Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Ad Revenue Optimization in Live Broadcasting

Dana G. Popescu, Pascale Crama

MANAGEMENT SCIENCE
Vol. 62, No. 4, April 2016, pp. 1145–1164
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2015.2185

Ad Revenue Optimization in Live Broadcasting

Dana G. Popescu
Department of Technology and Operations Management, INSEAD, Singapore 138676,
dana.popescu@insead.edu

Pascale Crama
Department of Operations Management, Lee Kong Chian School of Business, Singapore Management University,
Singapore 178899, pcrama@smu.edu.sg

In live broadcasting, the break lengths available for commercials are not always fixed and known in advance (e.g., strategic and injury time-outs are of variable duration in live sports transmissions). Broadcasters actively manage their advertising revenue by jointly optimizing sales and scheduling policies. We characterize the optimal dynamic schedule in a simplified setting that incorporates stochastic break durations and advertisement lengths of 15 and 30 seconds. The optimal policy is a "greedy" look-ahead rule that accounts for the remaining number of breaks; in this setting, there is no value to perfect information at the scheduling stage, and hence knowing the duration of all breaks would not change the schedule. We present heuristics to help solve scheduling problems of even greater complexity. The performance of these heuristics under various scenarios is tested by running simulations calibrated using industry data. The simple greedy heuristic is shown to perform well except when revenues are concave in ad length, in which case the look-ahead aspect of the optimal schedule becomes more important. Finally, we recommend ways for broadcasters to balance their portfolio of booked ads by determining the optimal overbooking level and mix of ads as a function of their associated revenues generated and penalties incurred.

Keywords: live broadcasting; advertising; scheduling; random capacity
History: Received January 24, 2012; accepted February 2, 2015, by Yossi Aviv, operations management. Published online in *Articles in Advance* October 19, 2015.

1. Introduction

Broadcasters generate a large part of their revenue through advertising. For CBS, the most-watched U.S. broadcast network, TV advertising accounts for two thirds of total revenue (Fixmer 2010). Major live events—such as election night, 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 three times the rate of an average prime-time spot (Kacen 2010). However, the selling and scheduling of advertisements in such an environment is a challenging task. This is especially true 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.[1] The uncertainty about break durations creates an obvious problem for the broadcaster: how to schedule (in real time) the ads that have been sold while respecting the relevant constraints. The broadcaster must respect not only *capacity* constraints, under which the total duration of the ads scheduled during a break cannot exceed the length of that break, but also *diversity* constraints, under which two ads from the same advertiser (or from advertisers in the same industry) cannot be shown during the same break.

Suboptimal or infeasible schedules have many undesirable consequences for a broadcaster. If the schedule violates diversity constraints, then 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

[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 its players' positions in the field between successive overs, and ads can be shown during that time. As soon as the players have taken up their new positions, the game restarts and the broadcaster resumes its live coverage.

[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 also shown.

penalties.[3] For example, cricket broadcasting rights require the broadcaster to guarantee live coverage of every ball of every match.

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. When a break occurs, its duration is known to the scheduler. This paper illustrates how to optimize revenue from live events by deriving dynamic scheduling policies and improving the ad portfolio booking decisions.

Our research offers theoretical, practical, and managerial contributions. From a theoretical perspective, identifying the optimal scheduling policy allows us to specify the conditions under which uncertainty has no effect on the broadcaster's revenue (i.e., when perfect information (PI) has no value). In the absence of diversity constraints, the optimal policy is a *greedy look-ahead* rule that takes the remaining number of breaks into account irrespective of the probability distribution of break durations. We determine the minimum set of conditions under which perfect information holds no value, which helps us better understand when and why break duration uncertainty may pose a problem to the broadcaster. We find that one of the two conditions most likely to be violated in practice, a restriction on the ad lengths booked, is something that the broadcaster could affect—for example, by contractually limiting the ad lengths that can be submitted by advertisers. Doing so could help prevent uncertainty from creating losses in a wide range of scenarios, but it might not maximize the broadcaster's revenue.

We then expand our theoretical analysis to investigate a more realistic setting with multiple ad formats. We derive bounds for the value function and use different relaxations of the more general problem to craft simple yet efficient heuristics. We perform an extensive numerical analysis to compare the performance of these heuristics under scenarios characterized by various overbooking levels and revenue ratios for long and short ads. We calibrate our simulation parameters using industry data collected from a live cricket tournament. The most important lesson from the numerical analysis is that, under fairly general conditions and despite the problem's complexity, the simple greedy heuristic performs extremely well. Note that the factors affecting this heuristic's success—namely, the ratio of average revenue of short to long ads and the diversity profile—are at least partly under the broadcaster's control; hence, it can design portfolios that are robust to uncertainty.

Finally, our analysis extends to another fundamental concern: constructing the initial ad portfolio. When creating that portfolio, one can reduce the negative effects of uncertainty by making it *flexible*. On the one hand, the broadcaster must decide how much airtime to sell. Random capacity and high prices drive it to sell in excess of expected airtime capacity, which lowers the service level and leads to advertiser dissatisfaction or even penalties. Thus, the broadcaster will need to identify the level of *overbooking* that balances the trade-off between capacity utilization and penalty payments. At the same time, the broadcaster must also consider the ad portfolio's composition in terms of ad duration. Because short ads enhance scheduling flexibility, we find that the higher the variability in break duration, the higher the proportion of short ads in the portfolio and the greater the discount on short ads the broadcaster is willing to accept to retain scheduling flexibility. These insights on portfolio composition are of paramount importance to the first stage of contract negotiation and ad sales.

The analysis reveals that a broadcaster's initial question—namely, how best to schedule ads during breaks of uncertain duration—is a prelude to the more strategic issue of how to craft a profitable ad portfolio. Thus, the managerial contribution of our work is that it highlights the importance of devising a flexible portfolio with the right amount of overbooking. Encouraging advertisers to submit short ads through differential pricing of short and long ads, or even contractually requiring some short spots in the advertiser's mix, would be of great value to the broadcaster; the shorter the ad, the more flexibility it provides. Thus, the typical observed portfolio mix in our data (with its range of various ad lengths) is beneficial to the broadcaster even though this benefit arises serendipitously from market characteristics and not as a consequence of the broadcaster's own planning. If judiciously employed, proactive revenue management tools could generate more value than merely reactive changes in scheduling practices.

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 (2004) formulated a goal-oriented programming model to solve the scheduling problem. The emphasis in that model

[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. Show-cause notice for Ten Cricket channel. Accessed September 15, 2015, http://www.espncricinfo.com/).

is on satisfying as many product conflict constraints and ad position constraints as possible by choosing an appropriate penalty for each constraint that is violated with the objective of minimizing the total penalty cost incurred. Kimms and Müller-Bungart (2007) formulated an integer program that maximizes the broadcaster's revenue while accommodating product conflict constraints and specific scheduling requests; they proposed several heuristics and conducted extensive numerical analyses that compared the performance of different solution methods. In contrast with those papers, which assume that capacity is deterministic and so derive static scheduling policies, we derive dynamic scheduling policies that account for stochastic capacity and for the sequential arrival over time of breaks that must be filled immediately.

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 either ratings uncertainty (Araman and Popescu 2009) or demand uncertainty (Roels and Fridgeirsdottir 2009). In our case, uncertainty stems from the duration of breaks; this poses scheduling difficulties when ads are of various lengths and some of those lengths exceed the duration of some commercial breaks.

Much less attention has been given to the random yield caused by stochastic capacity. The work of Ciarallo et al. (1994) is the first to explore the effect 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 because it varies as a function of the current inventory level.

Our model differs in two important aspects from the multiperiod random capacity models advanced by the works just cited. First, those papers assume a single product, whereas we consider 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 consume. 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. In that scenario, if a machine breaks down during the processing of a job, then all work done on the job is lost and processing must begin again from scratch (Birge et al. 1990, Cai et al. 2009, Pinedo and Rammouz 1988). Birge et al. (1990) briefly discuss the preemptive repeat model in the context of jobs characterized by deterministic processing time. The objective is to minimize the total weighted completion time, and the optimal rule schedules jobs ranked in increasing order of a ratio of functions of processing time and weight. Cai et al. (2009) expand that 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 this paper, we simplify the job structure by restricting ourselves to deterministic ad durations of two lengths and by assuming that the duration of the *current* commercial break is known. This approach makes sense in our setting and allows us to focus on generating insights about which ads to air and how to select the ads to be considered for broadcasting (i.e., the ad portfolio) 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 the transportation industry, where airlines can ship cargo via freighter planes and also in the holds of their scheduled passenger flights. In the latter case, cargo capacity varies among 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 §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 (for the topology of cutting and packing problems, see Wäscher et al. 2007); 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 1990), 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. The packed items must not exceed the knapsack capacity, and the number of packed items across all knapsacks must be maximized. The generalization in our case results from three aspects of the knapsacks (i.e., the commercial breaks): they have stochastic capacities, they arrive sequentially over a period of time, and they must be filled immediately upon arrival. In broadcast advertising, the packed items must satisfy not only capacity constraints, but also diversity constraints. To the best of our knowledge, the literature on cutting and packing problems has addressed neither how the properties of the items to be packed affect the algorithm's performance, nor which such properties would be desirable.

3. Model (Base Case)

Consider a television network that has a random capacity of advertising airtime during a live event. We take as given the portfolio of ads to air during this event. The ad revenues and durations are variable. For the base case model, we assume that the length of each ad is either S or $2S$ seconds. Let $\mathbf{s}^o = \{s_1^o, s_2^o, \ldots\}$ and $\mathbf{l}^o = \{l_1^o, l_2^o, \ldots\}$ be the sets of revenues generated by ads of length S and $2S$, respectively. We use superscript "o" to denote the initial set of long and short ads. Later on, we will drop this superscript to refer to the sets of short and long ads not yet aired at the time a break occurs. For simplicity of exposition (and without loss of generality (w.l.o.g.)), we assume that the cardinality of sets $\mathbf{l}^o$ and $\mathbf{s}^o$ is infinite. Note that the revenue from many of these ads will be zero. We also standardize the length of the short ads to $S = 1$. Moreover, we order sets $\mathbf{l}^o$ and $\mathbf{s}^o$ such that $l_i^o \geq l_{i+1}^o$ and $s_i^o \geq s_{i+1}^o$ for all positive integers i.

The airtime capacity consists of N commercial breaks of random duration. We assume that when a break occurs, its duration is known to the scheduler. We use $\{b_1, \ldots, b_N\}$ to denote the set of breaks, where b_n represents the duration of break n. Then the total advertising air time capacity is given by $B = \sum_{i=1}^n b_i$. Let $D = \{d_1, \ldots, d_K\}$ be the set of all possible break durations (in seconds), and let $p_{n,j}$ be the probability of break n having duration d_j. For the base case model with two ad lengths, we allow the duration of the breaks to vary in multiples of S seconds. This setup is equivalent to a general nonnegative distribution because, if ads must run in their entirety, then any break length that is not an integer multiple of S seconds must be rounded down to the nearest such integer.

As mentioned in the introduction, networks typically face restrictions regarding which ads can air during a given break. For instance, two ads from the same advertiser cannot be shown in the same break (*advertiser diversity* constraint), and two ads for products in the same industry cannot be shown in the same break (*industry diversity* constraint). For now we shall ignore these diversity constraints and study optimal scheduling for the simpler problem with only capacity constraints. In our numerical results in §5, we will analyze the impact of these constraints on the performance of the proposed scheduling policies. We analyze an extension of the model with advertiser diversity constraints in the electronic companion (available as supplemental material at http://dx.doi.org/10.1287/mnsc.2015.2185).

Let $\mathbf{l}$ and $\mathbf{s}$ denote, respectively, the *ordered* sets of long and short ads not yet aired. To avoid excessive notation, we do not explicitly index sets $\mathbf{l}$ and $\mathbf{s}$ with the break number n; however, $\mathbf{l}$ and $\mathbf{s}$ always represent the sets of remaining long and short ads not yet aired before the current break occurs. The (expected) *revenue to go* at stage $1 \leq n \leq N$, given ad vectors $\mathbf{l}$ and $\mathbf{s}$, can be written as follows:

$$V_n(\mathbf{l},\mathbf{s}) = \sum_{j=1}^{K} p_{n,j} \max_{\lambda,\,\theta\in\mathbb{N}} \left\{ \sum_{i=1}^{\lambda} l_i + \sum_{i=1}^{\theta} s_i + V_{n+1}(\mathbf{l}\backslash\{l_1,\ldots,l_\lambda\}, \mathbf{s}\backslash\{s_1,\ldots,s_\theta\}) \mid 2\lambda+\theta \leq d_j \right\};$$

$$V_{N+1}(\mathbf{l},\mathbf{s}) = 0.$$

We study the optimal dynamic scheduling policy under this setting. Despite simplifying the original problem somewhat, this formulation still captures the problem's main characteristics: (i) stochastic capacity; (ii) a heterogeneous assortment of ads of different values that can be classified into "short ads" and "long ads"; (iii) a heterogeneous assortment of break durations, some of which cannot be maximally utilized using long ads only (i.e., commercial breaks with duration a multiple of $2S$ are the only ones that can be filled using long ads only); and (iv) limited airtime capacity, so that not all ads can be accommodated.

3.1. Model Discussion

We have made a number of assumptions to achieve a tractable version of the broadcaster's original problem. These assumptions are needed to generate closed-form theoretical results and clear insights into the cost of uncertainty, the optimal scheduling policy, and the portfolio selection question. Here we briefly discuss how these assumptions compare with the real-world problem.

First, we assume that the break length is known at the start of the break. This assumption is based on the broadcaster's deep knowledge of the live event. In cricket, for instance, broadcasters spend considerable effort studying the break distributions per type of bowler transition (i.e., fast or slow bowler) and for different styles of play. The person calling the length of the break to the ad scheduler takes these considerations—as well as other team- and match-specific factors—into account when judging the length of each individual break. Additionally, live broadcasting frequently has some amount of delay or "timeline shift" upon start-up. Such intentional delay reflects either organizational concerns (e.g., so that programs begin exactly on the minute or half-minute) or the desire to maintain a short time delay to guard against the broadcasting of inappropriate material (Redhead 2007). These short delays help with accurately estimating the current commercial break's length. It therefore seems reasonable to assume that scheduling challenges in a live broadcast arise less from uncertainty over the current break length than from uncertainty over future break lengths.

Second, most of our analytical results are obtained for two ad lengths, S and $2S$. This is a valid assumption in many parts of the world and across various sports. In the United States, 90% of ads are sold as 15- or 30-second spots (Green 2006). There is some variation for special events; for example, ads that air during the Super Bowl are either 30 or 60 seconds (yet thus still satisfy the "S or $2S$" assumption) (Stampler 2013). In Japan, most ads are 15 seconds; purchasing two consecutive 15-second slots commands a large premium, which explains the dearth of 30-second ads there (Kawashima 2006).[4] The European market deviates from this norm; for instance, ads in the United Kingdom are 10, 20, or 30 seconds long with respective market shares of 20%, 20%, and 45% (Green 2006). The ad lengths in our data set exhibit a similar range, with 90% of them lasting 10, 15, 20, or 30 seconds. As a consequence, we have extended our analysis to define some properties of the scheduling problem with multiple ad lengths. We use the insights from the optimal schedule for two ad lengths along with the upper bounds derived for the problem with multiple ad lengths to propose heuristics in §4.

Third, we assume that the number of commercial breaks in the broadcast transmission is known. This assumption is motivated by the lack of variation observed in our data, most likely due to the structure of the game, where breaks occur naturally at certain stages (e.g., every time a wicket falls or after an "over"). However, we perform numerical experiments to establish that this assumption is not critical for our results.

Fourth, in our treatment of the penalties for not airing an ad (i.e., penalties for ads that were booked but not aired because of insufficient airtime), we assume that the penalty associated with each ad is known ex ante. It is actually quite difficult to quantify these penalties; however, the broadcaster is well aware of the need to avoid excessive overbooking because it damages the relationship with advertisers. Although such penalties are not always specified in the contract between advertiser and broadcaster, they sometimes materialize in the form of lost business and/or "make-goods": ads that are shown for free during a different (nonpremium) program.

Finally, the analytical results ignore both advertiser and industry diversity constraints. In an extension of the model reported in the electronic companion, the optimal schedule of the problem with advertiser diversity constraint is shown to be a modification of the optimal schedule without such a constraint that aims to preserve diversity in the remaining ad portfolio. A numerical study of the problem with industry diversity constraints—which also imply advertiser diversity constraints—shows that, at the level of diversity observed in the real-life portfolio, the existence of diversity constraints does not strongly affect the performance of our heuristics.

3.2. Optimal Scheduling Policy

The central property of the optimal scheduling policy that we devise is a strong result concerning the value of perfect information. Therefore, we begin by characterizing the optimal solution under PI, when all future break durations are known. Under PI, the scheduler must solve the following integer problem:

$$\max_{\lambda,\,\theta\in\mathbb{N}}\left\{\sum_{i=1}^{\sum_{n=1}^{N}\lambda_n} l_i^o+\sum_{i=1}^{\sum_{n=1}^{N}\theta_n} s_i^o \;\middle|\; 2\lambda_n+\theta_n\le b_n,\ \forall\, n=1,\dots,N\right\}, \quad (1)$$

[4] Given the media industry's nontransparent nature, it is seldom straightforward to assess the premiums (or discounts) associated with longer ad lengths. The premium or discount varies from country to country and from broadcaster to broadcaster. 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 that can be placed in any show at the network's discretion), the 15-second types 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%–80% as much as the cost of a 30-second ad (Newstead and Romaniuk 2010). In the United States, however, an even shorter ad typically sells at a discount (e.g., a 10-second ad is sometimes only 15%–20% the cost of a 30-second ad). In the United Kingdom, where 10-second ads are more popular than 15-second ads (Newstead and Romaniuk 2010), the Channel 5 network reports that a 30-second commercial costs twice as much as a 10-second ad and half as much as a 60-second ad (How to advertise FAQ—How much does it cost to advertise on television?" Accessed September 15, 2015, http://about.channel5.com/faqs/how-to-advertise). In South Asia, discussions between the authors and media industry professionals indicate that the price of an advertisement tends to be a linear function of the ad's length.

where λ_n and θ_n are the respective numbers of long and short ads scheduled during commercial break n. These constraints ensure that the total duration of the ads scheduled in a break does not exceed the duration of that break.

Since all elements of $\mathbf{l}^o$ and $\mathbf{s}^o$ are nonnegative, and since the cardinality of the two sets is infinite, it follows that the scheduler will always (weakly) prefer to air an ad in an empty slot, and thus to earn nonnegative revenue, over the option of allowing the slot to remain empty and thus earning zero revenue. Thus, the N constraints will always be tight; that is,

$$\theta_n = b_n - 2\lambda_n. \tag{2}$$

If we put $\lambda = \sum_{n=1}^{N} \lambda_n$, then (1) can be reduced to an optimization problem over a single decision variable as follows:

$$\max_{\lambda\in\mathbb{N}} \left\{ \sum_{i=1}^{\lambda} l_i^o + \sum_{i=1}^{B-2\lambda} s_i^o \;\middle|\; \lambda \le \sum_{n=1}^{N} \lfloor b_n/2 \rfloor \right\}. \tag{3}$$

Proposition 1. *Let $\lambda^* \le \sum_{n=1}^{N} \lfloor b_n/2 \rfloor$ be the greatest integer such that $l_\lambda^o \ge s_{B-2\lambda+1}^o + s_{B-2\lambda+2}^o$. If no such integer exists, let $\lambda^* = 0$. Then, λ^* solves* (3).

From Proposition 1 it follows that, given N (remaining) breaks of length $b_1, \ldots, b_N$, the optimal set of ads to be scheduled under perfect information consists of the first λ^* long ads and the first $B - 2\lambda^*$ short ads from the respective sets $\mathbf{l}^o$ and $\mathbf{s}^o$. (If $\lambda^* = 0$, then only short ads will be scheduled.)

One way to interpret this PI solution is as follows. The optimal policy will first set aside the $u = B - 2\sum_{n=1}^{N} \lfloor b_n/2 \rfloor$ highest-revenue short ads to schedule in each of the breaks with odd duration, and the remaining airtime will be filled (in accordance with a greedy policy) from the remaining set of ads. First, l_1^o will be compared to ads $s_u^o + s_{u+1}^o$ and then scheduled if it is higher. Subsequently, if ads $l_1^o, \ldots, l_{i-1}^o$ and $s_{u+1}^o, \ldots, s_{u+j-1}^o$ have already been scheduled, then l_i^o will be scheduled provided $l_i^o \ge s_{u+j}^o + s_{u+j+1}^o$ (and so forth).

However, such a schedule requires the broadcaster to know how many breaks are of odd duration, which in turn requires knowledge of all break durations $b_1, \ldots, b_N$. A surprising result is that, even without perfect information about the *duration* of future breaks, one can still construct a schedule that achieves the optimal PI revenue. This schedule is described in the following proposition.

Proposition 2 (Optimal PI Schedule). *The following rule generates an optimal PI schedule. In break n ($n = 1, \ldots, N$), schedule the highest-paying unaired λ_n long ads from set $\mathbf{l}^o$ and the highest-paying unaired $b_n - 2\lambda_n$ short ads from set $\mathbf{s}^o$, where $\lambda_n \le \lfloor b_n/2 \rfloor$ is recursively defined as the largest integer such that $l^o_{\sum_{i=1}^{n}\lambda_i} \ge s^o_{\sum_{i=1}^{n}(b_i-2\lambda_i)+N-n+1} + s^o_{\sum_{i=1}^{n}(b_i-2\lambda_i)+N-n+2}$; if no such integer exists, then $\lambda_n = 0$.*

Note that the schedule for break n does not depend on the duration of breaks $n+1, \ldots, N$. The only information required to construct the schedule for break n is the identity of the ads aired in the previous breaks. Thus, the rule described in Proposition 2 can be directly applied to a *dynamic setting with uncertainty*, in which the duration of future breaks is unknown to the scheduler. Under the optimal policy, then, 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. However, this does not mean that uncertainty about capacity has no effect on the broadcaster's revenue. Although such uncertainty makes no difference at the scheduling stage, it naturally plays a role at the (prior) selling stage, during which the broadcaster decides how many short and long ads to sell to different advertisers.

This result requires our assumption that ads can only be of length S or $2S$. If that assumption is relaxed, then the optimal policy will no longer be independent of the distribution of break durations. Consequently, if there are more than two possible ad lengths (or if there are only two ad lengths but one is not twice as long as the other), then we can expect any scheduling policy to generate *less* revenue than the PI schedule.[5] Another crucial assumption is knowing the number of breaks that will occur in the program.[6] We show in §5 that, for our data set, typically the value of perfect information is extremely low even when these two assumptions are relaxed.

Although the optimal policy is not the greedy policy, it has the structure of a greedy *look-ahead* policy: At stage n, given that $N-n+1$ breaks remain unfilled, for a current break of length $b_n > 1$, the revenue from the most lucrative long ad remaining is compared with the combined revenue from the $N - n + b_n - 1$

[5] Suppose ads can be either 15 or 45 seconds long and that the duration of a commercial break can be 15, 30, or 45 seconds with respective probability α_1, α_2, or $1 - \alpha_1 - \alpha_2$. Under this scenario, if break $N-1$ has a duration of 45 seconds, then, with one break remaining, clearly the optimal policy is to schedule l_1 (i.e., the highest-paying unaired 45-second ad) whenever $(1-\alpha_1)l_1 \ge \alpha_1 s_2 + (\alpha_1 + \alpha_2)(s_3 + s_4) + \alpha_2 s_5$; in all other cases, it is optimal to schedule (s_1, s_2, s_3) (i.e., the three highest-paying unaired 15-second ads). Thus, the scheduling policy now depends on the probability distribution of the remaining capacity.

[6] Suppose ads are either 15 or 30 seconds long and that the last break can have duration of 0, 15, or 30 seconds with respective probability α_1^N, α_2^N, or $1 - \alpha_1^N - \alpha_2^N$. Now if break $N-1$ has a duration of 30 seconds, then, with one break remaining, the optimal policy is to schedule l_1 (i.e., the highest-paying unaired 30-second ad) whenever $(\alpha_1^N + \alpha_2^N)l_1 \ge \alpha_1^N(s_1 + s_2) + \alpha_2^N(s_2 + s_3)$; otherwise, it is optimal to schedule (s_1, s_2) (i.e., the two highest-paying unaired 15-second ads). Once again, the policy is therefore no longer independent of the probability distribution of the remaining capacity.

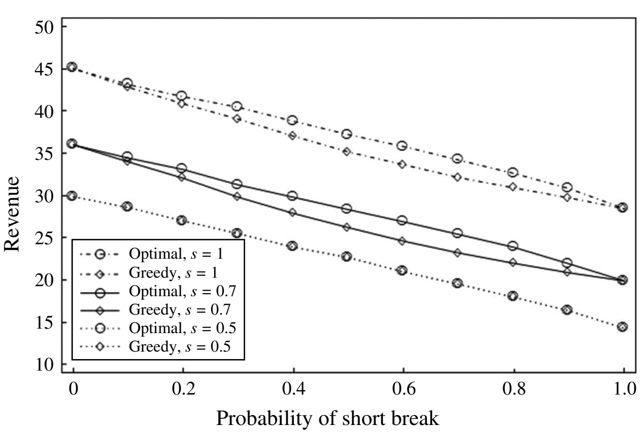
and $N-n+b_n$ most lucrative short ads remaining. If the former is higher, then the long ad is selected. This step is repeated until the break has been entirely filled with a combination of long and short ads. Of course, the broadcaster may air a long ad in the current break—despite it being less profitable than the greedy combination of short ads—to avoid scheduling less profitable short ads in the event that all future breaks have odd durations and thus require a short ad. In other words, the optimal schedule accepts an initial revenue loss at stage n to safeguard revenue at future stages; this feature explains the "look-ahead" aspect (Atkinson 1994).

The optimal policy will coincide with the greedy policy whenever short ads sell at a discount relative to long ads (i.e., if $s_i^o \le l_j^o/2$ for all nonzero $s_i^o \in \mathbf{s}^o$ and $l_j^o \in \mathbf{l}^o$) because the greedy heuristic will always schedule as many long ads as possible in a given break, as is optimal. If short ads sell at a premium ($s_i^o > l_j^o/2$ for all nonzero $s_i^o \in \mathbf{s}^o$ and $l_j^o \in \mathbf{l}^o$), then the greedy heuristic will be suboptimal because it always prefers to schedule short ads and so might run out of short ads before the end of the horizon.

We illustrate with a simple example in which there are only two possible break durations: either S or $2S$ seconds. In Figure 1 we plot the revenue of the greedy and optimal policies, for constant short and long ad revenues and under different ratios of short-to-long ad revenue, as a function of the probability α of a short break occurring. The heuristic's performance is affected by the distribution of break durations. As expected, the greedy and optimal policies coincide at the extremes—that is, when either all breaks are long or all breaks are short—and the performance gap is largest for intermediate probabilities of short breaks.

3.2.1. Penalties for Unaired Ads. It is likely that advertisers incur a disutility whenever their ads are not shown. As a consequence, contracts may specify a penalty that the network must pay to the advertiser when a booked ad is not actually aired. In that case, if ad l_i^o (s_i^o) does not air during any commercial break, then the broadcaster will not earn revenue l_i^o (s_i^o) and also must pay a penalty of β_i^o (γ_i^o) to the client whose ad l_i^o (s_i^o) was not aired. It is easy to see that if we rearrange the sets of long and short ads in decreasing order of the sum of revenue and penalty for each ad, then we can use the optimal scheduling policy described in Proposition 2.

Figure 1 Base Case: Performance of Optimal vs. Greedy Policy



Notes. We assume that $N=30$, break durations are independent and identically distributed, the ad set consists of 20 short ads and 20 long ads, and each long ad generates revenue $l=1$. We perform 1,000 simulation runs.

4. The Problem with Multiple Ad Lengths

Let $(\mathbf{r}^o, \mathbf{x}^o)$ be the initial portfolio of ads with nonzero revenue, where r_i^o is the revenue generated by ad i and x_i^o is the duration (in seconds) of ad i. In this section, we assume that ads can have more than two possible lengths, so that x_i^o is no longer restricted to be in the set $\{S, 2S\}$. As before, we use the notation $(\mathbf{r}, \mathbf{x})$ to denote the portfolio of ads not yet aired when the current break occurs. The value function for the dynamic scheduling problem with multiple ad length can be computed recursively via the Bellman optimality equations; we obtain

$$V_n(\mathbf{r},\mathbf{x}) = \sum_{j=1}^{K} p_{n,j} \max\left\{ \sum_{i=1}^{|\mathbf{r}|} r_i y_i + V_{n+1}\left(\mathbf{r}\backslash\{r_i\}_{I_{y_i=1}}, \mathbf{x}\backslash\{x_i\}_{I_{y_i=1}}\right) \,\middle|\, \sum_{i=1}^{|\mathbf{r}|} x_i y_i \le d_j, y_i \in \{0,1\} \right\}, \quad (4)$$

$$V_{N+1}(\mathbf{r},\mathbf{x}) = 0;$$

here $y_i = 1$ if and only if (iff) ad with revenue r_i is scheduled in break n. An optimal policy for (4)—at stage n when the remaining ad set is $\mathbf{r}$ and a break of duration b_n occurs—is to schedule a subset of ads $\mathbf{r}^* \subset \mathbf{r}$ that fit in the current break and satisfy

$$\sum_{r_i \in \mathbf{r}^*} r_i + V_{n+1}(\mathbf{r}\backslash\mathbf{r}^*, \mathbf{x}\backslash\mathbf{x}^*) \ge \sum_{r_i \in \mathbf{r}'} r_i + V_{n+1}(\mathbf{r}\backslash\mathbf{r}', \mathbf{x}\backslash\mathbf{x}') \quad (5)$$

for all subsets of ads $\mathbf{r}' \subset \mathbf{r}$ such that $\sum_{x_i \in \mathbf{x}'} x_i \le b_n$, where $\mathbf{x}^* = \{x_i \in \mathbf{x} \mid r_i \in \mathbf{r}^*\}$ and $\mathbf{x}' = \{x_i \in \mathbf{x} \mid r_i \in \mathbf{r}'\}$.

Solving problem (5) requires storing all possible values of $V_n(\mathbf{r},\mathbf{x})$ for all $n = 1, \ldots, N$ and all $2^{|\mathbf{r}|}$ subsets of ads in $\mathbf{r}$. Given the large ad set and the large number of breaks, there are significant computational challenges—in terms of both computing times and storage requirements—to solving this dynamic program. Any method with practical appeal must be solvable in real time during the live event. For this reason, we are not concerned with solving the dynamic program and instead study relaxations of the problem that can be used to design practical heuristics.

4.1. Relaxations and Bounds

For any fixed sample path $\mathbf{b}=\{b_1,\dots,b_N\}$, the perfect information solution to (4) is given by the following 0–1 integer program:

$$\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,\mathbf{b}) = \max\left\{\sum_{n=1}^{N}\sum_{i=1}^{|\mathbf{r}^o|} r_i^o y_{i,n} \,\middle|\, \sum_{i=1}^{|\mathbf{r}^o|} x_i^o y_{i,n} \le b_n,\ \sum_{n=1}^{N} y_{i,n} \le 1,\ y_{i,n}\in\{0,1\},\ \forall i=1,\dots,|\mathbf{r}^o|,\ \forall n=1,\dots,N\right\}, \quad (6)$$

where $y_{i,n}=1$ iff ad with revenue r_i^o is scheduled in break n.

Problem (6) is a 0–1 multiple knapsack problem that has been thoroughly studied in the literature (see, e.g., Martello and Toth 1990). A straightforward relaxation of (6) is given by the simple knapsack problem with only one long break of duration $B=\sum_{n=1}^{N} b_n$ (in the literature, this is often referred to as the "surrogate" relaxation of the multiple knapsack problem):

$$\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,B) = \max\left\{\sum_{i=1}^{|\mathbf{r}^o|} r_i^o y_i \,\middle|\, \sum_{i=1}^{|\mathbf{r}^o|} x_i^o y_i \le B,\ y_i\in\{0,1\},\ \forall i=1,\dots,|\mathbf{r}^o|\right\}, \quad (7)$$

where $y_i=1$ iff ad with revenue r_i^o is in the optimal schedule.

There are two main advantages of studying the single-break relaxation. First, the single-break problem is easier to solve. Despite its non-polynomial time complexity, many instances of the 0–1 knapsack can be solved quickly. Second, thinking ahead to this problem's stochastic version, it is easier to obtain a more accurate forecast of the total airtime in a game B than it is to accurately forecast the duration of each break individually.

Let Δ and Ω be the optimal set of ads for problems (6) and (7), respectively. An immediate observation, which will prove useful for designing one of the heuristics described in the next section, is as follows. If there is a feasible schedule for (6) that uses all ads in Ω, then

$$\Delta=\Omega \quad \text{and} \quad \mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,\mathbf{b})=\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,B). \quad (8)$$

This follows immediately because (7) is a relaxation of (6), and so if the solution to (7) is feasible for (6), then it is also optimal. One circumstance in which the two sets differ is when the ads in Ω are not granular enough to utilize the airtime of the N breaks.[7]

Consider now the continuous relaxations of problems (6) and (7) such that the variables $y_{i,n}$ and y_i are not restricted to be integer. Then,

$$c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,\mathbf{b})) = \max\left\{\sum_{n=1}^{N}\sum_{i=1}^{|\mathbf{r}^o|} r_i^o y_{i,n} \,\middle|\, \sum_{i=1}^{|\mathbf{r}^o|} x_i^o y_{i,n} \le b_n,\ \sum_{n=1}^{N} y_{i,n} \le 1,\ 0\le y_{i,n}\le 1,\ \forall i=1,\dots,|\mathbf{r}^o|,\ \forall n=1,\dots,N\right\}; \quad (9)$$

$$c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,B)) = \max\left\{\sum_{i=1}^{|\mathbf{r}^o|} r_i^o y_i \,\middle|\, \sum_{i=1}^{|\mathbf{r}^o|} x_i^o y_i \le B,\ 0\le y_i\le 1,\ \forall i=1,\dots,|\mathbf{r}^o|\right\}. \quad (10)$$

Proposition 3.

$$V_1(\mathbf{r}^o,\mathbf{x}^o) \le c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,E[\mathbf{b}])) = c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,E[B])).$$

It follows from this proposition that $c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,E[B]))$ is an upper bound for (4). As shown in Martello and Toth (1990), other continuous relaxations for the 0–1 multiple knapsack problem yield the same objective value as $c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,E(B)))$.

4.2. Heuristics for the Multiple-Ad-Length Problem

In this section we describe several heuristics that will be tested in §5: (a) a myopic greedy heuristic that maximizes the revenue in the current break while ignoring future breaks; (b) a greedy look-ahead (GLA) heuristic that rations the use of shorter ads, as suggested by Proposition 2; and (c) an upper-bound heuristic with look-ahead (UBLA) that uses the solution of $\mathcal{V}(\mathbf{r},E[B])$ in trying to recreate the PI schedule.

4.2.1. Greedy Heuristic. The greedy heuristic selects the combination of ads (to show in the current break) with revenues from the set $\mathbf{r}$ that maximizes current revenue. Toward this end, the greedy heuristic solves a knapsack problem for each break as soon as that break's duration b_n is revealed:

$$\max\left\{\sum_{i=1}^{|\mathbf{r}|} r_i y_i \,\middle|\, \sum_{i=1}^{|\mathbf{r}|} x_i y_i \le b_n,\ y_i\in\{0,1\}\ \forall i=1,\dots,|\mathbf{r}|\right\},$$

where $y_i=1$ iff ad with revenue r_i, and duration x_i is scheduled in the current break. After each break, the sets $\mathbf{r}$ and $\mathbf{x}$ are updated to reflect the ads that have already been scheduled. Because the size of each individual knapsack problem is small, this heuristic can efficiently solve the scheduling problem in real time.

[7] This does not mean that ads must be greater in length. Suppose, for example, that there are only two breaks left. Suppose also that the set Ω contains three 20-second ads worth \$8 each (for a total revenue of \$24) and that there are two 30-second ads in the portfolio worth \$10 each. If it turns out that the current break and the subsequent break are each of a 30-second duration, then the PI schedule will try to schedule the \$10 30-second ads in each break (for a total revenue of \$20).

4.2.2. Greedy Heuristic with Look-Ahead. The greedy look-ahead heuristic is based on the insight from the optimal scheduling policy in Proposition 2, which offers a rule that favors long ads over short ads by setting a lower threshold (than would a greedy policy) for scheduling long ads. To apply that insight to the case of multiple ads, we split the different ad lengths into two groups—"short" and "long" ads. Let $\mathbf{x}_s$ be the category of short ad durations, and let $\mathbf{x}_l$ be the category of long ad durations. Then, for every pair of such durations of the form $\{x, 2x\}$ with $x \in \mathbf{x}_s$ and $2x \in \mathbf{x}_l$, the GLA heuristic's first step consists of applying the rule of Proposition 2 *as if* only ads of those two durations exist in the portfolio; we use this outcome as an upper bound q_x on the number of ads of duration x that can air in a particular break. The second step consists of choosing a subset of ads from the whole portfolio that maximizes the revenue and also satisfies the upper-bound constraints on the number of short ads. Thus,

$$\max\left\{\sum_{i=1}^{|\mathbf{r}|} r_i y_i \,\middle|\, \sum_{i=1}^{|\mathbf{r}|} x_i y_i \le b_n, \sum_{i=1}^{|r|} I_{x_i = x} y_i \le q_x, \; \forall x \in \mathbf{x}_s, y_i \in \{0,1\}, \; \forall i = 1, \ldots, |\mathbf{r}|\right\},$$

where $y_i = 1$ iff ad with revenue r_i is scheduled in the current break. The sets $\mathbf{r}$ and $\mathbf{x}$ are then updated to account for the ads that have been scheduled.

4.2.3. Upper-Bound Heuristic with Look-Ahead. Upper-bound heuristics have been widely studied in the literature (see, e.g., Amaruchkul et al. 2007, Zhang and Cooper 2005). We propose an upper-bound look-ahead (UBLA) heuristic built on the insight derived from (8): if it is possible to schedule all ads from Ω within the N breaks, then the PI revenue will be automatically achieved. So this heuristic starts by approximating the set Ω via $\hat{\Omega}$, which is obtained by solving the following deterministic single-break problem:

$$\max\left\{\sum_{i=1}^{|\mathbf{r}^o|} r_i^o z_i^o \,\middle|\, \sum_{i=1}^{|\mathbf{r}^o|} x_i^o z_i^o \le E[B], \; z_i^o \in \{0,1\}, \; \forall i = 1, \ldots, |\mathbf{r}^o|\right\}.$$

If $z_i^o = 1$, then $r_i^o \in \hat{\Omega}$. This knapsack problem is only solved once before the start of the live event. As before, let $(\mathbf{r}, \mathbf{x})$ be the portfolio of ads not yet aired at the time break n occurs, and let $\mathbf{z} = \{z_1, z_2, \ldots\}$ be such that $z_i = 1$ iff $r_i \in \hat{\Omega}$. Then, the heuristic will try to schedule the subset of ads that maximizes the weighted sum of (a) the current revenue in a break and (b) a reward for using ads from $\hat{\Omega}$:

$$\max\left\{\sum_{i=1}^{|\mathbf{r}|} r_i y_i + w \sum_{i=1}^{|\mathbf{r}|} v(x_i) z_i y_i \,\middle|\, \sum_{i=1}^{|\mathbf{r}|} x_i y_i \le b_n, y_i \in \{0,1\}, \; \forall i = 1, \ldots, |\mathbf{r}|\right\};$$

here $y_i = 1$ iff ad with revenue r_i is scheduled in the current break. The sets $\mathbf{r}$, $\mathbf{x}$, and $\mathbf{z}$ are then updated to account for the ads that have been scheduled.

The "reward" function $v(x)$, which depends on the ad's length, is designed to skew the objective function in favor of longer ads in $\hat{\Omega}$; the reason is that longer ads have less scheduling flexibility and can therefore fit only in a subset of the breaks.[8] The tuning parameter w should be chosen in such a way that the greedy objective of maximizing the current break's revenue is balanced with the objective of using ads exclusively from the set $\hat{\Omega}$. For extremely small w, this policy reduces to the greedy heuristic. If w is large, however, then the heuristic will seek to schedule the ads from $\hat{\Omega}$ first—starting with the longest ads—and will go outside that set only if there is extra airtime left. This policy is similar in structure to the greedy look-ahead policy first described by Atkinson (1994). In prioritizing the use of ads from $\hat{\Omega}$, the policy is similar also to the heuristic proposed by Fisk and Hung (1979) for the deterministic 0–1 multiple knapsack problem. In our numerical experiments, we choose a value of w high enough to differentiate this heuristic from the greedy heuristic.

Industry Diversity Constraints. Observe that when there is an industry diversity constraint, each of these heuristics will incorporate a condition that prohibits two or more ads from the same industry from being shown in the same break.

5. Numerical Results

In practice, the advertisement scheduling problem may display features that make it analytically intractable, in which case perfect information *would* be of value. We perform an extensive numerical analysis to assess the performance of heuristics commonly used in the literature and of heuristics based on our theoretical results. As a measure of heuristic quality, we use the percentage gap from the revenue under perfect information, which is obtained by solving the optimization problem (6).

We use real (but appropriately disguised) data from the broadcaster, which feature not only multiple break lengths, but also multiple ad lengths and random numbers of breaks. We examine the problem with all of its attributes: diversity constraints, multiple break

[8] This intuition comes from the problem of two ad lengths. There, the main consideration for a scheduler when implementing the PI solution is to ensure that all the long ads from the set $\{l_1^o, \ldots, l_{\lambda^*}^o\}$ not yet aired (during the first n commercial breaks) will fit in the remaining $N - n$ breaks. A straightforward way of satisfying that condition is to schedule, in each break, as many long ads as possible until all λ^* long ads have been scheduled. If a break has a duration that is not a multiple of $2S$ seconds and/or if all λ^* long ads have already been aired, then the remaining airtime is filled with short ads.

Table 1 Percentage Revenue Gap *Without* Diversity as a Function of Overbooking Level, Ad Distribution, and Pricing

Overbooking level (%)	Ad distrib.	Linear pricing: Greedy	Linear pricing: GLA	Linear pricing: UBLA	Concave pricing: Greedy	Concave pricing: GLA	Concave pricing: UBLA
0	3:1	**0.01** ± 0.00	0.20 ± 0.02	0.44 ± 0.03	0.88 ± 0.05	**0.39** ± 0.03	1.08 ± 0.06
	1:1	**0.01** ± 0.00	0.36 ± 0.02	0.27 ± 0.02	1.55 ± 0.08	**0.86** ± 0.05	0.90 ± 0.05
	1:3	**0.05** ± 0.01	0.59 ± 0.04	0.07 ± 0.01	3.09 ± 0.12	1.87 ± 0.09	**1.68** ± 0.09
	Real life	**0.01** ± 0.00	0.39 ± 0.03	0.30 ± 0.00	1.50 ± 0.08	0.87 ± 0.05	**0.64** ± 0.04
20	3:1	**0.00** ± 0.00	0.37 ± 0.01	0.25 ± 0.01	0.64 ± 0.03	**0.23** ± 0.01	1.68 ± 0.03
	1:1	**0.00** ± 0.00	0.65 ± 0.02	0.18 ± 0.01	1.12 ± 0.05	**0.66** ± 0.03	0.97 ± 0.03
	1:3	**0.01** ± 0.00	1.01 ± 0.01	0.06 ± 0.00	2.99 ± 0.03	2.07 ± 0.02	**1.04** ± 0.03
	Real life	**0.00** ± 0.00	0.66 ± 0.02	0.23 ± 0.01	1.17 ± 0.04	0.66 ± 0.03	**0.63** ± 0.03

Note. Entries in bold show the best-performing heuristic for each category.

lengths, multiple ad lengths, and a random number of breaks. The simulation uses data from a recent cricket tournament that included 1,234 commercial breaks in 27 matches. In the simulation, the average number of breaks is set to 50, with N drawn from a discrete uniform distribution defined on the integer set $\{45, 46, \ldots, 55\}$. Breaks were of random length, with a mean of 65 seconds and a standard deviation of 22 seconds. We selected a log-normal distribution with appropriate parameters as the closest fit. The most commonly found ad lengths were 10, 15, 20, and 30 seconds. The revenues for 30-second spots, assigned via a spread-preserving scaling (for confidentiality purposes), follow the continuous uniform distribution $l_i^o \sim \mathcal{U}(7{,}000; 10{,}000)$. As before, we analyze both linear and concave pricing.[9]

The ad portfolio's composition—in terms of airtime sold as short (10 and 15 seconds) versus long (20 and 30 seconds) ads—is also varied and reflects four different settings: (i) thrice as much time sold to short ads (3:1 ratio); (ii) equal time sold to short and long ads (1:1); (iii) thrice as much time sold to long ads (1:3); and (iv) actual ("real-life") sales, which fall in between the second and third scenarios. The number of each ad type is equally split in each of these four settings. We consider two overbooking levels (0% and 20%) consisting of 15 and 18 advertisers, respectively, with each advertiser booking a total of 200 seconds. There are seven industries in our diversity setting.[10] We shall investigate the effectiveness of the heuristics described in §4.2 by comparing their expected revenue with the results obtained under perfect information.

[9] Under *linear* pricing we apply factors $(1/3, 1/2, 2/3, 1)$ to calculate revenues for respective ad lengths of $(10, 15, 20, 30)$ seconds (e.g., if the revenue from a 30-second ad of advertiser 1 is 9,000, then the revenue from a 10-second ad of this advertiser is 3,000, from a 15-second ad is 4,500, and from a 20-second ad is 6,000). Under *concave* pricing we apply the factors $(\sqrt{1/3}, \sqrt{1/2}, \sqrt{2/3}, 1)$.

[10] In the case without overbooking, advertisers are allocated to the seven industries as $\{3, 3, 3, 3, 1, 1, 1\}$; in the case with overbooking, they are allocated as $\{4, 4, 4, 3, 1, 1, 1\}$.

Table 1 reports results for the case with no diversity constraints; Table 2 reports results for the case with diversity constraints. We shall discuss each case separately. As expected from the analytical results, the greedy heuristic performs outstandingly well in the simple case of linear revenues without diversity constraints and obtains (near) optimal revenue. Yet when we introduce concave revenues in the absence of a diversity constraint, we are able to propose simple heuristics that outperform the greedy heuristic. For example, if short ads are plentiful, then the greedy look-ahead heuristic performs better, whereas if short ads are scarce, the upper-bound heuristic with look-ahead dominates. Heuristics with the look-ahead feature avoid overutilizing short ads early on, thus prolonging the portfolio's scheduling flexibility. That being said, the optimality gap in comparison with the greedy heuristic remains small in all scenarios, provided short ads are not extremely scarce (scenario 1:3). This result also indicates that broadcaster revenue is driven more by airtime utilization than by ad selection per se. The greedy heuristic generally performs well in terms of utilizing airtime, unless short ads are priced at a premium (i.e., concave pricing) and relatively few of them were sold, in which case the greedy heuristic quickly runs out of short ads.

Adding diversity constraints is expected to reduce the performance of all heuristics, because scheduling becomes more difficult. Although the greedy heuristic typically remains the best-performing one in the linear pricing case, it no longer strongly dominates: the UBLA heuristic occasionally outperforms the greedy heuristic, and their confidence intervals sometimes overlap. Interestingly, we remark that the diversity constraint improves the greedy heuristic's performance under concave revenues; indeed, the optimality gap shrinks when diversity constraints are added. Diversity constraints have the unintended side effect of reducing the number of short ads shown early on in the match because the greedy heuristic can no longer schedule only short ads in longer breaks, since there are only seven industries from which to

Table 2 Percentage Revenue Gap *With* Diversity as a Function of Overbooking Level, Ad Distribution, and Pricing

Overbooking level (%)	Ad distrib.	Linear pricing			Concave pricing		
		Greedy	GLA	UBLA	Greedy	GLA	UBLA
0	3:1	3.11±0.13	**3.05**±0.12	2.77±0.12	**0.89**±0.05	0.96±0.05	1.64±0.07
	1:1	**1.82**±0.10	2.06±0.09	1.99±0.09	1.17±0.06	1.35±0.06	**1.03**±0.05
	1:3	**1.15**±0.08	1.63±0.08	1.26±0.08	2.71±0.11	2.31±0.09	**1.57**±0.07
	Real life	1.80±0.10	1.97±0.09	**1.70**±0.08	1.16±0.06	1.32±0.06	**0.95**±0.06
20	3:1	**1.72**±0.11	1.90±0.11	2.03±0.11	**1.06**±0.04	1.16±0.05	1.77±0.05
	1:1	**0.67**±0.06	1.21±0.06	0.95±0.07	**0.68**±0.03	0.74±0.02	1.07±0.03
	1:3	**0.40**±0.04	1.34±0.04	0.56±0.05	2.54±0.06	2.26±0.04	**1.12**±0.05
	Real life	**0.51**±0.05	1.08±0.04	0.75±0.05	0.82±0.03	**0.78**±0.03	0.91±0.03

Note. Entries in bold show the best-performing heuristic for each category.

choose. The other heuristics do not benefit from this boost because they were already preferring longer to shorter ads in the match's early breaks; thus, they only suffer the drawback of an additional scheduling constraint. If we take the case featuring overbooking, concave revenues, and diversity, the greedy is the top-performing heuristic or sees its confidence interval overlap with the best-performing heuristic in three out of the four ad distribution scenarios considered, and it only fails when short ads are in extremely short supply. This confirms that even in complex settings—with multiple ad lengths, diversity constraints, and concave pricing—the greedy heuristic performs reasonably well. Hence, there is no need to design elaborate heuristics that would be difficult for the broadcaster to understand and implement.

6. Portfolio Composition

The random capacity characteristic of live television events, when combined with the high yield of advertising during such events, induces the broadcaster to sell in excess of expected airtime. Doing so lowers the service level (i.e., many of the booked ads will not be aired), which can lead to advertiser dissatisfaction (i.e., goodwill loss) and also to penalties in the case of contractual guarantees. If there are contractual penalties for unaired ads, then the broadcaster must choose a level of overbooking that balances the trade-off between capacity utilization and penalty payments.

At the same time, the broadcaster must balance the trade-offs involved when splitting the booking levels between short and long ads. Short ads give the broadcaster more scheduling flexibility because they can air during any type of break. So for a given overbooking level, increasing the number of short ads sold raises average capacity utilization. Yet if long ads sell at a *premium* (i.e., if the revenue per second generated by an ad is increasing in its length), then short ads become less profitable. The premium or discount as a function of ad length varies among countries and also among broadcasters. It is seldom straightforward to assess the premium (or discount) associated with longer ads because of the media industry's nontransparent nature. Even so, 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.

Taking ad prices as an exogenous input to the model, we look for the ideal *mix* and *quantity* of short and long ads to sell conditional on (a) their respective revenues and penalties for nonairing and (b) implementing the optimal policy at the scheduling stage.

Much as before, we assume that there is an exogenous set of ads of short (S) and long ($2S$) durations, and that from these the broadcaster must choose a subset to accept for possible airing during the live broadcast. (Recall that booking does not guarantee transmission.) Let δ be the overbooking level, defined as the percentage of airtime sold in excess of expected capacity. We assume there is a penalty—proportional to the revenue that airing the ad would generate—that the broadcaster must pay to the advertiser if an ad is booked but not actually aired; let $\beta \in [0,1]$ be this penalty. To simplify matters, we first abstract from variability in ad prices and also from the dynamic aspect of demand realization during the booking period. We shall later discuss the effect of variability on the optimal policy and perform a numerical sensitivity analysis. In addition, we will present heuristics for solving the more general problem with multiple ad lengths.

For now, let us assume that all long (short) ads generate the same revenue and that the demand for each ad category is large enough. This is a reasonable assumption for live events with tight capacity and high demand. Without loss of generality, we standardize the revenue and duration of short ads to 1. Then a long ad will generate revenue equal to $2(1+\varepsilon)$. We say there is a *premium* for long ads if $\varepsilon > 0$ and that there is a *discount* for long ads if $\varepsilon \leq 0$.

The discrete nature of our problem precludes the use of differential techniques to study the properties of the optimal booking policy, so instead we use lattice-theoretic methods. Denote by $\Pi(n_L, n_S)$: $\mathbb{N}^2 \to \mathbb{R}$ the total expected revenue of a broadcaster given a booking portfolio consisting of n_L long ads and n_S short ads. That same revenue can be expressed as a function of the number of long ads and the overbooking level: $\breve{\Pi}(n_L, \delta)$: $\mathbb{L} \to \mathbb{R}$, where $\delta = (n_S + 2n_L)/E[B] - 1$ denotes the expected overbooking level, $E[B]$ denotes the total expected capacity, and $\mathbb{L} = \{(n_L, \delta) \in \mathbb{N} \times \mathbb{R} \mid \exists n_S \in \mathbb{N} \text{ s.t. } \delta = (n_S + 2n_L)/E[B] - 1\}$. There is a one-to-one correspondence between these two representations of revenue, and it is easy to see that $\mathbb{L}$ is a lattice.[11]

6.1. Premium for Long Ads

If there is a premium for long ads, then by Proposition 2 we know that the optimal scheduling policy is the greedy policy: 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. For simplicity of exposition and w.l.o.g, we focus on the two break durations S and $2S$.

Let k_L denote the random number of long breaks, and let $p_i := \Pr(k_L = i)$ be the probability that the total number of long breaks in a game is i for $i = 1, \ldots, N$. In this case, the expected revenue of a broadcaster is given by

$$\begin{aligned}\Pi(n_L, n_S) = \sum_{i=0}^{n_L} \big[&2i(1+\varepsilon) + \min\{n_S, N-i\} \\ &- 2(n_L - i)\beta(1+\varepsilon) \\ &- \max\{0, n_S - N + i\}\beta\big] p_i \\ + \sum_{i=n_L+1}^{N} \big[&2n_L(1+\varepsilon) + \min\{n_S, N+i-2n_L\} \\ &- \max\{0, n_S - N - i + 2n_L\}\beta\big] p_i.\end{aligned}$$

Similarly, the expected revenue can be expressed as a function of the number of long ads and the overbooking level:

$$\begin{aligned}\breve{\Pi}(n_L, \delta) = \sum_{i=0}^{n_L} \big[&2i(1+\varepsilon) + \min\big\{(\delta+1)E[B] - 2n_L, N-i\big\} \\ &- 2(n_L - i)\beta(1+\varepsilon) \\ &- \max\big\{0, (\delta+1)E[B] - 2n_L - N + i\big\}\beta\big] p_i \\ + \sum_{i=n_L+1}^{N} \big[&2n_L(1+\varepsilon) \\ &+ \min\big\{(\delta+1)E[B] - 2n_L, N+i-2n_L\big\} \\ &- \max\{0, (\delta+1)E[B] - N - i\}\beta\big] p_i.\end{aligned}$$

[11] The set $X \subset \mathbb{R}^n$ is a *lattice* if, for every $x, y \in X$, we have $x \vee y \in X$ and $x \wedge y \in X$; here $x \vee y$ is the componentwise maximum of x and y, and $x \wedge y$ is their componentwise minimum. Let (n_L^i, δ^i), $(n_L^j, \delta^j) \in \mathbb{L}$; then $(n_L^i, \delta^i) \vee (n_L^j, \delta^j) \in \mathbb{L}$ and $(n_L^i, \delta^i) \wedge (n_L^j, \delta^j) \in \mathbb{L}$. To see the former consequence, assume that $n_L^i \ge n_L^j$ and $\delta^i \le \delta^j$ (all other cases are either similar to this one or trivial). Let $n_S^i = (\delta^i + 1)C - 2n_L^i$, and let $n_S^j = (\delta^j + 1)C - 2n_L^j$. Then there exists an $n_S^{ij} \in \mathbb{N}$ such that $\delta^j = (n_S^{ij} + 2n_L^i)/C - 1$. Now simply take $n_S^{ij} = n_S^j + 2(n_L^j - n_L^i)$ and note that, since $\delta^j \ge \delta^i$, we have $n_S^{ij} \ge n_S^i$. One can similarly show that $(n_L^i, \delta^i) \wedge (n_L^j, \delta^j) \in \mathbb{L}$ by constructing $n_S^{ji} = n_S^i + 2(n_L^i - n_L^j)$.

In the following proposition we characterize some properties of the revenue function.

Proposition 4 (Revenue Function Properties).
1. The function $\breve{\Pi}(n_L, \delta)$: $\mathbb{L} \to \mathbb{R}$ is supermodular.
2. *Let* $\Delta\breve{\Pi}(n_L, \delta) = \breve{\Pi}(n_L, \delta) - \breve{\Pi}(n_L - 1, \delta)$. *Then* $\Delta\breve{\Pi}(n_L, \delta)$ *satisfies the following properties*:
(a) $\Delta\breve{\Pi}(n_L, \delta^2) \ge \Delta\breve{\Pi}(n_L, \delta^1)$ *for all* $(n_L, \delta^1) \in \mathbb{L}$ *and* $\delta^2 \ge \delta^1$;
(b) $\Delta\breve{\Pi}(n_L + 1, \delta) \le \Delta\breve{\Pi}(n_L, \delta)$ *for all* $(n_L + 1, \delta) \in \mathbb{L}$.

Here supermodularity is equivalent to increasing differences: the higher the level of overbooking, the greater the effect (on the broadcaster's expected revenue) of adding one more long ad. Therefore, the optimal number of long ads ($n_L^*(\delta)$) is weakly increasing in the level of overbooking. Finding the optimal (n_L^*, δ^*) amounts to maximizing a supermodular function over a lattice (see, e.g., Topkis 1998). Note that the supermodularity of the expected revenue function implies complementarity between its arguments. One can easily show that $\Pi(n_L, n_S)$ is submodular, which implies that its arguments are substitutes.

From part 2 of Proposition 4 we see that, for a given number of long ads, the marginal expected revenue is increasing in the level of overbooking; moreover, for a given overbooking level, the marginal expected revenue is decreasing in the number of long ads. The first property follows directly from the expected revenue function's supermodularity. The second property is intuitive and simplifies the control policy when the level of overbooking is given. For instance, if the overbooking level is fixed at δ, then the optimal number of long ads, n_L^*, is the smallest integer such that $\Delta\breve{\Pi}(n_L^*, \delta) \le 0$.

6.2. Discount for Long Ads

If there is a discount for long ads then, by Proposition 2, short ads will be scheduled in a long break unless not enough short ads are available. As before, let k_L be the number of long breaks. If $k_L \le (n_S - N)^+$, then no long ads will be scheduled (i.e., realized capacity is $B = 2k_L + (N - k_L) = k_L + N$), given that $1 \le 2(1+\varepsilon) < 2$; however, if $k_L \ge (n_S - N)^+ + 1$, then $u(k_L, n_L, n_S) = \min\{n_L, k_L - \lfloor (n_S - N + k_L)^+/2 \rfloor\}$ long ads will be scheduled, and $v(k_L, n_S) = \min\{N - k_L, n_S\} + 2\lfloor (n_S - N + k_L)^+/2 \rfloor$ short ads will be scheduled. The expected revenue function will now be

$$\begin{aligned}&\Pi(n_L, n_S) \\ &= \sum_{i=0}^{(n_S-N)^+} \big[\min\{n_S, N+i\} - (n_S - i - N)^+\beta - 2n_L\beta(1+\varepsilon)\big] p_i \\ &\quad + \sum_{i=(n_S-N)^+ + 1}^{N} \big[2u(i, n_L, n_S)(1+\varepsilon) + v(i, n_S) \\ &\qquad - 2(n_L - u(i, n_L, n_S))\beta(1+\varepsilon) - (n_S - v(i, n_S))\beta\big] p_i.\end{aligned}$$

Proposition 5. *Under the optimal booking policy, $n_L^*=0$, and n_S^* is the smallest integer such that*

$$\Pr(k_L \le n_S^* - N) \ge \frac{1}{1+\beta}.$$

This result leads us to make two observations. First, if there is a discount for long ads (i.e., if $\varepsilon \le 0$), then the optimal number of long ads is zero. This finding 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. Second, choosing the right number of short ads to book can be reduced to a simple newsvendor inventory problem in which the overage cost is β and the underage cost is the revenue of a short ad $s=1$.

6.3. A Heuristic for the Multiple-Ad-Length Portfolio Problem

The foregoing discussion clearly shows the similarity between this problem and the multiclass airline booking problem. The main difference here is that the capacity is not homogenous as in the case of airlines, in which a high-fare customer will consume the same unit of capacity as a low-fare customer. Here, two 15-second ads can displace a 30-second ad, but the reverse need not hold (unless the 15-second ads occupy consecutive time slots). Thus, the well-known heuristics used by airlines (such as the Expected Marginal Seat Revenue (EMSR)-a,b algorithms; see Belobaba 1987) cannot be directly applied to this problem. Nevertheless, they can form a starting point from which customized heuristics can be devised for this setting. The purpose, of course, is not to solve the problem with only two ad lengths, but rather to derive algorithms that can be easily generalized to multiple ad durations.

Similarly to the EMSR algorithm, the heuristic described here aims to determine optimal booking levels in a sequential fashion, from the longest to shortest ads, by approximating the multiple ad length problem via a series of simpler newsvendor inventory problems.

Let $\{x_1, x_2, \dots, x_m\}$ be the set of ad lengths ordered from longest to shortest, and let ε_j be the premium/discount associated with ad length x_j. As before, w.l.o.g. we normalize both the revenue and the duration of the shortest ads to 1; hence, the revenue resulting from an ad of length x_j is $x_j(1+\varepsilon_j)$.

First, for each possible break duration d_i ($i=1, \dots, K$), we find the revenue-maximizing combination of ads to schedule in such a break while assuming unlimited availability of ads of all lengths:

$$\max_{n_{i,1},\dots,n_{i,m}\in\mathbb{N}} \left\{ \sum_{j=1}^{m} n_{i,j} x_j (1+\varepsilon_j) \,\middle|\, \sum_{j=1}^{m} n_{i,j} x_j \le d_i \right\}, \tag{11}$$

where $n_{i,j}$ is the number of ads of length x_j to be scheduled in a break of duration d_i.

Second, for each possible ad length x_j ($j=1,\dots,m$) we want to estimate the underage cost—in other words, the opportunity cost of not having enough ads of this length in the portfolio. We estimate this underage cost by finding the best combination of ads to replace an ad of length x_j in a break:

$$\max_{q_{j+1,j},\dots,q_{m,j}\in\mathbb{N}} \left\{ \sum_{i=j+1}^{m} q_{i,j} x_i (1+\varepsilon_i) \,\middle|\, \sum_{i=j+1}^{m} q_{i,j} x_i \le x_j \right\}, \tag{12}$$

where $(q_{i,j})$ is the number of ads of length $x_i < x_j$ ($i=j+1,\dots,m$) in the combination. The underage cost is therefore given by

$$x_j(1+\varepsilon_j) - \sum_{i=j+1}^{m} q_{i,j} x_i (1+\varepsilon_i). \tag{13}$$

Third, we compute the booking level $\tilde{n}_j$ for each ad length x_j ($j=1,\dots,m$), starting with $\tilde{n}_1$, in two stages. In the first stage, we find F_j, the distribution function of X_j, which is the number of ads of length x_j needed in all of the N breaks of the broadcast, provided that $\tilde{n}_i$ ads of length x_i have been already booked, where $i=1,\dots,j-1$. Thus, we obtain

$$X_j = \sum_{n=1}^{N} X_j^n + \sum_{i=1}^{j-1} q_{i,j} \left(\sum_{n=1}^{N} X_i^n - \tilde{n}_i \right)^+,$$

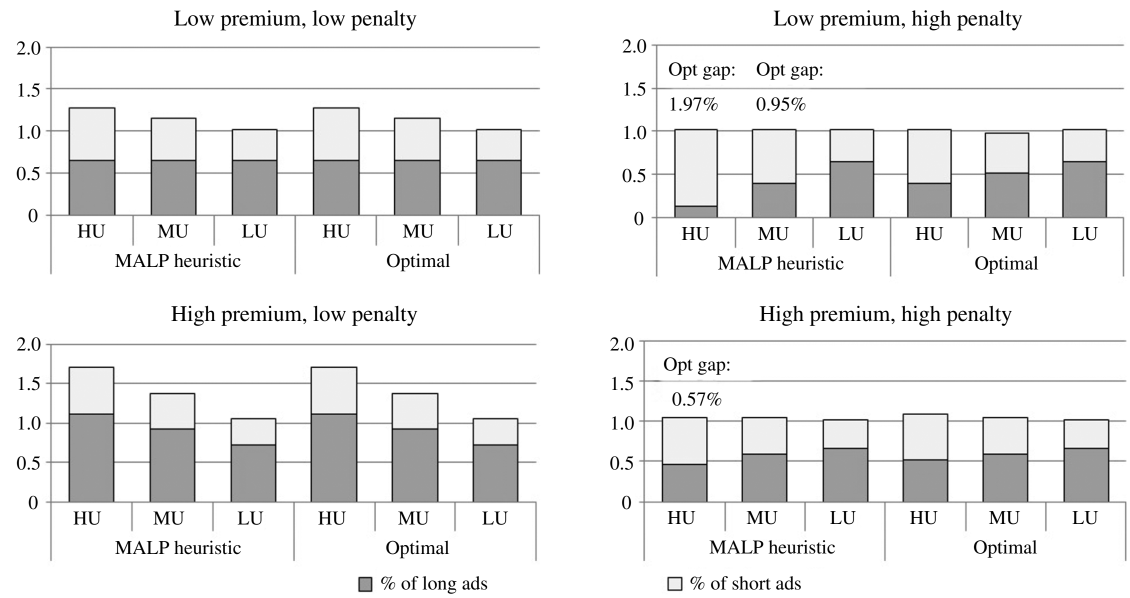
where X_j^n is the random number of ads of duration x_j needed in break n if there is unlimited availability of ads of other lengths (i.e., $X_j^n = n_{i,j}$ if $b_n = d_i$).[12] In the second stage, we solve a newsvendor inventory problem while assuming that (a) the "demand" distribution is F_j, (b) the underage cost is given by (13), and (c) the overage cost is equal to $\beta x_j(1+\varepsilon_j)$. We repeat this third step for all m ad lengths.

Example 1. We illustrate the steps of this heuristic for the two-ad-length problem with a premium on long ads. Assume that the duration and revenue of a short ad are both standardized to 1 (i.e., $S=1$ and $s=1$). Assume further that the duration of a long ad is $L=2$ and that its revenue is $l=2(1+\varepsilon)$, where $\varepsilon > 0$.

- *Step* 1. The optimal combination of ads to schedule in a break of duration d_i is $n_{i,L} = \lfloor d_i/2 \rfloor$ long ads and $n_{i,S} = d_i - 2\lfloor d_i/2 \rfloor$ short ads ($i=1,\dots,K$).
- *Step* 2. The opportunity cost for long ads is $l - 2s = 2\varepsilon$; for short ads, this opportunity cost is 1.

[12] Recall that $p_{n,i}$ is the probability that break n has duration d_i ($n=1,\dots,N$, $i=1,\dots,K$). Then $\Pr(X_j^n \le x) = \sum_{i=1}^{K} I_{n_{i,j}\le x}\, p_{n,i}$, where $I_{n_{i,j}\le x}$ is the indicator function that takes the value 1 only if $n_{i,j} \le x$.

Figure 2 **Percentages of Long and Short Ads from Total Expected Airtime Capacity Under the Optimal Policy and the MALP Heuristic**



Notes. We use HU, MU, and LU to denote high, medium, and low levels of uncertainty in capacity, respectively. We assume $N=20$ and the following distributions for the total number of long breaks: distribution HU, $\Pr(k_L=i)=1/20$, $k_L\in\{1,\ldots,20\}$; distribution MU, $\Pr(k_L=i)=1/10$, $k_L\in\{6,\ldots,15\}$; distribution LU, $\Pr(k_L=i)=1/2$, $k_L\in\{10,11\}$. We use the following values for premiums and penalties: low premium and low penalty, $\varepsilon=0.1$ and $\beta=0.1$; high premium and high penalty, $\varepsilon=1$ and $\beta=1$.

- *Step* 3. Compute the distribution function of the number of long ads needed ($X_L=\sum_{n=1}^N X_L^n$, where $X_L^n=\lfloor b_n/2\rfloor$). Then, solve a simple newsvendor inventory problem while assuming that (a) the demand distribution is the distribution function of X_L, (b) the unit opportunity cost of underbooking is 2ε, and (c) the unit cost of overbooking is $2\beta(1+\varepsilon)$. Let $\tilde{n}_L$ be this booking level.

Now repeat Step 3 to determine $\tilde{n}_S$, the number of short ads to book. First, compute the distribution function of the number of short ads needed, given that $\tilde{n}_L$ long ads have already been booked ($X_S=\sum_{n=1}^N X_S^n+2(X_L-\tilde{n}_L)^+$, where $X_S^n=b_n-2\lfloor b_n/2\rfloor$). Then, solve a newsvendor inventory problem while assuming that (a) the demand distribution is the distribution function of X_S, (b) the unit opportunity cost of underbooking is 1, and (c) the unit cost of overbooking is β. Using an argument similar to the one in Proposition 5, it is easy to show that $\tilde{n}_S$ is the smallest integer such that

$$\Pr(N-\tilde{n}_S\le k\le \tilde{n}_S+2\tilde{n}_L-N)\ge\frac{1}{1+\beta}.$$

If long ads sell at a discount (i.e., if $\varepsilon<0$), then the multiple-ad-length problem (MALP) heuristic will coincide with the optimal policy for the two-ad-length problem.

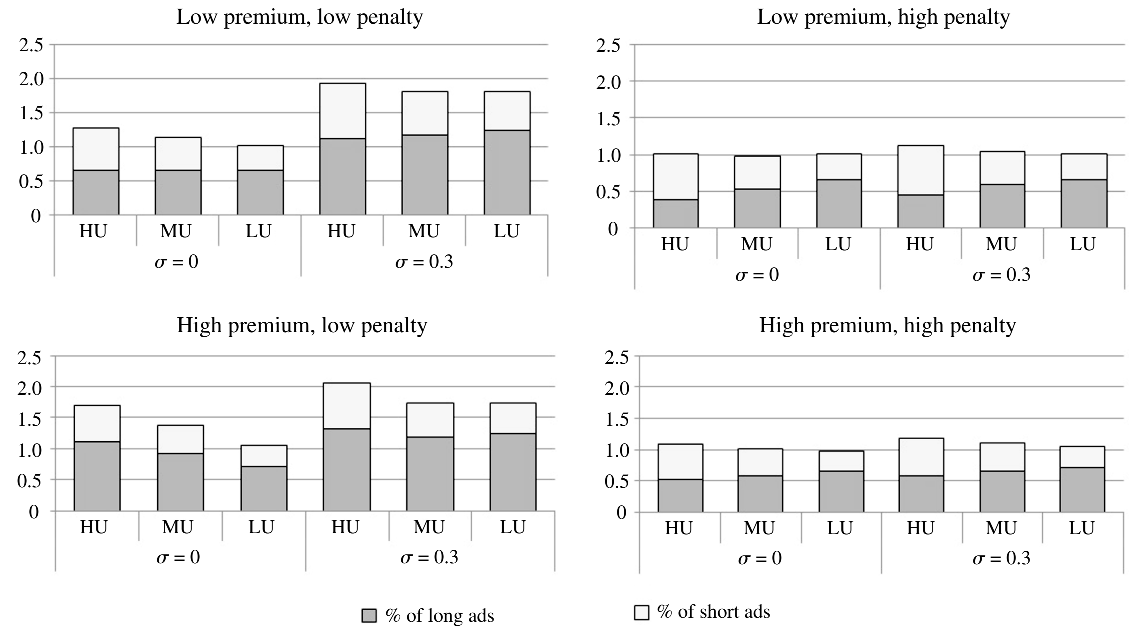
In Figure 2 we compare the booking levels derived under the MALP heuristic with the optimal booking levels for different penalty and premium levels. In all but three scenarios the booking levels are the same. The only time the MALP heuristic gives significantly different recommendations is when the premium is low and the penalty is high. In that case, the MALP heuristic tends to underbook long ads and overbook short ads. For both the optimal policy and the MALP heuristic, we observe that the higher the uncertainty in break duration, the higher the percentage of short ads relative to long ads—even if premiums for long ads are high. So when there is high uncertainty in the break duration, the flexibility afforded by shorter ads is more valuable than the premium generated by longer ads.

6.4. Variable Ad Yields

So far we have assumed that ad revenues are constant for ads of the same length; in reality, however, ads have variable yields. In this case, the previously derived properties of the expected revenue function might not always hold.[13] For the general case with

[13] For example, consider the following sets of revenues for 30-second and 15-second ads: $\mathbf{l}=\{7,7,7,7\}$ and $\mathbf{s}=\{5,1,1,1,1,1,1,1\}$. Note that the average price of a 30-second spot is higher than twice the average price of a 15-second spot. Assume that there is a penalty $\beta=10\%$ and that there are two possible levels of overbooking: $P\delta_1=0$ (i.e., 0%) and $\delta_2=1$ (100%). For simplicity, we assume that the capacity is deterministic and consists of two 30-second breaks. Then it is easy to check that none of the properties derived in §6.1 hold. We have $\breve{\Pi}(4,1)=12.6$, $\breve{\Pi}(3,1)=12.7$, $\breve{\Pi}(2,1)=13.2$, $\breve{\Pi}(1,1)=12.6$, and $\breve{\Pi}(0,1)=7.6$; also, $\breve{\Pi}(2,0)=14$, $\breve{\Pi}(1,0)=13$, and $\breve{\Pi}(0,0)=8$. Observe that neither is the expected revenue function supermodular in this case, since $\Delta\breve{\Pi}(2,\delta^2)=1.2>\Delta\breve{\Pi}(2,\delta^1)=1$. Moreover, for a given level of overbooking, the marginal return need not be decreasing in the number of long ads: $\Delta\breve{\Pi}(4,1)=-0.1>\Delta\breve{\Pi}(3,1)=-0.5$.

Figure 3 **Percentages of Long and Short Ads from Total Expected Airtime Capacity Under the Optimal Policy When Ad Revenues Are Constant ($\sigma = 0$) or Variable ($\sigma = 0.3$) (Simulation Results)**



Notes. HU, MU, and LU denote high, medium, and low levels of uncertainty in capacity, respectively. The values for premiums and penalties are the same as for Figure 2.

variable ad yields, it is no longer possible to have a closed-form expression for the expected revenue function, since that would require an expression for the optimal schedule on every sample path given any sets of ads $\mathbf{l}^o$ and $\mathbf{s}^o$. Nonetheless, our assumption of constant ad yields employed so far can serve as the basis for a simple heuristic that the broadcaster can use to determine what proportion of short and long ads to book. Relative to the optimal policy, such a heuristic, which uses expected yield values for ads of the same length, will underbook, as shown in the following proposition.

Proposition 6. *Let* (n_L^*, n_S^*) *be the optimal booking policy when all long (resp., short) ads have constant yields* l *(resp.,* s*), and let* $\Pi(n_L^*, n_S^*)$ *be the corresponding expected revenue of such a portfolio. Consider now the sets of long ads* $\{l_1, \ldots, l_{n_L^*}\}$ *and short ads* $\{s_1, \ldots, s_{n_S^*}\}$ *with variable yields and such that* (i) $l_1 \geq \cdots \geq l_{n_L^*}$; (ii) $s_1 \geq \cdots \geq s_{n_S^*}$; (iii) $\sum_{i=1}^{n_L^*} l_i / n_L^* = l$; *and* (iv) $\sum_{i=1}^{n_S^*} s_i / n_S^* = s$. *Let* $\Pi(\{l_1, \ldots, l_{n_L^*}\}, n_S^*)$ *be the expected revenue from a portfolio consisting of* $\{l_1, \ldots, l_{n_L^*}\}$ *long ads and* n_S^* *short ads with constant yield equal to* s*. Similarly, let* $\Pi(n_L^*, \{s_1, \ldots, s_{n_S^*}\})$ *be the expected revenue from a portfolio consisting of* n_L^* *long ads with constant yield equal to* l *and* $\{s_1, \ldots, s_{n_S^*}\}$ *short ads. Then the following inequalities hold*:

1. $\Pi(n_L^*+1, n_S^*) - \Pi(n_L^*, n_S^*) \leq \Pi(\{l_1, \ldots, l_{n_L^*}\} \cup \{l\}, n_S^*) - \Pi(\{l_1, \ldots, l_{n_L^*}\}, n_S^*)$;
2. $\Pi(n_L^*, n_S^*+1) - \Pi(n_L^*, n_S^*) \leq \Pi(n_L^*, \{s_1, \ldots, s_{n_S^*}\} \cup \{s\}) - \Pi(n_L^*, \{s_1, \ldots, s_{n_S^*}\})$.

We know that if (n_L^*, n_S^*) is the optimal booking policy when ads have constant yields, then the marginal revenue from one additional long (resp., short) ad must be negative. Proposition 6 shows that this is not necessarily true for the case of variable yields, where the marginal revenue from an additional long (resp., short) ad is always higher than the corresponding marginal revenue when yields are constant. The intuition behind this result is as follows. The scheduling policy always airs the ads with higher revenue first; therefore, penalties will be paid only for the cheaper ads that are not aired, whereas revenue is gained from the more expensive ads that are aired.

Figure 3 summarizes the simulated optimal proportions of short and long ads (expressed as a percentages of total expected airtime capacity) and overbooking levels under both constant and variable ad yields.[14] We notice that variability in ad yields results in a higher overbooking level. The overbooking level is significantly higher when penalties are low, irrespective of the uncertainty in break duration. However, there seem to be no significant changes in the proportion of short and long ads. Table 3 shows the revenue gap obtained from applying a heuristic that

[14] Because we cannot derive the optimal booking policy for variable ad yields, this policy was obtained via a large-scale simulation. For this we used the booking policy (n_L^*, n_S^*) that yielded the maximum average revenue. The ad revenues were generated from the uniform distributions $s \sim \mathcal{U}[1 \pm \sigma]$ and $l \sim \mathcal{U}[2(1+\varepsilon) \pm 2\sigma]$, where $\sigma \in \{0, 0.3\}$. The optimal booking policy for the case of constant ad yields was determined using the results from the previous section and while assuming that each short ad generated revenue $s = 1$ and each long ad generated revenue $l = 2(1+\varepsilon)$.

Table 3 **Percentage Gap in Average Revenue Obtained Using Optimal Policy Under Constant Yields (Π_{OPCY}) Relative to Average Revenue Obtained Using Optimal (Simulation-Wise) Policy Under Variable Yields (Π_{OPVY})**

	Degree of uncertainty in break duration		
Penalty, premium	High uncertainty	Medium uncertainty	Low uncertainty
(β, ε)	$\frac{\Pi_{OPVY}-\Pi_{OPCY}}{\Pi_{OPVY}} \times 100$	$\frac{\Pi_{OPVY}-\Pi_{OPCY}}{\Pi_{OPVY}} \times 100$	$\frac{\Pi_{OPVY}-\Pi_{OPCY}}{\Pi_{OPVY}} \times 100$
(0.1, 0)	**0.49** ± 0.25	**1.78** ± 0.25	**4.08** ± 0.25
(1, 0)	**0.43** ± 0.28	**0.29** ± 0.28	**0.06** ± 0.28
(0.1, 0.1)	**2.56** ± 0.26	**3.96** ± 0.29	**5.53** ± 0.30
(1, 0.1)	**0.30** ± 0.31	**0.47** ± 0.33	**0.77** ± 0.37
(0.1, 1)	**0.88** ± 0.28	**1.69** ± 0.24	**4.69** ± 0.27
(1, 1)	**0.23** ± 0.33	**0.68** ± 0.28	**0.29** ± 0.30

Note. Entries in bold show the best-performing heuristic for each category.

assumes constant yields in a situation where yields are actually variable. The percentage gap is highest when the penalties for overbooking are low, and the lower the uncertainty in break duration, the higher the gap.

Our results suggest the following procedure for crafting a good ad portfolio for the general problem with multiple ad lengths and variable ad yields. First, use the MALP heuristic that assumes constant ad yields to determine what proportion of ads of different lengths should be in the portfolio; second, increase the overbooking level while maintaining the same proportion of ads of different lengths to account for the variability in ad yields. The lower the penalties, the higher this adjustment in the overbooking level should be.

7. Conclusion

The problem of how best to sell and 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, maximizing the use of commercial airtime during live events can greatly enhance a broadcaster's revenue.

In the absence of diversity constraints, we can determine the optimal scheduling rule when ads are of the two commonly found lengths (15 and 30 seconds): a greedy look-ahead policy that is easy to implement and achieves the perfect-information profit. The composition of the broadcaster's ad portfolio can be manipulated to increase expected profit. The stochastic nature of total capacity naturally leads to overselling, which must be balanced against possible no-show penalties. The balance between short and long ads in the portfolio is also important. Short ads are valued for the flexibility they lend to scheduling; in fact, if there were no premium on long ads, then the optimal portfolio would consist exclusively of short ads.

For more complex settings with industry diversity constraints within breaks and multiple ad lengths, we perform a large-scale numerical analysis to compare several scheduling heuristics. We find that the greedy heuristic performs well under many scenarios, in line with our theoretical results. The greedy heuristic is impaired in the presence of concave revenues, and we propose more elaborate heuristics (based on our analytical insights) to reduce the optimality gap further in that case.

At the ad-selling stage, we show how to build the optimal portfolio of short and long ads for constant revenues and two ad lengths. We build a heuristic to extend our portfolio recommendations for the case of multiple ad lengths. We find that the heuristic performs close to optimal in most scenarios of ad pricing, no-show penalties, and break duration uncertainty. When ad revenue is variable, we show that a higher level of overbooking—relative to the constant ad revenue scenario—is optimal.

Although our research yields preliminary results on the question how short and long ads should be priced to maximize the broadcaster's expected revenue, much remains to be done in this area. Future work could profitably look into the following issues: (i) manipulating the relative (sales) prices of short and long ads to shape an optimal ad portfolio under break uncertainty and diversity constraints and (ii) designing a range of price levels (and concomitant service guarantees) that would maximize the revenue extracted from advertisers.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2015.2185.

Acknowledgments

The authors gratefully acknowledge INSEAD and the Centre for Decision Making and Risk Analysis, which provided financial support. They are also indebted to their research assistants Ajay Aravamudhan and Zhi Chen, and are thankful for the comments of colleagues and seminar participants and, in particular, Ioana Popescu and Sridhar Seshadri. Finally, the authors thank Yossi Aviv and the review team for their valuable feedback on this paper.

Appendix

Proof of Proposition 1. Let $\pi(\lambda)=\sum_{i=1}^{\lambda} l_i^o+\sum_{i=1}^{B-2\lambda} s_i^o$. Then, for $\lambda<\lambda^*$, we have $\pi(\lambda^*)-\pi(\lambda)=l_{\lambda+1}^o+\cdots+l_{\lambda^*}^o-s_{B-2\lambda^*+1}^o-\cdots-s_{B-2\lambda}^o\geq 0$ because

$$l_{\lambda^*}^o \geq s_{B-2\lambda^*+1}^o+s_{B-2\lambda^*+2}^o \quad \text{and}$$

$$l_{\lambda+1}^o\geq\cdots\geq l_{\lambda^*-1}^o\geq l_{\lambda^*}^o\geq s_{B-2\lambda^*+1}^o+s_{B-2\lambda^*+2}^o\geq\cdots\geq s_{B-2\lambda-1}^o+s_{B-2\lambda}^o.$$

The first inequality follows from our choice of λ^*, whereas the second line of inequalities follows because the sets $\mathbf{s}^o$ and $\mathbf{l}^o$ are ordered and $\lambda+1\le\lambda^*$.

Similarly, for $\lambda^*<\lambda\le\sum_{n=1}^N\lfloor b_n/2\rfloor$, we have $\pi(\lambda^*)-\pi(\lambda)=s^o_{B-2\lambda+1}-\cdots-s^o_{B-2\lambda^*}-l^o_{\lambda^*+1}+\cdots+l^o_\lambda\ge 0$ because

$$s^o_{B-2\lambda^*-1}+s^o_{B-2\lambda^*}\ge l^o_{\lambda^*+1}\quad\text{and}$$
$$s^o_{B-2\lambda+1}+s^o_{B-2\lambda+2}\ge\cdots\ge s^o_{B-2\lambda^*-1}+s^o_{B-2\lambda^*}\ge l^o_{\lambda^*+1}\ge\cdots\ge l^o_\lambda.$$

The first inequality follows from our choice of λ^* (i.e., the greatest integer less than $\sum_{n=1}^N\lfloor b_n/2\rfloor$ such that $l^o_{\lambda^*}\ge s^o_{B-2\lambda^*+1}+s^o_{B-2\lambda^*+2}$, which implies that either $l^o_{\lambda^*+1}\le s^o_{B-2\lambda^*-1}+s^o_{B-2\lambda^*}$ or $\lambda^*=\sum_{n=1}^N\lfloor b_n/2\rfloor$; hence there exists no λ such that $\lambda^*<\lambda\le\sum_{n=1}^N\lfloor b_n/2\rfloor$). The second line of inequalities again follows because the sets $\mathbf{s}^o$ and $\mathbf{l}^o$ are ordered and $\lambda^*+1\le\lambda$. $\square$

Proof of Proposition 2. To prove optimality, it suffices to show that $\lambda^*=\sum_{i=1}^N\lambda_i$. It is easy to see that $\sum_{i=1}^N\lambda_i\le\lambda^*$, since

$$\sum_{i=1}^N\lambda_i\le\sum_{i=1}^N\lfloor b_i/2\rfloor\quad\text{and}$$
$$l^o_{\sum_{i=1}^N\lambda_i}\ge s^o_{B-2\sum_{i=1}^N\lambda_i+1}+s^o_{B-2\sum_{i=1}^N\lambda_i+2}.$$

Yet by definition we know that λ^* is the greatest integer satisfying these inequalities.

We now show that $\sum_{i=1}^N\lambda_i\ge\lambda^*$. Assume the contrary. Then $\lambda^*-\sum_{i=1}^N\lambda_i\ge 1$. Let t be the smallest index for which $\lambda^*-\sum_{i=1}^t\lambda_i\ge\sum_{i=t+1}^N\lfloor b_i/2\rfloor+1$, in which case $\lambda^*-\sum_{i=1}^{t-1}\lambda_i\le\sum_{i=t}^N\lfloor b_i/2\rfloor$.

First we show that $\lambda_t+1\le\lfloor b_t/2\rfloor$. From our choice of t it follows that

$$\lambda^*-\sum_{i=1}^t\lambda_i\ge\sum_{i=t+1}^N\left\lfloor\frac{b_i}{2}\right\rfloor+1$$
$$\Leftrightarrow\ \lambda_t+1\le\lambda^*-\sum_{i=1}^{t-1}\lambda_i-\sum_{i=t+1}^N\left\lfloor\frac{b_i}{2}\right\rfloor$$
$$\le\sum_{i=t}^N\left\lfloor\frac{b_i}{2}\right\rfloor-\sum_{i=t+1}^N\left\lfloor\frac{b_i}{2}\right\rfloor=\left\lfloor\frac{b_t}{2}\right\rfloor.$$

Next we show that

$$l^o_{\sum_{i=1}^t\lambda_i+1}\ge s^o_{\sum_{i=1}^t b_i-2(\sum_{i=1}^t\lambda_i+1)+N-t+1}+s^o_{\sum_{i=1}^t b_i-2(\sum_{i=1}^t\lambda_i+1)+N-t+2}.$$

To see this, note that

$$2\left(\lambda^*-\sum_{i=1}^t\lambda_i\right)\ge 2\sum_{i=t+1}^N\left\lfloor\frac{b_i}{2}\right\rfloor+2\ge 2\sum_{i=t+1}^N\frac{b_i-1}{2}+2$$
$$=B-\sum_{i=1}^t b_i-N+t+2;$$

this expression implies that

$$B-2\lambda^*+1\le\sum_{i=1}^t b_i-2\left(\sum_{i=1}^t\lambda_i+1\right)+N-t+1.$$

Therefore,

$$l^o_{\sum_{i=1}^t\lambda_i+1}\ge l^o_{\lambda^*}\ge s^o_{B-2\lambda^*+1}+s^o_{B-2\lambda^*+2}$$
$$\ge s^o_{\sum_{i=1}^t b_i-2(\sum_{i=1}^t\lambda_i+1)+N-t+1}+s^o_{\sum_{i=1}^t b_i-2(\sum_{i=1}^t\lambda_i+1)+N-t+2}.$$

Thus, λ_t is not the largest integer that satisfies inequalities $\lambda_t\le\lfloor b_t/2\rfloor$ and $l^o_{\sum_{i=1}^t\lambda_i}\ge s^o_{\sum_{i=1}^t(b_i-2\lambda_i)+N-t+1}+s^o_{\sum_{i=1}^t(b_i-2\lambda_i)+N-t+2}$—a contradiction. Hence, $\sum_{i=1}^N\lambda_i\ge\lambda^*$, which implies that $\sum_{i=1}^N\lambda_i=\lambda^*$. $\square$

Proof of Proposition 3. It is shown in Martello and Toth (1990) that $c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,E[\mathbf{b}]))=c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,E[B]))$. Also, because $c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,\mathbf{b}))$ is a concave function, by Jensen's inequality we have $E[c(\mathcal{V}(\mathbf{r}^o,\ \mathbf{x}^o,\ \mathbf{b}))]\le c(\mathcal{V}(\mathbf{r}^o,\ \mathbf{x}^o,\ E[\mathbf{b}]))$. But $V_1(\mathbf{r}^o,\ \mathbf{x}^o)\le E[\mathcal{V}(\mathbf{r}^o,\ \mathbf{x}^o,\ \mathbf{b})]\le E[c(\mathcal{V}(\mathbf{r}^o,\ \mathbf{x}^o,\ \mathbf{b}))]$; thus $V_1(\mathbf{r}^o,\mathbf{x}^o)\le c(\mathcal{V}(\mathbf{r}^o,\mathbf{x}^o,E[\mathbf{b}]))$. $\square$

Proof of Proposition 4. *Part* 1. Note that if there is a strict premium for long ads, then it is always optimal to schedule the highest-paying long ad available whenever a long break occurs.

To prove that the profit function is supermodular, we must show that $\breve\Pi(x_1)+\breve\Pi(x_2)\le\breve\Pi(x_1\vee x_2)+\breve\Pi(x_1\wedge x_2)$ for all $x_1=(n_L^1,\delta^1)\in\mathbb{L}$ and all $x_2=(n_L^2,\delta^2)\in\mathbb{L}$. We analyze the case of $n_L^1\ge n_L^2$ and $\delta^1\le\delta^2$; the case $n_L^2\ge n_L^1$ and $\delta^2\le\delta^1$ is analogous, and the inequality is trivially satisfied for the two other cases (i.e., when $n_L^1\ge n_L^2$ and $\delta^1\ge\delta^2$ and when $n_L^2\ge n_L^1$ and $\delta^2\ge\delta^1$). We have $n_L^1\vee n_L^2=n_L^1$, $\delta^1\vee\delta^2=\delta^2$, $n_L^1\wedge n_L^2=n_L^2$, and $\delta^1\wedge\delta^2=\delta^1$. We need to show that $\breve\Pi(n_L^1,\delta^2)+\breve\Pi(n_L^2,\delta^1)\ge\breve\Pi(n_L^1,\delta^1)+\breve\Pi(n_L^2,\delta^2)$ or, equivalently, that

$$\mathbf{A}:=\breve\Pi(n_L^1,\delta^2)-\breve\Pi(n_L^1,\delta^1)\ge\mathbf{B}:=\breve\Pi(n_L^2,\delta^2)-\breve\Pi(n_L^2,\delta^1).$$

Let $\breve\Pi(n_L^1,\delta^1)=\Pi(n_L^1,n_S^1)$, $\breve\Pi(n_L^2,\delta^2)=\Pi(n_L^2,n_S^2)$, $\breve\Pi(n_L^1,\delta^2)=\Pi(n_L^1,n_S^3)$, and $\breve\Pi(n_L^2,\delta^1)=\Pi(n_L^2,n_S^4)$. We claim that $n_S^1\le n_S^3$ and $n_S^4\le n_S^2$. To see this, note that

$$2n_L^1+n_S^1=(1+\delta^1)E[B],\tag{14}$$
$$2n_L^2+n_S^2=(1+\delta^2)E[B],\tag{15}$$
$$2n_L^1+n_S^3=(1+\delta^2)E[B],\tag{16}$$
$$2n_L^2+n_S^4=(1+\delta^1)E[B].\tag{17}$$

Then (14) and (17) yield $n_S^1=n_S^4+2(n_L^2-n_L^1)\le n_S^4$, whereas (15) and (16) yield $n_S^2=n_S^3+2(n_L^1-n_L^2)\ge n_S^3$. By (14) and (16) we have $n_S^1=E[B](\delta^1-\delta^2)+n_S^3\le n_S^3$, and by (15) and (17) we have $n_S^2=E[B](\delta^2-\delta^1)+n_S^4\ge n_S^4$. Therefore, $n_S^1\le n_S^3$ and $n_S^4\le n_S^2$. Observe also that $n_S^1+n_S^2=n_S^3+n_S^4$. Therefore,

$$\begin{aligned}&\breve\Pi(n_L^1,\delta^2)-\breve\Pi(n_L^1,\delta^1)\\&=\Pi(n_L^1,n_S^3)-\Pi(n_L^1,n_S^1)\\&=\sum_{i=0}^{n_L^1}\big([\min\{n_S^3,N-i\}-\min\{n_S^1,N-i\}]\\&\qquad-[\max\{0,n_S^3-N+i\}-\max\{0,n_S^1-N+i\}]\beta\big)p_i\\&\quad+\sum_{i=n_L^1+1}^{N}[\min\{n_S^3,N+i-2n_L^1\}-\min\{n_S^1,N+i-2n_L^1\}]p_i\\&\quad-\sum_{i=n_L^1+1}^{N}[\max\{0,n_S^3-N-i+2n_L^1\}-\max\{0,n_S^1-N-i+2n_L^1\}]\beta p_i.\end{aligned}$$

Define

$$\gamma_{i,j}(a)=\min\{n_S^i,a\}-\min\{n_S^j,a\}\quad\text{and}$$
$$\psi_{i,j}(a)=\max\{0,n_S^i-a\}-\max\{0,n_S^j-a\}.$$

Then

$$\begin{aligned}\mathbf{A} = &\sum_{i=0}^{n_L^1}[\gamma_{3,1}(N-i)-\psi_{3,1}(N-i)\beta]p_i\\ &+\sum_{i=n_L^1+1}^{N}[\gamma_{3,1}(N+i-2n_L^1)-\psi_{3,1}(N+i-2n_L^1)\beta]p_i,\end{aligned} \tag{18}$$

and

$$\begin{aligned}\mathbf{B} = &\sum_{i=0}^{n_L^2}[\gamma_{2,4}(N-i)-\psi_{2,4}(N-i)\beta]p_i\\ &+\sum_{i=n_L^2+1}^{N}[\gamma_{2,4}(N+i-2n_L^2)-\psi_{2,4}(N+i-2n_L^2)\beta]p_i.\end{aligned} \tag{19}$$

Define $\Gamma(a_1,a_2)=\gamma_{3,1}(a_1)-\gamma_{2,4}(a_2)$ and $\Psi(a_1,a_2)=\psi_{3,1}(a_1)-\psi_{2,4}(a_2)$. Then, by subtracting (19) from (18), we obtain

$$\begin{aligned}\mathbf{A}-\mathbf{B} = &\sum_{i=0}^{n_L^2}\Gamma(N-i,N-i)p_i+\sum_{i=n_L^2+1}^{n_L^1}\Gamma(N-i,N+i-2n_L^2)p_i\\ &+\sum_{i=n_L^1+1}^{N}\Gamma(N+i-2n_L^1,N+i-2n_L^2)p_i\\ &-\beta\bigg[\sum_{i=0}^{n_L^2}\Psi(N-i,N-i)p_i\\ &\quad+\sum_{i=n_L^2+1}^{n_L^1}\Psi(N-i,N+i-2n_L^2)p_i\\ &\quad+\sum_{i=n_L^1+1}^{N}\Psi(N+i-2n_L^1,N+i-2n_L^2)p_i\bigg].\end{aligned} \tag{20}$$

Lemma. *If the following three conditions hold, then $\Gamma(a_1,a_2)\geq 0$ and $\Psi(a_1,a_2)\leq 0$:*
(i) $0\leq a_2-a_1\leq 2(n_L^1-n_L^2)$;
(ii) $n_S^1\leq n_S^3$ *and* $n_S^4\leq n_S^2$;
(iii) $n_S^4-n_S^1=n_S^2-n_S^3=2(n_L^1-n_L^2)$.

Proof. We analyze all possible scenarios as follows.

- If $a_1\leq n_S^1$ and $a_2\leq n_S^4$, then $\Gamma=0\geq 0$ and $\Psi=0\leq 0$.
- If $a_1\leq n_S^1$ and $a_2> n_S^4$, then $a_2-a_1>n_S^4-n_S^1$, so this scenario does not satisfy condition (i).
- If $a_1\geq n_S^1$ and $a_2\leq n_S^4$, then $\Gamma=\min\{n_S^3,a_1\}-n_S^1\geq 0$ and $\Psi=\max\{0,n_S^3-a_1\}+n_S^4-n_S^2$. If $n_S^3\geq a_1$, then $\Psi=n_S^3-a_1+n_S^4-n_S^2=n_S^1-a_1\leq 0$; if $n_S^3<a_1$, then $\Psi=n_S^4-n_S^2\leq 0$.
- If $a_1\geq n_S^1$ and $a_2\geq n_S^4$, then $\Gamma=\min\{n_S^3,a_1\}-n_S^1-\min\{n_S^2,a_2\}+n_S^4$.

— If $a_1\leq n_S^3$, then $a_2\leq n_S^2$ (else conditions (i) and (iii) would be violated), so $\Gamma=a_1-n_S^1-a_2+n_S^4\geq 0$ by (i) and (iii). Also, $\Psi=n_S^3-a_1-n_S^2+a_2=-2(n_L^1-n_L^2)+a_2-a_1\leq 0$ by (i).

— If $a_1>n_S^3$, then $\Gamma=n_S^3-n_S^1-\min\{n_S^2,a_2\}+n_S^4=n_S^2-\min\{n_S^2,a_2\}\geq 0$; also, $\Psi=-\max\{0,n_S^2-a_2\}\leq 0$.

When combined with (20), this lemma yields $\breve{\Pi}(n_L^1,\delta^2)-\breve{\Pi}(n_L^1,\delta^1)-\breve{\Pi}(n_L^2,\delta^2)+\breve{\Pi}(n_L^2,\delta^1)\geq 0$. Therefore, $\breve{\Pi}(n_L,\delta)$ is supermodular. □

Part 2. The first inequality follows directly from the supermodularity of $\breve{\Pi}(n_L,\delta)$. To prove the second inequality, let $\breve{\Pi}(n_L,\delta)=\Pi(n_L,n_S)$. We have

$$\begin{aligned}\Delta\breve{\Pi}(n_L+1,\delta) &= \Pi(n_L+1,n_S-2)-\Pi(n_L,n_S)\\ &=\sum_{i=0}^{n_L}[\min\{n_S-2,N-i\}-\min\{n_S,N-i\}]p_i\\ &\quad-2\beta(1+\epsilon)\sum_{i=1}^{n_L}p_i\\ &\quad-\sum_{i=0}^{n_L}[\max\{0,n_S-2-N+i\}\\ &\qquad-\max\{0,n_S-N+i\}]\beta p_i\\ &\quad+2\epsilon\sum_{i=n_L+1}^{N}p_i;\end{aligned}$$

$$\begin{aligned}\Delta\breve{\Pi}(n_L,\delta) &= \Pi(n_L,n_S)-\Pi(n_L-1,n_S+2)\\ &=\sum_{i=0}^{n_L-1}[\min\{n_S,N-i\}-\min\{n_S+2,N-i\}]p_i\\ &\quad-2\beta(1+\epsilon)\sum_{i=1}^{n_L-1}p_i\\ &\quad-\sum_{i=0}^{n_L-1}[\max\{0,n_S-N+i\}\\ &\qquad-\max\{0,n_S+2-N+i\}]\beta p_i\\ &\quad+2\epsilon\sum_{i=n_L}^{N}p_i.\end{aligned}$$

Then

$$\begin{aligned}&\Delta\breve{\Pi}(n_L+1,\delta)-\Delta\breve{\Pi}(n_L,\delta)\\ &=\sum_{i=0}^{n_L-1}[\min\{n_S-2,N-i\}-2\min\{n_S,N-i\}\\ &\qquad+\min\{n_S+2,N-i\}]p_i\\ &\quad-\sum_{i=0}^{n_L-1}[\max\{0,n_S-2-N+i\}-2\max\{0,n_S-N+i\}\\ &\qquad+\max\{0,n_S+2-N+i\}]\beta p_i\\ &\quad-(2\epsilon+2\beta(1+\epsilon))p_{n_L}\\ &\quad+[\min\{n_S-2,N-n_L\}-\min\{n_S,N-n_L\}]p_{n_L}\\ &\quad-[\max\{0,n_S-2-N+n_L\}-\max\{0,n_S-N+n_L\}]\beta p_{n_L}.\end{aligned}$$

For all $a\geq 0$, we have

$$\min\{n_S-2,a\}-2\min\{n_S,a\}+\min\{n_S+2,a\}\leq 0,$$
$$\max\{0,n_S-2-a\}-2\max\{0,n_S-a\}+\max\{0,n_S+2-a\}\geq 0,$$
$$2\epsilon+2\beta(1+\epsilon)\geq 2\beta,$$

and

$$\begin{aligned}&\min\{n_S-2,N-n_L\}-\min\{n_S,N-n_L\}\\ &\quad-[\max\{0,n_S-2-N+n_L\}-\max\{0,n_S-N+n_L\}]\beta\leq 2\beta.\end{aligned}$$

Hence, $\Delta\breve{\Pi}(n_L+1,\delta)\leq\Delta\breve{\Pi}(n_L,\delta)$. □

Proof of Proposition 5. One can easily show that it is suboptimal to have fewer short ads than the number of breaks N when short ads sell at a premium (i.e., when long ads sell at a discount), since the short ads will necessarily be aired (because capacity is always greater than N S-second spots). For brevity we focus our attention on the case $n_S \geq N$ and $s=1 \leq l=2(1+\epsilon) \leq 2s=2$. Let n_L^* be the optimal number of long ads, and let δ^* be the optimal overbooking level. We show that if $n_L^* \geq 1$, then $\breve{\Pi}(n_L^*-1, \delta^*) > \breve{\Pi}(n_L^*, \delta^*)$, which would contradict the optimality of n_L^*. Therefore, $n_L^*=0$.

It is clear that

$$\begin{aligned}\Pi(n_L-1, n_S+2)-\Pi(n_L, n_S)&\\ =-2\varepsilon\big[\Pr(k_L \geq n_S-N+1)-\beta\Pr(k_L \leq n_S-N)\big],& \quad (21)\\ \Pi(0, n_S+1)-\Pi(0, n_S)&\\ =\Pr(k_L \geq n_S-N+1)-\beta\Pr(k_L \leq n_S-N).& \quad (22)\end{aligned}$$

Let $n_S^*=\arg\max_{n_S}\Pi(0, n_S)$. (If there are multiple optima, we choose the lowest of them.) Then

$$\Pr(k_L \geq n_S^*-N+1)-\beta\Pr(k_L \leq n_S^*-N) \leq 0, \quad (23)$$

$$\Pr(k_L \geq n_S^*-N)-\beta\Pr(k_L \leq n_S^*-N-1) > 0. \quad (24)$$

Moreover, it follows trivially that $\Pr(k_L \geq n_S-N+1)-\beta\Pr(k_L \leq n_S-N) \leq 0$ for all $n_S \geq n_S^*$ (since $\Pr(k_L \geq n_S-N+1) \leq \Pr(k_L \geq n_S^*-N+1)$ and $\Pr(k_L \leq n_S-N) \geq \Pr(k_L \leq n_S^*-N)$ for all $n_S \geq n_S^*$) and that $\Pr(k_L \geq n_S-N+1)-\beta\Pr(k_L \leq n_S-N) > 0$ for all $n_S < n_S^*$. Also, from (22) we have

$$\Pi(0, n_S+2n_L)-\Pi(0, n_S) = \sum_{i=0}^{2n_L-1}\big[\Pr(k_L \geq n_S-N+1+i)-\beta\Pr(k_L \leq n_S-N+i)\big]. \quad (25)$$

Next, by (21),

$$\begin{aligned}\Pi(0, n_S+2n_L)-\Pi(n_L, n_S) &= -2\varepsilon\sum_{i=0}^{n_L-1}\big[\Pr(k_L \geq n_S-N+1+2i)\\ &\qquad -\beta\Pr(k_L \leq n_S-N+2i)\big]. \quad (26)\end{aligned}$$

From (25) and (26) we derive

$$\begin{aligned}\Pi(n_L, n_S)-\Pi(0, n_S) &= \sum_{i=0}^{n_L-1}\big[-\Pr(n_S-N+1+2i \leq k_L \leq n_S-N+1+2i+1)\\ &\qquad -\beta\Pr(n_S-N+2i \leq k_L \leq n_S-N+2i+1)\big]\\ &\quad +2(1+\varepsilon)\sum_{i=0}^{n_L-1}\big[\Pr(k_L \geq n_S-N+1+2i)\\ &\qquad -\beta\Pr(k_L \leq n_S-N+2i)\big], \quad (27)\end{aligned}$$

which is negative for all $n_S \geq n_S^*$.

For all $n_S \leq n_S^*-1$ and $n_L \geq 1$, the broadcaster's revenue $\Pi(n_L, n_S)$ cannot be optimal because $\Pi(n_L-1, n_S+2) > \Pi(n_L, n_S)$ (by (21) and (24)). From (23) and (27) it follows that $\Pi(n_L, n_S) \leq \Pi(0, n_S) \leq \Pi(0, n_S^*)$ for all $n_S \geq n_S^*$. Thus, we cannot have $n_L^* \geq 1$. Consequently, $n_L^*=0$, and $\Pi(0, n_S^*)$ is the maximum revenue. From (23) we now obtain the desired result; namely, n_S^* is the smallest integer such that $\Pr(k_L \leq n_S^*-N) \geq 1/(1+\beta)$. □

Proof of Proposition 6. (a) We know from Proposition 5 that if $l \leq 2s$, then $n_L^*=0$, and the result follows trivially. We now analyze the case $l > 2s$. When ads have constant yields, the marginal revenue from booking one extra long ad, while keeping the same number of short ads, is given in the second column of Table A.1, for all the various cases. When long ads have variable yields, then the marginal revenue from booking one extra long ad with yield l, while keeping the number of short ads constant, is given in the last column of Table A.1, for all the various cases.

The reasoning is as follows.

- If $k_L > n_L^*$ and $n_S^* \geq N+k_L-2n_L^*$, then all capacity will be utilized. Moreover, there would be a long break with two short ads in it. By booking one extra long ad with revenue $l > 2s$, we can replace the two short ads with the long ad and get a higher revenue. We would, however, incur a penalty on the two short ads, which will no longer be aired. When long ads have variable yields, it is also possible that the long break with two short ads in it is not the long break with the lowest revenue. There could be a long ad $l_j < 2s$, in which case the extra ad l would be used to replace that long ad in the break. In this case, the marginal revenue from booking one extra long ad is $l-l_j-\beta l_j > l-2s-2\beta s$.
- If $k_L > n_L^*$ and $n_S^*=N+k_L-2n_L^*-1$, then all breaks except one long break will be fully utilized. That long break will have only one short ad in it. By booking one extra long ad, we can replace that short ad with the long ad. The marginal revenue will be $l-s-\beta s$. As before, when long ads have variable yields, it is also possible that the long break with one short ad in it is not the long break with the lowest revenue. There could be a long ad $l_j < s$, in which case the extra ad l would be used to replace that long ad in the break. In this case, the marginal revenue from booking one extra long ad is $l-l_j-\beta l_j > l-s-\beta s$.
- If $k_L > n_L^*$ and $n_S^* \leq N+k_L-2n_L^*-2$, then there will be at least one long break that is empty. By booking one extra long ad we can fill the empty break and generate a marginal revenue of l.
- If $k_L \leq n_L^*$, then not all long ads will be aired. By booking one extra long ad, four scenarios are possible: the ad does not get aired (in which case we incur a penalty of βl);

Table A.1 Marginal Revenue from Booking One Extra Long Ad When Ads Have Constant Yields

Case	$\Pi(n_L^*+1, n_S^*)-\Pi(n_L^*, n_S^*)$	$\Pi(\{l_1,\ldots,l_{n_L^*}\}\cup\{l\}, n_S^*)-\Pi(\{l_1,\ldots,l_{n_L^*}\}, n_S^*)$
$k_L > n_L^*$, $n_S^* \geq N+k_L-2n_L^*$	$l-2s-2\beta s$	$l-2s-2\beta s$ or $l-l_j-\beta l_j$, with $l_j < 2s$
$k_L > n_L^*$, $n_S^*=N+k_L-2n_L^*-1$	$l-s-\beta s$	$l-s-\beta s$ or $l-l_j-\beta l_j$, with $l_j < s$
$k_L > n_L^*$, $n_S^* \leq N+k_L-2n_L^*-2$	l	l
$k_L \leq n_L^*$	$-\beta l$	$-\beta l$ or $l-2s-2\beta s$ or $l-s-\beta s$ or $l-l_j-\beta l_j$ with $l_j < 2s$

the ad gets aired and replaces a long ad with a lower revenue in a long break (in which case the marginal revenue will be $l - l_j - \beta l_j > -\beta l$, because $l_j < l$); the ad gets aired and replaces two short ads in a long break (in which case the marginal revenue will be $l - 2s - 2\beta s > -\beta l$, because $2s < l$); finally, the ad gets aired and replaces one short ad in a long break (in which case the marginal revenue will be $l - s - \beta s > -\beta l$, because $s < l$).

Note now that under all scenarios, the marginal revenue from booking one extra long ad when long ads have variable yields is at least as high as the marginal revenue from booking one extra long ad when ads have constant yields.

The proof for part (b) is similar, and we omit it for brevity. □

References

Amaruchkul K, Cooper WL, Gupta D (2007) Single-leg air-cargo revenue management. *Transportation Sci.* 41:457–469.

Araman VF, Popescu I (2009) Media revenue management with audience uncertainty: Balancing upfront and spot market sales. *Manufacturing Service Oper. Management* 12:190–212.

Atkinson LB (1994) A greedy, look-ahead heuristic for combinatorial optimization: An application to vehicle scheduling with time windows. *J. Oper. Res. Soc.* 45:673–684.

Belobaba PP (1987) Air travel demand and airline seat inventory management. Ph.D. thesis, Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambridge.

Birge J, Frenk JBG, Mittenthal J, Rinnooy Kan AHG (1990) Single-machine scheduling subject to stochastic breakdown. *Naval Res. Logist.* 37:661–677.

Bollapragada S, Garbiras M (2004) Scheduling commercials on broadcast television. *Oper. Res.* 52:337–345.

Bollapragada S, Bussieck MR, Mallik S (2004) Scheduling commercial videotapes in broadcast television. *Oper. Res.* 52:679–689.

Cai X, Wu X, Zhou X (2009) Stochastic scheduling subject to preemptive-repeat breakdowns with incomplete information. *Oper. Res.* 57:1236–1249.

Ciarallo F, Akella R, Morton T (1994) A periodic review, production planning model with uncertain capacity and uncertain demand optimality of extended myopic policies. *Management Sci.* 40:320–332.

Fisk JC, Hung MS (1979) A heuristic routine for solving large loading problems. *Naval Res. Logist. Quart.* 26:643–650.

Fixmer A (2010) CBS loss narrows as advertising revenue rebounds. *Bloomberg Business* (May 5), http://www.bloomberg.com/news/articles/2010-05-05/cbs-reports-first-quarter-loss-narrows-as-advertising-revenue-rebounds.

Ghodsi R, Sassani F (2005) Real time optimum sequencing of wood cutting process. *Internat. J. Production Res.* 43:1127–1141.

Gilmore PC, Gomory RE (1961) A linear programming approach to the cutting stock problem. *Oper. Res.* 9:849–859.

Gilmore PC, Gomory RE (1965) Multistage cutting stock problems of two and more dimensions. *Oper. Res.* 13:94–120.

Green A (2006) Are 60-sec ads twice as effective as 30-sec ads? *Warc Media FAQ* (WARC Ltd., London, UK), http://www.warc.com.

Grosfeld-Nir A, Gerchak Y (2004) Multiple lotsizing in production to order with random yields: Review of recent advances. *Ann. Oper. Res.* 126:43–69.

Hwang J, Singh M (1998) Optimal production policies for multi-stage systems with setup costs and uncertain capacities. *Management Sci.* 44:1279–1294.

Kacen J (2010) Super Bowl vs. Winter Olympics: Which is a better buy for advertisers? Bauer College of Business, University of Houston (February 1), http://bauerticker.uh.edu/media-alerts/week-of-feb-1-2010.

Kawashima N (2006) Advertising agencies, media and consumer market: The changing quality of TV advertising in Japan. *Media Culture Soc.* 28:393–410.

Khang D, Fujiwara O (2000) Optimality of myopic ordering policies for inventory model with stochastic supply. *Oper. Res.* 48: 181–184.

Kimms A, Müller-Bungart M (2007) Revenue management for broadcasting commercials: The channel's problem of selecting and scheduling the advertisements to be aired. *Internat. J. Revenue Management* 1:28–44.

Martello S, Toth P (1990) *Knapsack Problems: Algorithms and Computer Implementations* (John Willey & Sons, Chichester, UK).

Newstead K, Romaniuk J (2010) Cost per second: The relative effectiveness of 15- and 30-second television advertisements. *J. Advertising Res.* 50:68–76.

Pinedo M, Rammouz E (1988) A note on stochastic scheduling on a single machine subject to breakdown and repair. *Probab. Engrg. Informational Sci.* 2:41–49.

Redhead S (2007) Those absent from the stadium are always right. *J. Sport Soc. Issues* 31:226–241.

Roels G, Fridgeirsdottir K (2009) Dynamic revenue management for online display advertising. *J. Pricing Revenue Management* 8:452–466.

Scull D (1981) A stochastic cutting stock procedure: Cutting rolls of insulating tape. *Management Sci.* 27:946–952.

Stampler L (2013) UPDATED: Here's everything we know about the Super Bowl ad lineup. *Business Insider* (January 25), http://www.businessinsider.com/2013-super-bowl-ads-everything-we-know-2012-12?IR=T&op=1.

Tomlin B (2009) Impact of supply learning when suppliers are unreliable. *Manufacturing Service Oper. Management* 11:192–209.

Topkis DM (1998) *Supermodularity and Complementarity* (Princeton University Press, Princeton, NJ).

Wang Y, Gerchak Y (1996) Periodic review production models with variable capacity, random yield, and uncertain demand. *Management Sci.* 42:130–137.

Wäscher G, Haußner H, Schumann H (2007) An improved typology of cutting and packing problems. *Eur. J. Oper. Res.* 183: 1109–1130.

Yano C, Lee H (1995) Lot sizing with random yields: A review. *Oper. Res.* 43:311–334.

Zhang D, Cooper WL (2005) Revenue management for parallel flights with customer choice behavior. *Oper. Res.* 53:415–431.