INSEAD

The Business School for the World

Faculty & Research Working Paper

Stochastic Revenue Management Models for Media Broadcasting

Victor F. ARAMAN
Ioana POPESCU
2007/73/DS (revised version of 2007/05/DS)

Stochastic Revenue Management Models for Media Broadcasting

by

Victor F. Araman*

and

Ioana Popescu**

*	Information, Operations and Management Science Stern School of Business, New York, University, New York, NY 10012, USA, varaman@stern.edu

**	The Booz Allen Hamilton Term Chaired Professor in Strategic Revenue Management, Associate Professor of Decision Sciences at INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France, ioana.popescu@insead.edu

Stochastic Revenue Management Models for Media Broadcasting

Victor F. Araman • Ioana Popescu[1]

Information, Operations and Management Science Stern School of Business, New York University, New York, NY 10012, USA

Decision Science Area, INSEAD, Fontainebleau, 77300, France

varaman@stern.edu • ioana.popescu@insead.edu

An important challenge faced by media broadcasting companies is how to allocate limited advertising space between upfront contracts and the spot market (referred to in advertising as the scatter market), in order to maximize profits and meet contractual commitments. We develop stylized optimization models of airtime capacity planning and allocation across multiple clients under audience uncertainty. In a short term profit maximizing setting, our results suggest that broadcasting companies should prioritize upfront clients according to marginal revenue per contracted audience unit, also known as CPM (cost per thousand viewers). For capacity planning purposes, accepted upfront market contracts can be aggregated across clients. The upfront market capacity should then be allocated to clients in proportion to contracted audience. Closed form solutions are obtained in a static setting. These results remain valid in a dynamic setting, when considering the opportunity to increase allocation by airing make-goods during the broadcasting season. Our structural results characterize the impact of contracting parameters, time and audience uncertainty on profits and capacity decisions. The results hold under general audience and spot market profit models, as well as under service constrained models.

1. Introduction

Optimally managing and valuing limited advertising space is one of the most relevant problems faced by media companies today. This paper provides formal models and solutions for the problem of managing broadcast advertising capacity. We begin by describing how advertising is sold, in order to illustrate the complexities of this problem, and motivate the focus of our work. This introduction further highlights our contribution and related literature, and introduces the media jargon.

[1]This author's research was partially supported by the Booz Allen Hamilton Term Chair on Revenue Management at INSEAD.

1.1 The Media Broadcasting Advertising Market

In the 1930s, the sponsorship of radio serials by makers of household-cleaning products led to the soap opera. Listeners were enthralled by episodic, melodramatic storylines, and advertisers were guaranteed a big audience. According to TNS media intelligence, 2006 U.S. advertising spending exceeded USD 150 billion. Of this, television accounts for 44%, with advertisers ready to pay up to half a million dollars for a 30 second commercial in a popular show such as *Friends*, and new-show campaigns reaching up to USD 10 Million (source adage.com). According to the A.C. Nielsen Co., the average American watches more than 4 hours of TV each day, 30% of which is advertising.

Media broadcasting companies such as TV channels, cable networks or radio stations collect most of their revenues from selling impressions, or "eye-balls", through advertisement space (30 second commercial slots) during various programmes. Typically, in north America and several European countries, the bulk of advertising space (about 80%) is sold during an *upfront market*, following the announcement of program schedules and prices for the year. In the US, the upfront market occurs during less than a couple of weeks in May, much before the broadcasting season starts in mid-September. During this period, a few (about 20%) major advertisers (MA) buy long term contracts (called *media plans*, or *campaigns*) from the broadcasting company (BC) at relatively low margins.[2] Upfront contracts stipulate a budget to be spent throughout the year, and a cost per thousand viewers (CPM), the ratio of the two providing a target level of audience (or ratings) to be reached by the campaign. The remaining advertising space is sold throughout the broadcasting season on a spot market, called *scatter market* in the industry, on a price-per-slot basis, usually at a higher margin and with no audience guarantee.

This paper investigates the problem of optimally managing media advertising capacity. This problem resembles in many ways the standard capacity control paradigm in revenue management (see Talluri and van Ryzin 2004 for a comprehensive treatment and Elmaghraby and Keskinocak 2003 for a review), where a service provider allocates limited capacity (e.g. seats on a flight or rooms in a hotel) between customer classes with different valuations and arrival patterns (e.g. business and leisure customers). Similarly, in the media revenue management problem, the broadcasting company allocates limited advertising space, called

[2]In reality, sales are either direct or managed by intermediary buying and selling houses which represent their respective clients' interests. Our results and analysis are relevant to the party in charge of the capacity planning and revenue management function, for simplicity referred to as broadcasting company (BC).

airtime or media capacity,[3] between two customer classes: *upfront* (market) clients, who buy early during the upfront market at high discounts, and *scatter* (market) clients, who buy later on the spot (or scatter) market at higher prices. Revenue management is particularly relevant when capacity is tight, which is the case for example with prime-time advertising.

Several specificities differentiate the problem of managing media capacity from the standard revenue management setup. First, unlike traditional revenue management, in the media problem *the value of advertising capacity is uncertain* at the time when capacity allocation decisions are made. This is because upfront contract pricing is state contingent: the amount paid by an upfront client for a 30 second advertisement is determined by the audience reached, multiplied by the negotiated rate per viewer (CPM). Audience (i.e. the number of viewers, or impressions) is unknown ex-ante, when capacity allocation is made, but provided ex-post by media rating agencies. Therefore, in managing advertising capacity, the broadcasting company (BC) bears the risk of audience uncertainty in the upfront market. The value of an advertising slot depends both on the uncertain audience level, and on the opportunity cost of selling it later on the scatter market. Our goal is to develop a revenue management framework for valuing and managing limited advertising capacity under audience uncertainty.

Other important differences from traditional revenue management stem from the media B2B setting: transactions are contract based, upfront clients hold a strong bargaining power and retention of key accounts is an important strategic issue.[4] This paper focuses on the short term profit maximization problem of the firm in managing advertising capacity between upfront and scatter markets, with given contracts and prices, during one broadcasting season. We refer to this as the media revenue management, or capacity planning problem.

The media revenue management problem is a highly complex multi-level problem. Industry practice, organizational and technical considerations, all argue for a hierarchical planning approach, separating strategic, operational and tactical layers of decision making. First, during the upfront market, the BC negotiates contract allocation and estimates overall capacity requirements for the upfront clients; this is known in practice as the *strategic* or *upfront planning* phase. Subsequently, account executives allocate an initial portion of advertising

[3]Advertising capacity is fixed, by physical time limitation and regulations. In Europe for example, EU directives restrict advertising to 15% of total airtime, with at most 12 advertising minutes per hour, and not more than one break per movie. American regulations are less strict (about 30% of airtime).

[4]This partly explains the large volume sold at low margins upfront. Other arguments include stock market signaling, account executive incentives (sales volume), BC's risk aversion.

capacity to the client (known as a media plan or sales plan), with the provision that additional capacity, so-called *make-goods* (or audience deficiency units, ADU), will be allocated during the scatter market if the plan under-performs. Sequential make-goods allocation decisions are made periodically during the rest of the year; this is called *operational planning*. A detailed description of this process is provided in Section 2.2.

This paper focuses on the capacity planning decisions upfront, and during the broadcasting season. Other relevant elements of the media problem that are not addressed here include rate card pricing and contracting (both in monopolistic and competitive settings), allocation of clients to programs and scheduling. The scheduling phase uses planning recommendations (such as the ones provided by our results) to determine the exact location of an ad in a break, accounting for product conflict and other scheduling constraints.

1.2 Contribution and Structure

This paper focuses on the strategic and operational levels of the media capacity planning problem, across upfront and scatter markets. We propose stylized models that take into account key elements of the media revenue management problem and provide insights for the various layers of decision making. From a modeling standpoint, an important contribution is that we specifically account for prior uncertainty about audience ratings. We also specifically model the trade-off between spot and upfront allocation to multiple clients. Before detailing the contributions and structure of the paper, it is important to keep in mind that current practice in the industry is to make decisions qualitatively (see e.g. Bollapragada and Mallik 2007). To our knowledge, no mathematical model is currently being used for capacity planning or make-goods allocation, and, in general, for quantifying the value of capacity under audience uncertainty.

The next section provides further background about the media problem and business model, allowing to set up the terminology, model ingredients and main assumptions. We discuss models of audience uncertainty and describe the upfront and scatter market revenue models.

Section 3 derives relevant insights from a simple static, aggregate model for upfront capacity planning under audience uncertainty. In particular, we investigate the impact of audience uncertainty on the model, profits and decisions. In the media problem, the *value* of supply is a priori uncertain, measured by audience ratings. A transformation of audience (i.e. supply value) uncertainty into uncertainty in client demand, measured in capacity units

(30 second ad slots), allows to map the media problem into a traditional newsvendor problem (see e.g. Porteus 2000) and sheds light on the impact of audience uncertainty on optimal ad space allocation. We further provide static upfront planning recommendations across multiple periods (airdates or quarters) with varying audiences and performance targets.

Section 4 focuses on how capacity decisions are operationalized. We propose dynamic models for "make-goods" vs. scatter market allocation during the broadcasting season, when audience, modeled as a stochastic process, is revealed periodically. Our structural results show that initial capacity commitment to the upfront market should be minimal (known in the industry as "gapping"), and the optimal make-goods allocation policy is a monotone threshold type policy. Such models allow the BC to dynamically monitor capacity and performance, and make optimal trade-offs between scatter market profitability and fulfilment of upfront market contracts. Finally, several intuitive heuristics that approximate the optimal dynamic policy are proposed and evaluated.

The results described so far focus on aggregate models of capacity allocation for upfront vs. scatter markets. In Section 5 we analyze capacity allocation decisions across multiple customers. We do so subject to a common service level constraint across customers, which is consistent with industry practice. Our main result is that upfront clients can be aggregated for capacity planning and make-goods allocation. We show that the total capacity dedicated for the upfront market can be determined by solving a single aggregate allocation problem, akin to the model studied in Section 3. Individual client allocations should be in proportion to their contracted performance targets. These results remain valid for make-goods provisioning. The dynamic multi-client make-goods allocation problem reduces to an aggregate, one-dimensional dynamic program (studied in Section 4), which is much simpler to analyze. These results provide the BC with insights that significantly simplify the upfront contract negotiation and capacity planning task. For modeling purposes, these results allow us to simplify the analysis by focusing on aggregate planning models.

Section 6 reveals the robustness of our insights under alternative service constrained models, and discusses service differentiation and long term profit maximizing strategies. The last section concludes, outlining opportunities for further research in media revenue management.

1.3 Related Literature

The marketing literature has extensively investigated the impact of TV advertising on the consumer and sales (staring with Metheringham 1964, see also Kanetkar et al. 1992 and Lodish et al. 1995), but largely ignored the issue of airtime capacity planning. Despite its richness and complexity, the media revenue management problem has received limited attention in the operations literature. Chapter 10.5 of Talluri and van Ryzin (2004) provides a brief account of the media revenue management problem (this book is the most complete reference on revenue management to date). To our knowledge, the only published works on broadcast inventory and revenue management are Bollapragada et al. (2002) and Bollapragada and Garbiras (2004), referring to models successfully implemented at NBC, and recently Kimms and Müller-Bungart (2007). These papers provide deterministic solutions for capacity management in absence of audience uncertainty. Specifically, the first two papers study deterministic versions of the models in Section 3.4, that account for scheduling constraints (e.g. show mix and product conflict), but ignore spot market opportunity costs and audience uncertainty. Kimms and Müller-Bungart (2007) also consider the problem of client acceptance in a deterministic model that forces all client targets to be met.

A recent working paper by Bollapragada and Mallik (2007) investigates how a risk averse BC should allocate rating points between aggregate upfront and scatter markets, when audience and scatter market revenues are uncertain and independent (mainly modeled as uniform). Target revenue and value (revenue) at risk objectives are optimized in a static one period model that aggregates demand from each market, similar to our aggregate model in Section 3. They compare the optimal solution resulting from their models with a risk neutral one, and provide comparative statics with respect to audience parameters. Their sole decision variable is the total number of rating points sold in the upfront market. There is no specification of how this translates into actual capacity allocation, and how this allocation is "operationalized" across clients and over time. Our work complements theirs by answering these questions in a risk neutral context.

There are several industries where capacity is sold partly in a forward (advance purchase) market and partly on a spot market, such as electricity markets, cargo shipping, manufacturing etc. A growing body of literature, reviewed by Kleindorfer and Wu (2003), investigates inventory management in such settings. Wu and Kleindorfer (2005) develop a two-stage framework that integrates spot market transactions with supply chain contracting. In the first stage, multiple sellers compete and bid on capacity. In the second stage, the buyer decides how much to exercise from the contracts and how much to procure from the spot market. In a multi-period model, Araman and Ozer (2005) study optimal inventory allocation between a long term sales channel and a spot market. Work in this area focuses mainly on production models under demand uncertainty and supply contracts.

One feature that distinguished the media problem from the above literature, and much of the operations literature, is the uncertain value of supply: audience is contracted for, but only realized after airtime capacity allocation is made. This aspect makes our problem similar in spirit to production planning models under random yield, for which Yano and Lee (1995) provide an excellent review. Our aggregate dynamic model in Section 4 falls in the class of multiple lotsizing in production to order (MLPO) with random yield problems, surveyed by Grosfeld-Nir and Gerchak (2004) (see also Yano and Lee 1995, page 321). Such models aim to satisfy a fixed initial demand target through a pre-specified number of production runs to minimize inventory (holding and penalty) costs. In each production run, a random output is produced, that is a function of the lot size decision. In our case, the input decision is the number of ads to air for a client, and output is the audience (or ratings) for these ads. Our aggregate model differs from such random yield models in several respects, including the non-storable nature of capacity and absence of holding costs, the non-linear scatter market profit (corresponding to production costs), the general audience distribution (vs. specific discrete, e.g. binomial yield) and the multi-client model. While our insights focus exclusively on the media setting, several of our results, such as the multiple client analysis, can be viewed as a contribution to the random yield literature.

2. Audience and Revenue Models

This section provides further background for the media revenue management problem and business model. We describe the main ingredients of our models, including upfront and scatter market contracting terms, as well as audience and performance measures.

2.1 Audience Models, Performance Metrics and Forecasts

An important contribution of this paper is to provide tools for valuing advertising capacity by specifically considering the impact of audience uncertainty in the media planning task. Hence an important prerequisite for our model and results is to understand ex-post audience metrics and motivate ex-ante forecast models of audience variability. We begin by discussing existing metrics and possible levels of forecasted uncertainty.

Audience is the gross sum of all media exposures (the number of impressions, or "eyeballs" watching a given show), regardless of duplication. This is unknown ex-ante, but provided ex-post by media rating agencies, such as Nielsen Media Research in the US, or Médiamétrie in France. Another popular metric is *GRP* (gross rating point), also known as *rating*, the percentage of the target audience reached by an advertisement.[5] Because the ratio of audience to GRP is a constant (the number of TV viewers/households), the two terms are used interchangeably at no loss of generality.[6]

To obtain ratings forecasts, the audience (or equivalently GRP) that will be viewing an ad is modeled ex-ante as a positive random variable ξ with mean μ, standard deviation σ, distribution F and density f. Its realization, as well as relevant metrics, are provided ex-post by media rating agencies. For example, a subjective probability q that each of M potential viewers watches a given show suggests a binomial model $\xi \sim Bin(M, q)$, which for large market sizes M could be approximated by a normal distribution. All our results are distribution independent.

The performance of a media plan consisting of x ads targeted to an uncertain audience ξ is modeled as $\Psi(x, \xi) = x\xi$. This multiplicative performance model essentially assumes that a constant, but *a priori* unknown fraction of the population views a constant fraction of the ads. Here ξ captures the high level prior uncertainty about show popularity (or strength), arguably the main driver of audience uncertainty. (Once a show is on, audience is relatively more predictable: more or less the same people watch *Friends* every week.) The multiplicative performance model is simple to work with and fairly general.[7]

In reality there is more variability to audience than just show popularity; for example

[5] For example, during the week of November 20, 2006, the ABC show "Desperate Housewives" topped the US household ratings at 13.5, meaning that 13.5% of the estimated 110.2M TV households in the US watched the show that week.

[6] Other relevant metrics include *reach* (the percentage of individuals within a targeted market that receive the marketing message at least once) and *frequency* (the number of times the target consumer is exposed to the marketing message during a campaign); see e.g. Metheringham (1964).

[7] Our sensitivity results for upfront and dynamic make-goods allocation extend for more general performance measures of the form $\Psi(x, \xi)$, with Ψ an increasing function of both arguments and concave in x. The concavity assumption is justified by "repetition wearout", i.e. the diminishing marginal benefit of repeatedly reaching the same individuals. Models of the form $\Psi(x, \xi) = h(x) \cdot g(\xi)$ reduce to a multiplicative model, by a simple change of variable; for example *reach* can be modeled as $\Psi(x, \xi) = \rho(x)\xi = (1 - (1 - q)^x)\xi$, where ρ is increasing and concave in x.

audience is different from week to week, or for new rather than replay episodes of a show.[8] The next example discusses how the multiplicative model can be viewed as a first order approximation of a more complex model that captures audience variability across air-dates.

Example 1. The total number of (possibly duplicated) impressions from airing x commercials with audiences $\xi_1, \ldots, \xi_x$ on a show is the convolution $\xi^{(x)} = \xi_1 + \cdots + \xi_x$. If the number of ads x is large and audiences are i.i.d., the central limit theorem insures that $\xi^{(x)} \sim N(x\mu, x\sigma^2)$. However, audiences for a given show are usually highly correlated, especially within a narrow time-frame. Assuming they are exchangeable (i.e. their joint distribution is not affected by permutations of the individual random variables), jointly normally distributed with mean μ, variance σ^2 and correlation ρ, we obtain that $\xi^{(x)} \sim N(x\mu, (1-\rho+x\rho)x\sigma^2)$. Hence, denoting $Z \sim N(0,1)$, we have:

$$\Psi(x,\xi) = \xi^{(x)} = x\mu + \sigma\sqrt{x(1-\rho+x\rho)}\, Z = x\left(\mu + (\xi-\mu)\sqrt{\frac{1-\rho}{x} + \rho}\right),$$

which is increasing concave in x. Because x is large and ρ is high, ignoring second order terms, we obtain $\Psi(x,\xi) \simeq x\left(\mu + \sqrt{\rho}(\xi-\mu)\right) = x\zeta$, where $\zeta = \sqrt{\rho}\xi + (1-\sqrt{\rho})\mu$. Hence we can approximate Ψ by a multiplicative model corresponding to a modified intensity ζ with additional point-mass at the mean. In particular, if audiences are perfectly correlated $\rho = 1$, we obtain $\xi^{(x)} \sim N(x\mu, x^2\sigma^2)$, which has the same distribution as the multiplicative model $x\xi$. $\square$

2.2 Revenue and Business Model, Terms and Decisions

This paper focuses on (short term) revenue maximization from selling limited advertising space (30 second commercials during breaks in programs, a.k.a. airtime inventory) to upfront and scatter markets. We next describe the business and revenue models governing these two markets. For further details regarding business processes see Bollapragada et al. (2002).

Upfront Market. Following the announcement of program schedules and prices in May, clients contact the BC with upfront market requests to purchase advertising space in bulk, for the entire season. A typical request consists of a budget B for the entire year, and a negotiated CPM (cost per thousand viewers) C.[9] This translates into a target

[8] Such models are further investigated in Section 3.4.

[9] For example, in 2005, US TV advertising budget for P&G, the largest US advertiser, was $2.5 Billion, whereas Toyota's was half a billion. American prime time CPM averaged between $20-$35 for TV, and $8-$10 for cable.

performance $N = B/C$, measured by the total audience of the campaign, in the client's desired demographic (e.g. men between 18 and 49 years old). Audience is unknown ex-ante, but provided ex-post by media rating agencies. Additionally, the MA may specify performance targets for specific periods (quarterly or weekly weighings) and programs (e.g. show mix, prime time targets).

In response to the sales request, the BC provides a sales plan or proposal, consisting of a list of x commercials to be aired, by show and airdate; the specific break location is decided later, close to broadcasting time.[10] Of course, this allocation cannot exceed the advertising capacity, denoted by Q (maximum number of 30 second slots available in a given period). An important quantity throughout the paper is the *GRP allocation*, defined by $w = N/x$. This corresponds to the ratings (i.e. realized audience level $\hat{\xi}$) for which the performance target N is exactly met by allocating x units of advertising space. Good rating forecasts are crucial for sales plan generation, and airtime capacity management.

Because upfront contracts offer performance guarantees, the BC bears the risk of audience uncertainty. If the plan exceeds contracted performance, i.e. ratings exceed GRP allocation, the BC receives no additional payment. On the other hand, the BC is penalized (e.g. by providing make-goods) if at the end of a season a plan has under-performed (e.g. if large ratings were committed upfront for a show that turns out to be a miss). Our models account for unmet performance targets via a constant unit penalty b, which amounts to a total penalty cost of: $b \cdot (N - N_T)^+$ where N_T is the total number of impressions delivered by the end of the season. Penalties are not contractual. They are a mechanism to model and control under-performance, and have been previously used as such in the literature (see Bollapragada et al. 2002, Bollapragada and Mallik 2007). Under-performance penalties are a natural interpretation of strategic service constraints, whereby the BC commits to satisfy (aka 'steward') client requests within a given probability, or expected fraction. Uniformly high strategic service levels (usually around 95%) accurately represent the current industry approach to satisfying client contracts, according to our discussions with practitioners (see also Bollapragada et al. 2002). Service constrained models are proved equivalent to penalty models in Section 6.

In summary, the direct expected profit from an upfront market client is the result of the client's budget B minus the expected penalty cost of failing to meet the audience target $b \cdot (N - N_T)^+$, realized at the end of the season. The model so far does not account for the

[10] For concrete examples of a plan request and proposal, see Figures 3 and 4 in Bollapragada et al. (2002).

opportunity cost of allocating capacity to the scatter market, discussed next.

Scatter Market. As opposed to upfront contracts which are audience/CPM based, scatter market pricing is per advertising slot, with no audience guarantee. Broadcasting companies typically have to commit scatter prices in advance.[11] On top of these baseline prices, the BC applies discounts based on buyer characteristics (bundle, volume, loyalty) and capacity status. The realized scatter market profit from y available slots with random audience ξ is denoted $\Pi(y,\xi)$. This model allows for scatter market price and demand to be implicitly audience dependent, e.g. $\Pi(y,\xi) = p(\xi)S(y,\xi)$, where S is the number of ad slots sold at price $p(\xi)$ given capacity level y. If x slots are allocated to the upfront market, the expected profit from the remaining $Q-x$ slots on scatter market is $\pi(x) = \mathbb{E}_\xi[\Pi(Q-x,\xi)]$.

All our structural results hold under a general scatter profit model $\pi(x)$ that is decreasing and concave in x, accounting for diminishing marginal returns to the scatter market. For simplicity of exposition, the first part of the paper focuses on linear models $\pi(x) = p(Q-x)$, corresponding to a constant scatter market price p per advertising slot. This allows for close form solutions and clearer insights at no loss of generality.

Decision Making Layers. The following hierarchical decision making approach reflects industry practice, and motivates our approach in this paper.

At the strategic planning phase, the BC must decide which client contracts to accept upfront. For this purpose, the firm needs to assess how much capacity is required to satisfy upfront market clients. This capacity is not committed to the client upfront; it serves only for (strategic) capacity planning purposes at the upfront stage.

At the operational level, the decisions of how much initial capacity to commit to the upfront clients/market are made, with the specific provision of subsequent make-goods allocation during the scatter market. During the broadcast season, the BC periodically decides how many additional make-goods to allocate to upfront clients (vs. the scatter market), given the current performance of a campaign.

We begin by investigating aggregate capacity planning and allocation decisions under audience uncertainty, for the upfront market as a whole, and then extend the results to handle multiple upfront clients.

[11] Such list prices or rate cards are typically public, see for example `http://www.ftv-publicite.fr`

3. Aggregate Planning with Uncertain Audience

This section derives insights from a simple, static aggregate model of upfront capacity planning under audience uncertainty; this is a risk neutral version of the model proposed by Bollapragada and Mallik (2007). This model turns out to be an important milestone for our multi-period and multi-client analysis in subsequent sections. We first show how the total capacity requirement for the upfront market can be expressed as a critical fractile, in the spirit of the classical newsvendor model, and compare this to the allocation under a deterministic audience model. This approach allows us to further analyze the impact of audience uncertainty on profits and managerial decisions. Some of our results challenge conventional wisdom and intuitive practices. In particular, we show that (stochastically) larger audience, or lower audience variability, generate higher profits but do not necessarily lead to lower allocations.

3.1 Aggregate Upfront Planning.

An important question faced by media revenue managers is how much capacity should be allocated to upfront vs. scatter markets. In particular, once a given budget and audience target have been contracted upfront, what is the corresponding capacity requirement for the entire broadcasting season (see e.g. Alvarado 2007)? Central to our derivations is a simple aggregate upfront planning model, which answers the latter question by collapsing all upfront client targets into one cumulative upfront market audience target N. The total cost of allocating x capacity units with uncertain audience ξ to the upfront market is given by:

$$c^*(N) = \min_{0 \le x \le Q} px + b\mathbb{E}[N - x\xi]^+. \tag{1}$$

The main trade-off captured by this model is between the opportunity cost p of allocating a slot to the spot market and the penalty cost of not meeting the performance target N of upfront market. In this section only, we focus on constant scatter market prices p, but our results extend for non-linear scatter market profit functions. Recall that F is the distribution of audience uncertainty ξ (as described in Section 2.1), and denote its left tail expectation by:

$$G(u) = \mathbb{E}[\xi I_{\xi \le u}] = \mathbb{E}[\xi; \xi \le u]. \tag{2}$$

This allows to write the cost objective in (1), i.e. the cost of allocating x units of capacity to the upfront, as:

$$c(N,x) = px + b\Big[NF(N/x) - xG(N/x)\Big]. \tag{3}$$

The next result follows from the first order conditions, by convexity of the cost function and monotonicity of G. Throughout the paper, increasing/decreasing refer to weak monotonicity.

Lemma 1 *The optimal solution to Problem (1) is $x^* = \min(Q, \bar{x})$, where the unconstrained optimum $\bar{x}$ satisfies:*

$$G(N/x) = p/b. \tag{4}$$

Furthermore, x^* is piecewise linear and increasing in the target performance N, increasing in the penalty b and decreasing in the spot price p, all else equal. The optimal cost $c^*(N)$ equals $bNF(G^{-1}(p/b))$, if $p/b \geq G(N/Q)$ and $c(N,Q)$ otherwise.

The above result captures the sensitivity of the optimal profit and upfront capacity provision to the various factors affecting the decision, including target audience, penalty and scatter price. Audience uncertainty also affects the system and corresponding optimal decisions. It seems natural to expect that more popular (i.e. higher audience) shows will require a lower capacity provision. Similarly, one would expect that the introduction of TIVO, leading to an overall decrease in audience for advertisements, would call for higher capacity allocation. This is obviously true under deterministic audience models (for $\xi \equiv \mu$, $\bar{x} = N/\mu$), most common in the literature. Interestingly however, this is not necessarily the case under audience uncertainty, as illustrated by the following counterexample:

Counterexample 1. Consider two shows such that audience uncertainty for the first show, ξ_1 (measured in millions of viewers) is uniformly distributed on $[1,3]$ (so average audience is $\mu_1 = 2$). Uncertainty about the audience for the second show, ξ_2 is such that, with probability $\alpha = 0.5$, ξ_2 is uniform on $[1.5, 2]$, otherwise, it is uniform on $[2,3]$; that is, the show is a miss with probability α. It is easy to see that $\xi_1 <_{st} \xi_2$, suggesting that the second show is more popular, in a stochastic sense. Consider the same penalty level $b = 10$ and scatter market $CPM = p/\mu = 5$, i.e. scatter pricing correctly adjusts for popularity effects (see Section 3.3 for details). Eq. (4) implies that the less popular show requires less capacity allocation: $x_1^* = 22 < x_2^* = 23$. $\square$

In order to better understand the impact of show popularity on capacity decisions (and, in particular, the above example), the next section introduces an alternative model of the uncertainty underlying the system. This model is akin to the standard newsvendor model, and allows us to assess the value of modeling audience uncertainty, as well as revisit sensitivity to audience uncertainty in Section 3.3.

3.2 Supply vs. Demand Uncertainty. An Inventory Formulation

In the media revenue management problem, demand, measured in audience units, is deterministic (N), whereas the value of allocated supply is uncertain ($x\xi$); this makes the problem akin to a random yield model (see e.g. Yano and Lee 1995). An alternative perspective on the problem of optimally managing media capacity emerges by translating all model ingredients, in particular demand, in terms of inventory units (i.e. 30 second advertising slots). This is particularly appealing because the BC's operational decision is necessarily in terms of advertising space. Consider the following transformation of uncertainty in our original model (1).

Model Transformation.

1. Define the random variable ζ with cdf

$$H(u) = G(u)/\mu = \mathbb{E}[\xi; \xi \geq u]/\mu. \tag{5}$$

   This distribution is well defined[12], and captures left tail variability of the audience ξ.

2. Define the random variable $\theta = N/\zeta$ with cdf

$$\mathbb{P}(\theta \leq x) = \mathbb{P}(\zeta \geq N/x) = 1 - H(N/x). \tag{6}$$

   This measures demand uncertainty in capacity units (ad slots).

3. Recast the objective in Problem (1) in terms of these random variables as follows:

$$c(N, x) = px + b\mu\mathbb{E}[\xi/\mu(N/\xi - x)^+] = px + b\mu\mathbb{E}_\zeta[N/\zeta - x]^+ = px + b\mu\mathbb{E}_\theta[\theta - x]^+. \tag{7}$$

□

By converting audience uncertainty ξ into demand uncertainty θ, measured in inventory units, the media problem is reduced to a traditional inventory model (see e.g. Porteus 2000)

[12] This is because the left tail audience expectation $G(u)$ is increasing and bounded in $[0, \mu]$. The random variable ζ has density $h(u) = u/\mu f(u)$, and the likelihood ratio of the two measures is ξ/μ.

given by equation (7), with $cost = p$ and $price = b\mu$. The corresponding critical fractile solution, equivalent to (4), is given by

$$\mathbb{P}(\theta \geq x) = \frac{p}{\mu b}, \tag{8}$$

or in terms of GRP allocation $w = N/x$,

$$\mathbb{P}(\zeta \leq w) = \frac{p}{\mu b}. \tag{9}$$

Inventory considerations suggest writing $x^* = x_0 + SS_\theta$, where SS_θ is the safety stock, and $x_0 = \mathbb{E}\theta = \mathbb{E}\big[N/\xi \cdot \xi/\mu\big] = N/\mu$ is the optimal allocation corresponding to a deterministic audience $\xi \equiv \mu$. Practical considerations suggest that penalties much exceed scatter market CPM, $b \gg p/\mu$, so the right hand side in (8) is positive, and close to 1. This indicates a positive safety stock.[13] We conclude that *a deterministic model prediction would typically underestimate the upfront allocation* (i.e. overestimate GRP allocation) required to hedge for audience uncertainty, i.e. usually $x^* \geq x_0 = N/\mu$.

Example 2. For a binomial audience distribution $\xi \sim Bin(M, q)$, we have $\mu = Mq$ and

$$G(u) = \sum_{k=0}^{u} k \binom{M}{k} q^k (1-q)^{M-k} = Mq \sum_{k=1}^{u} \binom{M-1}{k-1} q^{k-1}(1-q)^{M-k} = \mu \sum_{k=0}^{u-1} \binom{M-1}{k} q^k (1-q)^{M-1-k}.$$

Because the cdf of ζ is defined as $H(u) = G(u)/\mu$, we obtain that $\zeta \sim 1 + Bin(M-1, q)$. Remark that ζ has (slightly) higher mean and lower variance than ξ, and both distributions stochastically increase with M and q. In practice, the market size M is very large, so ζ can be approximated by a normal distribution with mean $1 + (M-1)q$ and variance $(M-1)q(1-q)$. The first order condition (9) amounts to inverting the beta function that computes the binomial cdf, or the normal cdf for the approximation model. The latter allows to approximate the optimal GRP allocation as follows

$$w^* = N/x^* \approx \mathbb{E}\zeta + \sigma_\zeta \alpha = 1 + (M-1)q + \alpha\sqrt{(M-1)q(1-q)}, \tag{10}$$

where α is such that $\mathbb{P}(Z \leq \alpha) = \dfrac{p}{\mu b} = \dfrac{p}{Mqb}$ and Z is a standard normal distribution. In particular, $p \ll \mu b$ implies α is negative and sufficiently small, so $w^* < Mq$, as long as market size M is sufficiently large relative to viewing probability q. Hence a positive safety stock is required, $x^* = N/w^* > N/(Mq) = N/\mu = x_0$. □

[13] Formally, the safety stock is positive ($x^* > x_0$) if and only if $b > p/G(\mu)$, i.e. $H(\mu) > p/(\mu b) \approx 0$.

The transformation proposed in this section is different from transformations of random yield models to newsvendor models proposed in the literature (see Sepheri et al. 1986, Bollapragada and Morton 1999), in that our transformed random variable remains exogenous, i.e. independent of the decision variable.

3.3 Sensitivity to Audience Uncertainty

In order to model the impact of audience uncertainty on profits and decisions, consider a "show popularity" factor z, such that scatter market price $p(z)$ is increasing in z and audience $\xi(z)$ is stochastically increasing in z with respect to first order dominance. This can be seen as a signal or available information on the current status of the audience and market. Consider the aggregate planning profit maximization version of Problem (1):

$$r^*(z) = \max_{0 \le x \le Q} r(x, z), \quad r(x, z) = (Q - x)p(z) - b\mathbb{E}[N - x\xi(z)]^+. \tag{11}$$

It follows that $r(x, z)$ and $r^*(z)$ increase with show popularity z. On the other hand, popularity does not necessarily decrease upfront allocation cost $c(x, z) = p(z)x + b\mathbb{E}[N - x\xi(z)]^+$. This is because spot market opportunity costs increase whereas penalty costs decrease.[14]

Counterexample 1 illustrated the counter-intuitive fact that a higher popularity show, in the sense of first order dominance, may actually require a higher allocation $x^*(z)$. Moreover, this cannot be attributed to inconsistent pricing; pricing is so-called consistent if scatter prices reflect popularity effects, by keeping scatter market CPM $p(z)/\mu(z)$ constant. To better understand this, consider the critical fractile solution (8). This indicates that optimal capacity allocation decreases in the popularity factor z, under a stochastically decreasing demand uncertainty measured in capacity units, $\theta(z)$ (a sufficient condition). The problem is that a stochastically increasing audience distribution $\xi(z)$ does not guarantee capacity unit demand $\theta(z)$ to decrease stochastically (as the above Counterexample 1 indicates).

One can prove that the result is true, and hence allocation decreases with popularity, for a variety of parametric classes of continuous audience distributions, such as uniforms, exponentials, gammas and lognormals, with the natural order of parameters that induces first order dominance (see e.g. Table 1.1 in Müller and Stoyan 2002). It is also true for a binomial audience distribution, in light of Example 2. In addition, the next result presents

[14]Gupta and Cooper (2005) investigate the stochastic monotonicity of the profit function in response to changes in the yield distribution in a random yield context. Their results are different from ours, due in particular to the absence of holding costs in our model.

two general classes of audience distributions which also insure natural comparative statics with respect to popularity effects.

Proposition 1 *Let ξ_0 be a positive random variable with fixed distribution (independent of z), and $\mu(z)$ an increasing function of z. The following alternative conditions insure that the optimal allocation $x^*(z)$ is decreasing in the popularity factor z.*

(a) Audience is multiplicative $\xi(z) = \mu(z)\,\xi_0$, and scatter market CPM, $p(z)/\mu(z)$ is constant (or does not decrease with z).

(b) Audience is additive $\xi(z) = \mu(z) + \xi_0$, where ξ_0 is unimodal with mean 0 and non-negative mode. Moreover, penalties are sufficiently large to insure a positive safety stock, i.e. $b > p(z)/G(\mu(z))$.

Audience Variability. The factor z in (11) can alternatively be defined as a measure of audience variability, instead of show popularity. Similar arguments show that revenues decrease with variability in audience, i.e. a show with more variable audience, as measured by the concave or increasing concave order,[15] yields stochastically lower profits. Revenue monotonicity holds for these orders, provided that $p(z)$ is also concave in z. Again, it may appear surprising that higher audience variability does not necessarily require larger allocation, as the following example shows:

Counterexample 2. Consider Example 2, where $\xi(z) \sim Bin(M(z), q(z))$, with decreasing viewing probability $q(z)$ but a constant expected audience $M(z)q(z) \equiv \mu$, so z increases the variance of the audience distribution ξ. Eq. (10) can be written as $w^*(z) \approx 1 + \mu - q(z) + \alpha\sqrt{(\mu - q(z))(1 - q(z))}$. Assuming that $p(z) \equiv p$ (constant scatter market CPM), Example 3 shows that indeed $\alpha(z) \equiv \alpha$. Hence GRP allocation $w^*(z)$ is increasing in z, i.e. the optimal allocation $x^*(z)$ decreases with higher audience variance. $\square$

In general, we cannot say anything about monotonicity of the optimal allocation with respect to the standard deviation of the audience distribution. This is essentially because variance is not an appropriate measure of down-side risk, which is the relevant audience risk in the media problem.

[15] These are standard variability orders, defined as follows: X dominates Y in the (increasing) concave order if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all u (increasing) concave such that the expectation exists (see Müller and Stoyan 2002 p.16).

3.4 Multi-Stage Upfront Planning

During the upfront market planning stage, in addition to planning the total capacity requirement for the upfront market for the entire season (as determined in Section 3.1), broadcasters need to plan the distribution of this capacity across periods, e.g. quarters or weeks. This section briefly extends our basic aggregate upfront capacity planning model (1) to address such issues. Let $\boldsymbol{\xi}$, $\mathbf{p}$, $\mathbf{Q}$ denote the vectors of audiences, prices and capacities for each of K periods, and $\mathbf{x}$ the corresponding allocations. Throughout the paper, bold characters are reserved for vectors.

When a single cumulative target N is demanded for multiple air-dates with varying audiences, we obtain the following static aggregate multi-stage capacity planning problem:

$$(CT) \quad \min_{\mathbf{x}\in[\mathbf{0},\mathbf{Q}]} \quad \mathbf{p}'\mathbf{x} + b\mathbb{E}[N - \mathbf{x}'\boldsymbol{\xi}]^{+} \tag{12}$$

Upfront clients may require different performance targets in each period, $\mathbf{N}$, also known as weekly/quarterly weighings.[16] The corresponding static multi-stage capacity planning problem with multiple (periodic) targets N_i, can be formulated as:

$$(MT) \quad \min_{\mathbf{x}\in[\mathbf{0},\mathbf{Q}]} \quad \mathbf{p}'\mathbf{x} + b\sum_{i\in K}\mathbb{E}[N_i - x_i\xi_i]^{+}. \tag{13}$$

Proposition 2 *Assume that audience distributions ξ_i are exchangeable, scatter prices are uniform $p_i \equiv p$, and ad space capacity is the same $Q_i \equiv Q$ across periods.*

(a) The cumulative target problem (CT) reduces to solving an aggregate model (A) with audience $\xi = (\sum_{i=1}^{k}\xi_i)/k$, and allocating this capacity equally among periods.

(b) The multi-target problem (MT) reduces to solving an aggregate model (A) with cumulative target $N = \sum_i N_i$, and allocating this capacity among periods in proportion to the corresponding targets N_i.

The above result shows that, if audience is homogeneous (but not necessarily independent) across air-dates, and consistently priced on scatter (uniform scatter market CPM p/μ), then capacity should be allocated so as to balance GRP allocation across periods. Incidentally, uniform allocation is standard industry practice, supported by clients' preference

[16] See e.g. Table 10.7 in Talluri van Ryzin 2004 for an example of quarterly weighings.

against burstiness. The result is technically appealing, as it reduces to solving one aggregate model of type (1).[17] These results also extend for general scatter market profit models.

Remember that upfront capacity planning occurs before the start of the season (during the upfront market in May), and before audience uncertainty is resolved, hence the inherently static (no feedback) nature of the above upfront planning models. The next section considers the *operational* capacity allocation decisions during the broadcast season; in this case additional allocation to the upfront market (so called make-goods) is decided recursively, by sequentially incorporating the realization of audience uncertainty.

4. Operational Decisions. Make-Goods Allocation

Our results so far focused on static models that provide BCs with decision support during the upfront market decision process. These capacity provisions are for strategic planning purposes; they are not actually committed to the client upfront. This section shows how capacity allocation is operationalized, by extending our basic aggregate capacity planning model of Section 3 to a multi-period setting. First, the BC decides how much capacity should be *committed* to clients upfront, with the provision that this allocation cannot be reduced, but can be increased by subsequently allocating make goods. Second, as the season starts, and audience ratings unfold, the BC further decides when and whether to increase initial allocation to under-performing upfront market client campaigns, by airing make-goods. This decision is continuously traded-off against immediate scatter market profit opportunities.

The goal of this section is to provide the BC with a decision support tool to dynamically monitor capacity and profits, and adjust (aggregate) upfront market allocation in the most profitable way. At any point in time, given the dedicated capacity and achieved performance for the upfront market, our results diagnose potential under-performance. In this case, our model suggests how many make-goods should be aired, and provides projections of future expected profits.

This section, and the remainder of the paper, work with a general scatter market profit model $\pi(x) = \mathbb{E}_\xi[\Pi(Q - x, \xi)]$, as described in Section 2.2.[18] For simplicity of exposition, we

[17] One can obtain similar results in terms of budget variables $v_i = p_i x_i$, if scatter market CPM distributions are homogeneous (ξ/p_i exchangeable) across air-dates, but audiences and prices may not be. This is particularly relevant when dealing with show mix constraints. We obtain that the firm should balance GRP allocation, prorated by scatter market price.

[18] All comparative statics results extend under a general performance measure $\Psi(x, \xi)$ that is increasing and concave in x (accounting for repetition wearout).

assume audiences are i.i.d over time $\xi_t \equiv \xi$, but our results extend for Markovian processes. Consistent with the broadcasting business cycle, the planning horizon T is assumed to be one year, discretized into weeks, months or quarters.

4.1 Reversible Allocation

Let x_0 denote the initial "irreversible" allocation initially committed to the upfront market. In each period $t \geq 1$, given the remaining target performance N_t, the BC decides how many additional make-goods $x_t - x_0$ to allocate to the upfront market in order to maximize scatter market profits (realized each period) net of penalties for unmet performance, calculated at the end of the horizon. The resulting profit (net of contracted upfront client budgets) is denoted $V_t(x_t, N_t)$, and its optimal value $J_t(N_t)$. For an aggregate contracted upfront market target N, this leads to a dynamic programming model, that optimizes $J_0(N) = \max_{x_0} J_0(x_0, N)$, given recursively by the following Bellman equation:

$$\begin{array}{lll} J_t(N_t) = & \max\limits_{x_0 \leq x_t \leq Q} & V_t(x_t, N_t) \\ \text{where} & V_t(x_t, N_t) = & \pi(x_t) + \mathbb{E}J_{t+1}(N_t - \xi x_t) \\ \text{and} & J_T(N_T) = & -bN_T^+. \end{array} \tag{14}$$

Because initial allocation x_0 is irreversible, it can only limit the firm's flexibility. It is easy to see that $J_0(x_0, N)$ is decreasing in x_0, hence minimizing x_0 is optimal for $J_0(N)$. This common industry practice, known as *gapping*, is practically implemented by overstating performance ratings (i.e. overselling ξ projections). Technically, this allows us to set without loss of generality $x_0 = 0$ in the corresponding Problem (14). In reality, however, excessive gapping can negatively affect client relationships on the long run.

The next result characterizes relevant structural properties of the BC's expected profit, formalizing the following statements: Expected profit increases with achieved performance, but its marginal value decreases. Moreover, the value of allocating an extra make-good decreases with achieved performance, and there is a diminishing marginal rate of substitution between current make-goods allocation and achieved performance $(N - N_t)$. Finally, the marginal value of airing an additional make-good (or lowering the performance target by one unit) is higher later in the horizon.

Lemma 2 *The value function has the following properties: (a) J_t is decreasing and concave in N_t; (b) V_t is jointly concave and has increasing differences in (x_t, N_t); (c) $V_t(x, N) =$*

$V(x, N, t)$ *has increasing differences in* (x, t); *(d)* V_t *has increasing differences in* $(x, -N)$ *and* t.

Recall that a bivariate function $g(x, y)$ has increasing (decreasing) differences in (x, y), if for all $x \le x'$, $g(x', y) - g(x, y)$ is increasing (decreasing) in y.[19] The proof (provided in the Appendix) relies on a standard comparative statics result, *Topkis' Lemma*, stating that if $g(x, y)$ has increasing (decreasing) differences then $x^*(y) = \max \operatorname{argmax}_{x \in X} g(x, y)$ is increasing (decreasing) in y (Topkis 1998).

A consequence of Topkis' Lemma and Lemma 2(c), the next result shows that, at any point in time, the better the achieved performance, the lower the make-goods allocation, all else equal. To achieve the same given level of performance, more make-goods need to be aired later in the horizon. This is because there are less opportunities to make up for under-performance in the future.

Proposition 3 *The optimal make-goods allocation* $x_t^*(N)$ *is increasing in the remaining performance target* N *at any time* t, *and decreasing in the remaining horizon* $T - t$ *for any* $N \ge 0$, *all else equal.*

In particular, there exist threshold levels $\bar{N}_t$ so that additional make-goods are aired at time t only if the remaining target exceeds the current threshold $N_t > \bar{N}_t$. These thresholds are increasing with achieved performance $N - N_t$, and over time, all else equal.

4.2 Irreversible Allocation

Depending on the client's bargaining power, in certain markets make-goods allocation can be irreversible (e.g. one additional P&G ad will be aired in *Friends* every week until the end of the season). This section such an irreversible make-goods allocation process. Technically, if x_t is the total number of slots dedicated to a client at time t, irreversible allocation means that $x_{t+1} \ge x_t$, and effective available capacity at time $t + 1$ is $Q - x_t$.[20] In practice, sellers do not necessarily resent this limited flexibility, as it often simplifies their decision making task and induces a more uniform allocation. This intuition is supported by our numerical results in Section 4.3.

[19] In two dimensions, increasing differences is equivalent to supermodularity. If X is a set of integers, monotone differences amounts to monotonicity of $g(x + 1, y) - g(x, y)$.

[20] A conceptually related stream of literature in manufacturing considers firms' capacity expansion strategies as irreversible investments (see e.g. Oksendal 2000).

The aggregate objective is $\max J_0(N)$, with the value function given by the recursive Bellman equation:[21]

$$\begin{aligned} J_t(x_t, N_t) = &\max_{x_t \le x_{t+1} \le Q} V_t(x_{t+1}, N_t) \\ \text{where} \quad V_t(x_{t+1}, N_t) = &\ \pi(x_{t+1}) + \mathbb{E} J_{t+1}(x_{t+1}, N_t - \xi x_{t+1}) \\ \text{and} \quad J_T(x_T, N_T) = &\ -b N_T^+. \end{aligned} \tag{15}$$

Given a pre-committed allocation x and a remaining target N, the optimal allocation policy is $x_t^* = x_t^*(x, N) = \text{argmax}_{x_t \in [x,Q]} V_{t-1}(x_{t+1}, N)$; the unconstrained optimum is denoted by $\bar{x}_t = \bar{x}_t(N) = \text{argmax}_y V_{t-1}(y, N)$.

Proposition 4 *All the results of Section 4.1 under reversible allocation extend to the irreversible allocation case. Moreover, the optimal irreversible allocation policy is given by* $x_t^*(x, N) = \max(x, \min(\bar{x}_t, Q))$, *which is also increasing in the previously committed allocation* x.

The result indicates that the BC should schedule no additional make goods unless the amount already committed x_t falls below a certain level, $\bar{x}_{t+1}$. Equivalently, the optimal policy is characterized by threshold levels $\bar{N}_{t+1} = N_t - \bar{x}_{t+1}\mu$, so that additional make-goods are aired only if the remaining target exceeds the current threshold. These thresholds are increasing with past allocation x, achieved performance $N - N_t$, and over time, all else equal. Section 4.3 (Figure 1) illustrates these results numerically, contrasting them with the reversible allocation case.

Such structural properties, albeit intuitive, are typically difficult to preserve in a dynamic and stochastic setting. In particular, the monotonicity of the optimal allocation over time is a tricky result in the irreversible case, complicated by the state dependence of the action set $\mathcal{S}_x = [x, Q]$.[22]

4.3 Heuristics and Numerical Results

This section provides heuristics for dynamic make-goods allocation, and analyzes their performance compared to the optimal policy. Numerical results are illustrated for an isoelastic scatter market demand function $d(p) = ap^{-\eta}$, with elasticity $\eta > 1$. This leads to a profit

[21] Again, gapping considerations suggest that initial upfront allocation should be minimized, so w.l.o.g. we set $x_0 = 0$.

[22] This excludes the use of standard structural results for dynamic programs with action independent sets, including Smith and McCardle (2002, e.g. their Proposition 5, p.806).

model $\pi(x) = ap^{1-\eta} = p_0(Q-x)^{1-1/\eta}$, where $p_0 = a^{1/\eta}$ corresponds to the profit from one unit capacity available to the scatter market. We set $p_0 = 5$ (e.g. in tens of thousand of dollars), and $\eta = 1.5$. Audience is modeled as truncated normal distribution with $\mu = 4$ and $\sigma = 2$, measured in millions of eyeballs (this would correspond to a Binomial model with market size $M = 8$ million and the viewing probability $p = .5$). Target performance N is also measured in millions (generally in the order of a few hundreds). Total ad space capacity Q is set to 30, corresponding to the number of 30 second spots in a 15 minute break. We report results for penalty levels b of 10 and 70, corresponding approximately to 70% and 90% service levels (see Figure 3 of Section 6). Penalties b have the same order of magnitude as p_0, so for an ad price of $p_0 = \$50,000$, the penalties considered are $b = \$100$, respectively \$700 per thousand eyeballs. As expected, this value is much higher than scatter market CPM, which in our case is at most $p_0/\mu = \$12.5$. Our extensive numerical studies, for a wide range of meaningful numerical values, indicate that the results presented here are representative and robust.

Figure 1 supports our theoretical findings by illustrating the behavior of the optimal allocation policy and expected profits with respect to time and remaining target, for both the reversible and irreversible models. As predicted, expected profit is decreasing and concave in the remaining target N, and decreasing with respect to remaining horizon length $T - t$. Capacity allocation x is higher, the higher the remaining target N, and the shorter the remaining horizon. Figure 1 also illustrates the suboptimality of irreversible allocation, which postpones allocation to later periods in order to avoid being "stuck" early with too high commitments to the upfront.

The practical inconvenience of solving the dynamic program to optimality (due to the curse of dimensionality) motivates us to study several intuitive and efficient heuristics for make-goods allocation, described next

The *myopic* policy reduces the visibility of the decision maker to a shorter horizon, by introducing *sub*-targets for corresponding *sub*-horizons. For clarity, we consider the extreme myopic situation where each period is a sub-horizon and the target N is initially divided uniformly across periods. The myopic policy allocates a constant amount of capacity each period, corresponding to a per period target N/T, as long as this does not exceed the actual remaining target. Concretely, in each period t, the myopic allocation x_{myopic} solves the static model (A) with target $\tilde{N}_t = \min(N/T, N_t)$, where N_t is the total actual remaining target. Alternatively, we also consider an *updated myopic* policy, *u-myopic*, which updates the target

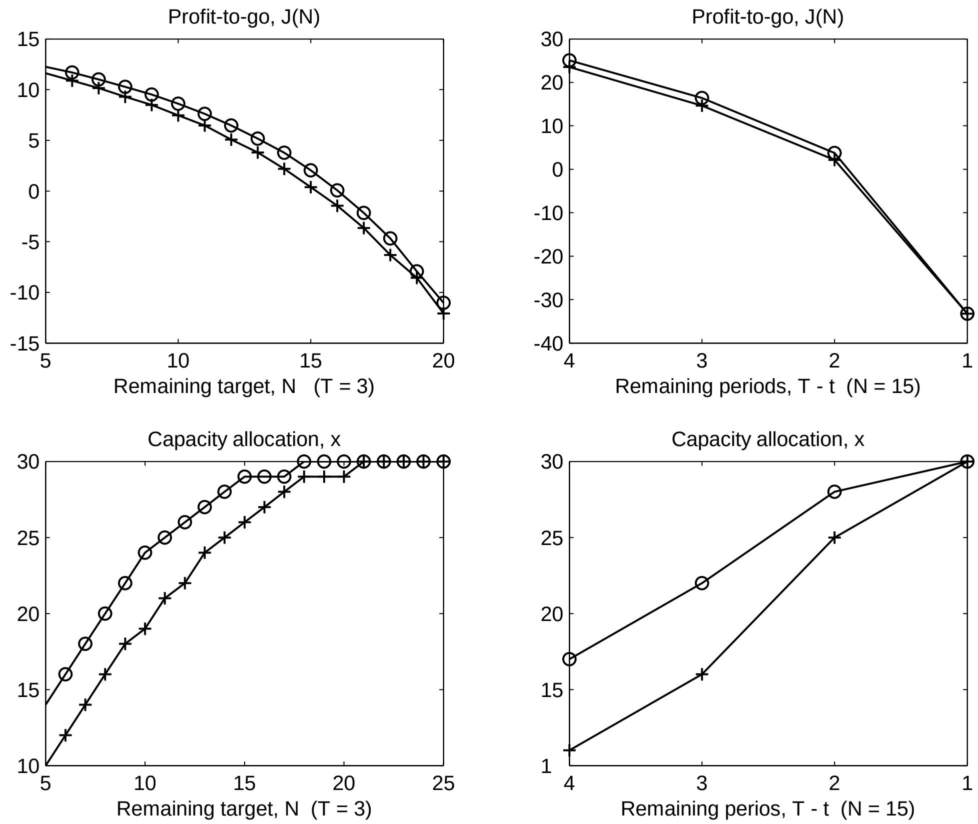

Figure 1: Optimal profit-to-go and capacity allocation for reversible (—∘—∘) and irreversible (—∗—) models as a function of remaining target N and remaining time $T-t$; $b = 10$.

for the next period uniformly, based on the actual remaining target, $\tilde{N}_t = N_t/(T-t)$.

The *static* policy provides the corresponding optimal open-loop policy, deciding allocations for each period before audience realizations are revealed. This model is analogous to the multi-stage model (CT) of Section 3.4 with non-linear scatter market profit. Proposition 2 (b) insures that the static policy allocates the same amount of make-goods x_{static} in each period. Specifically, the static solution solves (the non-linear scatter market version of) model (1), with target N and cumulative audience distribution equal to the convolution $\xi^{(T)} = \xi_1 + \cdots + \xi_T$, where $\xi_i,\ i = 1, \ldots, T$ are i.i.d. with the same distribution as ξ. That is, x_{static} solves:

$$\max_{x \in [0,Q]} T\pi(x) - b\mathbb{E}\left[N - x\xi^{(T)}\right]^+. \tag{16}$$

By design, the static policy is implicitly also a valid heuristic for the irreversible model.

We also propose a *minimal postponement* policy, which prioritizes allocation to the upfront market, while approximating future audience by its mean. Only once the client target is met, the policy starts selling to the scatter market. Specifically, it sets $x_t = \min(Q, N_t/\mu)$, as long as $N_t > 0$. One can define an analogous *maximal postponement* policy, but this turns out to perform relatively poorly, hence is not reported. Another policy which is not reported due to poor performance (too low upfront allocation) is the *Certainty Equivalent Control* (CEC) heuristic, which approximates the profit-to-go by setting audience in all remaining periods to its expected value $\mathbb{E}\xi = \mu$ (see e.g. Bertsekas 2000).[23]

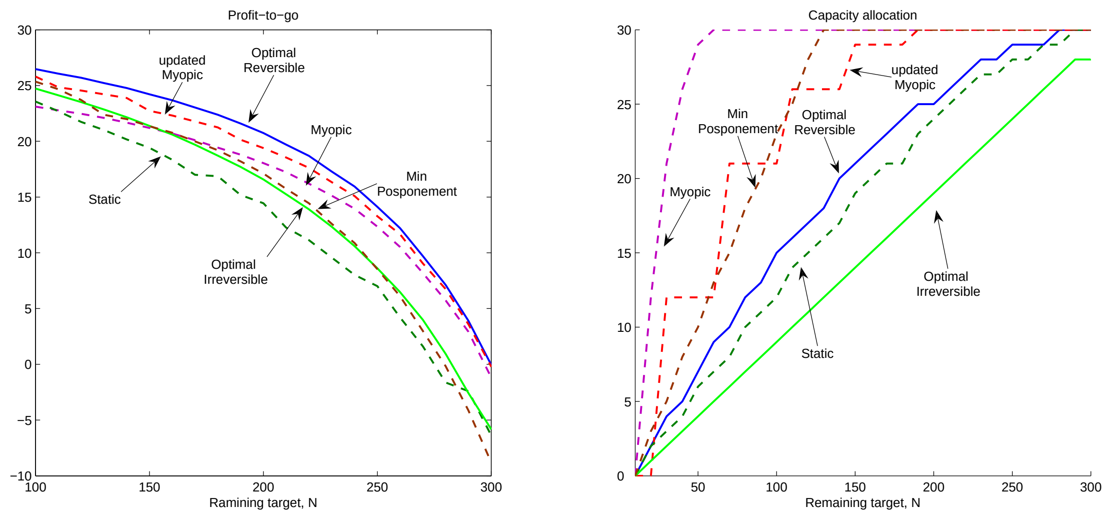

Figure 2: The value function and allocation policy at time $t = 1$ as functions of the target N, for $T = 4$ and $b = 70$.

Figure 2 compares numerically the performance of these policies in terms of total expected profits and allocation (calculated by Monte Carlo simulation) at the beginning of a four quarter horizon $T = 4$. Updated myopic policies achieved the best value function approximation, with a relative error with respect to the optimal reversible policy in the order of 0.04 (relative error is defined as the percentage optimality gap, here $(J_{rev} - J_{u-myopic})/J_{rev}$).[24] The good

[23]In each period, the CEC solves the deterministic concave problem: $\max\{\sum_{i=t}^{T} \pi(x_i) - bZ \mid Z \geq N_t - (\sum_{i=t}^{T} x_i)\mu, Z \geq 0 \text{ and } 0 \leq x_i \leq Q, \quad t \leq i \leq T\}$. Concavity of π implies that, at an optimal solution, the allocation is uniform and equal to $x_t^* = \max\{\text{argmax}(\pi(x) + xb\mu), \frac{N_t}{(T-t+1)\mu}\}$. This is the amount allocated by the CEC policy in period t. In each period, the policy re-solves the corresponding deterministic approximation, after updating the remaining target based on realized audience.

[24]Relative *profit* performance is actually much better, because it should also consider the upfront budget

performance of the updated myopic policy seems to be robust with respect to changes in the remaining target N, and improves significantly as the latter increases (independent of the value of b). For low and moderate targets N, the myopic policy is slightly outperformed by the minimum postponement. The policy ranking depicted in Figure 2 is robust for all practically relevant service levels (i.e. penalty values $b \geq 70$, corresponding to service levels of 90% or better; see Section 6, Figure 3). The updated myopic systematically outperforms all others, except for intermediate penalty values ($b \in [15, 70]$, i.e. service levels between $75 - 90\%$), where it competes with the minimum postponement. Because both myopic and minimum postponement policies favor the upfront market, their performance improves as the penalty value increases (higher service level). The static policy is the only heuristic among those considered here that is also feasible for the irreversible allocation settings, relative to which it achieves an error in the order of .09.

The second graph in Figure 2 shows the corresponding allocation of each heuristic at the beginning of a four period horizon. Interestingly, the static allocation policy comes closest to the optimal policy, but its corresponding revenue is outperformed by both minimal postponement and myopic policies; this is due to the asymmetry of the value function. Our numerical results suggest that allocation ordering depicted in Figure 2 is consistent, that is $x_{irrev} \leq x_{static} \leq x_{rev} \leq x_{u-myopic} \leq x_{myopic}$, and $x_{rev} \leq x_{MinPost}$, for all practical levels of b.[25] We provide an intuitive argument for this. Because the irreversible policy can (only) increase initial allocation in the future, it is expected to set a lower initial allocation than the static one. Relative to the reversible regime, the static policy does not have the opportunity to decrease make-goods allocation over time, so it sets a lower first stage allocation (to avoid unrewarded over-performance). Minimum postponement and myopic policies neglect the opportunity of correcting performance in future allocations, hence over-allocate relative to the optimal reversible policy. The updated myopic policy is more conservative than the myopic because its target is lowered upon audience realization. By design, the minimum postponement policy starts by allocating maximum capacity to the upfront (based on an average audience) $x_{MinPost} = \min(Q, N/\mu)$, which is expected to exceed the optimal reversible allocation. We also often observe numerically that $x_{MinPost} \leq x_{myopic}$, which is obvious for

B, omitted by our value function calculations. To give an idea of the magnitude of B, in 2005, the six biggest networks in the US alone collected over $9 billion in upfront revenues, that is 1/7 of the $63 billion TV advertising market, according to TNS. So, accounting for B, relative profit performances would be at most $6/7 = 86\%$ of reported values.

[25]For impractically low values of b, we observe that $x_{rev} \leq x_{u-myopic}$.

reasonably large targets that exceed average one-period ratings capacity, $N > Q\mu$.

We conclude that the updated-myopic policies, based on the simple static upfront allocation model (1), presented in Section 3 provide a surprisingly good approximation for the complex dynamic make-goods optimization problem. [26] In addition, intuitive minimal postponement and static policies also provide good approximations for reversible, respectively irreversible regimes.

5. Multiple Clients

Our models so far have considered aggregate capacity planning and allocation decisions for the upfront market, where the requirements of all upfront clients were cumulated into one single aggregate upfront market target. This section extends our aggregate planning models to a multi-client setting. Our results prescribe individual allocations to multiple clients, both upfront and during the broadcasting season. We begin by studying the optimal capacity planning problem for a set of contracted clients during the upfront market. We then solve the dynamic make-goods allocation problem for multiple clients during the broadcasting season. These results are further used to provide recommendations as to which client contracts the BC should accept in the first place, and how thse should be prioritized.

During the upfront market, multiple clients (say k of them) approach the BC around the same time (during a couple of weeks in May in the U.S.) and each negotiates advertising plans for the entire season. Requests consist of a budget $B_i, i \in K$ and CPM (cost per thousand viewers) $C_i, i \in K$, resulting in a target performance $N_i = B_i/C_i$. We let $\mathbf{N} = (N_1, \ldots, N_k)$. During this brief upfront market period, contracts are negotiated in parallel by account executives. The latter report to a strategic planning group, who oversees the process, and considers all client requests and proposals in order to provide strategic sales guidelines. This process motivates a joint optimization model (as opposed to a dynamic, sequential approach) for simultaneous contracting and capacity planning for multiple clients.

Our main result is that upfront clients can be aggregated for capacity planning and make-goods allocation, provided that they are not differentiated in terms of service level, i.e. penalties are uniform across clients. This is precisely the assumption we make in what follows. While the restriction to a common service level may seem to pass up profit opportunities from

[26] Myopic policies have also been noted to perform well for certain random yield models (see e.g. Sepheri et al. 1986, Bollapragada and Morton 1999), but this is a less robust result, particularly for periodic review (see e.g. Rajaram and Karmakar (2002)).

service differentiation across customers, according to our discussions with industry experts,[27] this assumption is consistent with current industry practice. Clients are differentiated based on price (CPM) and volume buys, but a uniformly high service level is maintained across customers. Partly this is because of the difficulty of assuring ex ante the timing of make-goods necessary to honor specific service differentials ex post. Moreover, once service levels depart measurably from targets, buyers have increasing difficulties planning their marketing campaigns.[28] The ultimate consequence is that differentiating penalties (i.e. service levels) across customers, while technically feasible, is not consistent with industry practice or norms. We therefore assume in what follows that penalty costs are uniform among different clients ($b_i \equiv b$). This reflects a uniform service target at the operational level (see Section 6), consistent with industry practice.

5.1 Multi-Client Upfront Planning

Suppose that the BC has signed contracts stipulating performance targets $N_i, i \in K$ with a set of k clients. In order to estimate the amount of capacity $X_i, i \in K$ necessary to satisfy each client's request at minimal cost to the firm, the following multi-client capacity planning problem is solved:

$$(M) \qquad r_M^*(\mathbf{N}) = \max_{X_i \geq 0, \sum X_i \leq Q} \pi(\sum_{i \in K} X_i) - b \sum_{i \in K} \mathbb{E}[N_i - X_i \xi]^+. \tag{17}$$

Let $N = \sum_i N_i$ be the aggregate performance contracted upfront, and $X = \sum_i X_i \leq Q$, the total capacity provisioned for the upfront market for planning purposes. Consider the profit maximization version of the aggregate model (1), under general scatter profit π:

$$(A) \qquad r^*(N) = \max_{0 \leq X \leq Q} \pi(X) - b\mathbb{E}[N - X\xi]^+. \tag{18}$$

This problem has a unique solution, whose unconstrained value solves $G(N/x) = -\pi'(x)/b$.

The next result shows that the aggregate model (A) actually produces the optimal recommendation for total upfront market allocation. This is surprising, because the aggregate

[27] at JDA Software, Mereo, NBC, Prorize, Rapt, TNT and Zilliant, among others.

[28] The situation is similar to the yield problem in semi-conductor production, where the customer is shipped a fixed quantity of semi-conductors over a year, but in varying shipment sizes depending on foundry yields. Even with "make-goods" made up by varying the wafer starts to make up for past yield problems, so that the ultimate agreed quantity of semi-conductors is shipped in total over a year, customers are not keen on having wide variations in monthly shipments. They would prefer to have a very high yield level maintained so that they can plan their own production schedules in a stable fashion. In the semi-conductor context, this also aligns well with the incentives of the manufacturer to achieve high yield. We thank Paul Kleindorfer for suggesting this example and analogy.

performance target induces additional pooling effects that may not occur when penalties are incurred at the client level.

Proposition 5 *The optimal solution to Problem (M) equates GRP allocation across clients:*

$$\frac{N_i}{X_i^*} = \frac{N}{X^*}, i \in K, \text{ where } X^* = \sum_{i \in K} X_i^* \text{ and } N = \sum_{i \in K} N_i.$$

That is, the total upfront allocation X^* solves the aggregate Problem (A) with target performance N. Moreover, the optimal profit under the two problems are the same $r_M^*(\mathbf{N}) = r^*(N)$.

Proof: For any feasible multi-client allocation $\mathbf{X} = (X_1, \ldots, X_k)$, the corresponding aggregate allocation $X = \sum_{i \in K} X_i$ is feasible for model (A), and the corresponding penalty cost can not be higher in the aggregate model (because of pooling effects $\sum a_i^+ \geq (\sum a_i)^+$). Hence,

$$r_M^*(\mathbf{N}) = \max_{X_i \geq 0, \sum X_i \leq Q} \pi(\sum_{i \in K} X_i) - b \sum_{i \in K} \mathbb{E}[N_i - X_i \xi]^+ \leq \max_{0 \leq x \leq Q} \pi(x) - b\mathbb{E}[N - x\xi]^+ = r^*(N). \tag{19}$$

Consider now the optimal solution x^* of the aggregate Problem (A), and define $X_i^* = N_i x^*/N, i \in K$. This is feasible to Problem (M), and satisfies $\sum_i X_i^* = x^*$ and $\sum_{i \in K} \mathbb{E}[N_i - X_i^* \xi]^+ = \mathbb{E}[N - x^* \xi]^+$. So $\mathbf{x}^*$ achieves the same profit as the optimal aggregate model profit $r^*(N)$, hence it is optimal for model (M) (and unique by concavity of the objective). Moreover, the optimal profits of the multi-client and aggregate problems are equal. ■

The argument relies on the uniform penalty assumption, motivated by strategic service level operations. The result also extends when penalties are non-linear but convex in GRP allocation x/N, i.e. marginal penalty increases with higher percentages of unmet audience.[29]

In summary, given a set of clients and their respective requirements, the firm only needs to determine the upfront market GRP allocation, i.e. contracted audience per capacity unit $w = N/X^*$. This should be set equally across clients i.e. $N_i/X_i^* = w$. In particular, for linear scatter profit $\pi(x) = p(Q - x)$, client allocation can be determined in closed form because $X^* = \min(Q, N/G^{-1}(p/b))$. Bollapragada and Mallik (2007) use GRP allocation as a single decision variable in an aggregate upfront market model; our results in this section validate their approach.

[29] That is, the linear penalty $b \cdot (N - x\xi)^+ = Nb \cdot (1 - \frac{x}{N}\xi)^+$ is replaced by $Ng(1 - \frac{x}{N}\xi)$, where the marginal penalty $g(\cdot)$ is convex increasing in the percentage of unmet audience $\frac{N - x\xi}{N}$.

5.2 Make-Goods Allocation for Multiple Clients

This section investigates the airtime capacity planning problem for multiple clients during the broadcasting season, extending the aggregation result to operational make-goods allocation decisions. Let $\mathbf{x}_0$ denote the initial "irreversible" allocation committed to clients upfront. In each period $t \geq 1$, given the remaining client performance targets $\mathbf{N}_t$, the BC decides how many additional make-goods $\mathbf{x}_t - \mathbf{x}_0$ to allocate to each upfront advertiser in order to maximize scatter market profits (realized each period) net of penalties for unmet performance, calculated at the end of the horizon. The resulting profit (net of contracted upfront client budgets $\mathbf{B}$) is denoted $V_t^M(\mathbf{x}_t, \mathbf{N}_t)$, and its optimal value $J_t^M(\mathbf{N}_t)$. For a contracted target $\mathbf{N}$, this leads to a dynamic programming model, that optimizes $J_0^M(\mathbf{N}) = \max_{\mathbf{x}_0} J_0^M(\mathbf{x}_0, \mathbf{N})$, given recursively by the following Bellman equation:

$$\begin{aligned} J_t^M(\mathbf{N}_t) = & \max_{\mathbf{x}_0 \leq \mathbf{x}_t; \mathbf{1}'\mathbf{x}_t \leq Q} & V_t^M(\mathbf{x}_t, \mathbf{N}_t) \\ \text{where} \quad & V_t^M(\mathbf{x}_t, \mathbf{N}_t) = & \pi(\mathbf{1}'\mathbf{x}_t) + \mathbb{E} J_{t+1}^M(\mathbf{N}_t - \xi \mathbf{x}_t) \\ \text{and} \quad & J_T^M(\mathbf{N}_T) = & -b\mathbf{1}'\mathbf{N}_T^+, \end{aligned} \tag{20}$$

where $\mathbf{1}$ denotes the vector of ones.

Our next result shows that the multi-client dynamic make-goods allocation model (20) reduces to solving the corresponding aggregate model (14), resulting in a dynamic program with a one-dimensional state variable. This further allows us to easily characterize the structure of the optimal profit and make-goods allocation based on the results of Section 4.

Proposition 6 *The dynamic-make goods multi-client allocation Problem (20) is equivalent to the corresponding aggregate Problem (14) with target performance* $N_t = \mathbf{1}'\mathbf{N}_t^+$*, in that: (1) the value functions of the two problems are equal,* $J_t^M(\mathbf{N}_t) = J_t(N_t)$*, and (2) the total optimal make-goods allocation under model (20) equals the optimal make-goods allocation under the aggregate model (14),* $\mathbf{1}'\mathbf{x}_t^* = x_t^*$*. Furthermore, in each period, make-goods* $x_{i,t}^*$ *are optimally allocated to clients in proportion to their remaining performance targets* $N_{i,t}^+$*, i.e. by balancing GRP allocation.*

We conclude that in order to describe the multiple client solution, it is sufficient to characterize the optimal revenue and allocation corresponding to the aggregate model.

5.3 Multi-Client Contracting and Priority Heuristic

In this section, we step back to investigate which upfront clients the BC should accept, among a given set of client requests, consisting of a target performance N_i and budget B_i, at the negotiated CPM rate $C_i = B_i/N_i$. We provide a simple heuristic for upfront market contract negotiation and client prioritization, which is consistent with industry practice. Our model is limited to the firm's short term profit maximization problem, and thus ignores the impact of current decisions on customer retention in the context of the firm's long term profit maximization problem (see Section 6 for a discussion).

We consider a setup where the BC may offer partial fulfilment of client requests, at the negotiated CPM level C_i. Let $y_i \in Y = [0, 1]$ be the satisfied fraction of client i's demand,[30] and x_i the amount of capacity provisioned for client i. Accepting client i increases revenues by $B_i y_i$, less potential penalties for unmet performance, but generates scatter market opportunity cost. Consider the upfront planning profit maximization model:

$$P = \max_{y_i \in Y} r^*_M(\mathbf{N} \circ \mathbf{y}) + \mathbf{B}'\mathbf{y}, \tag{21}$$

where $\mathbf{N} \circ \mathbf{y} = (N_1 y_1, \ldots, N_k y_k)$ denotes the componentwise product vector of $\mathbf{N}$ and $\mathbf{y}$. The first term stems from the multi-client upfront planning model (M) with contracted client performance targets $N_i y_i$. Proposition 5 reduces this to an aggregate model (A) with cumulative target $\mathbf{N}'\mathbf{y}$. So $r^*_M(\mathbf{N} \circ \mathbf{y}) = r^*(\mathbf{N}'\mathbf{y})$, and the problem can be restated as:

$$\begin{aligned} P &= \max\nolimits_{y_i \in Y} r^*(\mathbf{N}'\mathbf{y}) + \mathbf{B}'\mathbf{y} \\ &= \max\nolimits_N \left\{ r^*(N) + \max_{\mathbf{N}'\mathbf{y} = N; y_i \in Y} \mathbf{B}'\mathbf{y} \right\}. \end{aligned} \tag{22}$$

For any given contracted audience $N = \mathbf{N}'\mathbf{y}$, the inner problem is a knapsack model, whose optimal fractional solution amounts to serving clients in decreasing order of their marginal profitability, or CPM, $C_i = B_i/N_i$.[31] Hence this will also be true for the optimal solution of Problem (22). Client demands should be served as long as they are profitable (i.e. CPM exceeds marginal penalty, so under linear scatter profit $C_i \geq bF(G^{-1}(p/b)))$ and capacity is available. Our results have the following implications for client contracting:

Proposition 7 *Client contracts should be accepted sequentially in decreasing order of marginal revenue, or CPM $C_i = B_i/N_i$, as long as they are profitable and capacity is available.*

[30] Restricting $Y = \{0, 1\}$ corresponds to an all-or-nothing contracting model.

[31] Under all-or-nothing contracting, this is known as the greedy heuristic for the corresponding $0-1$ knapsack model.

These insights remain valid when modeling the recourse provided by the opportunity of allocating additional capacity (make-goods) to multiple client during the broadcasting season. In this case, model (21) is replaced by:

$$\max_{y_i \in Y} \mathbf{B}'\mathbf{y} + \max_{\mathbf{x}_0} J_0^M(\mathbf{x}_0, \mathbf{N} \circ \mathbf{y}). \tag{23}$$

The last term above is given by the dynamic programming model (20), and gapping consideration (see Section 4) imply $\mathbf{x}_0 = \mathbf{0}$. The above argument for Problem (21) applies, with $r_M^*(\mathbf{N} \circ \mathbf{y})$ replaced by $J_0^M(\mathbf{0}, \mathbf{N} \circ \mathbf{y})$, using the make-goods aggregation result of Proposition 6. Our results in this section can be summarized as follows:

Optimal procedure for client contracting, planning and make-goods allocation:

1. Client Contracting: Accept client contracts in decreasing CPM order, as long as they are profitable and capacity is available.

2. Upfront Capacity Planning:

   (a) Given the fractions of contracted client requests $\mathbf{y}$, solve aggregate model (A) with cumulative performance target $\mathbf{N}'\mathbf{y}$ to obtain the total capacity requirement for the upfront market X^*.

   (b) Provision capacity requirements for each client by dividing X^* in proportion to contracted performance targets $y_i N_i$, i.e. by equating average GRP allocation $N_i/X_i^* = N/X^*$.

3. Sequential Capacity Allocation:

   (a) Allocate minimal initial capacity to clients upfront (gapping);

   (b) Obtain total make-goods allocation by solving the (one-dimensional state) dynamic programming Problem (14) with aggregate target performance $N_t = \mathbf{1}'\mathbf{N}_t^+$; its value function represents the expected year-end profit (net of client budgets).

   (c) Allocate make-goods $x_{i,t}^*$ to clients in proportion to their remaining performance targets, $N_{i,t}^+$, i.e. by balancing GRP allocation.

□

The proposed sequential approach is practically appealing due to its simplicity and transparency, relative to the overall complexity of the media problem. Moreover, our results suggest how to implement a central planning procedure based on aggregate upfront and scatter markets, which further provides consistent capacity planning and make-goods allocation guidelines for account executives.

6. Service Constrained Models

Our models in this paper used a linear penalty to account for unmet performance under uncertain audience. In practice, under-performance penalties are implicit, as they must indirectly account for retention factors such as loss of goodwill, potential loss of client etc. This section briefly highlights how key results obtained under the penalty model are preserved under alternative service constrained models. According to our discussions with practitioners, strategic (uniform) service constraints are an accurate representation of current business practices.

Type 1 Service. Suppose that the BC wants to insure that the probability of meeting the contracted target performance exceeds a preset type 1 service level $1-\epsilon$. The following aggregate service model obtains:

$$(S) \qquad \begin{array}{ll} \max\limits_{0 \le x \le Q} & \pi(x) \\ \text{s.t.} & \mathbb{P}(x\xi \le N) \le \epsilon. \end{array} \tag{24}$$

This problem is feasible provided that $N/F^{-1}(\epsilon) \le Q$. The resulting allocation is $x^* = N/F^{-1}(\epsilon)$, i.e. the optimal GRP allocation satisfies the critical fractile condition $w^* = F^{-1}(\epsilon)$. Note, this is actually independent of the scatter market profit model. In particular, for linear scatter profit, the aggregate service model (S) and penalty model (A) are consistent, i.e. result in the same optimal allocation x^*, if and only if $G^{-1}(p/b) = F^{-1}(\epsilon)$.[32] Figure 3 illustrates this correspondence between service and penalty levels.

By imposing a uniform strategic service level $1-\epsilon$ across clients one can formulate the service-counterpart of the multi-client contracting and capacity planning penalty model (21), and an aggregation result analogous to Proposition 7 holds. The one-stage service constrained model (S) can also be extended to handle dynamic make-goods allocation (the Bellman recursions only differ in the terminal value function). This yields the same structural results as those obtained for penalty models in Section 4.

Type 2 Service. All our insights remain valid under type 2 service models, which impose a rigid (strategic) bound $\delta > 0$ on the expected fraction of unmet performance:

[32] In this case, feasibility of the service model implies that the unconstrained solution is optimal for the penalty model.

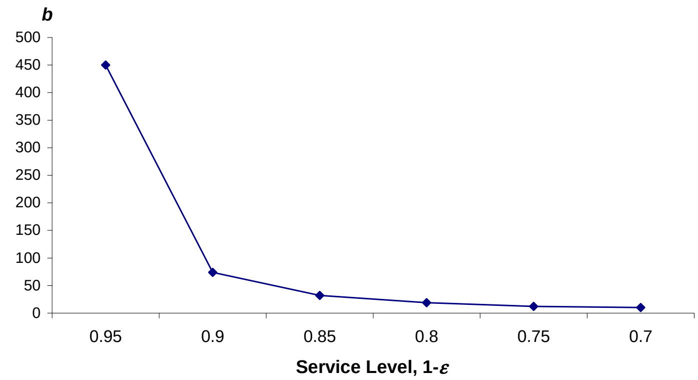

Figure 3: Penalty b vs. service level $1-\epsilon$; ξ is truncated $N(4,2)$ and $p=5$.

$\mathbb{E}[N_i - x_i\xi]^+ \leq \delta N_i$. The results of penalty and type 1 service models are reproduced with optimal GRP allocation $w^* = L^{-1}(\delta)$, where $L(w) = \mathbb{E}[1-\xi/w]^+ = F(w) - G(w)/w$. Such models provide a realistic formalization of the approach adopted in practice (see e.g. Bollapragada et al. 2002, p.54).

Service Differentiation and Long Term Profits. Our multi-client models assume a strategic service level $1-\epsilon$ for all clients, captured by the uniform penalty b for unmet performance. Differentiating service levels $1-\epsilon_i$ across clients at the strategic contracting phase results in the recommendation to prioritize clients according to service-adjusted CPM $C_i F^{-1}(\epsilon_i)$. In particular, among clients of similar immediate profitability C_i, the BC will favor easier to fulfil contracts (hence potentially less important clients). This motivates the need for more complex models of service differentiation, that (optimally) set different service levels $1-\epsilon_i$ for each client, while capturing the endogenous impact of service level on customer lifetime value $L_i(\epsilon_i)$. Such recourse models are relevant, but beyond the scope of the current paper. Aflaki and Popescu (2007) model the endogenous problem of a service provider when demand responds to the cumulative service experience.

7. Conclusions

This paper provides new, stylized models for media revenue management in the presence of audience uncertainty. Our models provide a tool for valuing airtime capacity under audience uncertainty. Our results offer several levels of decision support for broadcasting companies to optimize revenues from advertising space. We propose a simple rule for accepting upfront client contracts and estimating their overall capacity requirements during the upfront market. Operationally, we indicate how much capacity should be initially committed to upfront clients before the start of the season, with the provision that additional make-goods can be aired subsequently. Finally, we make recommendations for dynamic make-goods allocation during the scatter market.

The following procedure for upfront client contracting, capacity planning and make-goods allocation is optimal under strategic (non-differentiated) service level operations :

1. Serve clients in decreasing CPM order; plan capacity requirements for upfront clients in proportion to requested audience (total capacity requirement for the upfront market can be estimated using model (A)).

2. Allocate minimal initial capacity to clients upfront (gapping);

3. Obtain total make-goods allocation by solving the (one-dimensional state) dynamic programming Problem (14) (or e.g. a myopic heuristic) with aggregate target performance; its value function represents the expected year-end profit (net of client budgets).

4. Allocate make-goods to clients in proportion to their remaining performance targets, i.e. by balancing GRP allocation.

Given that current industry practice is to make such decisions qualitatively, this procedure can provide systematic guidance for capacity planning in a centralized decision process, thereby simplifying the media planning task. We also provide structural results and efficient heuristics to support high-level managers and account executives in this complex decision process. The impact of audience uncertainty on capacity decisions is assessed, leading to surprising findings. Our results are robust, in that they hold under general models of audience uncertainty and scatter market profit, as well as under alternative service constrained models.

Our models can also be useful in other settings where the value of supply is uncertain, and/or the firm serves dual markets. For instance, in manufacturing with random yield, a firm with limited resources and uncertain production rate (e.g. machine or workforce reliability) honors both key accounts with long term contracts and small clients with opportunistic contracts (alternatively, our scatter market opportunity cost corresponds to direct production cost). Similarly, non-profit organizations (NPO) afford to sustain pro bono service to a mission market by often offering similar paid services to distinct markets, that share the same limited facilities (e.g. hospitals) powered by uncertain resources (e.g. volunteers). Here, the mission market corresponds to our upfront market and the paying market to our scatter market. de Vericourt and Lobo (2006) investigate revenue management for NPOs.

There are multiple facets of the media revenue management problem that this paper leaves to be explored, including pricing, scheduling and contract design. Banciu et al. (2007) investigate bundle pricing strategies for TV ads on the scatter market. A classic reference for tariff design and non-linear pricing is Wilson (1993). Allocating clients to programs based on socio-demographic audience requirements involves the solution of an appropriate assignment problem, that is beyond our scope. A static deterministic integer programming model for spot scheduling is proposed by Bollapragada and Garbiras (2004). Contract design is an interesting area, where the industry is quite sophisticated, managing a variety of flexible products, such as option-cutbacks (revised planning due to client's budget cuts) or callable products (lower rates with the option of recalling the slot for a higher paying customer). The latter, and their value for standard B2C revenue management are investigated in Gallego et al. (2004).

Finally, it is an irreversible fact that audience is moving online, where capacity is more complex, dynamic and customizable, leading to different contractual terms (e.g. pay per click vs CPM – pay per view). Araman and Fridgeirsdottir (2006) investigate a queuing model for online advertising revenue management.

Acknowledgments

The authors would like to thank Paul Kleindorfer for insightful comments and suggestions. The following industry experts are gratefully acknowledged for insightful exchanges about practical aspects of media planning and revenue management: Srinivas Bollapragada, Michael Chigirinskiy, Monique Garnier, Ahmet Kuyumcu, Andreea Popescu, Antoine Taconnet, Abraham Toledano, Loren Williams and Graham Young.

Appendix (Proofs)

8. Proofs for Section 3

Proof of Proposition 1 (Audience sensitivity)

(a) The first order condition can be restated as $H_{\zeta(z)}(N/x) - bC(z) = 0$. The assumptions imply that the left hand side is decreasing in z, hence so is $x^*(z)$.

(b) Lemma 1 suggests that the optimal (unconstrained) allocation solves:

$$G(N/x, z) = p(z)/b, \text{ where } G(N/x, z) = \mathbb{E}[\xi(z); \xi(z) \leq N/x] = \int_0^{N/x} y f(y, z) dy. \tag{25}$$

Here $f(y, z) = f_0(y - \mu(z))$ denotes the density function of $\xi(z)$. This is decreasing in z for $y \leq \mu(z)$, because f_0 is unimodal with non-negative mode. Thus, $G(N/x, z)$ is decreasing in z for $x \geq N/\mu(z)$.

The condition on b implies $x^*(z) > N/\mu(z)$. Monotonicity of $x^*(z)$ follows because $p(z)$ is increasing in z and $G(N/x, z)$ is decreasing x, and decreasing in z for $x \geq N/\mu(z)$. ■

Proof of Proposition 2 (Multi-stage planning)

Part (b) related to model (MT) follows the same lines as Proposition 5, proved in the text. Our proof of part (a) relies on the following lemma (see Theorem 8.2.3 in Müller and Stoyan 2002), concerning the problem:

$$(U) \qquad \min_{\mathbf{1}'\mathbf{y}=q,\ \mathbf{y}\geq 0} \mathbb{E}u(\mathbf{y}'\omega). \tag{26}$$

Lemma 3 *Suppose that u is a decreasing convex function and ω_i are exchangeable. The optimal solution $\mathbf{y}^* \in \mathbb{R}^n$ to Problem (U) above satisfies $y_i^* = q/n$ for all i.*

Problem (CT) can be written as

$$(P_Q) \qquad \min_{\mathbf{x}\in[\mathbf{0},\mathbf{Q}]} p\mathbf{1}'\mathbf{x} + b\mathbb{E}[N - \mathbf{x}'\boldsymbol{\xi}] = \min_{q\geq 0} pq + bZ(q), \tag{27}$$

$$\text{where} \qquad Z(q) = \min_{\mathbf{x}\in[\mathbf{0},\mathbf{Q}]} \ \mathbb{E}[N - \mathbf{x}'\boldsymbol{\xi}]^+ \quad \mathbf{1}'\mathbf{x} = q. \tag{28}$$

We prove that for any q, the solution $\mathbf{x}^*(q)$ of Problem (28) has the desired structure, which implies the result for the solution of the original problem $\mathbf{x}^*(q^*) = \mathbf{x}^*$.

Problem (28) without capacity constraints has the structure required by Lemma 3. So, in particular, if the optimal solution $\bar{\mathbf{x}}$ of this problem satisfies $\bar{\mathbf{x}} \leq \mathbf{Q}$, then it is optimal, so the result of Lemma 3 holds. Otherwise, consider an optimal solution $\mathbf{x}^* = \mathbf{x}^*(q)$ to Problem (CT). Consider the *reduced problem* obtained by fixing the variables that are equal to their upper boundary ($\mathbf{x}_I = \mathbf{x}_I^* = \mathbf{Q}_I$) in the given solution, and optimizing the objective over the remaining subset of variables $\mathbf{x}_J$, in absence of capacity constraints $\mathbf{Q}_J$. For a vector $\mathbf{y} \in \mathbb{R}^k$, we denoted its projection on the $L \subseteq \{1, \ldots, K\}$ coordinate set by $\mathbf{y}_L = (y_l)_{l \in L}$. Letting $\eta_J = N - \mathbf{x}_I^* \boldsymbol{\xi}_I$ (random) and $q_J = q - \mathbf{1}'\mathbf{Q}_I$, the reduced problem is:

$$\min_{\mathbf{v}_J \geq \mathbf{0}} \quad \mathbb{E}[\eta_J - \mathbf{x}_J' \boldsymbol{\xi}_J]^+ \tag{29}$$
$$\mathbf{1}'\mathbf{x}_J = q_J.$$

This reduced problem has the structure required by Lemma 3, hence its solution $\bar{\mathbf{x}}_J$ satisfies the corresponding properties. Moreover $\mathbf{x}^*$ given by $x_J^* = \bar{\mathbf{x}}_J, \mathbf{x}_I^* = \mathbf{Q}_I$ is optimal to Problem (CT) (see e.g. Proposition 3.4.2 in Bertsekas, 1995). This proves the desired results. ■

9. Proofs for Section 4

Proof of Lemma 2:

The terminal value function J_T is decreasing and concave in N_T. Together with concavity of π_t, this implies that V_{T-1} is jointly concave and has increasing differences. By induction we obtain that V_t is jointly concave and has increasing differences in (x, N) and J_t is decreasing concave in N_t implying (a) and (b).

Parts (c) and (d) follow by writing $V_{t-1}(x, N) - V_t(x, N) = \mathbb{E}[J_t(N - x\xi) - J_{t+1}(N - x\xi)]$. Concavity (hence decreasing differences) of J_{t+1} implies that, for any realization z of ξ, the function inside this expectation:

$$\max_{x_0 \leq y \leq Q} \pi(y) + \mathbb{E}[J_{t+1}(N - xz - y\xi) - J_{t+1}(N - xz)] \tag{30}$$

is increasing in $(N, -x)$. ■

Preliminary Technical Lemmas for Section 4.2

We present a set of general technical lemmas that are used to prove the results in this section. The generality of the presentation allows to easily extend the proofs for general performance metrics $\Psi(x,\xi)$, validating the claims made in the Introduction.[33]

Lemma 4

(a) f concave implies $g(x,y) = f(x+y)$ is concave and has decreasing differences.

(b) f concave and h increasing implies $g(x,y) = f(h(x)+y)$ has decreasing differences.

(c) f increasing concave and h concave implies $g(x,y) = f(h(x)+y)$ joint concave.

(d) If $f(x,y)$ has decreasing differences in (x,y) and is concave in y and $h(x)$ is increasing then $g(x,y) = f(x, h(x)+y)$ has decreasing differences in (x,y).

Analogous conditions for increasing differences obtain from the fact that $g(x,y)$ has increasing differences if and only if $g(-x,y)$ has decreasing differences if and only if $g(x,-y)$ has decreasing differences.

Proof:

(a) Follows from the definitions.

(b) For $x \le x'$ and $y \le y'$, concavity of f implies:

$$\begin{aligned} g(x',y') - g(x',y) &= f(y'+h(x')) - f(y+h(x')) \\ &\le f(y'+h(x)) - f(y+h(x)) = g(x,y') - g(x,y). \end{aligned}$$

(c) Denote $x_\alpha = \alpha x + (1-\alpha)x'$, $y_\alpha = \alpha y + (1-\alpha)y'$. We have

$$g(x_\alpha, y_\alpha) = f(h(x_\alpha)+y_\alpha) \ge f(\alpha(h(x)+y)+(1-\alpha)(h(x')+y')) \ge \alpha f(h(x)+y)+(1-\alpha)f(h(x')+y'),$$

where the first inequality is by concavity of h and monotonicity of f, and the second by concavity of f.

(d) For $x \le x'$ and $y \le y'$ we have:

$$\begin{aligned} g(x',y') - g(x',y) &= f(x', y'+h(x')) - f(x', y+h(x')) \\ &\le f(x', y'+h(x)) - f(x', y+h(x)) \\ &\le f(x, y'+h(x)) - f(x, y+h(x)) \\ &= g(x,y') - g(x,y), \end{aligned}$$

where the first inequality follows by concavity (hence decreasing differences) of f in the second argument, and the second part by decreasing differences of f in (x,y). ■

The next lemma requires the following definition:

[33] The proof for $\Psi(x,\xi)$ uses Lemma 4 b) and c). The induction in Proposition 4 should be conducted simultaneously for *all* increasing concave functions $\Psi(x,\xi)$, in order to achieve the last step of the proof.

Definition 1 *The set function $x \mapsto \mathcal{S}_x$, or in short $\mathcal{S}_x$, is said to be:*

(a) decreasing if $\mathcal{S}_{x'} \subseteq \mathcal{S}_x$ for $x \leq x'$; increasing if $\mathcal{S}_{x'} \subseteq \mathcal{S}_x$ for $x \geq x'$.

(b) convex if $y_i \in \mathcal{S}_{x_i}$, $i = 1, 2$ implies $\alpha y_1 + (1-\alpha)y_2 \in \mathcal{S}_{\alpha x_1 + (1-\alpha)x_2}$ for all $\alpha \in [0, 1]$.

Lemma 5 *(a) The function $g(x) = \max_{y \in \mathcal{S}_x} f(y)$ is decreasing in x if $\mathcal{S}_x$ is decreasing in x.*

(b) The function $g(x) = \max_{\mathbf{y} \in \mathcal{S}_x} f(x, y)$ is concave in x if $f(x, y)$ is jointly concave in (x, y) and $\mathcal{S}_x$ is convex in x.

(c) The function $g(x, y) = \max_{v \in [x, Q]} f(v, y)$ has decreasing (increasing) differences in (x, y) if $f(v, y)$ has decreasing (increasing) differences in (v, y) and is concave in v.

Proof: (a) Trivial

(b) Given x_1, x_2 and $\alpha \in (0, 1)$, denote $x_\alpha = \alpha x_1 + (1-\alpha)x_2$. We have

$$\begin{aligned} \alpha g(x_1) + (1-\alpha)g(x_2) &= \alpha f(x_1, y_1^*) + (1-\alpha) f(x_2, y_2^*) && (31) \\ &\leq f(x_\alpha, \alpha y_1^* + (1-\alpha)y_2^*) && (32) \\ &\leq \max_{y \in \mathcal{S}_{x_\alpha}} f(x_\alpha, y) = g(x_\alpha) && (33) \end{aligned}$$

where the first inequality holds by joint concavity of f, and the second uses the fact that $\alpha y_1^* + (1-\alpha)y_2^* \in \mathcal{S}_{x_\alpha}$ for $y_1^* \in \mathcal{S}_{x_1}$ and $y_2^* \in \mathcal{S}_{x_2}$ by convexity of the correspondence $\mathcal{S}_x$.

(c) We only prove the decreasing differences result here; the other is analogous. By concavity of f in v, we can write $\text{argmax}_{x \leq v \leq Q} f(v, y) = \max(x, \bar{x}_y)$, where $\bar{x}_y = \text{argmax}_{0 \leq v \leq Q} f(v, y)$. By Topkis' Lemma, $\bar{x}_y$ is decreasing in y.

Decreasing differences of g amounts to proving for any $x \leq x'$ and $y \leq y'$:

$$\begin{aligned} g(x', y') + g(x, y) = &\max_{x' \leq v' \leq Q} f(v', y') + \max_{x \leq v \leq Q} f(v, y) \\ = \; & f(\max(x', \bar{x}_{y'}), y') + f(\max(x, \bar{x}_y), y) \\ \leq \; & f(\max(x, \bar{x}_{y'}), y') + f(\max(x', \bar{x}_y), y) \\ = \; & \max_{x \leq v \leq Q} f(v, y') + \max_{x' \leq v' \leq Q} f(v', y) = g(x, y') + g(x', y). \end{aligned} \quad (34)$$

If $X = \max(x, \bar{x}_y) \leq \max(x', \bar{x}_{y'}) = X'$, then (34) follows by decreasing differences of f:

$$f(X', y') + f(X, y) \leq f(X, y') + f(X', y) \leq \max_{x \leq v \leq Q} f(v, y') + \max_{x' \leq v' \leq Q} f(v', y).$$

It remains to show it for $X' \leq X$. Because $\bar{x}_{y'} \leq \bar{x}_y$ (from $y' \geq y$) and $x' \geq x$, we obtain $x \leq x' \leq \bar{x}_y$. There are three cases, depending on where $\bar{x}_{y'}$ falls.

- For $x \leq x' \leq \bar{x}_{y'} \leq \bar{x}_y$, relation (34) obviously holds with equality.

- For $x \le \bar{x}_{y'} \le x' \le \bar{x}_y$, relation (34) becomes $f(x', y') + f(\bar{x}_y, y) \le f(\bar{x}_{y'}, y') + f(\bar{x}_y, y)$ which is obvious by unconstrained maximality of $\bar{x}_{y'}$ for $f(\bar{x}_{y'}, y')$.

- For $\bar{x}_{y'} \le x \le x' \le \bar{x}_y$, relation (34) becomes $f(x', y') + f(\bar{x}_y, y) \le f(x, y') + f(\bar{x}_y, y)$. Indeed $f(x', y') \le f(x, y')$ because $x' \ge x \ge \bar{x}_{y'}$, which is the unconstrained maximizer of the concave function $f(\cdot, y')$.

■

The next result extends Lemma 2 under irreversible commitment. In addition, it shows that revenue to go decreases with the number of committed make-goods, at a decreasing marginal rate. Moreover, there is a diminishing marginal rate of substitution between committed allocation and remaining target. The opportunity cost of a committed make-good is higher for lower performance targets (or the better the achieved performance). Finally, the marginal value of an additional make-good increases over time.

Lemma 6 *The value function has the following properties:(a) J_t is decreasing in N_t and in x_t; (b) J_t is jointly concave in (x_t, N_t); (c) J_t has increasing differences in (x_t, N_t) and V_t has increasing differences in (x_{t+1}, N_t); (d) $V_t(x, N)$ and $J_t(x, N)$ have increasing differences in (x, t).*

Proof: Proof: (a) Monotonicity is easily proved by induction and Lemma 5 (a). The base case is trivial, transitions are linear and the profit per stage is state-independent.

(b) We show by backward induction that J_t is jointly concave in (x_t, N_t). The base case is trivial. Suppose J_{t+1} is jointly concave, so $J_{t+1}(x_{t+1}, N_t - \xi x_{t+1})$ is jointly concave in (x_{t+1}, N_t), hence so is $V_t(x_{t+1}, N_t)$. By Lemma 5 (b), J_t is jointly concave in (x_t, N_t).

(c) We prove both statements in parallel by induction. The base case is trivial. Assume for $t+1$ and show for t. Because J_{t+1} has increasing differences, by Lemma 4 (d), we obtain that $\mathbb{E}J_{t+1}(x_{t+1}, N_t - \xi_{t+1}x_{t+1})$ has increasing differences in (x_{t+1}, N_t), hence the same holds for V_t. Finally, increasing differences and concavity of V_t implies increasing differences of J_t by Lemma 5 (c).

(d) The result follows independently from the proof of Proposition 4 below. ■

Proof of Proposition 4:

Monotonicity of the optimal policy in N and x follows from Lemma 6 (c) together with Topkis' Lemma. It remains to show monotonicity with time. We show by backwards induction the following three statements in parallel:

[J-t] $J_t(x, N - xz) - J_{t+1}(x, N - xz)$ decreasing in x for all $N, z \geq 0$.

[V-t] $V_{t-1}(x, N) - V_t(x, N)$ is decreasing in x for all $N \geq 0$.

[x-t] $x_t^*(x, N) \leq x_{t+1}^*(x, N)$ for all $N, x \geq 0$.

The base case [J-(T-1)] is obviously true. The following is true for all t:

[J-t] $\Rightarrow$ [V-t] because $V_{t-1}(x, N) - V_t(x, N) = \mathbb{E}[J_t(x, N - x\xi) - J_{t+1}(x, N - x\xi)]$.

[V-t] $\Rightarrow$ [x-t] by Topkis' Lemma.

It remains to show [x-t] $\Rightarrow$ [J-(t-1)], i.e. the following function is decreasing in x for all $N, z \geq 0$:

$$\begin{aligned} j(x) &= J_{t-1}(x, N - xz) - J_t(x, N - xz) \\ &= V_{t-1}(x_t^*(x, N - xz), N - xz) - J_t(x, N - xz) \qquad (35) \\ &= \max_{x \leq y \leq Q} \pi(y) + \mathbb{E}[J_t(y, N - xz - y\xi) - J_t(x, N - xz)]. \qquad (36) \end{aligned}$$

We have

$$J_t(x, N - xz) = \begin{cases} J_t(0, N - xz), & \text{if } x_{t+1}^*(x, N - xz) > x \quad (case\ 1) \\ V_t(x, N - xz), & \text{if } x_{t+1}^*(x, N - xz) = x \quad (case\ 2). \end{cases} \qquad (37)$$

Case 1: To show that $j(x)$ is decreasing in x, it is enough to show that the following function (inside the expectation in (36)) is decreasing in x for all $y, z, v \geq 0$:

$$\begin{aligned} g(x) &= J_t(y, N - xz - yv) - J_t(x, N - xz) = J_t(y, N - xz - yv) - J_t(0, N - xz) \\ &= [J_t(y, N - xz - yv) - J_t(y, N - xz)] + [J_t(y, N - xz) - J_t(0, N - xz)]. \qquad (38) \end{aligned}$$

Indeed, by Lemma 6, the first difference is decreasing in x by concavity of J_t in the second argument, and the second difference is decreasing in x by increasing differences of J_t.

Case 2: We have $x \leq x_t^*(x, N - xz) \leq x_{t+1}^*(x, N - xz) = x$ by the induction hypothesis [x-t]. Therefore $x_t^*(x, N - xz) = x$. This together with (37) allows to rewrite (35) as

$$j(x) = V_{t-1}(x, N - xz) - V_t(x, N - xz) = \mathbb{E}[J_t(x, N - xz - x\xi) - J_{t+1}(x, N - xz - x\xi)].$$

The right hand side is decreasing in x because the function inside the expectation $J_t(x, N - x(z + v)) - J_{t+1}(x, N - x(z + v))$ is so for any $z, v \geq 0$, by [J-t]. This concludes the proof.

10. Proofs for Section 5

Proof of Proposition 6 (Make-goods aggregation)

The proof is by induction. The base case is given by Proposition 5, so in particular $\mathbf{x}$ defined by $x_i/N_i = X^*/N$ is optimal to the multi-client model, and $J^M_{T-1}(\mathbf{N}) = J_{T-1}(N)$.

Now, assuming the result is true for $t+1$, we show that it also holds for t. Assume wlog that $\mathbf{N} \geq 0$; otherwise replace $\mathbf{N}$ by its positive part, because $J^M_t(\mathbf{N}) = J^M_t(\mathbf{N}^+)$. For $X = \mathbf{1}'\mathbf{x}$ we have:

$$
\begin{aligned}
V^M_t(\mathbf{x}, \mathbf{N}) = &\ \pi(\mathbf{1}'\mathbf{x}) + \mathbb{E}J^M_{t+1}((\mathbf{N} - \xi\mathbf{x})^+) \\
= &\ \pi(X) + \mathbb{E}J_{t+1}(\mathbf{1}'(\mathbf{N} - \xi\mathbf{x})^+) \quad \text{(induction step)} \\
\leq &\ \pi(X) + \mathbb{E}J_{t+1}((N - \xi X)^+). \quad \text{(monotonicity of } J_t\text{)}
\end{aligned} \tag{39}
$$

Therefore,

$$
\begin{aligned}
J^M_t(\mathbf{N}) = &\ \max_{\mathbf{0} \leq \mathbf{x}; \mathbf{1}'\mathbf{x} \leq Q} V^M_t(\mathbf{x}, \mathbf{N}) \\
= &\ \max_{0 \leq X \leq Q} \pi(X) + \mathbb{E}J_{t+1}(\mathbf{1}'(\mathbf{N} - \xi\mathbf{x})^+) \\
\leq &\ \max_{0 \leq X \leq Q} \pi(X) + \mathbb{E}J_{t+1}((N - \xi X)^+) \\
= &\ J_t(N).
\end{aligned} \tag{40}
$$

Consider the optimal allocation X^* that optimizes $J_t(N)$ above, and let $\mathbf{x}$ so that $x_i/N_i = X^*/N$, for all i with $N_i > 0$. This is feasible and achieves (40) with equality, hence it is an optimal solution to the multi-client model, and the two value functions are equal. ■

Extension for Irreversible Allocation:

The aggregation result follows the same lines as the reversible allocation case. The only additional condition to verify is feasibility of the multi-client solution provided by the equal GRP rule. We need to show that the lower bounds imposed by the irreversible allocation are satisfied. In a balanced GRP solution, $\dfrac{X^t_i}{N^t_i} = \dfrac{X^t_j}{N^t_j}$ for $N^t_i, N^t_j > 0$, so $\dfrac{X^t_i}{N^t_i - zX^t_i} = \dfrac{X^t_j}{N^t_j - zX^t_j}$, for any realization $z \geq 0$ of ξ, implying $\dfrac{X^t_i}{X^t_j} = \dfrac{N^{t+1}_i}{N^{t+1}_j} = \dfrac{X^{t+1}_i}{X^{t+1}_j}$. This shows that allocation increases proportionally over time for each unsatisfied client, so the irreversibility constraints are met. Hence the multi-client problem reduces to the aggregate case under irreversible commitment.

References

Aflaki, S., I. Popescu 2007. Service Dynamics under Repeated Interactions. *INSEAD Working paper.*

Alvarado, C. 2007., Revenue Management for the Inventory Optimization of Television Azteca, *Television Azteca external communication document*

Araman, V.F., K. Fridgeirsdottir 2006. Online advertisement: A revenue management approach. *Working paper.*

Araman, V.F., O. Ozer 2005. Capacity and inventory management in the presence of a long-term channel and a spot market, *Working paper, New York University.*

Banciu, M., E. Gal-Or, P. Mirchandani, 2007. *A Mixed-Bundling Pricing Strategy for the TV Advertising Market*, working paper, Katz Graduate School of Business, University of Pittsburgh, 2007 Bassok, Y., W. J. Hopp, M. Rohatgi. 2002. A simple heuristic for the service constrained random yield problem. *IIE Transactions* **34** 479-487.

Bertsekas, D. 1995. *Nonlinear Programming.* Athena Scientific.

Bertsekas, D. 2000. *Dynamic Programming and Optimal Control.* Athena Scientific.

Bollapragada, S., H. Cheng, M. Phillips, M. Garbiras. 2002. NBC's optimization systems increase revenues and productivity. *Interfaces* **32** (1) 47-60.

Bollapragada, S., H. Mallik, M. 2007. Managing the inventory of commercial airtimes in broadcast television. *Working paper.*

Bollapragada, S., M. Garbiras. 2004. Scheduling commercials on broadcast television. *Operations Research* **53** (3) 337-345.

Bollapragada, S., T.E. Morton. 1999. Myopic heuristics for the random yield problem. *Operations Research* **47** (5) 713-722.

de Vericourt, F., M. S. Lobo. 2006. Resource and Revenue Management in Nonprofit Operations. *Fuqua Business School working paper.*

Elmaghraby,K., P. Keskinocak. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices and future directions. *Management Science* **49**(10) 1287-1309.

Gallego, G., S.G. Kou, R. Phillips. 2004. Revenue Management of Callable Products. *working paper*

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

Gupta, D., W. L. Cooper. 2005. Stochastic comparisons in production yield management. *Operations Research* **53** (2) 377-384.

Kanetkar, V., C.B. Weinberg, D.L. Weiss. 1992. Price Sensitivity and Television Advertising Exposures: Some Empirical Findings. *Marketing Science* **11** (4) 359-371.

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

Kleindorfer, P and D.J. Wu. 2003. Integrating long-term and short-term contracting via business-to-business exchanges for capital-intensive industries. *Management Science* **49** (11) 1597-1615.

Lodish, L.M., M. Abraham, S. Kalmenson, J. Livelsberger, B. Lubetkin, B. Richardson, M.E. Stevens. 1995. How TV Advertising Works: A Meta-Analysis of 389 Real World Cable TV Advertising Experiments. *Journal of Marketing Research* **32** 125-139.

Metheringham, R. A., 1964, Measuring the Net Cumulative Coverage of a Print Campaign, *Journal of Advertising Research*, **4** 23-28

Müller, A., D. Stoyan. 2002. *Comparison Methods for Stochastic Models and Risks*. Wiley Publishers.

Porteus, E. L. 2000. *Foundations of Stochastic Inventory Theory.* Stanford Business Books.

Oksendal, A. 2000. Irreversible investment problems. *Finance Stochastics* **4** 223-250.

Rajaram, K., U.S. Karmakar. 2002. Product Cycling With Uncertain Yields: Analysis And Application To The Process Industry. *Operations Research* **50**(4): 680-691. Smith, J. E., K.E. McCardle 2002. Structural Properties of Stochastic Dynamic Programs, *Operations Research* **50** 796-809.

Sepehri, M., E. A. Silver, C. New. 1986. A Heuristic for multiple lotsizing for an order under variable yield. *IIE Transactions* 63-69.

Talluri, T. T., G. J. Van Ryzin. 2004. *The Theory and Practice of Revenue Management.* Kluwers Academic Publishers.

Topkis, D. M. 1998. *Supermodularity and Complementarity.* Princeton University Press.

Yano, C., H. L. Lee. 1995. Lot sizing and random yields: A review. *Oper. Res.* **43**(2) 311-334. Wilson, R. 1993. *Nonlinear Pricing.* Oxford University Press, Oxford.

Wu, D. J., P. Kleindorfer. 2005. Competitive options, supply contracts and electronic markets. *Management Science* **51** (3) 452-466.

Europe Campus
Boulevard de Constance,
77305 Fontainebleau Cedex, France
Tel: +33 (0)1 6072 40 00
Fax: +33 (0)1 60 74 00/01

Asia Campus
1 Ayer Rajah Avenue, Singapore 138676
Tel: +65 67 99 53 88
Fax: +65 67 99 53 99

www.insead.edu

INSEAD

The Business School
for the World