Advertising Revenue Optimization in Live Television Broadcasting

Pascale Crama[†], Dana G. Popescu[‡], Ajay S. Aravamudhan[†]
[†]Department of Operations Management, Singapore Management University
[‡]Department of Technology and Operations Management, INSEAD

In live broadcasting, the break lengths available for commercials may not always be fixed and known ex ante (e.g., strategic and injury time-outs are of variable duration in live sport transmissions). Because advertising represents a significant share of the broadcasters' revenue, broadcasters actively manage that revenue by jointly optimizing their advertising sales and scheduling policies. We characterize the optimal dynamic schedule in a simplified setting that incorporates stochastic break durations and advertisement lengths of 30 seconds and 15 seconds. The optimal policy is a greedy look-ahead rule that takes the remaining number of breaks into account. Under this setting, we find that there is no value to perfect information at the scheduling stage and knowing the duration of all the breaks will not change the schedule. When we incorporate diversity constraints (i.e., two ads from the same advertiser or for competing products cannot be shown during the same break), we characterize the optimal policy for a restricted set of stochastic break lengths. This policy combines the logic of the greedy look-ahead rule with the necessity to maintain an acceptable level of diversity in the ad portfolio. Finally, we also present heuristics that can be used to solve scheduling problems of greater complexity, and we recommend ways for broadcasters to balance their portfolio of booked ads. We run simulations to test the performance of the heuristics under various scenarios and find that two heuristic: myopic greedy and dynamic modified certainty equivalent (DMCE) perform close to optimal.

Key words: live broadcasting, advertising, scheduling, random capacity

---

1. Introduction

Broadcasters generate a large part of their revenue through advertising. At CBS, the most watched US broadcast network, TV advertising accounted for two thirds of the total revenue (Bloomberg Businessweek 2010). Major sporting events—such as the Super Bowl, the Olympics, and the FIFA World Cup—strongly boost such revenues because advertisers are willing to pay a premium for their ads to air during the live broadcast of these events. In 2010, for instance, the cost of a 30-second spot during the Super Bowl was between $2.5 and $2.8 million, or 18 times higher than the corresponding prime-time advertising rates. Similarly, a 30-second spot during the Winter Olympics in the same year generated between $360,000 and $490,000, which was about 3 times the rate of an average prime-time spot (Bauer Insight 2010).

While live broadcasting of major sporting events can significantly boost revenues, selling and scheduling advertisements in that environment can be a challenging task, especially for sports events that involve unpredictable breaks during which ads can be shown. A case in point is cricket, a major sport in South Asia, whose matches have breaks of random duration in the action.[1] The uncertainty about the duration of breaks creates an obvious problem for the broadcaster, namely how to schedule (live) the ads that have been sold while respecting the constraints on the schedule. The *diversity* constraints, i.e., two ads from the same advertiser or for competing products cannot be shown during the same break, which are commonly found in advertisement scheduling, are augmented by *capacity* constraints, i.e., the total duration of the ads scheduled during a break must not exceed the length of that break.

Suboptimal or infeasible schedules have many undesirable consequences for a broadcaster. If the schedule does not allow an ad to be shown in its entirety or if the schedule violates diversity constraints, no revenue will be earned and capacity will be wasted.[2] A schedule that violates capacity constraints could lead to rescinding of the broadcast rights or other costly penalties[3], e.g., cricket broadcasting rights require the broadcaster to guarantee live coverage of every ball of every match. Moreover, showing an excessive number of ads at inopportune times will displease viewers and lead to lower future ratings.[4] Thus, to generate the maximum possible revenue from live events, we look for the optimal dynamic scheduling policy under various scenarios of capacity and diversity constraints.

We model a television network that has a stochastic capacity of advertising airtime during a live event. This capacity consists of a number of commercial breaks of random duration. Breaks occur sequentially over a period of time and must be filled immediately upon arrival. Once a break occurs, its duration becomes known to the scheduler. We take as given the portfolio of booked ads which are to be aired during the live event.[5] The ads have variable length and yields. For tractability, but also for practical relevance, we analyze a setting with ads of two lengths, 15 and 30 seconds.[6]

[1] In cricket, two batsmen attempt to score runs against the fielding team. The fielding team's bowlers throw six balls in succession, called an 'over', from opposite ends of the field. The fielding team can rearrange the players' positions in the field between every over, and ads can be shown during that time. As soon as the players have taken up their new positions, the game re-starts and the broadcaster resumes the live coverage of the game.

[2] Contracts between broadcaster and advertiser typically specify that the advertiser will pay only if its ad is shown in full and not in a commercial break during which the same ad—or one for a competing product—is shown.

[3] For instance, in 2011 the Indian government issued a show-cause notice to the Ten Cricket channel for violating the country's advertising codes during its coverage of India's tour of South Africa, claiming that the broadcaster's ads had interfered with the program (ESPN Cricket Info, http://www.espncricinfo.com/).

[4] This occurred during the 2008 Summer Olympics: the Australian network Seven's coverage was widely criticized on these grounds.

[5] For a general description of the pricing and ad sales process in the US television advertising market see Bollapragada et al. (2002) and Phillips and Young (2010).

[6] This assumption reflects the US market, in which more than 90% of the ads sold are in one of these two formats.

In the base case, absent any diversity constraints, the optimal policy is a *greedy look-ahead* rule that takes the remaining number of breaks into account. Two surprising characteristics of the optimal policy are interesting to note. First, the optimal policy does not depend on the probability distribution of the break duration. This is counterintuitive, as one might expect the distribution of the remaining capacity to play a role. Second, and most importantly, perfect information is of no value; in other words, advance knowledge of the duration of all future commercial breaks does not change the network's revenue or schedule.[7] Finally, we find that the optimal scheduling policy is not affected by service level penalties that are proportional to the ad yield or length.

Incorporating diversity constraints into the scheduling problem substantially complicates the scheduling algorithm. However, we are able to derive the optimal policy when the break durations mirror the ad lengths, and show that there is no value to perfect information and the policy does not depend on the break length distribution. When the break durations are distributed over more than two values, the problem of optimal scheduling with diversity constraints becomes analytically intractable. Because ad scheduling transpires in real time during a live event, we therefore seek to derive simple and efficient heuristics that are fast and easy to implement. In Section 7 we propose several heuristics and compare their performance under scenarios characterized by various revenue ratios for long and short ads and overbooking levels.

A comparison of the expected revenue under the optimal policy (or perfect information, if optimal policy is intractable) and the greedy heuristic shows that the latter performs commendably well in many situations. Together with a clear understanding of the circumstances in which the greedy heuristic might fail, this result shows that the broadcaster does not lose much value by applying this simple algorithm. The greedy heuristic is adversely affected when short ads are selling at a *premium* (i.e., the yield of a short ad is, on average, higher than half the yield of a long ad): the revenue under the greedy heuristic might even decline as the premium on short ads increases and the total value of the portfolio increases. This results from the suboptimality in the scheduling which outweighs the benefit from the increased value of the short ads. In the presence of a diversity constraint, the performance of the greedy heuristic is further adversely affected when the ad portfolio displays a high *concentration* in the low-priced short ads, i.e., the low-priced short ads belong mostly to one advertiser or product category.

Finding the optimal solution to the scheduling problem described above, also allows to consider two more fundamental questions. First, the broadcaster has to decide how much airtime

[7] Nevertheless, such perfect knowledge would allow the broadcaster to improve its ad portfolio's composition in terms of the relative proportions of short and long advertisements.

to sell. Random capacity and high prices push the broadcaster to sell in excess of airtime capacity: this lowers the service level (i.e., the ratio of ads aired to ads sold), which will lead to advertiser dissatisfaction and—in the case of contractual guarantees—to penalties. Selling less than the available airtime capacity, however, causes underutilization and a loss of revenue. In the presence of penalties, the broadcaster will have to choose his level of overbooking carefully to balance the trade-off between expected benefit and penalty payments. Second, the broadcaster must also consider the ad portfolio's diversity in terms of ad duration and number of advertiser or product categories. The portfolio composition plays a role when the broadcaster is scheduling ads for a live event because a judicious composition of the portfolio can help the scheduling policy perform better under capacity and diversity constraints. Taking ad prices as exogenous input to the model, we look for the ideal mix of short and long ads to sell depending on their respective revenue and conditional on implementing the optimal policy at the scheduling stage. We also investigate when high advertiser concentration (i.e., high percentage of ads sold to the same advertiser) becomes detrimental to revenue.

These insights on the portfolio composition are of paramount importance to the first stage of contract negotiation and ad sales. First, we find that the level of concentration can have a substantial effect on the total revenue. High advertiser concentration increases scheduling difficulties; with insufficient diversity, it may be impossible to schedule short ads in long breaks causing low service levels *for* and revenues *from* for such ads. This effect is particularly pronounced for high concentration of high-paying short ads.

Second, we consider the composition of the portfolio in terms of long and short ads. Because short ads enhance scheduling flexibility, we find that, in the absence of overbooking, a broadcaster should consider selling more short ads than the expected number of short breaks (or conversely, fewer long ads than the expected number of long breaks), even if short ads generate, on average, significantly less than half the revenue of long ads. The higher the variability in break duration, the larger the discount on short ads the broadcaster is willing to accept in order to retain the scheduling flexibility afforded by a higher number of short ads.

It is interesting to note the opposite effect that the ratio of long to short ad *revenues* has on the simplicity of the scheduling and portfolio composition problems. A high ratio ($\gg 2$) reduces the optimal scheduling policy to a simple *myopic greedy* algorithm, but it makes the portfolio composition problem more challenging as it is not clear how many of each ad types to sell; that is, the optimal sales ratio is not self-evident: long ads are profitable but reduce the scheduling flexibility. For a low ratio ($\ll 2$), the optimal scheduling policy is non-trivial and the myopic greedy algorithm will perform poorly, in terms of total revenue, but the optimal sales ratio is evident: the broadcaster's goal is then to sell as many short ads as possible.

The rest of this paper is organized as follows. Section 2 reviews the literature. In Section 3, we set up the model (with two ad durations) and in Section 4 we derive and interpret the optimal scheduling policy. Section 5 presents extensions of the model that address break duration, contractual penalties and advertiser diversity. In Section 6, we address the portfolio composition problem and in Section 7 we describe several heuristics and test their performance relative to the case of perfect information. We conclude in Section 8. All (nontrivial) proofs are in the Appendix.

2. Literature Review

Previous work on media revenue management has examined the joint problem of scheduling and order acceptance while assuming deterministic break lengths (Bollapragada et al. 2004, Bollapragada and Garbiras 2004, Kimms and Müller-Bungart 2007). For example, Bollapragada and Garbiras formulate a goal-programming model to solve the scheduling problem. In their model, the emphasis is on satisfying as many product conflict and ad position constraints as possible, by choosing an appropriate penalty for each constraint that is violated (a much higher penalty for product conflict constraints than for position constraints) with the objective of minimizing the total penalty cost incurred. Kimms and Müller-Bungart formulate an integer program that maximizes the broadcaster's revenue while taking into account nonconflicting product constraints and specific scheduling requests; they propose several heuristics and conduct extensive numerical analyses that compare performance across different solution methods. In contrast with these papers, which assume capacity to be deterministic and derive static scheduling policies, we derive dynamic scheduling policies that take into account the stochastic capacity and the fact that breaks arrive sequentially over a period of time and must be filled immediately upon arrival.

Our problem is related to revenue management under conditions of random yield, a research field whose results are typically applied to production planning (for reviews of the literature see Grosfeld-Nir and Gerchak 2004, Yano and Lee 1995) or supply chain management (e.g., Tomlin 2009). Applications in the field of media revenue management focus on random yield due to uncertainty in ratings (Araman and Popescu 2009) or uncertainty in demand (Roels and Fridgeirsdottir 2009), whereas in our case uncertainty comes from the breaks' duration; this poses scheduling difficulties when ads have various lengths which exceed the duration of some breaks.

Much less attention has been given to the random yield caused by stochastic capacity. Ciarallo et al. (1994) is the the first work to explore the impact of random capacity. These authors find that a so-called order-up-to policy is optimal for minimizing production costs. Khang and Fujiwara (2000) establish the conditions under which the myopic order-up-to policy is optimal in a multiperiod setting. Hwang and Singh (1998) extend the analysis to a multistage production process and find that the optimal policy is characterized by a sequence of two critical numbers for each stage: a minimum input level, below which no production takes place; and a maximum desired production level. Wang and Gerchak (1996) incorporate randomness in both yield and capacity while showing that the optimal policy is characterized by a single reorder point in each period; that critical point is not constant and instead depends on the current inventory.

Our model differs in two important aspects from the multiperiod random capacity models advanced in the papers just cited. First, those previous works assume a single product whereas we engage a multiproduct setting with varying prices and production costs; this means that the products (i.e., ads sold) must be scheduled based on their profitability and the amount of capacity they use. Second, we assume integer units, which entails orders of fixed size. Hence the broadcaster cannot simply "max out" its capacity and hold inventory (i.e., airtime) to complete an order across multiple periods. In other words, each order must be entirely processed within a single production period.

This observation points to another related stream of literature, job scheduling with stochastic machine breakdowns and preemptive repeat, i.e., if a machine breakdown occurs during the processing of a job, all work done on the job is lost and processing has to start from scratch (Birge et al. 1990, Cai et al. 2009, Pinedo and Rammouz 1988). Birge et al. briefly discuss the preemptive repeat model with jobs of deterministic processing time. The objective is to minimize the total weighted completion time and the optimal rule schedules the jobs by increasing ratio of a function of processing time by weight. More recently, Cai et al. expand the model to include jobs with stochastic processing time, incomplete information and more general objective functions. The optimal schedules under different conditions are similarly based on appropriately tailored rankings that relate the weight and expected processing time of the jobs. In our paper, we simplify the structure of the jobs by restricting ourselves to two deterministic ad lengths and our assumption that the realization of the current break is known. These assumptions are justified in our setting and allow us to focus on generating insights about which ads to air and how to select the ads to be considered for broadcasting in the first place.

Finally, the optimization problem addressed in this paper has much in common with both the stochastic *cutting stock* problem and the dynamic stochastic *knapsack* problem, which have applications in such industries as materials (wood, steel, paper) and transportation. Consider, for example, the transportation industry, where airlines can ship cargo through freighter planes but also in the hold of their scheduled passenger flights. For the latter, cargo capacity varies between flights as a function of the booking level and the amount of passengers' checked-in luggage. Thus airlines face a problem similar to the one described in Section 1 as they seek to maximize revenue under stochastic capacity for a given a set of transport requests.

The cutting stock problem originated as a knapsack problem and involves minimizing unused capacity or waste (see Wäscher et al. 2007 for a discussion on the topology of cutting and packing problems); it was introduced by Gilmore and Gomory (1961), who later proposed a set of specialized solution techniques (Gilmore and Gomory 1965). This problem has been extended to address stochastic capacity and quality while minimizing waste (Ghodsi and Sassani 2005, Scull 1981). Our problem is a generalization of the multiple heterogeneous knapsack problem (see Martello and Toth 1999), where a heterogeneous set of small items characterized by a given weight and yield must be packed into a set of knapsacks of different capacities. For each knapsack, the packed items must not exceed the available capacity and the expected total yield of the packed items across all knapsacks has to be maximized. The generalization, in our case, comes from the fact that the knapsacks (i.e., the breaks) have stochastic capacities, arrive sequentially over a period of time and must be filled immediately upon arrival. Moreover, apart from capacity constrains, additional constraints (i.e., diversity constraints) have to be satisfied by the packed items. To the best of our knowledge, the stream of literature on cutting and packing problems has not tried to answer how the orders affect the performance of the cutting stock algorithm and what are desirable properties of the set of orders.

3. Model

Consider a television network that has a random capacity of advertising airtime during a live event. This capacity consists of N commercial breaks of random duration.[8] We use $\{b_1, \ldots, b_N\}$ to denote the set of breaks, where b_n represents the duration of break n. Let $f_n(\cdot)$ be the ex-ante distribution of b_n and let $[S, U_n]$ be its support. We assume that once a break occurs, its duration is known to the scheduler.

Let M be the total number of booked ads at the start of the live event and $\mathbf{r} = \{r_1, r_2, \cdots, r_M\}$ the set of ad revenues; here r_i stands for the revenue generated by *fully* airing ad i. In case the live transmission is resumed before the ad finishes airing, then the network receives no revenue for that ad. Let d_i be the duration (in seconds) of ad i.

As mentioned in the introduction, networks typically face restrictions regarding which ads can air during a given break. For instance, two ads belonging to the same advertiser cannot be shown in the same break (*advertiser diversity* constraint) and two ads for products in the same category cannot be shown in the same break (*product diversity* constraint). For brevity, in this paper we focus solely on advertiser diversity constraints and merely note that an analysis of

[8] For some live events, the number of breaks can be also random, but generally the distribution is likely to be very tight. For more tractability, we consider a model with deterministic number of breaks and note that numerical experiments show robustness of our results even when the number of breaks is stochastic.

product constraints would lead to similar insights. Let H denote the number of advertisers. The set of all possible combinations of ads in $\mathbf{r}$ that can air in a break of length b is then

$$\Gamma(\mathbf{r},b)=\{\{r_{i_1},...,r_{i_k}\}\subseteq\mathbf{r}:\sum_{j=1}^{k}d_{i_j}\le b;\ |\{i_1,...,i_k\}\cap A_h|\le 1\ \forall h=1,H\},\tag{1}$$

where A_h is the subset of ads belonging to advertiser h and $|A|$ denotes the cardinality of set A. The (expected) revenue-to-go at stage $1\le n<N$, given a vector of ad revenues $\mathbf{r}$ and a realized break duration b, can be written as follows:

$$V_n(\mathbf{r},b)=\max_{\{r_{i_1},...,r_{i_k}\}\ \subseteq\ \Gamma(\mathbf{r},b)}\left\{\sum_{j=1}^{k}r_{i_j}+E_{b_{n+1}}[V_{n+1}(\mathbf{r}-\{r_{i_1},...,r_{i_k}\},b_{n+1})]\right\}.\tag{2}$$

We further assume that each ad is of length L seconds or S seconds, where $L=2S$; then $d_i\in\{L,S\}$ for all $i=1,M$. In practice, this translates into $L=30$ seconds and $S=15$ seconds. Let $\mathbf{r}=\{\mathbf{l}\cup\mathbf{s}\}$, where $\mathbf{l}=\{l_1,\ l_2,\ \cdots,\ l_{n_L}\}$ is the set of revenues generated by each ad of length L and $\mathbf{s}=\{s_1,\ s_2,\ \cdots,\ s_{n_S}\}$ is the set of revenues generated by each ad of length S; also, n_L is the number of long ads, n_S is the number of short ads, and $n_L+n_S=M$. We assume that $l_i\ge l_{i+1}$ and $s_i\ge s_{i+1}$ for all positive integers i.

3.1. Definitions

In order to facilitate the analysis in the subsequent sections, we propose the following definitions for overbooking, service levels, and capacity utilization in the context of live television advertising.

Definition 1. The *overbooking level*, δ, is defined as the percentage of airtime sold in excess of expected airtime capacity:

$$\delta=\frac{\sum_{i=1}^{M}d_i}{\sum_{n=1}^{N}E[b_n]}-1.\tag{3}$$

Definition 2. The *service level*, ρ, is defined as the ratio of the number of ads aired to the number of ads sold:

$$\rho=\frac{\sum_{i=1}^{M}\sum_{n=1}^{N}x_i^n}{M},\tag{4}$$

where $x_i^n=1$ if advertisement i aired during break n (and $x_i^n=0$ otherwise).

The service level can also be computed for specific categories of ads or per advertiser (e.g., service level for short/long ads, service level for advertiser l). In this paper, we mainly refer to the *expected* service level, defined as a ratio of the expected number of ads aired to number of ads sold. This ratio depends, of course, on the scheduling policy in place.

Definition 3. The *capacity utilization level*, μ, is defined as the ratio of the combined duration of all ads aired to the total realized airtime capacity:

$$\mu = \frac{\sum_{i=1}^{M}\sum_{n=1}^{N} d_i x_i^n}{\sum_{n=1}^{N} b_n}. \tag{5}$$

Observe that, whereas the overbooking level is an ex ante measure defined in terms of expected capacity, the capacity utilization is an ex post measure defined in terms of realized airtime. Thus, a given overbooking level can yield different levels of capacity utilization depending on the relative predominance of the two ad (length) categories and the realized break durations.

4. Base Case Analysis

For now we shall ignore the diversity constraints and study the optimal scheduling policy for the unconstrained problem. To simplify the exposition, we assume that there are only two possible break durations: either L or S seconds; thus $b_n \in \{L, S\}$ for all $n = 1, N$. The break durations are independent and identically distributed (i.i.d.), with probability α that the break is short (i.e., S) and $1 - \alpha$ that the break is long (i.e., L). We assume that once a break occurs, the duration of that break is known.

We study the optimal dynamic scheduling policy under this setting. Although this is a simplification of the original problem, it captures the main characteristics: (1) stochastic capacity; (2) a heterogeneous assortment of ads of different values which can be classified into "short ads" and "long ads"; (3) a heterogeneous assortment of break durations which can be classified into "short breaks" and "long breaks"; (4) limited availability of breaks such that not all ads can be accommodated; and last but not least, (5) not every ad fits in every break (i.e., the long ads only fit in long breaks). In Section 5 we will relax some of the simplifying assumptions and derive the optimal policy under more general settings. As we shall see, the main structure of the optimal policy remains unchanged.

There are N stages, each corresponding to a particular break. Let $V_n(\mathbf{r}, b) = V_n(\mathbf{l}, \mathbf{s}, b)$ be the revenue-to-go function at stage n, with $N - n$ stages remaining, when break n of duration b occurs, given the ordered sets $\mathbf{l}$ and $\mathbf{s}$ of remaining ads of respective length L and S. In the absence of diversity constraints and with only two possible break durations, the revenue-to-go function described in (2) reduces to

$$\begin{cases} V_n(\mathbf{l}, \mathbf{s}, L) = \max\left\{ l_1 + E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})],\ s_1 + s_2 + E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_{n+1})] \right\}, \\ V_n(\mathbf{l}, \mathbf{s}, S) = s_1 + E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - s_1, b_{n+1})]. \end{cases}$$

The optimal dynamic scheduling policy is given in our first proposition.[9]

[9] For simplicity of exposition of the optimal policy (and w.l.o.g.) we assume that the number M of sold ads is infinite, which implies that an infinite number of ads will have revenue zero.

PROPOSITION 1 **(Optimal Scheduling Policy)**. *The optimal dynamic scheduling policy at stage n, when break n of duration b occurs, given the ordered sets $\mathbf{l}$ and $\mathbf{s}$ of remaining long and short ads, is as follows.*

1. *When $b = S$, schedule ad s_1.*
2. *When $b = L$, schedule ads (s_1, s_2) if $l_1 < s_{N-n+1} + s_{N-n+2}$; else schedule ad l_1.*

A first observation is that the policy does not depend on the probability of a short break, α. This is counterintuitive, as one might expect the distribution of the remaining capacity to play a role. It is important to identify the smallest set of assumptions that drive this result. On one hand, relaxing the assumption that the length of a long ad is twice the length of a short ad, i.e. $L = 2S$, will not change the result and the optimal policy will remain the same. As a matter of fact, the proof of Proposition 1 is constructed independently of this assumption. Similarly, relaxing the assumption that the break duration is either S or L, will not change the result either, as shown in Section 5.1. However, if we simultaneously relax these two assumptions, then the threshold policy will no longer be independent of α, as shown in the Example 1 below.

EXAMPLE 1. Suppose ads can have two lengths: 15 seconds and 45 seconds and breaks can have durations of 15, 30, and 45 seconds with respective probabilities α_1, α_2 and $1 - \alpha_1 - \alpha_2$. If break $N - 1$ has a duration of 45 seconds then, with one break remaining, it is easy to see that the optimal policy is to schedule l_1 (i.e., the highest yield 45-second ad), whenever $(1 - \alpha_1)l_1 \geq \alpha_1 s_2 + (\alpha_1 + \alpha_2)(s_3 + s_4) + \alpha_2 s_5$. In all other cases, it is optimal to schedule (s_1, s_2, s_3) (i.e., the three highest yield 15-second ads). Hence, the policy is no longer independent of the probability distribution of the remaining capacity.

An important and surprising result follows immediately from our first observation: for the base case setting, the optimal scheduling policy has the property that the expected value of perfect information is zero. Therefore, uncertainty regarding the break duration does not lead to a loss of efficiency for the broadcaster. This result is formalized in the following proposition.

PROPOSITION 2 **(Expected Value of Perfect Information)**. *Under the optimal scheduling policy, the expected value of perfect information is zero.*

So even though capacity is uncertain, the broadcaster—given a portfolio of booked ads—is guaranteed to obtain the same revenue as if capacity were known at the start of the live event. This does not mean, however, that uncertainty in capacity has no effect on the broadcaster's revenue. Although such uncertainty makes no difference at the scheduling stage, it would, of course, play a role at the prior selling stage, during which the broadcaster decides how many short and long ads to sell to different advertisers.

A second observation is that the optimal policy so described has the structure of a *greedy look-ahead* policy in the following sense: the revenue from the most lucrative long ad (l_1) is compared to the combined revenue from the $N-n+1$ and $N-n+2$ most lucrative short ads ($s_{N-n+1}+s_{N-n+2}$) given that $N-n$ stages remain. It is instructive to contrast this policy to a *myopic greedy* (a.k.a. "greedy") heuristic, one that schedules the most lucrative ad(s) at each stage by comparing l_1 to s_1+s_2. Although the two policies are similar in nature, the threshold for scheduling l_1 is reduced under the greedy look-ahead rule because it is beneficial to place long ads—even if they are less valuable in the short run—in order to avoid a scenario in which all the remaining breaks are short and the comparatively less profitable short ads s_{N-n+1} and s_{N-n+2} must be scheduled. That is, the optimal schedule accepts an immediate revenue loss at stage n in order to safeguard revenues from future stages; this is what makes it a "greedy look-ahead" policy. The impact on the respective service levels for short and long ads follows directly from this observation: the optimal scheduling rule is more likely to air long ads than the myopic greedy, so the former achieves a better service level for long ads but a lower service level for short ads.

We remark that the myopic greedy heuristic does not always perform suboptimally. Depending on the premium (or discount) for short ads, the greedy policy can perform almost as well as the optimal rule. In Figure 1, we plot the performance (in revenue terms) of the greedy myopic and optimal policies as a function of the probability of a short break α, for various short-to-long ad revenue ratios (i.e., $\frac{l}{s}=1$, $\frac{l}{s}=\frac{1}{0.7}$, $\frac{l}{s}=\frac{1}{0.5}$). If short ads sell at a discount (i.e., if $l \geq 2s$) then the greedy heuristic performs as well as the optimal one; if short ads sell at a premium (i.e., if $l < 2s$) then the greedy heuristic will perform suboptimally. The distribution of break durations also has an impact on the greedy heuristic's performance. As the probability of a short break increases, the performance shortfall of the greedy over the optimal policy increases up to a point; it then starts to decrease and eventually approaches zero as that probability approaches unity. At the extremes (i.e., when either all breaks are long or all breaks are short) the two policies coincide, as expected. An interesting observation is that the revenue under the greedy heuristic might even decline as the premium on short ads increases and the total value of the portfolio increases. This results from the suboptimality in scheduling which outweighs the benefit from the increased value of short ads. In Figure 1, for example, the average revenue under the greedy heuristic when $\alpha = 0.6$ and $l = \$10$, $s = \$5$ (i.e., ratio=0.5) is higher than the average revenue corresponding to ad prices $l = \$10$, $s = \$7$ (i.e., ratio=0.7). Thus, even though the broadcaster earns a higher yield on short ads, overall the broadcaster's revenue decreases due to the suboptimality of the (greedy) scheduling heuristic.

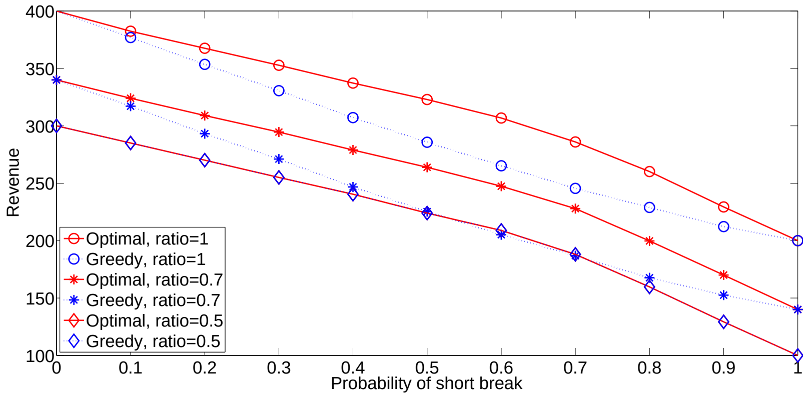

Figure 1 Performance of Optimal versus Greedy Policy

(1000 simulation runs; 20 short ads and 20 long ads; N = 30; l = 10)

5. Extensions

5.1. Generalized Break Duration

In this section we relax the assumption that breaks can be only of S or L seconds in duration; instead, we allow the duration to vary in multiples of S seconds. The results are equivalent to a general nonnegative distribution because, if ads are not to be cut off, then any break length that is not an integer multiple of S seconds must be rounded down to the nearest integer multiple.[10]

Proposition 3 **(Optimal Scheduling Policy for Generalized Break Duration).** *The optimal dynamic scheduling policy at stage n is to select $(l_1,\ldots,l_\lambda,s_1,\ldots,s_{m-2\lambda})$ for $b = mS$, where $l_\lambda \geq s_{m-2\lambda+n+1} + s_{m-2\lambda+n+2}$ and $2\lambda \leq m$.*

Similar to the case of only two break durations, the expected value of perfect information in this setting is zero. Hence, again, there is no efficiency loss due to uncertain capacity at the scheduling stage.

Example 2. To illustrate how this policy works, we discuss the scenario in which break lengths can be of four types: $\{S, 2S, 3S, 4S\}$ (e.g., the maximum break length cannot exceed one minute when $S = 15$ seconds). Then, the optimal dynamic scheduling policy at stage n is as follows:

- if $b_n = 2S$ and $l_1 \geq s_{N-n+1} + s_{N-n+2}$, schedule l_1;
- if $b_n = 3S$ and $l_1 \geq s_{N-n+2} + s_{N-n+3}$, schedule l_1 and s_1; else, schedule (s_1, s_2, s_3);
- if $b_n = 4S$ and $l_2 \geq s_{N-n+1} + s_{N-n+2}$, schedule (l_1, l_2); else,

  —if $b_n = 4S$ and $l_1 \geq s_{N-n+3} + s_{N-n+4}$, schedule (l_1, s_1, s_2);

  —if $b_n = 4S$ and $l_1 < s_{N-n+3} + s_{N-n+4}$, schedule (s_1, s_2, s_3, s_4).

[10] If f_n is a continuous probability density function of b_n with support $[S, U_n]$, then we can find the corresponding discrete probability distribution by setting $\Pr(b_n = mS) = \int_{mS}^{(m+1)S} f_n(x)\,dx$.

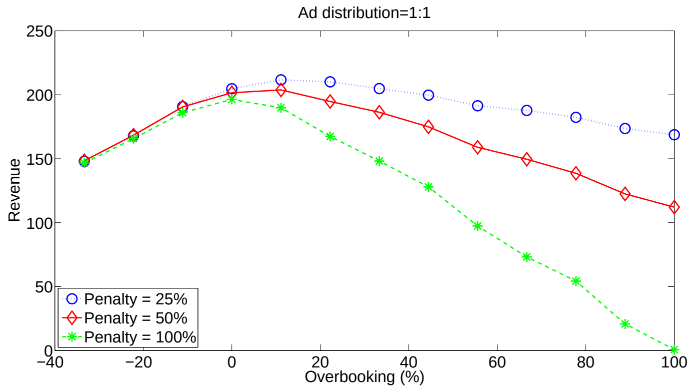

Figure 2 **Performance of Optimal Policy with Different Penalty Rates**
1000 simulation runs; N = 30; l = 10, s = 5

5.2. Penalties for Unaired Ads

Advertisers incur a disutility whenever their ads are not shown. Sometimes, contracts between advertisers and the network specify a penalty that the network must pay in case a specific ad is not shown. If the penalty is proportional to the revenue the ad generates, or if there is a constant penalty per long (short) ad, proportional to the length of the ad, then the optimal policy described in Section 4 will remain the same. The following proposition states this result.

PROPOSITION 4 **(Optimal Scheduling Policy with Penalties for Unaired Ads)**. *The optimal dynamic scheduling policy at stage n, when break n of duration b occurs, given the ordered sets $\mathbf{l}$ and $\mathbf{s}$ of remaining long and short ads and no-show penalties proportional to either the ad value (i.e., $\beta\mathbf{l}$ ($\beta\mathbf{s}$) for long (short) ads) or length (i.e., βL (βS) for a long (short) ad) , is as follows.*

1. *When $b = S$, schedule ad s_1.*
2. *When $b = L$, schedule ads (s_1, s_2) if $l_1 < s_{N-n+1} + s_{N-n+2}$; else schedule ad l_1.*

In the absence of a penalty for unaired ads, the broadcaster will overbook up to a level corresponding the the maximum possible airtime. Figure 2 shows that even a relatively moderate penalty sharply reduces the optimal level of overbooking. However, we also observe that even with a substantial penalty rate the broadcaster still books a number of ads of total duration adding up to nearly the expected airtime. Note that this result is in nature similar to the newsvendor's problem of choosing the optimal ordering quantity, which has to balance the overage and underage cost. Thus the insights from the newsvendor problem remain valid in our context.

5.3. Advertiser Diversity within Break

The broadcaster's problem typically includes diversity constraints as well. Advertisers often purchase several ads in the same live event and do not want their ads to be shown twice within the same break. We extend our analysis to include multiple advertisers $h = 1, H$ purchasing long and short ads at a constant, advertiser-specific price. All the other assumptions remain.

Advertiser h pays r_l^h and r_s^h for long and short ads, respectively. We define the mappings $h^l : \{1,2,\ldots\} \to \{1,\ldots,H\}$ and $h^s : \{1,2,\ldots\} \to \{1,\ldots,H\}$ to indicate, respectively, the advertiser that purchased the ith long and short ad. The number of long and short ads bought by advertiser h are denoted m_l^h and m_s^h, respectively. A portfolio shows high concentration if a large proportion of ads are bought by the same advertiser h.

The diversity constraint forces the broadcaster to choose short ads more carefully in order to preserve an inventory diverse enough that allowable combinations of short ads are available for any long break during which it is suboptimal to show a long ad. These considerations render nontrivial the decision on which short ad to show in a short break.

Proposition 5 **(Optimal Scheduling Policy with Diversity)**. *The optimal dynamic scheduling policy at stage n (i.e., when there are $N-n$ breaks remaining) is as follows.*

1. *If $b_n = S$ then schedule one short ad s_i such that $i \le N-n-t+1$ and $m_s^{h^s(i)} = \max_{h\in\{h^s(1),\ldots,h^s(N-n-t+1)\}}\{m_s^h\}$ with the largest $t \le N-n$ such that $l_t \ge s_{N-n-t+2} + s_{N-n-t+3}$.*
2. *If $b_n = L$ then schedule l_1 if $l_1 \ge s_{N-n+1} + s_{N-n+2}$. Otherwise: (i) if $l_1 \ge s_{N-n+2} + s_{N-n+3}$, schedule (s_1, s_i) such that $i \le N-n+1$ and $m_s^{h^s(i)} = \max_{h\in\{h^s(1),\ldots,h^s(N-n+1)\}}\{m_s^h\}$ if that set is nonempty; (ii) else schedule (s_1, s_i) such that $i \le N-n+2$ and $m_s^{h^s(i)} = \max_{h\in\{h^s(1),\ldots,h^s(N-n+2)\}}\{m_s^h\}$.*

Note that, after introducing the diversity constraint, we no longer systematically schedule the most profitable short ad in a short break. It may be preferable to choose a less profitable short ad if there are many short ads of lower value from the same advertiser and not enough high-value short ads to avoid exhausting their supply. This avoids creating an overly concentrated inventory to enable pairs of short ads for future long breaks. The scheduling rule for long breaks combines the logic of the optimal scheduling rule without diversity and the need to avoid creating a highly concentrated inventory of short ads. Thus, a long ad that is less profitable than the most valuable possible combination of short ads may be scheduled nonetheless simply to avoid revenue loss in future breaks and when two short ads are shown, they are selected not to maximize current revenue but instead to minimize concentration of remaining short ads.

Corollary 1. *Given H advertisers, two break lengths, S and $L = 2S$, and a within-break diversity constraint, the expected value of perfect information is zero.*

The diversity constraint does not create value for information. The same ads are shown in the optimal schedules obtained under Proposition 5 as in the optimal schedule with full information, though not necessarily in the same order. This result does not extend to the general case with arbitrary break lengths S and tS for $t \in N$.

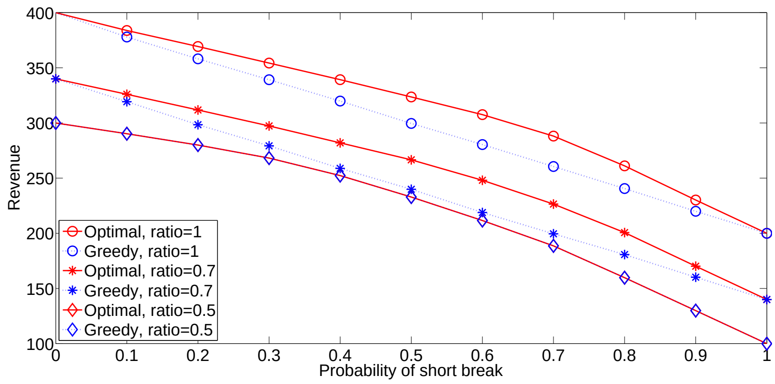

Figure 3 **Performance of Optimal Policy vs. Greedy Heuristic**

1000 simulation runs; 30 short ads and 30 long ads; N = 30; l = 10

Example 3. Assume a case with two breaks, $t = 3$, and four short ads in inventory: one from advertiser 1 and 2 respectively, and two ads from advertiser 3. Under perfect information, the optimal schedule chooses s_1 and s_2 if both breaks are short, in no particular order, but s_3 in the first break and (s_1, s_2, s_3) in the second break if the first break is short and the second break is long. Thus the perfect information profit cannot be achieved without knowledge of future breaks.

The greedy heuristic performs as well as the optimal policy when long ads sell at a premium, but it underperforms when short ads are more lucrative (see Figure 3). The greedy heuristic and optimal policy are impacted differently by the initial portfolio's concentration. The revenue from the optimal policy is stable across different starting portfolios. The greedy heuristic performs close to the optimal policy except when (i) short ads are relatively more valuable than long ads but few have been sold or (ii) short ads are plentiful and cheap but the lowest value short ads are concentrated in the hands of one advertiser. In the first case, the greedy heuristic schedules the valuable short ads early on but runs out of (all types of) short ads at the end of the scheduling horizon because few have been sold; this dynamic is relatively unaffected by the starting portfolio. In the second case, short ads are scheduled in long breaks once the long ads are exhausted, but the greedy heuristic chooses from the high value short ads first, and the concentration in the portfolio increases to the extent that the broadcaster may run out of pairs of short ads for long breaks.

6. Portfolio Composition

As discussed in Section 5.2, the random capacity in live television broadcasting pushes the broadcaster to sell in excess of expected airtime. This lowers the service level (i.e., many of the ordered ads will not be aired), which will lead to advertiser dissatisfaction and—in the case of contractual guarantees—to penalties. In the presence of contractual penalties for unaired ads,

Overbooking Level	Short Ads	Long Ads	Average Capacity Utilization
0%	36%	64%	87.35%
	50%	50%	94.63%
	61%	38%	97.18%
8.33%	37%	63%	90.98%
	50%	50%	98.14%
	61%	39%	99.61%
16.67%	38%	62%	93.66%
	50%	50%	99.22%
	60%	40%	99.98%

Table 1 Overbooking Level, Ad Distribution, and Capacity Utilization ($\alpha = 0.5$)

the broadcaster will have to choose his level of overbooking carefully to balance the trade-off between expected benefit and penalty payments. The determination of optimal overbooking levels can be made via analysis of the traditional newsvendor type, which estimates the cost of lost business associated with low service levels (as well as potential contractual penalties) and weighs these costs against the opportunity cost of underutilized capacity. The overbooking problem is not unique to live broadcast advertising and has been well studied in other contexts, so we will not address that issue and instead focus on the characteristics unique to this context, the most important of which is the trade-offs involved when splitting the booking levels between short and long ads.

First, note that for a given overbooking level and a scheduling policy, there can be different levels of capacity utilization depending on the relative predominance of the two ad (length) categories (see Table 1). Short ads give the broadcaster more scheduling flexibility because they can air in any type of break, hence the greater the number of short ads sold, the higher the average capacity utilization.

However, depending on the market characteristics, long ads could potentially generate more revenue whenever they are sold at a premium – that is, the price of an ad is an increasing, convex function of the length of the ad. The convexity or concavity of the ad price function with respect to the length of the ad varies from country to country and from broadcaster to broadcaster.[11] In general, it is not straightforward to asses the premium or the discount associated with longer

[11] In the United States and Australasia, where the 15-second and 30-second formats are predominant, for ROS ads (i.e., run-off-schedule ads which can be placed in any show, depending of the network's preference), the 15-second ads generally sell at a premium. The media cost of a 15-second ad, although half the length of a 30-second ad, is typically 60 to 80% the cost of a 30-second ad (Newstead and Romaniuk 2010). However, an even shorter ad (e.g., 10-second ad) typically sells at a discount in the United States (e.g., a 10-second ad is sometimes 15%-20% the cost of a 30-second ad - http://www.brucemedia.com/10s-why.html). In the United Kingdom, where the 10-second ads are more popular than the 15-second ads (Newstead and Romaniuk 2010), Channel 5 network reports that a 30-second commercial will cost twice as much as a 10-second ad and half as much as a 60-second ad (http://about.channel5.com/faqs/how-to-advertise).

ads because of the non-transparent nature of the media industry.[12] Nevertheless, it is important to quantify the trade-off between yield and scheduling flexibility as reflected in the optimal sales ratio of short to long ads.

Thus, taking ad prices and the level of overbooking as exogenous input to the model, we look for the ideal mix of short and long ads to sell, depending on their respective revenue and *conditional* on implementing the optimal policy at the scheduling stage.

Assume, for simplicity, that long (resp., short) ads have constant price l (resp., s, with $s \leq l$) and the broadcaster sets an overbooking level of δ. Let $l = 2s(1+\epsilon)$, where ϵ denotes the premium (i.e., $\epsilon \geq 0$) or the discount (i.e., $\epsilon < 0$) corresponding to a long ad.

If $\epsilon \geq 0$, then the broadcaster will always schedule a long ad in a long break, unless there are no more long ads left. Absent any diversity constraints, the short ads will be scheduled in the remaining long and short breaks. In this case, the expected revenue of a broadcaster is a concave function of the number of long ads and is given by:

$$\begin{aligned} V(n_L, n_S) = & \sum_{i=0}^{n_L} [i * l + \min\{n_S, N-i\} * s] \Pr(k=i) + \\ & + \sum_{i=n_L+1}^{N} [n_L * l + \min\{n_S, N+i-2n_L\} * s] \Pr(k=i), \end{aligned} \tag{6}$$

where k denotes the random number of long breaks, n_S is the number of (booked) short ads, n_L is the number of (booked) long ads. Since the level of overbooking is fixed, then n_S and n_L must also satisfy the equation $n_s + 2n_L = \lfloor (1+\delta)(2-\alpha)N \rfloor$.

If $\epsilon < 0$, then the broadcaster will schedule two short ads in a long break, while still retaining enough short ads to cover the short breaks. In this case, the expected revenue is a decreasing function of the number of long ads:

$$\begin{aligned} V(n_L, n_S) = & \sum_{i=0}^{\max\{N-n_S, 0\}} (\min\{i, n_L\} * l + \min\{N-i, n_S\} * s) \Pr(k=i) + \\ + & \sum_{i=\max\{N-n_S+1, 0\}}^{N} [\min\{n_L, i-u\} * l + (N-i+2u+v) * s] \Pr(k=i), \end{aligned} \tag{7}$$

where $u = \min\left\{i, \left\lfloor \frac{n_S-N+i}{2} \right\rfloor\right\}$ and $v = \min\left\{\left\lceil \frac{n_S-N+i}{2} \right\rceil - \left\lfloor \frac{n_S-N+i}{2} \right\rfloor, i-u-n_L\right\}$.

The following proposition summarizes the optimal number n^* of long and short ads that a broadcaster needs to sell in order to maximize expected revenue.

[12] In the media industry, advertisers buy bundles of airtime spots of different lengths. The price of the bundle is negotiated between an advertiser and the broadcaster and often depends on quantity discounts and the number and type of advertiser constraints, but also on more qualitative factors such as the bargaining power of each party, the length of time the advertiser has done business with the network, the quality of their business relationship, etc.

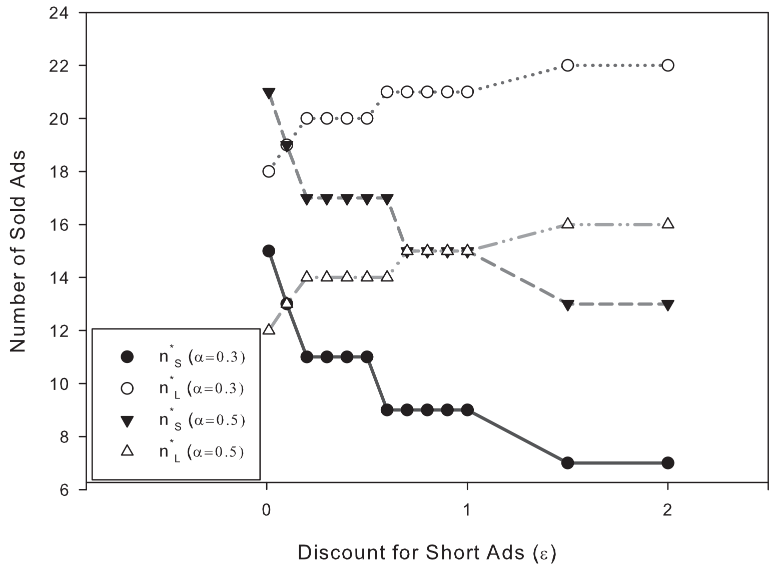

Figure 4 **Optimal Number of Short and Long Ads as a Function of the Discount**
δ=0, N = 30; l = 10

PROPOSITION 6 **(Optimal Portfolio Mix).** *If $l = 2s(1+\epsilon)$, then n_L^* is the smallest integer such that*

$$2\epsilon \Pr(k \geq n_L + 1) \leq 2\Pr\left(k \leq \min(N - \lfloor (1+\delta)(2-\alpha)N - 2n_L \rfloor, n_L)\right) + \\ + \Pr\left((k = N - \lfloor (1+\delta)(2-\alpha)N - 2n_L \rfloor + 1) \wedge (k \leq n_L)\right),$$

and $n_S^* = \lfloor (1+\delta)(2-\alpha)N - 2n_L^* \rfloor$.

This proposition states that the optimal number of long ads is the number for which the opportunity cost of selling one extra long ad at a premium is outweighed by the opportunity cost of unutilized capacity. It is important to note that if there is no premium for long ads (i.e., if $\epsilon \leq 0$) then the optimal number of long ads is zero. This fact is intuitive because only short ads offer the benefit of flexibility. So absent a premium for long ads, the broadcaster will prefer to sell only short ads and thereby reduce the risk of unutilized airtime capacity.

Figure 4 plots the optimal number of short and long ads as a function of the discount for short ads ($\epsilon \geq 0.01$), when there is no overbooking (i.e., $\delta = 0$), for two values of α, specifically $\alpha \in \{0.3, 0.5\}$. Not surprisingly, as the discount increases, the optimal number of short ads decreases (and eventually approaches zero) and the optimal number of long ads increases (and eventually approaches N). Note that even if the discount is significant, e.g., $\epsilon = 0.5$, it is still optimal to sell more short ads than the expected number of short breaks, and conversely fewer long ads than the expected number of long breaks. The higher the variability in break duration (i.e., the closer α is to 0.5), the larger the discount on short ads the broadcaster is willing to accept in order to retain the scheduling flexibility afforded by a higher number of short ads.

7. Numerical Example

7.1. Data

In practice, the advertisement scheduling problem displays features—such as diversity constraints and multiple break lengths— that make it analytically intractable. We therefore perform a numerical analysis of the problem in a more realistic setting. The simulation parameters are listed in the following table.

Parameter	**Value**
Ad lengths	15, 30 seconds
Revenue distribution	$\mathcal{U}(7{,}000, 10{,}000)$
Break distribution	$\mathcal{U}(15, 75)$
Number of breaks	50

We consider two overbooking levels, 0% and 16.67%. The former offers a near guarantee that all the booked ads will be shown during the tournament whereas the latter provides the broadcaster with a reasonable safety margin to prevent unutilized break time. The ad portfolio's composition—in terms of airtime sold as short versus long ads—is also varied and reflects three different settings: twice as much time sold to short ads (2 : 1 ratio), equal time sold to short and long ads (1 : 1), and twice as much time sold to long ads (1 : 2). Finally, we allow for linear and nonlinear pricing with a discount for long ads.

7.2. Heuristics

We investigate the effectiveness of several heuristic approaches to the scheduling problem by comparing their expected revenue with the results obtained under the optimal policy (no diversity case) or with perfect information (diversity case). The simulation and greedy heuristic were run on Matlab. Our routine calls "CPLEX" in GAMS to solve the knapsack problem for the class of certainty equivalent heuristics and the perfect-information schedule.

7.2.1. Certainty Equivalent Heuristics This class of simple heuristic forms the basis of the broadcaster's current scheduling decisions. The certainty equivalent heuristic (or "CE" for short) schedules ads in advance based on the *average* break length. Thus the ads are aired as already scheduled irrespective of the break length actually realized; however, the broadcaster earns revenue only for ads that are aired in full. The modified CE (MCE) is an intuitive improvement that prepares ad bundles of varying length based on the *distribution* of break lengths. As soon as a break's length is known, the scheduler selects a correspondingly sized bundle of ads to air. This prevents crashed ads or unutilized air time—unless the supply of bundles for a needed break length has run out. The broadcaster currently uses the MCE algorithm to prepare bundles of ads, but also allows human intervention during the live broadcast so that bundles can be rescheduled on the fly. When refined in this way, the algorithm is referred to as the dynamic MCE (DMCE). This heuristic further decreases the probability of crashed ads or unutilized airtime.

Overbooking Level	Ad Distribution	l/s = 1/0.5 CE	MCE	DMCE	Greedy	l/s = 1/0.75 CE	MCE	DMCE	Greedy
0%	2 : 1	23.6%	2.7%	1.3%	0.0%	20.8%	2.4%	1.2%	4.1%
	1 : 1	26.0%	3.1%	1.5%	0.1%	22.5%	3.0%	1.4%	6.7%
	1 : 2	30.9%	3.4%	1.9%	0.6%	28.1%	3.1%	1.8%	8.6%
16.67%	2 : 1	24.9%	4.7%	2.1%	0.0%	21.7%	4.1%	2.0%	4.1%
	1 : 1	26.5%	5.1%	2.2%	0.0%	23.0%	4.2%	2.0%	7.5%
	1 : 2	31.4%	5.1%	2.3%	0.0%	27.9%	4.6%	2.3%	10.6%

Table 2 Overbooking Level, Ad Distribution, and Revenue Gap without Diversity

7.2.2. Greedy Heuristic The greedy heuristic maximizes current revenue by solving a knapsack problem for each break of known duration. Because the size of each individual knapsack problem is small, the greedy heuristic solves the scheduling problem quite efficiently.

7.2.3. Perfect Information We test the algorithms described so far against the revenue observed under perfect information (i.e., when all break lengths are known in advance). (Recall that, in the absence of diversity constraints, there is no advantage to having perfect information and we can implement the optimal policy.) The perfect information knapsack problem with diversity constraints is solved with a minimum precision of 99.925%.

7.3. Simulation Results

Table 2 and Table 3, respectively, report results for the case without and with diversity constraint. We find that DMCE offers a significant revenue increase over MCE owing to the former's better utilization of break time. Yet even the DMCE markedly underperforms the perfect information schedule irrespective of the existence of diversity constraints.

When short ads are priced linearly or at a discount, the greedy heuristic outperforms the DMCE and yields close to the optimal revenue. When short ads sell at a premium, however, the greedy heuristic's performance falls sharply and may even underperform that of the DMCE. The greedy heuristic's underperformance is exacerbated when there are fewer short ads. Yet we also find that the greedy heuristic's performance is *not* significantly affected by the diversity constraint.

Performance differences among heuristics arise from the confluence of two facts. First, only the greedy heuristic bases the length of its ad bundles on the actual (observed) break length; the use by the class of CE heuristics of predetermined bundles can lead to unutilized airtime. Second, airtime utilization is the single largest driver of broadcaster revenue. The complexity introduced by diversity constraints, which one might expect to be most strongly reflected in the performance of the greedy heuristic, has a minor impact as long as the broadcaster's ad inventory is reasonably well balanced. Therefore, except in cases of extreme advertiser concentration or

Overbooking Level	Ad Distribution	l/s = 1/0.5 CE	MCE	DMCE	Greedy	l/s = 1/0.75 CE	MCE	DMCE	Greedy
16.67%	2 : 1	26.0%	8.1%	3.0%	0.0%	20.3%	2.3%	0.2%	2.3%
	1 : 1	26.8%	4.9%	2.2%	0.0%	22.5%	3.7%	1.5%	7.0%
	1 : 2	31.4%	5.2%	2.3%	0.3%	27.9%	4.7%	2.3%	10.6%

Table 3 **Overbooking Level, Ad Distribution, and Revenue Gap with Diversity**

high premiums on short ads, the greedy heuristic's inability to anticipate the future need for short ads will not lead to significant revenue loss and the greedy heuristic performs well.

8. Conclusion

The problem of how best to schedule television advertising has been widely studied in the literature for the case of deterministic commercial breaks. In live broadcasting, however, the duration of commercial breaks is often unknown at the time of scheduling. Many live events such as sports and election coverage enjoy high ratings, and their advertising slots command a significant premium over regular shows. For this reason, appropriate scheduling to maximize the utilization of commercial airtime can greatly enhance a broadcaster's revenue.

In the absence of diversity constraints, we can determine the optimal scheduling rule: an easy-to-implement, look-ahead greedy policy that achieves the perfect-information profit. Even though overselling naturally follows from the stochastic nature of the total capacity, it is important to study how a broadcaster's ad portfolio should be balanced between short and long ads. Short ads are valued for the flexibility they lend to scheduling; in fact, the optimal portfolio would consist exclusively of short ads if there were no premium on long ads.

If the broadcaster must respect advertiser-specific diversity constraints, then the optimal policy cannot be found analytically without assuming there are only two ad lengths and two break lengths. In this case there is no value to perfect information, just as in the optimal scheduling case. In order to facilitate workable combinations of short and long ads, the broadcaster should avoid excessive concentration of advertisers. The concentration pattern of advertisers influences the revenue of the broadcaster, and increased concentration of high-paying short ads is detrimental to total revenue. Because optimal scheduling cannot be determined in the more general case of many different break lengths, we assess the performance of the greedy heuristic against the optimal policy not only for the case of two break lengths but also for a more general set of values as generated by some numerical experiments. The greedy heuristic's performance is mainly affected by concentration patterns and the portfolio composition in terms of short and long ads. In particular, either concentration at the low-value end of the broadcaster's portfolio or a premium on short ads will impair the greedy heuristic's performance.

Finally, we used a large-scale numerical analysis comparing several scheduling heuristics and found that the greedy heuristic performs well under many scenarios. Revenue generated by the greedy heuristic is close to the optimal value except when short ads sell at a premium. We have seen that overselling levels, portfolio mix, and diversity constraints all have little impact on the relative performance of the different heuristics considered. It follows that the broadcaster can seldom improve on using the greedy (or modified greedy) heuristic to schedule ads in their live transmissions.

Future research on this topic should explore the type of contracts that broadcasters could write. Along these lines, two promising avenues that could be pursued by the broadcaster are (i) manipulating the relative (sales) price of short and long ads and (ii) designing a range of price levels (and concomitant service guarantees) that will extract maximum value from the advertisers.

Appendix

PROOF OF PROPOSITION 1: For simplicity and w.l.g. we can assume that the number of sold ads, M, is infinite. This will imply that an infinite number of ads will have revenue and duration zero. Also recall that the set of revenues for long and short ads, $\mathbf{l} = \{l_1, ..., l_k, ...\}$ and $\mathbf{s} = \{s_1, ..., s_k, ...\}$, are ordered such that $l_i \geq l_{i+1}$ and $s_i \geq s_{i+1}$ for all positive integers i.

We will prove a more general statement of the proposition which assumes $L = kS$ where $k \geq 2$ is integer. We must show that, for any L-second break, we will schedule the longer ad whenever the inequality $l_1 \geq s_{N-n+1} + s_{N-n+2} + ... + s_{N-n+k}$ holds. In all other cases, we will schedule the first k shorter ads (i.e., $s_1, s_2, ..., s_k$) during that break. We will prove, by backward induction over $n = 1, N$.

At the last stage N, the policy is straightforward: if the break is short then we schedule s_1; else we schedule l_1 if $l_1 \geq s_1 + s_2 + ... + s_k$ or schedule $(s_1, s_2, ..., s_k)$ if $l_1 < s_1 + s_2 + ... + s_k$.

At stage $N-1$, if the break is short then we schedule s_1; if the break is long and if $l_1 \geq s_2 + s_3 + ... + s_{k+1}$, then the following inequality holds:

$$l_1 + E_{b_N}[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N)] \geq s_1 + s_2 + ... + s_k + E_{b_N}[V_N(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_N)]. \tag{8}$$

In other words, $V_{N-1}(\mathbf{l}, \mathbf{s}, L) = l_1 + E_{b_N}[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N)]$. To see this, note that $E_{b_N}[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N)] = \alpha s_1 + (1-\alpha)\max(l_2, s_1 + s_2 + ... + s_k) \geq s_1 + (1-\alpha)(s_2 + ... + s_k)$ and $E_{b_N}[V_N(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_N)] = \alpha s_{k+1} + (1-\alpha)\max(l_1, s_{k+1} + ... + s_{2k}) = \alpha s_{k+1} + (1-\alpha) l_1$. Then

$$E[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N) - V_N(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_N)] \geq s_1 + ... + s_k - l_1 - \alpha(s_2 + ... + s_{k+1} - l_1). \tag{9}$$

Thus, inequality (9) together with $l_1 \geq s_2 + ... + s_{k+1}$ implies (8). This means that if $l_1 \geq s_2 + ... + s_{k+1}$ then l_1 is the optimal scheduling policy at stage $N-1$.

We now show that if $l_1 < s_2 + ... + s_{k+1}$ then it is optimal to schedule $(s_1, ..., s_k)$ when the break is long. For this it is sufficient to show that

$$l_1 + E[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N)] < s_1 + ... + s_k + E[V_N(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_N)]. \tag{10}$$

To see this, observe that

$$E[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N)] = \alpha s_1 + (1-\alpha)\max(l_2, s_1 + ... + s_k) = s_1 + (1-\alpha)(s_2 + ... + s_k), \tag{11}$$

$$E[V_N(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_N)] = \alpha s_{k+1} + (1-\alpha)\max(l_1, s_{k+1} + ... + s_{2k}) \geq \alpha s_{k+1} + (1-\alpha) l_1. \tag{12}$$

Equality (11), inequality (12), and $l_1 < s_2 + ... + s_{k+1}$ together imply (10). As a result, if $l_1 < s_2 + ... + s_{k+1}$ then $s_1, ..., s_k$ is the optimal scheduling policy at stage $N-1$.

Inductive Step: We assume that the policy is optimal at stages $n+1$, ..., N and then prove that it is also optimal at stage n. At stage n, if the break is short then we schedule s_1. If the break is long and if $l_1 \geq s_{N-n+1} + ... + s_{N-n+k}$, then the following inequality holds:

$$l_1 + E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] \geq s_1 + ... + s_k + E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_{n+1})]. \tag{13}$$

To see this, note that

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] = \alpha(s_1 + A) + (1-\alpha)\max(l_2 + B, s_1 + ... + s_k + C), \tag{14}$$

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_{n+1})] = \alpha(s_{k+1} + D) + (1-\alpha)\max(s_{k+1} + ... + s_{2k} + F, l_1 + C), \tag{15}$$

where $A = E_{b_{n+2}}[V_{n+2}(\mathbf{l} - l_1, \mathbf{s} - s_1, b_{n+2})]$, $B = E_{b_{n+2}}[V_{n+2}(\mathbf{l} - \{l_1, l_2\}, \mathbf{s}, b_{n+2})]$, $C = E_{b_{n+2}}[V_{n+2}(\mathbf{l} - l_1, \mathbf{s} - \{s_1, ..., s_k\}, b_{n+2})]$, $D = E_{b_{n+2}}[V_{n+2}(\mathbf{l}, \mathbf{s} - \{s_1, s_2, ..., s_{k+1}\}, b_{n+2})]$, $F = E_{b_{n+2}}[V_{n+2}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_{2k}\}, b_{n+2})]$.

From (14) we have $E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] \geq \alpha(s_1 + A) + (1-\alpha)(s_1 + ... + s_k + C) = s_1 + ... + s_k + \alpha(A - s_2 - ... - s_k) + (1-\alpha)C$. From the induction hypothesis, since $l_1 \geq s_{N-n+1} + ... + s_{N-n+k}$ it follows that $l_1 + A \geq s_2 + ... + s_{k+1} + D$. Then $E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] \geq s_1 + ... + s_k + \alpha(s_{k+1} + D - l_1) + (1-\alpha)C$. From the induction hypothesis, since $l_1 \geq s_{N-n+1} + ... + s_{N-n+k}$ it follows that $l_1 + C \geq s_{k+1} + ... + s_{2k} + F$. Then we have $E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_{n+1})] = \alpha(s_{k+1} + D) + (1-\alpha)(l_1 + C)$. Therefore,

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1}) - V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_{n+1})] \geq s_1 + ... + s_k - l_1. \tag{16}$$

But (16) implies (13). Hence, we get that the optimal policy at stage n, is to schedule l_1 if $l_1 \geq s_{N-n+1} + ... + s_{N-n+k}$.

If $l_1 < s_{N-n+1} + ... + s_{N-n+k}$ then, by the induction hypothesis, $l_2 + B \leq l_1 + B < s_2 + ... + s_{k+1} + C$. To see this, note that by the induction hypothesis, at stage $n+1$ when sets $\mathbf{l} - l_1$ and $\mathbf{s}$ are available, then one should schedule $s_1, ..., s_k$ when $l_2 < s_{N-n} + ... + s_{N-n+k-1}$. But $l_2 \leq l_1 < s_{N-n+1} + ... + s_{N-n+k} < s_{N-n} + ... + s_{N-n+k-1}$. Then we have $E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] = \alpha(s_1 + A) + (1-\alpha)(s_1 + ... + s_k + C)$ and $E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_{n+1})] \geq \alpha(s_{k+1} + D) + (1-\alpha)(l_1 + C)$. Therefore,

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1}) - V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_{n+1})] < s_1 + ... + s_k - l_1. \tag{17}$$

We used the fact that from the induction hypothesis, since $l_1 < s_{N-n+1} + ... + s_{N-n+k}$ it follows that $l_1 + A < s_2 + ... + s_{k+1} + D$. From (17) it follows that the optimal policy, at stage n, is to schedule $(s_1, ..., s_k)$ if $l_1 < s_{N-n+1} + ... + s_{N-n+k}$. ■

Proof of Proposition 2: We show that even with perfect information of all break lengths, we cannot do a better schedule. To avoid additional complications, and w.l.o.g., we assume as before that there are infinite number of ads, some of them with yield 0. Thus, assume $(s_1^n, s_2^n, ...)$ and $(l_1^n, l_2^n, ...)$ are the revenues for the remaining short ads and long ads respectively at stage n, in decreasing order of magnitude.

Let us now show that our scheduling policy is pathwise optimal (i.e., it is optimal irrespective of the future realizations of break durations). Note that if the break is short (i.e., $b_n = S$), there is no choice and the scheduler will schedule the highest revenue generating short ad. This will be optimal, under perfect information, irrespective of the future sample path for the remaining breaks. If $b_n = L$ and there are $N-n$ breaks remaining, the scheduler will schedule the highest revenue long ad, provided that it generates more than the highest $N-n+1$ and $N-n+2$ short ads together.

Moreover, if break n is long and $l_1^n \geq s_{N-n+1}^n + s_{N-n+2}^n$, then l_1^n will be in the optimal schedule on any sample path $(b_n, b_{n+1}, ..., b_N)$ (under perfect information). To see this, note that the combined duration of all these breaks has a lower bound of $(N-n)S + L$, which is tight on the sample path where all remaining $(N-n)$ breaks are short. On that sample path, the optimal schedule is $(l_1^n, s_1^n, ..., s_{N-n}^n)$. On all other sample paths, the combined duration of the breaks exceeds $(N-n)S + L$, hence l_1^n will be scheduled on all other sample paths as well.

Similarly, if break n is long and $l_1^n < s_{N-n+1}^n + s_{N-n+2}^n$, then (s_1^n, s_2^n) will be in the optimal schedule on any sample path $(b_n, b_{n+1}, ..., b_N)$. To see this, note that the combined duration of all these breaks has a lower bound of $(N-n)S + L$, which is tight on the sample path where all remaining $(N-n)$ breaks are short. On that sample path, the optimal schedule is $(s_1^n, ..., s_{N-n+1}^n, s_{N-n+2}^n)$. On all other sample paths, the combined duration of the breaks exceeds $(N-n)S+L$, hence there will be for sure a long break with two short ads in it.

This argument can be used recursively at every stage to show that the schedule at that stage is pathwise optimal. ■

PROOF OF PROPOSITION 3: The proof proceeds exactly as in the case of two break durations for $k=2$. We omit the details for space considerations. ■

PROOF OF PROPOSITION 4: Suppose a penalty proportional to the ad's yield is charged in case of the ad is not shown on air, i.e., βr_i, for all $i=1,M$. Then at stage N, the revenue to go function becomes:

$$\begin{cases} V_N(\mathbf{l},\mathbf{s},L) = \max\left\{l_1 - \sum_{i=2}^{n_L}\beta l_i - \sum_{i=1}^{n_S}\beta s_i,\ s_1+s_2-\sum_{i=1}^{n_L}\beta l_i - \sum_{i=3}^{n_S}\beta s_i\right\}, \\ V_N(\mathbf{l},\mathbf{s},S) = s_1 - \sum_{i=1}^{n_L}\beta l_i - \sum_{i=2}^{n_S}\beta s_i \end{cases}$$

where $n_L = |\mathbf{l}|$ and $n_S = |\mathbf{s}|$ represent the remaining number of long ads and short ads. As before, we assume n_L and n_S to be infinite (and an infinite number of ads will have zero revenue, and consequently zero penalty in case of no-air). At stage N, if the break is short, the optimal policy is to schedule s_1. If the break is long, then the optimal policy is to schedule l_1 whenever $l_1 - \sum_{i=2}^{n_L}\beta l_i - \sum_{i=1}^{n_S}\beta s_i \geq s_1+s_2-\sum_{i=1}^{n_L}\beta l_i - \sum_{i=3}^{n_S}\beta s_i$, or equivalently $(1+\beta)l_1 \geq (1+\beta)(s_1+s_2)$, which is the same as $l_1 \geq (s_1+s_2)$.

At stage $N-1$, if the break is short then we schedule s_1; if the break is long and if $l_1 \geq s_2+s_3$, then the following inequality holds:

$$l_1 + E_{b_N}[V_N(\mathbf{l}-l_1,\mathbf{s},b_N)] \geq s_1+s_2+E_{b_N}[V_N(\mathbf{l},\mathbf{s}-\{s_1,s_2\},b_N)]. \tag{18}$$

To see this, note that $E_{b_N}[V_N(\mathbf{l}-l_1,\mathbf{s},b_N)] = \alpha(s_1 - \sum_{i=2}^{n_L}\beta l_i - \sum_{i=2}^{n_S}\beta s_i) + (1-\alpha)\max(l_2 - \sum_{i=3}^{n_L}\beta l_i - \sum_{i=1}^{n_S}\beta s_i, s_1+s_2-\sum_{i=2}^{n_L}\beta l_i - \sum_{i=3}^{n_S}\beta s_i) \geq s_1 - \sum_{i=2}^{n_L}\beta l_i - \sum_{i=3}^{n_S}\beta s_i - \alpha\beta s_2 + (1-\alpha)s_2$ and $E_{b_N}[V_N(\mathbf{l},\mathbf{s}-\{s_1,s_2\},b_N)] = \alpha(s_3 - \sum_{i=1}^{n_L}\beta l_i - \sum_{i=4}^{n_S}\beta s_i) + (1-\alpha)\max(l_1 - \sum_{i=2}^{n_L}\beta l_i - \sum_{i=3}^{n_S}\beta s_i, s_3+s_4-\sum_{i=1}^{n_L}\beta l_i - \sum_{i=5}^{n_S}\beta s_i) = \alpha(s_3-\beta l_1) + (1-\alpha)(l_1-\beta s_3) - \sum_{i=2}^{n_L}\beta l_i - \sum_{i=4}^{n_S}\beta s_i$. Then

$$\begin{aligned} & E[V_N(\mathbf{l}-l_1,\mathbf{s},b_N) - V_N(\mathbf{l},\mathbf{s}-\{s_1,s_2\},b_N)] \geq \\ & \geq s_1+s_2-l_1-\beta s_3-\alpha\beta s_2+\alpha\beta l_1+(1-\alpha)\beta s_3-\alpha(s_2+s_3-l_1) \\ & = s_1+s_2-l_1-\alpha(1+\beta)(s_2+s_3-l_1) \end{aligned} \tag{19}$$

Thus, inequality (19) together with $l_1 \geq s_2+s_3$ implies (18). This means that if $l_1 \geq s_2+s_3$ then l_1 is the optimal scheduling policy at stage $N-1$.

We now show that if $l_1 < s_2 + s_3$ then it is optimal to schedule (s_1, s_2) when the break is long. For this it is sufficient to show that

$$l_1 + E[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N)] < s_1 + s_2 + E[V_N(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_N)]. \tag{20}$$

To see this, observe that

$$E[V_N(\mathbf{l} - l_1, \mathbf{s}, b_N)] = s_1 - \sum_{i=2}^{n_L} \beta l_i - \sum_{i=3}^{n_S} \beta s_i - \alpha\beta s_2 + (1-\alpha)s_2, \tag{21}$$

$$E[V_N(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_k\}, b_N)] \geq \alpha(s_3 - \beta l_1) + (1-\alpha)(l_1 - \beta s_3) - \sum_{i=2}^{n_L} \beta l_i - \sum_{i=4}^{n_S} \beta s_i. \tag{22}$$

Equality (21), inequality (22), and $l_1 < s_2 + s_3$ together imply (20). As a result, if $l_1 < s_2 + s_3$ then s_1, s_2 is the optimal scheduling policy at stage $N-1$.

Inductive Step: We assume that the policy is optimal at stages $n+1$, ..., N and then prove that it is also optimal at stage n. The proof follows exactly the same steps as the proof of the inductive step in Proposition 1. For the sake of completion, we reproduce it below. At stage n, if the break is short then we schedule s_1. If the break is long and if $l_1 \geq s_{N-n+1} + s_{N-n+2}$, then the following inequality holds:

$$l_1 + E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] \geq s_1 + s_2 + E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_{n+1})]. \tag{23}$$

To see this, note that

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] = \alpha(s_1 + A) + (1-\alpha)\max(l_2 + B, s_1 + s_2 + C), \tag{24}$$

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_{n+1})] = \alpha(s_3 + D) + (1-\alpha)\max(s_2 + s_4 + F, l_1 + C), \tag{25}$$

where $A = E_{b_{n+2}}[V_{n+2}(\mathbf{l} - l_1, \mathbf{s} - s_1, b_{n+2})]$, $B = E_{b_{n+2}}[V_{n+2}(\mathbf{l} - \{l_1, l_2\}, \mathbf{s}, b_{n+2})]$, $C = E_{b_{n+2}}[V_{n+2}(\mathbf{l} - l_1, \mathbf{s} - \{s_1, s_2\}, b_{n+2})]$, $D = E_{b_{n+2}}[V_{n+2}(\mathbf{l}, \mathbf{s} - \{s_1, s_2, s_3\}, b_{n+2})]$, $F = E_{b_{n+2}}[V_{n+2}(\mathbf{l}, \mathbf{s} - \{s_1, ..., s_4\}, b_{n+2})]$.

From (24) we have $E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] \geq \alpha(s_1 + A) + (1-\alpha)(s_1 + s_2 + C) = s_1 + s_2 + \alpha(A - s_2 - s_2) + (1-\alpha)C$. From the induction hypothesis, since $l_1 \geq s_{N-n+1} + s_{N-n+2}$ it follows that $l_1 + A \geq s_2 + s_2 + D$. Then $E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] \geq s_1 + s_2 + \alpha(s_3 + D - l_1) + (1-\alpha)C$. From the induction hypothesis, since $l_1 \geq s_{N-n+1} + s_{N-n+2}$ it follows that $l_1 + C \geq s_3 + s_4 + F$. Then we have $E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_{n+1})] = \alpha(s_3 + D) + (1-\alpha)(l_1 + C)$. Therefore,

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1}) - V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_{n+1})] \geq s_1 + s_2 - l_1. \tag{26}$$

But (26) implies (23). Hence, we get that the optimal policy at stage n, is to schedule l_1 if $l_1 \geq s_{N-n+1} + s_{N-n+2}$.

If $l_1 < s_{N-n+1} + s_{N-n+2}$ then, by the induction hypothesis, $l_2 + B \leq l_1 + B < s_2 + s_3 + C$. To see this, note that by the induction hypothesis, at stage $n+1$ when sets $\mathbf{l} - l_1$ and $\mathbf{s}$ are available, then one should schedule s_1, s_2 when $l_2 < s_{N-n} + s_{N-n+1}$. But $l_2 \leq l_1 < s_{N-n+1} + s_{N-n+2} < s_{N-n} + s_{N-n+1}$. Then we have $E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1})] = \alpha(s_1 + A) + (1-\alpha)(s_1 + s_2 + C)$ and $E_{b_{n+1}}[V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_{n+1})] \geq \alpha(s_3 + D) + (1-\alpha)(l_1 + C)$. Therefore,

$$E_{b_{n+1}}[V_{n+1}(\mathbf{l} - l_1, \mathbf{s}, b_{n+1}) - V_{n+1}(\mathbf{l}, \mathbf{s} - \{s_1, s_2\}, b_{n+1})] < s_1 + s_2 - l_1. \tag{27}$$

We used the fact that from the induction hypothesis, since $l_1 < s_{N-n+1} + s_{N-n+2}$ it follows that $l_1 + A < s_2 + s_3 + D$. From (27) it follows that the optimal policy, at stage n, is to schedule (s_1, s_2) if $l_1 < s_{N-n+1} + s_{N-n+2}$.

Also, a similar proof can be constructed for the case where the penalty is proportional to the length of the ad (i.e., βL for all long ads and βS for all short ads). ■

Proof of Proposition 5: We will show that the scheduling rule described by the proposition achieves the same profit as under perfect information.

Stage n **(short break):** Take the maximum $t \leq N - n$ such that $l_t \geq s_{N-n-t+2} + s_{N-n-t+3}$. Let n_S and n_L denote the respective number of short and long breaks in a realized schedule going forward, where $n_S + n_L = N - n + 1$.

- $n_L \leq t$ The perfect information schedule uses the first n_L long ads in each long break and the first $n_S = N - n + 1 - n_L \geq N - n + 1 - t$ short ads in the short breaks. Thus, any realized schedule with $n_L \leq t$ shows the first $N - n + 1 - t$ short ads. Short ads are not shown in long breaks, and diversity does not matter. Any short ad can be selected from the first $N - n - t + 1$ short ads.
- $n_L = t + j$ with $0 < j \leq N - n - t$ The perfect information schedule uses the first t long ads. In the $N - n + 1 - (t + j)$ short breaks and at least $\lceil j/2 \rceil$ long breaks, the optimal schedule shows short ads subject to the diversity constraint. Let $k \geq 0$ be the largest such k for which $s_{N-n-t+1+k}$ belongs to the same advertiser as $s_{N-n-t+1}$. Then the ads $(s_{N-n-t+1}, \ldots, s_{N-n-t+1+k})$ earn the same revenue. Since $l_{t+1} < s_{N-n-t+1} + s_{N-n-t+2}$, it follows that l_{t+1} will not be scheduled before all $s_{N-n-t+1+k}$ short ads have been shown (again subject to the diversity constraint). Choose the short ad from $N - n - t + 1$ that maximizes the total number of combinations within the first $N - n - t + 1 + k$ short ads—in other words, take a short ad from the advertiser with the most ads in the first $N - n - t + 1 + k$ short ads.

Thus, the short ad shown at stage n should belong to the first $N - n - t + 1$ ads and should maximize diversity. This is accomplished by choosing an ad from the advertiser with the smallest i^* such that $m_s^{h^s(i)} = \max_{h^s(1) \leq h \leq h^s(N-n-t+1)} \{m_s^h\}$.

Stage n (long break): At this stage we need to choose between scheduling long ad l_1 or a combination of two short ads. If $l_1 \geq s_{N-n+1} + s_{N-n+2}$ then we schedule the long ad; otherwise, we schedule two short ads (see Proposition 1). The diversity constraint makes finding an acceptable combination of short ads a nontrivial task.

Case 1: $l_1 \geq s_{N-n+2} + s_{N-n+3}$.

- $n_L = 1$ The perfect information schedule selects from the first $N - n + 2$ short ads to form one pair. Because no single advertiser can place more than $N - n + 1$ ads, any pair of short ads can be scheduled in the current break.
- $n_L > 1$ The perfect information schedules for all possible realizations involving two or more long breaks (out of the $N - n + 1$ remaining breaks) will use at least $N - n + 1$ short ads, possibly in long breaks, as well as the long ad l_1.

Thus, the short ads chosen for the current break should belong to the first $N - n + 1$ short ads and should maintain the maximum possible future combinations. This means that the broadcaster should first schedule the most profitable short ad and then select the other ad from the advertiser with the highest number of ads in the first $N - n + 1$ ads.

Case 2: $l_1 < s_{N-n+2} + s_{N-n+3}$.

- $n_L = 1$ The schedule will be the same as under the same condition in Case 1.
- $n_L > 1$ The perfect information schedules for all possible realizations involving two or more long breaks (out of $N - n + 1$ remaining breaks) will use at least $N - n + 3$ short ads.

Thus, the short ads chosen for the current break should belong to the first $N - n + 2$ short ads and should maintain the maximum possible future combinations. That is, the broadcaster first schedules the most profitable ad and then selects the other ad from the advertiser with the highest number of ads in the first $N - n + 2$ ads.

At each stage n, the procedure schedules ad(s) that belong to all perfect information schedules for all possible realizations of remaining breaks. It therefore achieves the perfect information profit. ■

Proof of Corollary 1: See the proof of Proposition 5. ■

Proof of Proposition 6: If there is a premium for long ads, then the optimal scheduling policy will always schedule a long ad, if available, in a long break. The expected revenue, given the optimal scheduling policy, is the following:

$$V(n_L) = \sum_{i=0}^{n_L} [i * l + \min\{n_S, N - i\} * s] \Pr(k = i) + \\ + \sum_{i=n_L+1}^{N} [n_L * l + \min\{n_S, N + i - 2n_L\} * s] \Pr(k = i),$$

where k is the random number of long breaks, n_S is the number of (booked) short ads of revenue s, n_L is the number of (booked) long ads of revenue l and $l = 2s(1+\epsilon)$. Also, n_L and n_S must satisfy $n_s = (1+\delta)(2-\alpha)N - 2n_L$, where δ is the level of overbooking; note that we assumed, for simplicity (and w.l.o.g.), that $(1+\delta)(2-\alpha)N$ is integer.

The revenue from booking $n_L + 1$ ads will then be:

$$\begin{aligned} V(n_L+1) = & \sum_{i=0}^{n_L+1} [i * l + \min\{n_S - 2, N - i\} * s] \Pr(k=i) + \\ & + \sum_{i=n_L+2}^{N} [(n_L+1) * l + \min\{n_S - 2, N + i - 2n_L - 2\} * s] \Pr(k=i) \\ = & \sum_{i=0}^{n_L+1} [i * l + \min\{n_S - 2, N - i\} * s] \Pr(k=i) + \\ & + \sum_{i=n_L+2}^{N} [(n_L+1) * l + \min\{n_S, N + i - 2n_L\} * s - 2s] \Pr(k=i). \end{aligned}$$

The marginal revenue from booking one extra long ad is given by:

$$\begin{aligned} \Delta V(n_L) &= V(n_L+1) - V(n_L) \\ &= (l-2s)\Pr(k \geq n_L+1) - 2s\Pr(k \leq \min(N-n_S, n_L)) - s\Pr(k = N - n_S + 1 \wedge k \leq n_L) \\ &= 2\epsilon s \Pr(k \geq n_L+1) - 2s\Pr(k \leq \min(N-n_S, n_L)) - s\Pr(k = N - n_S + 1 \wedge k \leq n_L) \\ &= \begin{cases} 2\epsilon s \Pr(k \geq n_L+1) - 2s\Pr(k \leq N - n_s) - s\Pr(k = N - n_S + 1), & \text{if } N - n_S + 1 \leq n_L \\ 2\epsilon s \Pr(k \geq n_L+1) - 2s\Pr(k \leq n_L), & \text{if } N - n_S \geq n_L. \end{cases} \end{aligned}$$

The reasoning is as follows:

- If $k > n_L$ (i.e., the number of long breaks is strictly greater than the current number of long ads), we would have gained $2\epsilon s$ had we booked one extra long ad;
- If $k \leq n_L$, then

—if $n_S \leq N - k$, then we would have lost $2s$ had we booked one extra long ad; the reason is we would have had to book $n_S - 2$ short ads in order to maintain the same level of overbooking, δ. Since the current number of short ads, n_S, is lower than the number of short breaks $(N-k)$, then by booking $n_S - 2$ short ads we incur a revenue loss of $2s$ by airing two fewer short ads. No revenue will be gained from the extra long ad, because the number of long breaks k is already less than or equal to the current number of long ads.

—if $n_S = N - k + 1$, then we would have lost s had we booked one extra long ad;

—if $n_S \geq N - k + 2$, then there would be no impact on revenue.

It is easy to see that the marginal revenue is a decreasing function of n_L. Note that the first term is decreasing in n_L, while the second and last terms together are increasing in n_L, but have negative signs in the expression (we used the fact that for a discrete random variable k, we have

$2\Pr(k \le x) + \Pr(k = x+1) \ge 2\Pr(k \le x-1) + \Pr(k = x)$. That is because $2\Pr(k \le x) + \Pr(k = x+1) - 2\Pr(k \le x-1) - \Pr(k = x) = 2\Pr(k = x) - \Pr(k = x) + \Pr(k = x+1) = \Pr(k = x) + \Pr(k = x+1) \ge 0$).

Hence the optimal number of long ads to book, n_L^*, is the smallest integer such that

$$2s\epsilon\Pr(k \ge n_L + 1) \le 2s\Pr(k \le \min(N - n_S, n_L)) + s\Pr(k = N - n_S + 1 \wedge k \le n_L), \tag{28}$$

or equivalently

$$2\epsilon\Pr(k \ge n_L + 1) \le 2\Pr(k \le \min(N - (1+\delta)(2-\alpha)N + 2n_L, n_L)) + \\ + \Pr(k = N - (1+\delta)(2-\alpha)N + 2n_L + 1 \wedge k \le n_L). \tag{29}$$

■

References

[1] Araman, V. F. and I. Popescu. 2009. Media Revenue Management with Audience Uncertainty: Balancing Upfront and Spot Market Sales. *Manufacturing and Service Operations Management*. **12** 190-212.

[2] Bauer Insight. 2010. Advertising: The Super Bowl vs. Winter Olympics. *http://blogs.bauer.uh.edu/faculty/uncategorized/advertising-the-super-bowl-vs-winter-olympics/*.

[3] Birge, J., J. B. G. Frenk, J. Mittenthal and A. H. G. Rinnooy Kan. 1990. Single-machine scheduling subject to stochastic breakdown. *Naval Research Logistics*. **37** 661-677.

[4] Bloomberg BusinessWeek. 2010. CBS loss narrows as advertising revenue rebounds. May 5, 2010.

[5] Bollapragada, S., H. Cheng, M. Phillips, M. Garbiras, M. Scholes, T. Gibbs, M. Humphreville. 2002. NBC's Optimization Systems Increase Revenues and Productivity. *Interfaces* **32**(1) 47-60.

[6] Bollapragada, S., M.R. Bussieck, and S. Mallik. 2004. Scheduling Commercial Videotapes in Broadcast Television. *Operations Research*. **52** 679-689.

[7] Bollapragada, S. and M. Garbiras. 2004. Scheduling Commercials on Broadcast Television. *Operations Research*. **52** 337-345.

[8] Cai, X., X. Wu and X. Zhou. 2009. Stochastic Scheduling Subject to Preemptive-Repeat Breakdowns with Incomplete Information. *Operations Research*. **57** 1236-1249.

[9] Ciarallo, F., R. Akella, and T. Morton. 1994. A periodic review, production planning model with uncertain capacity and uncertain demand optimality of extended myopic policies. *Management Science*. **40** 320-332.

[10] Ghodsi, R. and F. Sassani. 2005. Real Time Optimum Sequencing of Wood Cutting Process. *International Journal of Production Research*. **43** 1127-1141.

[11] Gilmore, P.C. and R.E. Gomory. 1961. A Linear Programming Approach to the Cutting Stock Problem. *Operations Research*. **9** 849-859.

[12] Gilmore, P.C. and R.E. Gomory. 1965. Multistage Cutting Stock Problems of Two and More Dimensions. *Operations Research*. **13** 94-120.

[13] Grosfeld-Nir, A. and Y. Gerchak. 2004. Multiple lotsizing in production to order with random yields: Review of recent advances. *Annals of Operations Research*. **126** 43-69.

[14] Hwang, J. and M. Singh. 1998. Optimal production policies for multi-stage systems with setup costs and uncertain capacities. *Management Science*. **44** 1279-1294.

[15] Khang, D. and O. Fujiwara. 2000. Optimality of myopic ordering policies for inventory model with stochastic supply. *Operations Research*. **48** 181-184.

[16] Kimms, A. and M. Müller-Bungart. 2007. Revenue management for broadcasting commercials: the channel's problem of selecting and scheduling the advertisements to be aired. *International Journal of Revenue Management*. **1** 28-44.

[17] Martello, S. and P. Toth. 1990. Knaspack Problems - Algorithms and Computer Implementations. John Willey & Sons, Chichester.

[18] Newstead, K. and J. Romaniuk. 2010. Cost Per Second: The Relative Effectiveness of 15- and 30-Second Television Advertisements. *Journal of Advertising Research*. **50** 68-76.

[19] Phillips, R. and G. Young. 2010. Television Advertisement Pricing in the U.S. *Oxford Handbook of Pricing Management*. Ed. Ozalp Ozer and Robert Phillips. Oxford University Press, Oxford.

[20] Pinedo, M. and E. Rammouz. 1988. A Note on Stochastic Scheduling on a Single Machine Subject to Breakdown and Repair. *Probability in the Engineering and Informational Sciences*. **2** 41-49.

[21] Roels, G. and K. Fridgeirsdottir. 2009. Dynamic Revenue Management for Online Display Advertising. *Journal of Pricing and Revenue Management*. **8** 452-466.

[22] Scull, D. 1981. A stochastic cutting stock procedure: cutting rolls of insulating tape. *Management Science*. **27** 946-952.

[23] Tomlin, B. 2009. Impact of Supply Learning When Suppliers Are Unreliable. *Manufacturing & Service Operations Management* **11** 192-209.

[24] Wang, Y. and Y. Gerchak. 1996. Periodic review production models with variable capacity, random yield, and uncertain demand. *Management Science*. **42** 130-137.

[25] Yano, C. and H. Lee. 1995. Lot sizing with random yields: A review. *Operations Research*. **43** 311-334.