(k,y;\alpha = 2) + E_{f(\cdot|\theta_1)}u(n-1, k+1, m(k,y,X_2))\}, \tag{4.2}$$

$$u_1(n,k,y) = n\bar{c}(k,y;\alpha = 1), \tag{4.3}$$

$$u_2(n,k,y) = \bar{c}(k,y;\alpha=2) + E_{f(\cdot|\theta_1)} u(n-1,k+1,m(k,y,X_2)), \tag{4.4}$$

$$\Delta(n,k,y) = u_2(n,k,y) - u_1(n,k,y). \tag{4.5}$$

Then

$$u(n,k,y) = u_1(n,k,y) + \min\{0, \Delta(n,k,y)\}, \tag{4.6}$$

and from Lemma 3.11(c), Δ can be seen to satisfy the following recursion:

$$\begin{aligned}\Delta(n+1,k,y) &= \bar{c}(k,y;\alpha=2) + E_{f(\cdot|\theta_1)}[u(n,k+1,m(k,y,X_2))] \\ &\quad - (n+1)\bar{c}(k,y;\alpha=1) \\ &= q(k,y) + E_{f(\cdot|\theta_1)}[u(n,k+1,m(k,y,X_2)) \\ &\quad - n\bar{c}(k+1,m(k,y,X_2);\alpha=1)] \\ &= q(k,y) + E_{f(\cdot|\theta_1)}[\min\{0,\Delta(n,k+1,m(k,y,x))\}],\end{aligned} \tag{4.7}$$

where $q(k,y)$ is the difference of the transformed one-step regrets defined in Eq. (3.51). We shall prove simultaneously by induction on n the following:

(i) Equation $(4.1)_n$,

(ii) $\Delta(n,k,y)$ is increasing in y. $\quad (4.8)_n$

For $n=1$, Eq. $(4.8)_1$ is immediate from Lemma 3.11(b), because $\Delta(0,k,y)=0$ and $\Delta(1,k,y)=q(k,y)$. Let

$$y_1(k) = \inf\{y : \Delta(1,k,y) \geq 0\}, \tag{4.9}$$

where we define $\inf \varnothing = +\infty$.

For $y < y_1(k)$ $\Delta(1,k,y)$ is negative, whereas for $y \geq y_1(k)$ it is nonnegative. This completes the proof of Eq. $(4.1)_1$.

Suppose that Eqs. $(4.1)_n$ and $(4.8)_n$ hold. Then $\Delta(n,k+1,m(k,y,x)) < 0$ when $m(k,y,x) < y_n(k+1)$, which from Eq. (3.32) is equivalent to

$$x < x_n(k,y) \equiv (k+1)y_n(k+1) - ky. \tag{4.10}$$

Hence, Eq. (4.7) can be written as

$$\Delta(n+1,k,y) = q(k,y) + \int_{x<x_n(k,y)} \Delta(n,k+1,m(k,y,x)) \times f(x|\theta_1)\nu(dx). \tag{4.11}$$

To prove statement $(4.8)_{n+1}$, we have from Eq. $(4.8)_n$ that $\Delta(n, k+1, m(k,y,x))$ is increasing in m, while $m(k,y,x)$ is increasing in y, and so $\Delta(n,k+1,m(k,y,x))$ is increasing in y. Also, $\Delta(n,k+1, m(k,y,x)) < 0$ for $x < x_n(k,y)$. Furthermore, $x_n(k,y)$ is decreasing in y, so when y increases the range of integration decreases. Thus, Eq. $(4.8)_{n+1}$ follows from the preceding observations. Equation $(4.1)_{n+1}$ can be established using Eq. $(4.8)_{n+1}$ and defining $y_{n+1}(k)$ in the same way as that in Eq. (4.9). This completes the proof of Theorem 4.1(a).

(b) We will prove the following with simultaneous induction on n:

$$y_n(k) \geq y_{n-1}(k), \qquad \forall k, n \qquad (4.12)_n$$

$$\min\{0, \Delta(n,k,y)\} \leq \min\{0, \Delta(n-1,k,y)\}, \qquad \forall k, n, y. \qquad (4.13)_n$$

For $n = 1$ both Eq. $(4.12)_1$ and Eq. $(4.13)_1$ are immediate, because $\Delta(0,k,y) = 0$ and $y_0(k) = -\infty$. Assume they are true for n. Then, for $n+1$ we get from Eq. (4.7) and the induction hypothesis

$$\begin{aligned}\Delta(n+1,k,y) &- \Delta(n,k,y)\\ &= E_{f(\cdot|\theta_1)}[\min\{0, \Delta(n,k+1,m(k,y,x))\}\\ &\quad - \min\{0, \Delta(n-1,k+1,m(k,y,x))\}] \leq 0.\end{aligned}$$

Thus, for $y = y_n(k)$

$$\Delta(n+1,k,y_n(k)) \leq \Delta(n,k,y_n(k)) = 0, \qquad (4.14)$$

and because Δ is increasing in y,

$$y_{n+1}(k) \geq y_n(k),$$

which proves Eq. $(4.12)_{n+1}$. Equation $(4.13)_{n+1}$ also follows easily from Eq. $(4.12)_{n+1}$ and Eq. (4.14). ■

Corollary 4.2:

(a) *For each n, k, y the optimal action $\pi^*(n,k,y)$ has the form*

$$\pi^*(n,k,y) = \begin{cases} 1, & \text{if } E_{H(\cdot|(k,y))}[\mu(\theta_2)] - \epsilon(n,k,y) > \mu(\theta_1) \\ 2, & \text{if } E_{H(\cdot|(k,y))}[\mu(\theta_2)] - \epsilon(n,k,y) \leq \mu(\theta_1) \end{cases}, \qquad (4.15)$$

where $E_{H(\cdot|(k,y))}[\mu(\theta_2)]$ is the Bayes estimate of $\mu(\theta_2)$ given (k,y) and

$$\epsilon(n,k,y) = \frac{q(k, y_n(k))}{\Lambda(k,y)}. \qquad (4.16)$$

(b) *The quantities $\epsilon(n,k,y)$ are nonnegative and nondecreasing in n for all (k,y).*

PROOF:

(a) From Theorem 4.1(b) $y_n(k) \geq y_1(k)$. We also have from Theorem 4.1 and Lemma 4.2 that if $\pi^*(n, k, y) = 1$, then $y > y_n(k)$, and so $q(k, y) > q(k, y_n(k))$. By Eq. (3.51)

$$q(k, y) = \int_\Theta \delta(\theta) e^{kl(\theta, \theta_1 | y)} \, dH_0(\theta)$$

$$= (E_{H(\cdot|(k,y))}[\mu(\theta_2)] - \mu(\theta_1))\Lambda(k, y). \tag{4.17}$$

Thus,

$$y > y_n(k) \Leftrightarrow q(k, y) > q(k, y_n(k))$$

$$\Leftrightarrow E_{H(\cdot|(k,y))}[\mu(\theta_2)] - \frac{q(k, y_n(k))}{\Lambda(k, y)} > \mu(\theta_1), \tag{4.18}$$

and this completes the proof of Corollary 4.2(a).

(b) Because $y_n(k) \geq y_1(k)$, it is true that $q(k, y_n(k)) \geq 0$, and so $\epsilon(n, k, y) \geq 0$. From Eq. (4.13) we see that the dependence of $\epsilon(n, k, y)$ on n is due to $q(k, y_n(k))$. By Lemma 3.1, $q(k, y_n(k))$ is increasing in $y_n(k)$. Finally, because $y_n(k)$ is nondecreasing in n, we have that $\epsilon(n, k, y)$ is nondecreasing in n. ■

Remark 4.3:

(a) The "cutting point" $y_n(k)$ is related to the uncertainty due to the ignorance of parameter θ_2 and represents in some way the amount of immediate payoff that we can afford to sacrifice to get information about θ_2, which is valuable for our future decisions. Because of that, it is intuitively expected that $y_n(k)$ has the stated property, since further sampling from E_2 reduces the uncertainty.

(b) One interpretation of the quantities $\epsilon(n, k, y)$ is that they represent a correction, which we subtract from the current Bayes estimate of the job length from E_2, $\hat{\mu}_2 = E_{H(\cdot|(k,y))}[\mu(\theta_2)]$, to take account of the uncertainty associated with it. So the properties of $\epsilon(n, k, y)$ stated in Corollary 4.2(b) are intuitively expected. A related result in terms of indices is given in Theorem 5.3.1 of Berry and Fristedt [3] for the case of general underlying distributions.

5. ASYMPTOTIC APPROXIMATIONS

In this section we obtain properties of the optimal policy that are related to its behavior when the number of jobs of the unknown type is large (i.e., when $n \to \infty$). Before we proceed to the statement of the results, we shall make another assumption. Namely, we assume that the prior distribution of θ_2 is continuous in $[\underline{\theta}, \bar{\theta}]$; i.e., there is a prior p.d.f. denoted by $h_0(\theta)$. This assumption simplifies the following discussion without loss of generality, because the general case can be treated in an analogous way.

The main results of this section are given in Theorems 5.11 and 5.12, which provide us with upper and lower approximations to the optimal stopping regions. The proofs of these two theorems are based on a number of intermediate properties, which are stated later and are proved in the Appendix.

For a family of distributions $f(x|\theta)$, let $\mathbf{I}(\sigma,\tau)$ denote the Kullback–Leibler information number

$$\mathbf{I}(\sigma,\tau) = E_\sigma\left[\log\frac{f(X|\sigma)}{f(X|\tau)}\right]. \tag{5.1}$$

LEMMA 5.1: *For the one-parameter exponential family* $e^{\theta x-\psi(\theta)}$, $\mathbf{I}(\cdot,\cdot)$ *has the following properties*:

$$\mathbf{I}(\sigma,\tau) = (\sigma-\tau)\mu(\sigma) - (\psi(\sigma)-\psi(\tau)), \tag{5.2}$$

$$\mathbf{I}(\sigma,\tau) = \int_\tau^\sigma (\tau-\theta)\psi''(\theta)\,d\theta, \tag{5.3}$$

$$\zeta_1\frac{(\sigma-\tau)^2}{2} \le \mathbf{I}(\sigma,\tau) \le \zeta_2\frac{(\sigma-\tau)^2}{2}. \tag{5.4}$$

Lemma 5.2 indicates a useful relationship between the log-likelihood ratio $l(\theta,\theta_1|x)$ and the Kullback–Leibler information number.

LEMMA 5.2:

(a) $l(\theta,\theta_1|x)$ *is concave in* θ.

(b) $\forall x\in\mathbb{R}:\exists\theta^* = \theta^*(x)$, *such that*

$$l(\theta^*,\theta_l|x) = \max_{\theta\in\Theta} l(\theta,\theta_1|x), \tag{5.5}$$

where

$$\theta^*(x) = \begin{cases}\mu^{-1}(x), & \text{if } \mu^{-1}(x)\in\Theta\\ \bar{\theta}, & \text{if } \mu^{-1}(x)\notin\Theta \text{ and } l(\bar{\theta},\theta_1|x) > l(\underline{\theta},\theta_1|x)\\ \underline{\theta}, & \text{if } \mu^{-1}(x)\notin\Theta \text{ and } l(\bar{\theta},\theta_1|x) \le l(\underline{\theta},\theta_1|x)\end{cases} \tag{5.6}$$

Moreover, if $\mu^{-1}(x)\in\Theta$, *then*

$$l(\theta^*,\theta_1|x) = \mathbf{I}(\theta^*,\theta_1). \tag{5.7}$$

(c) *If* $x > \mu(\bar{\theta})$, *then* $l(\bar{\theta},\theta_1|x) > 0$.

LEMMA 5.3: *Let*

$$\gamma(k,y) = \min\{\bar{c}(k,y;\alpha=1), \bar{c}(k,y;\alpha=2)\}. \tag{5.8}$$

There exists $A > 0$ such that $\gamma(k,y) < A/k$, $\forall k = 1,2,\ldots, y \in \mathbb{R}$.

LEMMA 5.4: *Consider the following inequality:*

$$Mx^{-\alpha}e^{-\lambda x} < \frac{1}{n}, \tag{5.9}$$

for $x > 0$, and constants $\alpha, M, \lambda, n > 0$. Let $x(n)$ denote the solution of the equation resulting when the inequality sign in Eq. (5.9) is replaced by equality. Then

(a) (i) $x(n) > 0$, $\forall n \geq 1$,
 (ii) $\lim_{n\to\infty} x(n) = \infty$, *and*
 (iii) *inequality (5.9) holds for $x > x(n)$.*

(b) *There exists a function $\epsilon(n)$ such that $\lambda x(n) = \log n - \epsilon(n)$ and $\epsilon(n) \sim \alpha \log(\log n)$ as $n \to \infty$.*

The next lemma concerning the asymptotic expansion of integrals according to the Laplace method (cf. Erdélyi [8, p. 37]) will be used to obtain asymptotic expressions for the one-step regrets.

LEMMA 5.5: *Let $g_1(\theta)$ and $g_2(\theta)$ be real functions on the interval (α,β). Let $\epsilon, \eta > 0$ such that $g_2(\theta)$ is continuous at $\theta = \alpha$, continuously differentiable for $\alpha < \theta \leq \alpha + \eta$, $g_2'(\theta) < 0$ for $\alpha < \theta \leq \alpha + \eta$, and $g_2(\theta) \leq g_2(\alpha) - \epsilon$, for $\alpha + \eta \leq \theta \leq \beta$. Suppose that there exist numbers b, d, ν, λ, such that $\nu, \lambda > 0$ and $g_2'(\theta) \sim -b(\theta-\alpha)^{\nu-1}$, $g_1(\theta) \sim d(\theta-\alpha)^{\lambda-1}$, as $\theta \to \alpha$. Then*

$$\int_\alpha^\beta g_1(\theta)e^{kg_2(\theta)}\,d\theta \sim \frac{d}{\nu}\Gamma\left(\frac{\lambda}{\nu}\right)\left(\frac{\nu}{bk}\right)^{\lambda/\nu} e^{kg_2(\alpha)}, \quad \text{as } k \to \infty. \tag{5.10}$$

LEMMA 5.6: *If $h_0(\theta) > 0$, $\forall \theta \in \Theta$, and $y > \mu(\theta_1)$, then the following asymptotic relations hold.*

1. *For $\alpha = 1$,*

$$\bar{c}(k,y;\alpha=1) \sim \frac{h_0(\theta_1)\psi''(\theta_1)}{(y-\mu(\theta_1))^2k^2}, \qquad (k\to\infty). \tag{5.11}$$

2. *For $\alpha = 2$,*
 (a) *If $\mu(\theta_1) < y < \mu(\bar{\theta})$, then*

$$\bar{c}(k,y;\alpha=2) \sim \delta(\theta^*(y))h_0(\theta^*(y))e^{kI(\theta^*(y),\theta_1)}\sqrt{\frac{2\pi}{\psi''(\theta^*(y))k}}, \qquad (k\to\infty). \tag{5.12}$$

(b) *If* $y = \mu(\bar{\theta})$, *then*

$$\bar{c}(k, y; \alpha = 2) \sim \delta(\bar{\theta}) h_0(\bar{\theta}) e^{kI(\bar{\theta}, \theta_1 | y)} \sqrt{\frac{\pi}{\psi''(\bar{\theta}) k}}, \qquad (k \to \infty). \tag{5.13}$$

(c) *If* $y > \mu(\bar{\theta})$, *then*

$$\bar{c}(k, y; \alpha = 2) \sim \frac{\delta(\bar{\theta}) h_0(\bar{\theta})}{y - \mu(\bar{\theta})} \frac{e^{kI(\bar{\theta}, \theta_1 | y)}}{k}, \qquad (k \to \infty). \tag{5.14}$$

LEMMA 5.7: *If* $h_0(\theta) > 0$, $\forall \theta \in \Theta$, *and* $y \leq \mu(\theta_1)$, *then the following asymptotic relations hold.*

1. *For* $\alpha = 1$,

(a) *If* $y < \mu(\underline{\theta})$, *then*

$$\bar{c}(k, y; \alpha = 1) \sim \frac{\delta(\underline{\theta}) h_0(\underline{\theta})}{y - \mu(\underline{\theta})} \frac{e^{kI(\underline{\theta}, \theta_1 | y)}}{k}, \qquad (k \to \infty). \tag{5.15}$$

(b) *If* $y = \mu(\underline{\theta})$, *then*

$$\bar{c}(k, y; \alpha = 1) \sim -\delta(\underline{\theta}) h_0(\underline{\theta}) e^{kI(\underline{\theta}, \theta_1 | y)} \sqrt{\frac{\pi}{\psi''(\underline{\theta}) k}}, \qquad (k \to \infty). \tag{5.16}$$

(c) *If* $\mu(\underline{\theta}) < y < \mu(\theta_1)$, *then*

$$\bar{c}(k, y; \alpha = 1) \sim -\delta(\theta^*(y)) h_0(\theta^*(y)) e^{kI(\theta^*(y), \theta_1)} \sqrt{\frac{2\pi}{\psi''(\theta^*(y)) k}}, \qquad (k \to \infty). \tag{5.17}$$

(d) *If* $y = \mu(\theta_1)$, *then*

$$\bar{c}(k, y; \alpha = 1) \sim \frac{h_0(\theta_1)}{k}, \qquad (k \to \infty). \tag{5.18}$$

2. *For* $\alpha = 2$,

(a) *If* $y < \mu(\theta_1)$, *then*

$$\bar{c}(k, y; \alpha = 2) \sim \frac{h_0(\theta_1) \psi''(\theta_1)}{(y - \mu(\theta_1))^2 k^2}, \qquad (k \to \infty). \tag{5.19}$$

(b) *If* $y = \mu(\theta_1)$, *then*

$$\bar{c}(k, y; \alpha = 2) \sim \frac{h_0(\theta_1)}{k}, \qquad (k \to \infty). \tag{5.20}$$

Remark 5.8: In Lemmata 5.6 and 5.7 we made the assumption that the prior p.d.f. h_0 is positive in the whole parameter space $\Theta = [\underline{\theta}, \bar{\theta}]$. This ensures that the values for which the log-likelihood ratio attains its maximum value in the integration region are independent from $h_0(\theta)$. When this assumption is dropped, the same line of argument remains valid. However, the expansion of the integrals becomes more tedious, because one must consider separately cases such as $h_0(\theta) = 0$, for $\theta \leq \theta_1 + \epsilon$, $\theta \geq \theta_1 - \epsilon$, or $\theta_1 - \epsilon \leq \theta \leq \theta_1 + \epsilon$. According to each individual case examined, one must integrate in a neighborhood of a value θ, which is closest to the maximizing value and has positive prior p.d.f. The corresponding asymptotic expressions cannot be given in advance for the general case, but they can be derived following the same procedure as in the proof of the Lemma 5.6, for every possible prior p.d.f.

The next auxiliary result is immediate from Lemma 5.6. We state it separately because the expressions involved are used in the proof of Theorem 5.12.

LEMMA 5.9: *If $h_0(\theta) > 0$, $\forall \theta \in \Theta$, and $\mu(\theta_1) < y < \mu(\bar{\theta})$, then the following asymptotic relations hold, as $k \to \infty$:*

$$\frac{\bar{c}(k, y; \alpha = 1)}{\bar{c}(k, y; \alpha = 2)} \sim \frac{2h_0(\theta_1)\sqrt{\psi''(\theta^*(y))}}{(y - \mu(\theta_1))^2 \delta(\theta^*(y)) h_0(\theta^*(y)) k^{3/2} e^{kI(\theta^*(y), \theta_1)}} \tag{5.21}$$

and

$$\frac{(\bar{c}(k, y; \alpha = 1))^2}{\bar{c}(k, y; \alpha = 2)} \sim \frac{4h_0^2(\theta_1)\sqrt{\psi''(\theta^*(y))}}{(y - \mu(\theta_1))^4 \delta(\theta^*(y)) h_0(\theta^*(y)) k^{7/2} e^{kI(\theta^*(y), \theta_1)}}. \tag{5.22}$$

The next proposition is used in the proof of Theorem 5.11.

Proposition 5.10: For each state (n, k, y) with sufficiently large n there exists a scheduling rule $\tau^0 = \tau^0(n, k, y)$, which processes at least one job from type E_2 and its total regret $W(\tau^0)$ (transformed according to Proposition 3.10), satisfies

$$W(\tau^0) < \phi^*(n, k, y), \tag{5.23}$$

where

$$\phi^*(n, k, y) = 2\sqrt{A(n + k)\bar{c}(k, y; \alpha = 2)} \tag{5.24}$$

and A is the constant defined in Lemma 5.3.

PROOF: For each (n, k, y) define a class of scheduling rules $\tau(i) = \tau(i; n, k, y)$ as follows. Perform a fixed number $i\,(\leq n)$ of type E_2 jobs, and then the remaining $n - i$ before or after those from E_1, according to whether

$$\begin{aligned} &\bar{c}(k + i, m(k, y, y_{2,k+1}, \ldots, y_{2,k+i}; \alpha = 1) \\ &\quad \geq \bar{c}(k + i, m(k, y, y_{2,k+1}, \ldots, y_{2,k+i}; \alpha = 2) \end{aligned} \tag{5.25}$$

or

$$\bar{c}(k+i, m(k, y, y_{2,k+1}, \ldots, y_{2,k+i}; \alpha = 1)$$
$$< \bar{c}(k+i, m(k, y, y_{2,k+1}, \ldots, y_{2,k+i}; \alpha = 2), \tag{5.26}$$

respectively, where $m(k, y, y_{2,k+1}, \ldots, y_{2,k+i})$ denotes the new average after the i additional outcomes

$$m(k, y, y_{2,k+1}, \ldots, y_{2,k+i}) = \frac{ky}{k+i} + \frac{y_{2,k+1} + \cdots + y_{2,k+i}}{k+i}. \tag{5.27}$$

Because of the martingale property Eq. (3.52) of $\bar{c}(k, y; \alpha = 2)$, the total regret corresponding to rule τ, according to the cost structure and dynamics of Proposition 3.10, can be written as follows

$$W^{\tau(i)}(n, k, y) = i\bar{c}(k, y; \alpha = 2) + (n-i) E_{f(\cdot|\theta_1)}$$
$$\times [\gamma(k+i, m(k, y, Y_{2,k+1}, \ldots, Y_{2,k+i}))]. \tag{5.28}$$

Applying Lemma 5.3 we get

$$W^{\tau(i)}(n, k, y) < \phi(i; n, k, y), \qquad i = 0, 1, \ldots, n, \tag{5.29}$$

where

$$\phi(i; n, k, y) = i\bar{c}(k, y; \alpha = 2) + (n-i)\frac{A}{k+i}. \tag{5.30}$$

If we consider the extension of the function $\phi(i)$ in Eq. (5.30) in the real domain,

$$\phi(i; n, k, y) = i\bar{c}(k, y; \alpha = 2) + (n-i)\frac{A}{k+i}, \qquad 0 \le i \le n,\ i \in \mathbb{R}, \tag{5.31}$$

and differentiate with respect to i

$$\phi'(i) = \bar{c}(k, y; \alpha = 2) - A\frac{k+n}{(k+i)^2}, \tag{5.32}$$

$$\phi''(i) = 2A\frac{k+n}{(k+i)^3} > 0, \tag{5.33}$$

we see that $\phi(i)$ is convex. We also have

$$\phi'(0) = \bar{c}(k, y; \alpha = 2) - A\frac{k+n}{k^2}, \tag{5.34}$$

which is negative for n sufficiently large, and

$$\phi'(n) = \bar{c}(k, y; \alpha = 2) - \frac{A}{k+n}, \tag{5.35}$$

which is positive also for n sufficiently large. This means that, for fixed n, k, y, ϕ attains its minimum for $i = i^*$, $1 \le i^* < n$, such that $\phi'(i^*) = 0$; i.e.,

$$i^* = \sqrt{\frac{A}{\bar{c}(k, y; \alpha = 2)}(k+n)} - k. \tag{5.36}$$

Let $\lceil i^* \rceil = \min\{i \in \mathbb{N} : i \ge i^*\}$. Then $\lceil i^* \rceil < i^* + 1$, and, because ϕ is convex,

$$\phi(\lceil i^* \rceil) < \phi(i^* + 1). \tag{5.37}$$

But

$$\begin{aligned}\phi(i^*+1) &= (i^*+1)\bar{c}(k, y; \alpha = 2) + (n - i^* - 1)\frac{A}{k + i^* + 1} \\ &\le (i^*+1)\bar{c}(k, y; \alpha = 2) + (n - i^*)\frac{A}{k + i^*} \\ &= 2\sqrt{A(n+k)\bar{c}(k, y; \alpha = 2)} - (k-1)\bar{c}(k, y; \alpha = 2) - A \\ &< 2\sqrt{A(n+k)\bar{c}(k, y; \alpha = 2)} = \phi^*(n, k, y).\end{aligned} \tag{5.38}$$

Combining the preceding inequalities we have

$$W^{\tau(\lceil i^* \rceil)}(n, k, y) < \phi^*(n, k, y). \tag{5.39}$$

Taking $\tau^0(n, k, y) = \tau(\lceil i^*(n, k, y) \rceil)$ completes the proof. ■

We can now state the two main results of this section. Let

$$S_n = \{(k, y) : \pi^*(n, k, y) = 1\}. \tag{5.40}$$

THEOREM 5.11: *Under the exponential family assumption, when $n \to \infty$*

$$\underline{S}_n \subset S_n \subset \bar{S}_n, \tag{5.41}$$

where

$$\underline{S}_n = \{(k, y) : n\bar{c}(k, y; \alpha = 1) < \bar{c}(k, y; \alpha = 2)\}, \tag{5.42}$$

$$\bar{S}_n = \{(k, y) : n\bar{c}(k, y; \alpha = 1) < 2\sqrt{A(n+k)\bar{c}(k, y; \alpha = 2)}\}. \tag{5.43}$$

PROOF: From Eq. (3.46) and Lemma 3.11(a) it follows that $u(n, k, y) > 0$. So it is easy to see that if $n\bar{c}(k, y; \alpha = 1) < \bar{c}(k, y; \alpha = 2)$, then it is optimal to switch to type E_1, i.e., $\pi^*(n, k, y) = 1$. This proves the left inclusion relationship in Eq. (5.41).

To prove the right relationship, we get from Proposition 5.10 that, if $n\bar{c}(k, y; \alpha = 1) \ge \phi^*$, it is by no way optimal to switch to E_1, since applying rule $\tau^0(n, k, y)$ is a better policy. So $n\bar{c}(k, y; \alpha = 1) \ge \phi^*$ implies that $\pi^*(n, k, y) = 2$, or equivalently $\pi^*(n, k, y) = 1$ implies that $n\bar{c}(k, y; \alpha = 1) < \phi^*$, and the proof is complete. ■

Note that the arguments used in the proof of Theorem 5.11 are related to those used in the proof of Theorem 5.1 of Whittle [14, p. 252].

Based on Theorem 5.11 we now derive an asymptotic approximation of the optimal policy $\pi^*(n,k,y)$ as $n \to \infty$. Let

$$G(y,\theta_1) = \begin{cases} \mathbf{I}(\theta^*(y),\theta_1), & \text{if } \mu(\theta_1) < y < \mu(\bar{\theta}) \\ I(\bar{\theta},\theta_1 | y), & \text{if } y \geq \mu(\bar{\theta}) \end{cases}. \tag{5.44}$$

THEOREM 5.12: *If $h_0(\theta) > 0$, $\forall\theta \in \Theta$, then there exist functions $\epsilon_1(n,y)$, $\epsilon_2(n,y)$, $\epsilon(n,y)$, such that when $n \to \infty$*

$$\pi^*(n,k,y) = \begin{cases} 1, & \text{if } y > \mu(\theta_1) \quad \text{and} \quad kG(y,\theta_1) > \log n - \epsilon(n,y) \\ 2, & \text{otherwise} \end{cases}, \tag{5.45}$$

and

$$\epsilon_1(n,y) = \begin{cases} \frac{3}{2}\log\log n - o(\log\log n), & \text{if } \mu(\theta_1) < y < \mu(\bar{\theta}) \\ \log\log n - o(\log\log n), & \text{if } y \geq \mu(\bar{\theta}) \end{cases},$$

$$\epsilon_2(n,y) = \begin{cases} \frac{7}{2}\log\log n - o(\log\log n), & \text{if } \mu(\theta_1) < y < \mu(\bar{\theta}) \\ 2\log\log n - o(\log\log n), & \text{if } y \geq \mu(\bar{\theta}) \end{cases},$$

$$\epsilon_1(n,y) < \epsilon(n,y) < \epsilon_2(n,y).$$

Remark 5.13: An interpretation of Eq. (5.45) is due at this point, in relation to the results of Corollary 4.2. As discussed in Remark 4.3(b) the optimal policy for a finite horizon suggests that a nonnegative quantity should be subtracted from the Bayes estimate of $\mu(\theta_2)$ and the result be compared to $\mu(\theta_1)$. This correction has the effect of allowing jobs from E_2 to be further tried, even if their current estimate is unfavorable, so that more information about their distribution is acquired. The same is accomplished with the policy described in Eq. (5.45). Indeed, even if the observed average job length of E_2 is greater than $\mu(\theta_1)$, to switch to E_1 the additional condition $kG(y,\theta_1) > \log n - \epsilon(n,y)$ must be satisfied. Therefore, in this case, too, jobs from E_2 are given priority over those from E_1, for the purpose of learning.

PROOF OF THEOREM 5.12: We will show that, for large n, there exist sets $\bar{S}'_n$ and functions $\epsilon_1(n,y)$ and $\epsilon_2(n,y)$, such that

(i) $\underline{S}_n = \{(k,y) : y > \mu(\theta_1) \text{ and } kG(y,\theta_1) > \log n - \epsilon_1(n,y)\}$,

(ii) $S_n \subset \bar{S}'_n \subset \bar{S}_n$, and $\bar{S}'_n = \{(k,y) : y > \mu(\theta_1) \text{ and } kG(y,\theta_1) > \log n - \epsilon_2(n,y)\}$.

Then the assertion will follow in view of Theorem 5.11.

First consider $\underline{S}_n$, which is described by the following relation:

$$\frac{\bar{c}(k,y;\alpha=1)}{\bar{c}(k,y;\alpha=2)} < \frac{1}{n}. \tag{5.46}$$

The proof of Theorem 5.12(i) will follow four steps.

Step 1: For any fixed y, for Eq. (5.46) to hold when $n \to \infty$, it is necessary that $k \to \infty$. To see this, note that Eq. (5.46) holds when at least one of the following conditions is satisfied:

$$\bar{c}(k,y;\alpha=1) \to 0 \quad \text{or} \quad \bar{c}(k,y;\alpha=2) \to \infty,$$

because both quantities are positive. Any one of the preceding implies that $k \to \infty$.

Step 2: For Eq. (5.46) to hold when $n \to \infty$, it is also necessary that $y > \mu(\theta_1)$, because from Lemmata 5.6 and 5.7 we see that, when $k \to \infty$, the values of y for which the preceding ratio tends to 0 are those included in the range $y > \mu(\theta_1)$.

Step 3: For Eq. (5.46) to hold when $n \to \infty$, it is necessary and sufficient that

$$y > \mu(\theta_1) \quad \text{and} \quad k > \frac{1}{G(y,\theta_1)} \log n - \epsilon_1(n,y),$$

where $\epsilon_1(n,y) = o(\log n)$, $\forall y$. To prove this claim, we take $y > \mu(\theta_1)$, which is necessary from the previous step, and consider three cases corresponding to Lemmata 5.6.2(a)–5.6.2(c).

In Lemma 5.6.2(a), for which $\mu(\theta_1) < y < \mu(\bar{\theta})$, the left-hand side of Eq. (5.46) is asymptotically approximated by Eq. (5.21). From Eq. (5.21) we have that, for any fixed value of y, Eq. (5.46) is an inequality of the form Eq. (5.9) with $x = k$, $\alpha = \frac{3}{2}$, and $\lambda = I(\theta^*(y), \theta_1)$. Hence, there exists a function $\epsilon_1(n,y)$ such that, when $n \to \infty$, Eq. (5.46) is satisfied for

$$kI(\theta^*(y),\theta_1) > \log n - \epsilon_1(n,y), \tag{5.47}$$

and $\epsilon_1(n,y) = \frac{3}{2}\log(\log n) - o(\log \log n) = o(\log n)$, $\forall y \in (\mu(\theta_1), \mu(\bar{\theta}))$.

For both Lemma 5.6.2(b) and Lemma 5.6.2(c) we can show in the same way that the asymptotic solution of Eq. (5.46) is

$$kI(\bar{\theta},\theta_1 | y) > \log n - \epsilon_1(n,y), \tag{5.48}$$

where $\epsilon_1(n,y) = \log(\log n) - o(\log\log n) = o(\log n)$, $\forall y \geq \mu(\bar{\theta})$.

Step 4: The proof of Theorem 5.12(i) follows by combining the three cases considered in Step 3.

We now turn to the inequality which defines the set $\bar{S}_n$ in Eq. (5.43) and which can be rewritten as

$$\frac{(\bar{c}(k,y;\alpha=1))^2}{\bar{c}(k,y;\alpha=2)} < \frac{4A(n+k)}{n^2}. \tag{5.49}$$

Consider also the same relation with equality:

$$\frac{(\bar{c}(k,y;\alpha=1))^2}{\bar{c}(k,y;\alpha=2)} = \frac{4A(n+k)}{n^2}. \tag{5.50}$$

The proof of the second claim will follow six steps.

Step 1: For any fixed y, for Eq. (5.49) to hold when $n \to \infty$, it is necessary that $k \to \infty$, for the same reason as in Step 1 of the proof of the first claim.

Step 2: When $n \to \infty$, for Eq. (5.49) to hold, it is necessary that $y \geq \mu(\theta_1)$. Indeed, from Lemma 5.7 we see that, when $y < \mu(\theta_1)$ and $k \to \infty$, the left-hand side of Eq. (5.49) goes to ∞ with exponential rate in k (Eqs. (5.15)–(5.17) and (5.19)) and, thus, cannot be less than the linear function of k in the right-hand side.

Step 3: Although Step 2 implies that there may be points $(k,y) \in \bar{S}_n$ such that $y = \mu(\theta_1)$, these points need not be considered for the approximation of the stopping set S_n. Indeed, let $\bar{S}_n' = \bar{S}_n - \{(k,y) \in \bar{S}_n, y = \mu(\theta_1)\}$. It is proved next that $S_n \subset \bar{S}_n'$.

When $y = \mu(\theta_1)$ and $k \to \infty$, it is not uniquely optimal to switch to E_1. This property holds because from Lemma 5.7 we see that in this case $\bar{c}(k,y;\alpha=1)$ and $\bar{c}(k,y;\alpha=2)$ are asymptotically equal. Thus, performing all jobs from E_2 first has the same total expected regret as switching to E_1, therefore continuing for one more step and then optimally will be at least as good.

Step 4: When $n \to \infty$, $k \to \infty$, and $y > \mu(\theta_1)$, Eq. (5.50), considered as an equation in k, has a unique solution $k(n,y)$ and Eq. (5.49) holds for $k > k(n,y)$. Indeed, in this case the left-hand side takes the asymptotic form in Eq. (5.22), which is decreasing in k, while the right-hand side increases from 0 to ∞.

Step 5: For $y > \mu(\theta_1)$,

$$k(n,y)G(y,\theta_1) = \log n - \epsilon_2(n,y), \tag{5.51}$$

where $\epsilon_2(n,y)$ is of the form stated in the assertion of Theorem 5.12. To prove Eq. (5.51), we will show that $k_0(n,y) = 1/G(y,\theta_1)(\log n - \epsilon_2(n,y))$ satisfies Eq. (5.47). Then Eq. (5.51) will follow from the uniqueness of the solution discussed in Step 4.

First note that $(k_0(n,y))/n \to 0$, as $n \to \infty$; thus, the right-hand side of Eq. (5.50) is of the same order as $1/n$.

Also, because $k_0(n,y) \to \infty$ as $n \to \infty$, the left-hand side of Eq. (5.50) is asymptotically approximated by expressions corresponding to cases 2(a)–2(c) of Lemma 5.6.

In case 2(a) the left-hand side of Eq. (5.47) takes the form described in Eq. (5.22). Since in this case

$$k_0(n,y) = \frac{1}{\mathbf{I}(\theta^*(y),\theta_1)}(\log n - \epsilon_2(n,y)),$$

with $\epsilon_2(n,y) = \frac{3}{2}\log\log n - o(\log\log n)$, $\forall y \in (\mu(\theta_1), \mu(\bar{\theta}))$, we get that $k_0(n,y)$ as already defined satisfies Eq. (5.47).

In cases 2(b) and 2(c) the corresponding expressions

$$k_0(n,y) = \frac{1}{I(\bar{\theta},\theta_1|y)}(\log n - \epsilon_2(n,y)), \tag{5.52}$$

with $\epsilon_2(n,y) = 2\log\log n - o(\log\log n)$, $\forall y \geq \mu(\bar{\theta})$, can be shown to satisfy Eq. (5.50) with similar reasoning. Thus, Eq. (5.51) follows.

Step 6: The proof of the second claim and the whole theorem is completed by combining the cases of Step 5. ∎

Remark 5.14:

(a) From Theorem 5.12 we can see that, for n large enough, it is never optimal to stop performing jobs from E_2 when $y \leq \mu(\theta_1)$, even if the current information on θ_2 indicates that $E_H[\mu(\theta_2)] > \mu(\theta_1)$.

(b) The asymptotic policy derived in Theorem 5.12 is independent of the initial prior p.d.f. h_0, when $h_0(\theta) > 0$, $\forall\theta \in \Theta$. If the last condition fails to hold, we can still show that the general form of the asymptotically optimal policy is

$$\pi_1(n,k,y) = \begin{cases} 1, & \text{if } y > \mu(\xi) \quad \text{and} \quad kI(\tau,\theta_1|y) > \log n - \epsilon(n,y) \\ 2, & \text{otherwise} \end{cases}, \tag{5.53}$$

where

$$\xi = \sup\{\theta \leq \theta_1,\ h_0(\theta) > 0\} \tag{5.54}$$

and τ is the value of θ, which maximizes $I(\theta,\theta_1|y)$ in the support of the prior p.d.f. This can be shown by following the same line of reasoning as that for the proof of Theorem 5.12, the difference being that the asymptotic expressions of Lemmata 5.6 and 5.7 need to be modified (see also Remark 5.8).

(c) Consider the following policy:

$$\pi^0(n,k,y) = \begin{cases} 1, & \text{if } y > \mu(\theta_1) \quad \text{and} \quad kG(y,\theta_1) > \log n \\ 2, & \text{otherwise} \end{cases}. \tag{5.55}$$

π^0 is an approximation to policy π_1, because it increases the stopping threshold for $kG(y,\theta_1)$ by an amount equal to $\epsilon(n,y)$, which is of order $o(\log n)$. The ratio of the approximation error $\epsilon(n,y)$ to the correct value of the threshold, $\log n - \epsilon(n,y)$ tends to zero for large n.

We conclude this section showing some properties of the function $G(y,\theta_1)$ and a resulting graphical representation of the optimal policy approximation π^0.

Proposition 5.15:

(a) $G(y, \theta_1)$ is continuous, continuously differentiable, and increasing in y, $\forall \theta_1$, $y > \mu(\theta_1)$.

(b) $G(y, \theta_1)$ is decreasing in θ_1, $\forall y$, $\mu(\theta_1) < y$. If $y < \mu(\bar{\theta})$, then $G(y, \theta^*(y)) = 0$. If $y \geq \mu(\bar{\theta})$, then $G(y, \bar{\theta}) = 0$.

Figure 1 describes policy π^0 in terms of y, for some fixed value of θ_1. From Proposition 5.15(a) and Figure 1 we see that switching to E_1 is required when $y > y^0(n,k)$, where $y^0(n,k)$ is an asymptotic approximation to the cutting points $y_n(k)$ defined in Theorem 4.1.

$$y^0(n,k) = \inf\left\{y > \mu(\theta_1), G(y, \theta_1) > \frac{\log n}{k}\right\}. \tag{5.56}$$

Figures 2 and 3 describe the π^0 in terms of θ_1, for some fixed value of y below and above $\mu(\bar{\theta})$, respectively. If we define

$$\theta_1'(n,k,y) = \sup\left\{\theta_1 \leq \theta^*(y), G(y, \theta) > \frac{\log n}{k}\right\}, \tag{5.57}$$

then from Proposition 5.15(b) we see that an alternative interpretation is to switch to E_1 if $\theta_1 < \theta_1'(n,k,y)$, for given values of n, k, y.

The preceding discussion makes clear some interesting properties of the asymptotic policy of Theorem 5.12 that are intuitively expected. Namely, at each step, if the average length of the scheduled jobs from the unknown type E_2 does not exceed the expected job length from the known type E_1, i.e., $y \leq \mu(\theta_1)$, we continue scheduling jobs from E_2. Otherwise, we base our decision on the quantity $G(y, \theta_1)$, which represents in some sense the estimated distance

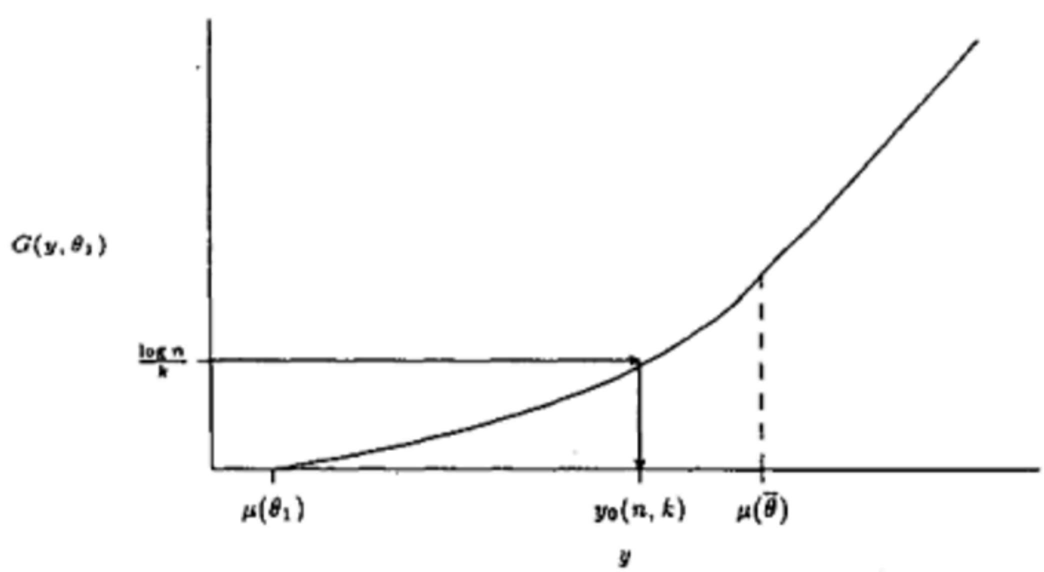

FIGURE 1. Policy π^0 in terms of y for some fixed value of θ_1.

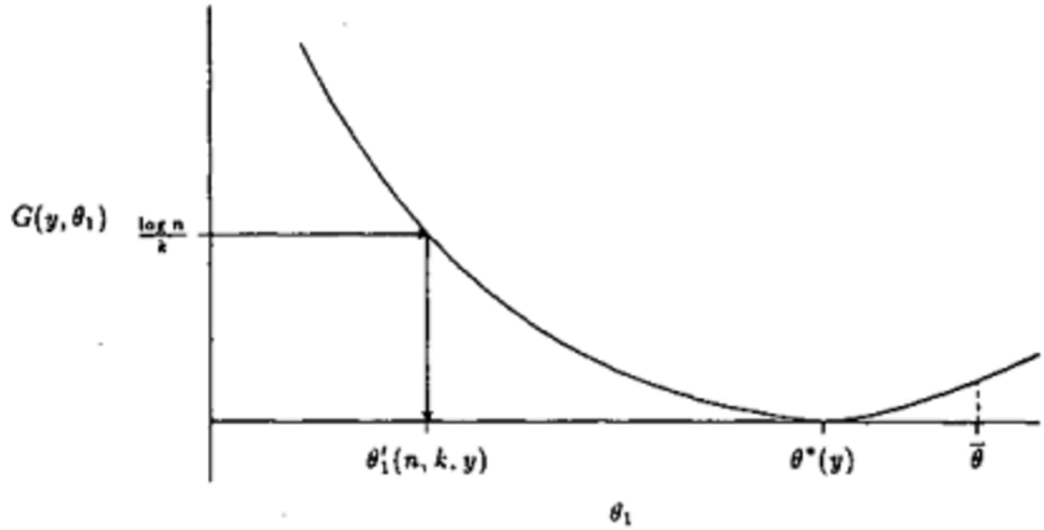

FIGURE 2. Policy π^0 in terms of θ_1 for some fixed value of y below $\mu(\bar{\theta})$.

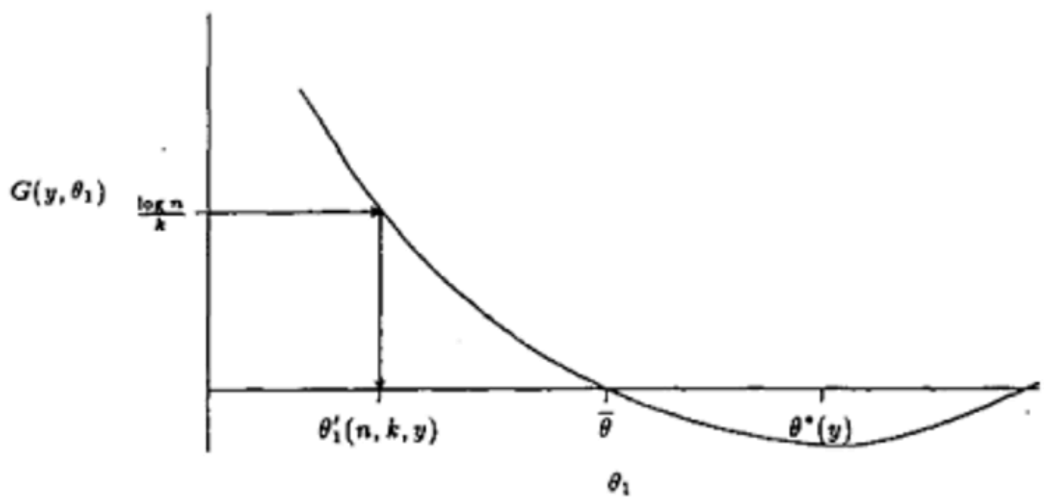

FIGURE 3. Policy π^0 in terms of θ_1 for some fixed value of y above $\mu(\bar{\theta})$.

between the distributions of the two types in the parameter space (Kullback and Leibler [10]) and provides a measure of our confidence that the true value of θ_2 is really larger than θ_1, when the sample average we have observed is larger than $\mu(\theta_1)$.

The policy described in Eq. (5.45) is analogous to that developed in the fundamental papers by Lai and Robbins [13] and Lai [12]. They refer to the general case where there are i unknown experiments to be compared, with the objective of maximizing the total expected reward for a large planning horizon. Their asymptotically optimal policy is based on the use of upper confidence bounds (which essentially estimate the unknown parameters) in the following way. If x_j is the average of T_j successive observations from experiment E_j, $j = 1, \ldots, i$, the upper confidence bound is defined as

$$U_j(T_j, x_j) = \inf\left\{\theta > \theta_{x_j}, \mathrm{I}(\theta_{x_j}, \theta) > \frac{g(T_j/N)}{T_j}\right\}, \tag{5.58}$$

where θ_{x_j} is the maximum likelihood estimate for θ_j given (T_j, x_j), and g is a function that satisfies certain assumptions (cf. Lai [12]), among which is that $g(t) \sim \log t^{-1}$ when $t \to 0$. Then the policy suggests sampling from the experiment with the largest upper confidence bound.

Here, we have seen that for every state (n, k, y) there is a number $\theta_1'(n, k, y)$, defined in Eq. (5.57), such that if the known parameter θ_1 of E_1 is greater than $\theta_1'(n, k, y)$, then it is optimal to continue from E_2; otherwise, it is optimal to switch. Hence, $\theta_1'(n, k, y)$ plays essentially the same role as the upper confidence bounds, if one considers the fact that $T_j/N \to 0$ in Eq. (5.58). The difference is that in our context θ_1' is a lower instead of upper confidence bound.

Acknowledgments

The authors are thankful to an anonymous referee for a number of useful suggestions that improved the presentation of the paper. The second author wishes to express his deep gratitude to Professor H. Robbins for introducing him to this area of research, and many useful discussions. H. Robbins, in 1991, was a source of inspiration and encouragement.

References

1. Agrawal, R., Hedge, M., & Teneketzis, D. (1988). Asymptotically efficient adaptive allocation rules for the multi-armed bandit problem with switching cost. *IEEE Transactions on Automatic Control* AC-33: 899–906.
2. Baker, K.R. (1985). *Introduction to sequencing and scheduling.* New York: John Wiley.
3. Berry, D.A. & Fristedt, B. (1985). *Bandit problems: Sequential allocation of experiments.* New York: Chapman and Hall.
4. Bertsekas, D.P. (1976). *Dynamic programming and stochastic control.* New York: Academic Press.
5. Bradt, R.N., Johnson, S.M., & Karlin, S. (1956). On sequential designs for maximizing the sum of n observations. *Annals of Mathematical Statistics* 27: 1060–1074.
6. Cox, D.R. & Hinkley, D.V. (1974). *Theoretical statistics.* New York: Chapman and Hall.
7. Dynkin, E.B. & Yushkevich, A.A. (1979). *Controlled Markov processes.* New York: Springer-Verlag.
8. Erdélyi, A. (1956). *Asymptotic expansions.* New York: Dover.
9. Gittins, J.C. & Glazebrook, K.D. (1977). On Bayesian models in stochastic scheduling. *Journal of Applied Probability* 14: 556–565.
10. Kullback, S. & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics* 22: 79–86.
11. Kumar, P.R. & Varayia, P. (1986). *Stochastic systems: Estimation, identification and adaptive control.* Englewood Cliffs, NJ: Prentice-Hall.
12. Lai, T.L. (1987). Adaptive treatment allocation and the multi-armed bandit problem. *Annals of Statistics* 15: 1091–1114.
13. Lai, T.L. & Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Advances in Applied Math* 6: 4–22.
14. Whittle, P. (1982). *Optimization over time,* Vol. 2. New York: John Wiley.

APPENDIX

PROOF OF PROPOSITION 3.2: We will show how Eqs. (3.9) and (3.10) can be derived from Eqs. (3.4)–(3.6). Subtracting $E_H F^*(\theta; n_1, n_2)$ from both sides of Eq. (3.4), we get

$$U(n_1, n_2, H) = \min\{U_1(n_1, n_2, H), U_2(n_1, n_2, H)\}, \tag{A.1}$$

where

$$U_i(n_1, n_2, H) = V_i(n_1, n_2, H) - E_H F^*(\theta; n_1, n_2), \qquad i = 1,2, \tag{A.2}$$

$$V_1(n_1, n_2, H) = (n_1 + n_2) r(H; \alpha = 1) + V(n_1 - 1, n_2, H), \tag{A.3}$$

$$V_2(n_1, n_2, H) = (n_1 + n_2) r(H; \alpha = 2) + E_{f(\cdot|H)} V\big(n_1, n_2 - 1, \bar{H}(\cdot|X_2)\big). \tag{A.4}$$

From Eq. (2.8) we can obtain the following recursive relations for $F^*(\theta; n_1, n_2)$:

$$F^*(\theta; n_1, n_2) - F^*(\theta; n_1 - 1, n_2) = n_1 \mu(\theta_1) + n_2 \mu^*(\theta), \tag{A.5}$$

$$F^*(\theta; n_1, n_2) - F^*(\theta; n_1, n_2 - 1) = n_2 \mu(\theta) + n_1 \mu^*(\theta). \tag{A.6}$$

Thus,

$$\begin{aligned} U_1(n_1, n_2, Y) &= V_1(n_1, n_2, H) - E_H F^*(\theta; n_1, n_2) \\ &= (n_1 + n_2)\mu(\theta_1) + V(n_1 - 1, n_2, H) \\ &\quad - E_H(n_1 \mu(\theta_1) + n_2 \mu^*(\theta) + F^*(\theta; n_1 - 1, n_2)) \\ &= n_2 c(H; \alpha = 1) + U(n_1 - 1, n_2, H), \end{aligned} \tag{A.7}$$

where

$$\begin{aligned} c(H; \alpha = 1) &= \mu(\theta_1) - E_H[\mu^*(\theta)] \\ &= \mu(\theta_1) - \int_{\mu(\theta) < \mu(\theta_1)} \mu(\theta)\, dH(\theta) - \int_{\mu(\theta) \geq \mu(\theta_1)} \mu(\theta_1)\, dH(\theta) \\ &= \int_{\mu(\theta) < \mu(\theta_1)} (\mu(\theta_1) - \mu(\theta))\, dH(\theta). \end{aligned} \tag{A.8}$$

Similarly,

$$\begin{aligned} U_2(n_1, n_2, H) &= V_2(n_1, n_2, H) - E_H F^*(\theta; n_1, n_2) \\ &= (n_1 + n_2) E_H \mu(\theta) + E_{f(\cdot|H)} V\big(n_1, n_2 - 1, \bar{H}(\cdot|X_2)\big) \\ &\quad - E_H(n_2 \mu(\theta) + n_1 \mu^*(\theta) + F^*(\theta; n_1, n_2 - 1)). \end{aligned} \tag{A.9}$$

Applying Lemma 2.1 with $g(\theta) = F^*(\theta; n_1, n_2 - 1)$, we get

$$U_2(n_1, n_2, H) = n_1 c(H; \alpha = 2) + E_{f(\cdot|H)} U\big(n_1, n_2 - 1, \bar{H}(\cdot|X_2)\big), \tag{A.10}$$

where

$$c(H;\alpha=2) = E_H[\mu(\theta)] - E_H[\mu^*(\theta)]$$

$$= \int_\Theta \mu(\theta)\,dH(\theta|(k,y)) - \int_{\mu(\theta)<\mu(\theta_1)} \mu(\theta)\,dH(\theta) - \int_{\mu(\theta)\ge\mu(\theta_1)} \mu(\theta_1)\,dH(\theta)$$

$$= \int_{\mu(\theta)>\mu(\theta_1)} (\mu(\theta)-\mu(\theta_1))\,dH(\theta). \tag{A.11}$$

■

LEMMA A.1: *For every function $g(k,y)$ such that $E_{f(\cdot|(k,y))}(|g(k,X_2)|) < \infty$ we have that*

$$E_{f(\cdot|(k,y))}[g(k+1,m(k,y,X_2))]$$
$$= \frac{1}{\Lambda(k,y)} E_{f(\cdot|\theta_1)}[g(k+1,m(k,y,X_2))\Lambda(k+1,m(k,y,X_2))]. \tag{A.12}$$

PROOF: From Eqs. (2.2), (3.43), and (3.45)

$$f(x|(k,y)) = \frac{1}{\Lambda(k,y)} \int_\Theta f(x|\theta_1)e^{I(\theta,\theta_1|x)}e^{kI(\theta,\theta_1|y)}\,dH_0(\theta)$$

$$= \frac{\Lambda(k+1,m(k,y,x))}{\Lambda(k,y)} f(x|\theta_1). \tag{A.13}$$

Thus,

$$E_{f(\cdot|(k,y))}[g(k+1,m(k,y,X_2))]$$

$$= \int g(k+1,m(k,y,x))f(x|(k,y))\nu(dx)$$

$$= \int g(k+1,m(k,y,x)) \frac{\Lambda(k+1,m(k,y,x))}{\Lambda(k,y)} f(x|\theta_1)\nu(dx)$$

$$= \frac{1}{\Lambda(k,y)} E_{f(\cdot|\theta_1)}[g(k+1,m(k,y,X_2))\Lambda(k+1,m(k,y,X_2))],$$

and the proof is complete. ■

PROOF OF PROPOSITION 3.10: Using Definition 3.8 and substituting Eq. (3.43) into Eq. (3.11)

$$c(k,y;\alpha=1) = c(H(\cdot|(k,y));\alpha=1)$$

$$= -\frac{1}{\Lambda(k,y)} \int_{\theta<\theta_1} \delta(\theta)e^{kI(\theta,\theta_1|y)}\,dH_0(\theta), \tag{A.14}$$

$$c(k,y;\alpha=2) = c(H(\cdot|(k,y));\alpha=2)$$

$$= \frac{1}{\Lambda(k,y)} \int_{\theta\ge\theta_1} \delta(\theta)e^{kI(\theta,\theta_1|y)}\,dH_0(\theta). \tag{A.15}$$

Also from Lemma A.1

$$E_{f(\cdot|(k,y))}[U(n-1,k+1,m(k,y,X_2))]$$
$$= \frac{1}{\Lambda(k,y)} E_{f(\cdot|\theta_1)}[U(n-1,k+1,m(k,y,X_2))\Lambda(k+1,m(k,y,X_2))]. \quad \textbf{(A.16)}$$

Thus, Eq. (3.35) is equivalent to

$$U(n,k,y) = \min\left\{n\frac{\bar{c}(k,y;\alpha=1)}{\Lambda(k,y)}, \frac{\bar{c}(k,y;\alpha=2)}{\Lambda(k,y)} + \frac{1}{\Lambda(k,y)}\right.$$
$$\left.\times E_{f(\cdot|\theta_1)}[U(n-1,k+1,m(k,y,X_2))\Lambda(k+1,m(k,y,X_2))]\right\}. \quad \textbf{(A.17)}$$

To complete the proof we only need to multiply both sides of Eq. (A.17) by $\Lambda(k,y)$ (where $\Lambda(k,y) > 0$). ■

PROOF OF LEMMA 3.11: Part (a) is immediate from the definition of $\bar{c}(k,y;\alpha)$. We will prove only the first assertion of part (b), because the other two are similar. From Eq. (3.44) we see that $l(\theta,\theta_1|y)$ is decreasing in y for $\theta < \theta_1$ and increasing in y for $\theta > \theta_1$. Also from Eq. (3.40), $\delta(\theta) < 0$ (>0) for $\theta < \theta_1$ ($\theta > \theta_1$). Therefore, $\delta(\theta)e^{kl(\theta,\theta_1|y)}$ is increasing in y for every $\theta \in \Theta$, $\theta \neq \theta_1$ and is equal to zero for $\theta = \theta_1$. The statement follows from this fact and the definition of $\bar{c}(k,y;\alpha=1)$.

Part (c) can be easily proved as follows. For $\alpha = 1$, we use Eqs. (3.45) and (3.49) to obtain

$$E_{\theta_1}[\bar{c}(k+1,m(k,y,X);\alpha=1)]$$
$$= \int \bar{c}(k+1,m(k,y,x);\alpha=1)f(x|\theta_1)\nu(dx)$$
$$= -\int_{\mathbb{R}}\int_{\theta<\theta_1} \delta(\theta)e^{(k+1)l(\theta,\theta_1|m(k,y,x))}\,dH_0(\theta)f(x|\theta_1)\nu(dx)$$
$$= -\int_{\mathbb{R}}\int_{\theta<\theta_1} \delta(\theta)e^{kl(\theta,\theta_1|y)+l(\theta,\theta_1|x)}\,dH_0(\theta)f(x|\theta_1)\nu(dx)$$
$$= -\int_{\theta<\theta_1} \delta(\theta)e^{kl(\theta,\theta_1|y)}\int_{\mathbb{R}} f(x|\theta)\nu(dx)\,dH_0(\theta)$$
$$= -\int_{\theta<\theta_1} \delta(\theta)e^{kl(\theta,\theta_1|y)}\,dH_0(\theta) = \bar{c}(k,y;\alpha=1). \quad \textbf{(A.18)}$$

The case $\alpha = 2$ can be proved similarly. ■

PROOF OF LEMMA 5.1: Equations (5.2) and (5.3) follow from Eqs. (5.1) and (3.28). Equation (5.4) is immediate from Eqs. (5.3) and (3.30). ■

PROOF OF LEMMA 5.2: Parts (a) and (b) follow from the expression for l in Eq. (3.44).

For part (c) we first note that every concave function has at most one maximum and at most two roots, lying on opposite sides with respect to its maximizing value. Hence,

$\forall x \in \mathbb{R}$, the equation $l(\theta, \theta_1 | x) = 0$, besides θ_1, has at most one more solution $\tilde{\theta}(x)$, possibly not in Θ, with the following property:

$$\tilde{\theta}(x) < \mu^{-1}(x) < \theta_1 \quad \text{if } x < \mu(\theta_1) \tag{A.19}$$

and

$$\theta_1 < \mu^{-1}(x) < \tilde{\theta}(x) \quad \text{if } x > \mu(\theta_1). \tag{A.20}$$

When $x > \mu(\bar{\theta}) > \mu(\theta_1)$ we get $\tilde{\theta}(x) > \mu^{-1}(x) > \bar{\theta} > \theta_1$, thus $l(\bar{\theta}, \theta_1 | x) > 0$. ∎

Proof of Lemma 5.3: We need to show that

$$\exists A > 0 : \gamma(k, y) < \frac{A}{k}, \forall k = 1, 2, \ldots, y \in \mathbb{R}. \tag{A.21}$$

It suffices to prove the following intermediate claim:

$$\exists A > 0 : \bar{c}(k, \mu(\theta_1); \alpha) < \frac{A}{k}, \forall k = 1, 2, \ldots, \alpha = 1, 2. \tag{A.22}$$

Indeed, suppose that Eq. (A.22) holds. Then we consider two cases.

Case a: $y \geq \mu(\theta_1)$. From Lemma 3.11(b),

$$\gamma(k, y) \leq \bar{c}(k, y; \alpha = 1) \leq \bar{c}(k, \mu(\theta_1); \alpha = 1) < \frac{A}{k}. \tag{A.23}$$

Case b: $y < \mu(\theta_1)$. In the same way,

$$\gamma(k, y) \leq \bar{c}(k, y; \alpha = 2) \leq \bar{c}(k, \mu(\theta_1); \alpha = 2) < \frac{A}{k}. \tag{A.24}$$

So $\gamma(k, y) < A/k$, $\forall k, y$, which proves the lemma.

Now we prove the claim stated in Eq. (A.22). Let $\bar{h}_0 = \sup_\Theta h_0(\theta)$. From Eq. (3.50),

$$\bar{c}(k, \mu(\theta_1); \alpha = 2) = \int_{\theta_1}^{\bar{\theta}} \delta(\theta) e^{k l(\theta, \theta_1 | \mu(\theta_1))} h_0(\theta)\, d\theta. \tag{A.25}$$

But $l(\theta, \theta_1 | \mu(\theta_1)) = (\theta - \theta_1)\mu(\theta_1) - (\psi(\theta) - \psi(\theta_1)) = -\mathbf{I}(\theta_1, \theta)$, and from Eq. (5.4)

$$-\zeta_2 \frac{(\theta - \theta_1)^2}{2} \leq l(\theta, \theta_1 | \mu(\theta_1)) \leq -\zeta_1 \frac{(\theta - \theta_1)^2}{2}. \tag{A.26}$$

From the mean value theorem

$$\delta(\theta) = \mu(\theta) - \mu(\theta_1) = \psi''(\xi)(\theta - \theta_1), \tag{A.27}$$

for some $\xi \in (\theta_1, \theta)$. So for $\theta \geq \theta_1$

$$\delta(\theta) \leq \zeta_2(\theta - \theta_1). \tag{A.28}$$

Now we use inequalities (A.26) and (A.28) to obtain

$$\bar{c}(k, \mu(\theta_1); \alpha = 2) \leq \zeta_2 \bar{h}_0 \int_{\theta_1}^{\bar{\theta}} (\theta - \theta_1) e^{-k\zeta_1[(\theta - \theta_1)^2/2]}\, d\theta. \tag{A.29}$$

Let $A = (\zeta_2 \bar{h}_0)/\zeta_1$. Then

$$\bar{c}(k, \mu(\theta_1); \alpha = 2) \le \frac{A}{k}\left(1 - e^{-k\zeta_1[(\theta-\theta_1)^2/2]}\right) < \frac{A}{k}. \tag{A.30}$$

Following the same procedure we can show that

$$\bar{c}(k, \mu(\theta_1); \alpha = 1) < \frac{A}{k}, \tag{A.31}$$

and Eq. (A.22) is proved. ■

Proof of Lemma 5.4:

(a) For $M, \lambda, \alpha > 0$ the left-hand side of Eq. (5.9) is continuous and decreasing from ∞ to zero.

(b) For $x = x(n)$, inequality (5.9), replaced by equation, can be rewritten

$$\lambda x(n) + \alpha \log x(n) = \log n + \log M. \tag{A.32}$$

Let $\epsilon(n) = \log n - \lambda x(n)$, $B = \alpha \log \lambda + \log M$. Then

$$-\epsilon(n) + \alpha \log x(n) = \log M \tag{A.33}$$

or, equivalently,

$$-\epsilon(n) + \alpha \log(\lambda x(n)) = B.$$

Substituting $\lambda x(n)$ from Eq. (A.32) we get

$$\epsilon(n) - \alpha \log\left(1 - \frac{\alpha \log x(n)}{\log n} + \frac{\log M}{\log n}\right) = \alpha \log(\log n) - B. \tag{A.34}$$

Note that Eq. (A.32) implies that $x(n) = O(\log n)$ and $\lim_{n\to\infty} x(n) = \infty$; therefore,

$$\lim_{n\to\infty} \frac{\log x(n)}{\log n} = \lim_{n\to\infty} \frac{\log x(n)}{x(n)} \lim_{n\to\infty} \frac{x(n)}{\log n} = 0.$$

Hence,

$$\lim_{n\to\infty} \log\left(1 - \frac{\alpha \log x(n)}{\log n} + \frac{\log M}{\log n}\right) = 0.$$

Thus, dividing both sides of Eq. (A.34) by $\log(\log n)$ and taking the limit as $n \to \infty$, we get

$$\lim_{n\to\infty} \frac{\epsilon(n)}{\log(\log n)} = \alpha,$$

which completes the proof. ■

Proof of Lemma 5.6: The proof is based on Lemma 5.5. From Eqs. (3.49) and (3.50) we get

$$\bar{c}(k,y;\alpha=1) = -\int_{\underline{\theta}}^{\theta_1} \delta(\theta) e^{kl(\theta,\theta_1|y)} h_0(\theta)\, d\theta, \tag{A.35}$$

$$\bar{c}(k,y;\alpha=2) = \int_{\theta_1}^{\bar{\theta}} \delta(\theta) e^{kl(\theta,\theta_1|y)} h_0(\theta)\, d\theta. \tag{A.36}$$

1. We first consider the case $\alpha = 1$. From Lemma 5.2, when $y > \mu(\theta_1)$, $l(\theta,\theta_1/y)$ attains its maximum in the interval $(\underline{\theta},\theta_1)$ for $\theta = \theta_1$, which is the upper limit of the integral. To bring the expression into the appropriate form for Lemma 5.5 to be applicable, change the variable of integration to $\tau = -\theta$. Then

$$\bar{c}(k,y;\alpha=1) = -\int_{\tau_1}^{\underline{\tau}} \tilde{\delta}(\tau) e^{k\tilde{l}(\tau,\tau_1|y)} \tilde{h}_0(\tau)\, d\tau, \tag{A.37}$$

where $\underline{\tau} = -\underline{\theta}$, $\tau_1 = -\theta_1$, $\tilde{\delta}(\tau) = \delta(-\tau)$, $\tilde{h}_0(\tau) = h_0(-\tau)$, and

$$\tilde{l}(\tau,\tau_1/y) = l(-\tau,-\tau_1/y) = (\tau_1 - \tau)y - (\psi(-\tau) - \psi(-\tau_1)). \tag{A.38}$$

Let $g_1(\tau) = \tilde{\delta}(\tau)\tilde{h}_0(\tau)$, $g_2(\tau) = \tilde{l}(\tau,\tau_1/y)$. Because $\tilde{\delta}(\tau) \sim \tilde{\delta}'(\tau)(\tau - \tau_1) = -\psi''(\theta_1)(\tau - \tau_1)$ and $g_2'(\tau) = -y + \mu(-\tau) \to \mu(\theta_1) - y$, as $\tau \to \tau_1$, we can see that g_1 and g_2 satisfy the conditions of Lemma 5.5 with $\alpha = \tau_1$, $\beta = \underline{\tau}$, $l(\alpha) = 0$, $b = y - \mu(\theta_1)$, $d = h_0(\theta_1)\psi''(\theta_1)$, $\nu = 1$, and $\lambda = 2$. Thus, as $k \to \infty$, Eq. (A.37) becomes

$$\bar{c}(k,y;\alpha=1) \sim h_0(\theta_1)\psi''(\theta_1)\Gamma(2)\left(\frac{1}{(y-\mu(\theta_1))k}\right)^2 = \frac{h_0(\theta_1)\psi''(\theta_1)}{(y-\mu(\theta_1))^2 k^2},$$

which proves Eq. (5.11).

2. We now consider the case $\alpha = 2$ and each one of the three subcases.

(a) $\mu(\theta_1) < y < \mu(\bar{\theta})$. In this case $l(\theta,\theta_1/y)$ attains its maximum in $(\theta_1,\bar{\theta})$ for $\theta = \theta^*(y)$, while from Lemma 5.2(b) $l(\theta^*(y),\theta_1|y) = \mathbf{I}(\theta^*(y),\theta_1)$. Because the maximizing point is interior to the area of integration, we break the integral in two parts

$$\bar{c}(k,y;\alpha=2) = \int_{\theta_1}^{\theta^*} \delta(\theta) e^{kl(\theta,\theta_1|y)} h_0(\theta)\, d\theta + \int_{\theta^*}^{\bar{\theta}} \delta(\theta) e^{kl(\theta,\theta_1|y)} h_0(\theta)\, d\theta. \tag{A.39}$$

A change of the integration variable similar to the previous is also needed for the first integral in the preceding expression. After this, Lemma 5.5 is again applicable to both terms in the right-hand side of Eq. (A.37), with $l(\alpha) = \mathbf{I}(\theta^*(y),\theta_1)$, $b = -\psi''(\theta^*(y))$, $d = \delta(\theta^*(y))h_0(\theta^*(y))$, $\nu = 2$, and $\lambda = 1$. Thus, for $k \to \infty$,

$$\bar{c}(k,y;\alpha=2) \sim 2\,\frac{\delta(\theta^*(y))h_0(\theta^*(y))}{2}\,\Gamma\left(\frac{1}{2}\right)\left(\frac{2}{\psi''(\theta^*(y))k}\right)^{1/2} e^{kI(\theta^*(y),\theta_1)}$$

$$= \delta(\theta^*(y))h_0(\theta^*(y))\sqrt{\frac{2\pi}{\psi''(\theta^*(y))k}}\, e^{kI(\theta^*(y),\theta_1)}.$$

Thus, we have established Eq. (5.12).

In the remaining cases (b) and (c), for which $y \geq \mu(\bar{\theta})$, $l(\theta, \theta_1 | y)$ is maximized in $(\theta_1, \bar{\theta})$ for $\theta = \bar{\theta}$, and a change of variable $\tau = -\theta$ is again necessary.

(b) $y = \mu(\bar{\theta})$. Then $(\partial l(\theta, \theta_1/y))/\partial y = 0$ for $\theta = \theta_1$, and the assumptions of Lemma 5.5 apply with $b = \psi''(\bar{\theta})$, $d = \delta(\bar{\theta})h_0(\bar{\theta})$, $\nu = 2$, and $\lambda = 1$. Thus, as $k \to \infty$,

$$\bar{c}(k, y; \alpha = 2) \sim \frac{\delta(\bar{\theta})h_0(\bar{\theta})}{2}\Gamma\left(\frac{1}{2}\right)\left(\frac{2}{\psi''(\bar{\theta})k}\right)^{1/2} e^{kl(\bar{\theta}, \theta_1 | y)}$$

$$= \delta(\bar{\theta})h_0(\bar{\theta})\sqrt{\frac{\pi}{2\psi''(\bar{\theta})k}}\, e^{kl(\bar{\theta}, \theta_1 | y)}, \tag{A.40}$$

which proves Eq. (5.13).

(c) $y > \mu(\bar{\theta})$. Then $b = y - \mu(\bar{\theta})$, $d = \delta(\bar{\theta})h_0(\bar{\theta})$, $\nu = 1$, and $\lambda = 1$. Thus, as $k \to \infty$,

$$\bar{c}(k, y; \alpha = 2) \sim \delta(\bar{\theta})h_0(\bar{\theta})\Gamma(1)\frac{1}{y - \mu(\bar{\theta})k}\, e^{kl(\bar{\theta}, \theta_1 | y)}$$

$$= \frac{\delta(\bar{\theta})h_0(\bar{\theta})}{y - \mu(\bar{\theta})}\,\frac{e^{kl(\bar{\theta}, \theta_1 | y)}}{k}. \tag{A.41}$$

■

PROOF OF LEMMA 5.7: The proof follows the same lines as that of Lemma 5.6. ■

PROOF OF PROPOSITION 5.15:

(a) From Lemma 5.1 we get

$$\frac{\partial \mathbf{I}(\theta^*(y), \theta_1)}{\partial y} = \frac{\partial \mathbf{I}(\theta^*, \theta_1)}{\partial \theta^*}\,\frac{d\theta^*(y)}{\partial y} = \theta^*(y) - \theta_1 > 0, \tag{A.42}$$

for $\mu(\theta_1) < y < \mu(\bar{\theta})$. From Definition 3.8 and Lemma 5.2,

$$\frac{\partial l(\bar{\theta}, \theta_1 | y)}{\partial y} = \bar{\theta} - \theta_1 > 0. \tag{A.43}$$

Thus G is increasing in both ranges of y. It remains to show that G and $\partial G/\partial y$ are continuous at $y = \mu(\bar{\theta})$. The former assertion is immediate from Lemma 5.2(b) and the latter from comparison of Eqs. (A.42) and (A.43) at the particular y.

(b) For $\mu(\theta_1) < y < \mu(\bar{\theta})$,

$$\frac{\partial G(y, \theta_1)}{\partial \theta_1} = \frac{\partial \mathbf{I}(\theta^*(y), \theta_1)}{\partial \theta_1} = -\mu(\theta^*(y)) + \mu(\theta_1) = -y + \mu(\theta_1) < 0, \tag{A.44}$$

while for $y \geq \mu(\bar{\theta})$

$$\frac{\partial G(y, \theta_1)}{\partial \theta_1} = \frac{\partial l(\bar{\theta}, \theta_1 | y)}{\partial \theta_1} = -y + \mu(\theta_1) < 0, \tag{A.45}$$

which completes the proof of monotonicity. The last two assertions are again immediate from Lemma 5.1 and Definition 3.8. ■