The set of equilibria is *identical* under the SC and the PF models. Moreover, each firm's equilibrium service level in *any* such equilibrium is *uniquely* determined as a function of *that* firm's characteristics only, and it is a dominant choice for this firm, i.e., with fixed prices, the equilibrium service level is the firm's optimal choice, regardless of what service levels are adopted by its competitors. In contrast, the equilibrium in the SF model differs from that in the other two competition models. Here, a firm's equilibrium service level *does* depend, in general, on the characteristics of the competitors. Assuming the SF model has a unique equilibrium, we derive a simple sufficient condition under which each firm adopts a *higher price* and a *higher service level* while enjoying a *higher* demand volume, compared to the other types of competition. In the presence of multiple equilibria, the same uniform ranking applies to the componentwise-smallest equilibria. Thus, if firms choose and announce their service levels before choosing their price, this will result in higher but more expensive service by all competitors. Because all firms' demand volumes increase as well, this type of competition appears to benefit the consumer. It also suggests that value is added to the consumer when government agencies, industry consortia, or independent organizations periodically report on service levels.

References

Cachon, P. C., P. T. Harker. 2002. Competition and outsourcing with scale economies. *Management Sci.* **48**(10) 1314–1333.

Chen, H., Y.-W. Wan. 2003. Price competition of make-to-order firms. *IIE Trans.* **35**(9) 817–832.

Levhari, D., I. Luski. 1978. Duopoly pricing and waiting lines. *Eur. Econom. Rev.* **11** 17–35.

Luski, I. 1976. On partial equilibrium in a queueing system with two servers. *Rev. Econom. Stud.* **43** 519–525.

Dynamic Assortment with Demand Learning for Short Life–Cycle Consumer Goods

Felipe Caro
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, fcaro@sloan.mit.edu

Advisor: Jérémie Gallien
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, jgallien@mit.edu

1. Introduction

Long development, procurement, and production leadtimes resulting in part from a widespread reliance on overseas suppliers have traditionally constrained fashion retailers to make supply and assortment decisions well in advance of the selling season, when only limited and uncertain demand information is available. With little ability to modify product assortments and order quantities after the season starts and demand forecasts can be refined, many retailers are seemingly cursed with simultaneously missing sales for want of popular products, while having to use markdowns in order to sell the many unpopular products still accumulating in their stores (see Fisher et al. 2000).

Recently, however, a few innovative firms, including Spain-based Zara, Mango, and Japan-based World Co. (sometimes referred to as "fast-fashion" companies), have gone substantially further, implementing product development processes and supply chain architectures that allow them to make *most* product design and assortment decisions *during* the selling season. Remarkably, their higher flexibility and responsiveness is partly achieved through an increased reliance on more costly local production relative to the supply networks of more traditional retailers.

At the operational level, leveraging the ability to introduce and test new products once the season has started motivates a new and important decision problem, which seems key to the success of these fast-fashion companies: Given the constantly evolving demand information available, which products should be included in the assortment at each point in time?

The problem just described seems challenging, in part because it relates to the classical trade-off known as exploration versus exploitation: In each period the retailer must choose between including products in the assortment that he has a "good sense" are profitable (exploitation) or products for which he would like to gather more demand information (exploration). That is, he must decide between being "greedy" based on his current information or trying to learn more about product demand (which might be more profitable in the future). In that respect, the dynamic assortment problem can be seen as a variant of the multiarmed bandit problem with finite horizon and several plays per stage. Each arm represents a product, and pulling an arm is equivalent to including the respective product in the assortment.

2. Model Definition

2.1. Supply and Demand

Consider a retailer selling products in a store during a limited selling season. The set of all products that the retailer may potentially sell is denoted by $\mathcal{S} = \{1, 2, \ldots, S\}$; this set includes both the products already available when the season starts and all the variants and new products that may be designed during the season. The net margin r_s of product $s \in \mathcal{S}$ is assumed to be exogenously given, positive, and constant. We assume that the selling season can be divided into T periods and that at the beginning of each of these periods the product assortment in the store may be revised; time is counted backward and denoted by the index t.

The store's limited shelf space is captured by the constraint that the assortment in each period may include at most N different products out of the S available; we are thus implicitly assuming that all products require the same shelf space. We also assume a perfect inventory replenishment process during each assortment period, so that there are no stockouts or lost sales. Consequently, in our model, realized sales equal total demand, and for each product we focus on assortment inclusion or exclusion as opposed to order quantity. Holding costs are ignored in our formulation.

The demand for each product in the assortment is exogenous and stationary, but stochastic, and we do not capture substitution effects. Specifically, we assume that customers willing to buy one unit of each product s in the assortment arrive at the store according to a Poisson process with an unknown but constant rate γ_s. That is, the underlying arrival rate γ_s is assumed to remain constant throughout the entire season, but the resulting actual demand for product s may only be observed in the periods when that product is included in the assortment. In addition, the arrival processes corresponding to different products are assumed to be independent.

We adopt a standard Gamma-Poisson Bayesian learning mechanism. The underlying demand rate γ_s for each product s is initially unknown to the retailer; however, he starts each period with a prior belief on the value of that parameter represented by a Gamma distribution with shape parameter m_s and scale parameter α_s (m_s and α_s must be positive, and m_s is assumed to be integer). Redefining time units if necessary, we can assume with no loss of generality that the length of each assortment period is 1; the predictive demand distribution under that belief for product s in the upcoming period is then given by a negative binomial distribution with parameters m_s and $\alpha_s(\alpha_s + 1)^{-1}$. If now product s is included in the assortment and n_s actual sales are observed in that period, it follows from Bayes's rule that the posterior distribution of γ_s has a Gamma distribution with shape parameter $(m_s + n_s)$ and scale parameter $(\alpha_s + 1)$.

2.2. Dynamic Programming Formulation

Given the discrete and sequential character of our problem, the natural solution approach is dynamic programming (DP); the state at time t is given in our model by the parameter vector $\mathbf{I}^t = (\mathbf{m}, \boldsymbol{\alpha})$, which summarizes all relevant information, including past assortments and observed sales. For ease of notation, we omit the dependence of $\mathbf{m}$ and $\boldsymbol{\alpha}$ on t. In each period, the decision to include product s in the assortment or not can be represented by a binary variable $u_s \in \{0, 1\}$, where $u_s = 1$ means that product s is included.

The optimal profit-to-go function $J_t^*(\mathbf{m}, \boldsymbol{\alpha})$ given state $(\mathbf{m}, \boldsymbol{\alpha})$ and t remaining periods must then satisfy the following Bellman equation:

$$J_t^*(\mathbf{m}, \boldsymbol{\alpha}) = \max_{\substack{\mathbf{u}\in\{0,1\}^S:\\ \sum_{s=1}^S u_s \le N}} \sum_{s=1}^{S} r_s \frac{m_s}{\alpha_s} u_s + \mathbb{E}_{\mathbf{n}}[J_{t-1}^*(\mathbf{m}+\mathbf{n}\cdot\mathbf{u}, \boldsymbol{\alpha}+\mathbf{u})],$$

where $\mathbf{v}\cdot\mathbf{u}$ represents the componentwise product of two vectors, and the terminal condition is $J_0^*(\mathbf{m}, \boldsymbol{\alpha}) = 0$ for all states.

Note that the only link between consecutive periods in this model is the information acquired about demand, and that different products are only coupled at a given period through the shelf space constraint $\sum_{s=1}^S u_s \le N$; this type of problem is known as a *weakly coupled* DP.

3. Analysis

3.1. The Dual Dynamic Program

The analysis of the model is based on Lagrangian relaxation and the decomposition of weakly coupled dynamic programs (see, for instance, Bertsimas and Mersereau 2004 and the references therein). Specifically, we relax the shelf space constraint, which leads to the definition of *dual policies* that can be shown to be useful in finding near-optimal *primal* policies and upper bounds for the optimal profit-to-go. Let $\lambda_t(\mathbf{m}, \boldsymbol{\alpha})$ denote any function associated with period t that maps the state space into the set of nonnegative real values; we define a *dual policy* to be a vector of functions $\boldsymbol{\lambda}_t = (\lambda_t(\cdot), \lambda_{t-1}(\cdot), \ldots, \lambda_1(\cdot))$.

For any dual policy $\boldsymbol{\lambda}_\mathbf{t}$ and any initial state $(\mathbf{m}, \boldsymbol{\alpha})$, the corresponding profit-to-go is obtained by solving the *dual dynamic program* given by:

$$H_t^{\boldsymbol{\lambda}_\mathbf{t}}(\mathbf{m}, \boldsymbol{\alpha}) = N\lambda_t(\mathbf{m}, \boldsymbol{\alpha}) + \max_{\mathbf{u}\in\{0,1\}^S} \sum_{s=1}^{S} \Big(r_s \frac{m_s}{\alpha_s} - \lambda_t(\mathbf{m}, \boldsymbol{\alpha})\Big) u_s + \mathbb{E}_{\mathbf{n}}[H_{t-1}^{\boldsymbol{\lambda}_{\mathbf{t-1}}}(\mathbf{m}+\mathbf{n}\cdot\mathbf{u}, \boldsymbol{\alpha}+\mathbf{u})],$$

with $H_0^{\boldsymbol{\lambda}_\mathbf{0}}(\mathbf{m}, \boldsymbol{\alpha}) = 0 \;\forall\, (\mathbf{m}, \boldsymbol{\alpha})$.

In words, a dual policy gives the price of a unit of shelf space for each period and each possible state. As expected, weak duality holds, and for any dual policy and initial state we have that $J_t^*(\mathbf{m}, \boldsymbol{\alpha}) \le H_t^{\boldsymbol{\lambda}_t}(\mathbf{m}, \boldsymbol{\alpha})$. By considering *open-loop* dual policies (i.e., a constant shadow price per period), one can calculate an upper bound for $J_t^*(\mathbf{m}, \boldsymbol{\alpha})$ using standard convex nondifferentiable optimization methods.

3.2. The Index Policy

It is well known that index policies are not optimal for our version of the multiarmed bandit problem (see Berry and Fristedt 1985), however they are still appealing given their simple structure. Through a sequence of intuitive approximations to the dual DP we derive a heuristic index policy for the dynamic assortment problem. The suggested rule is to include the N products with the highest indices in the assortment, where the index for product s at period t is given by the following formula:

$$\eta_{t,s} \approx r_s\mathbb{E}[\gamma_s] + z_t \frac{r_s\mathbb{V}[\gamma_s]}{\sqrt{\mathbb{V}[\gamma_s] + \mathbb{E}[\gamma_s]}}. \tag{1}$$

The factor z_t is the unique solution to the equation $(t-1)\cdot\Psi(z_t) = z_t$, where $\Psi(z) = \int_z^\infty (x-z)\phi(x)\,dx$ is the loss function of a standard normal. The values z_t, which are independent of the problem data, are increasing and concave in t.

The index $\eta_{t,s}$ represents the highest price at which one should be willing to rent some shelf space to display (and sell) product s there; it is thus a measure of the desirability of including each individual product in the assortment, and from that standpoint the rationale behind the suggested index policy is to fill all shelf space with the most desirable products. Note that the first term in the index expression (1) favors exploitation, and the second term favors exploration because it is increasing in both the variance of γ_s and the number of remaining periods (through z_t). Intuitively, when uncertainty about demand for a product s (captured by $\mathbb{V}[\gamma_s]$) is high, there is more benefit to learn from including s in the assortment because of the upside potential from future sales. However, one should increasingly favor exploitation over exploration as the remaining planning horizon (and opportunity for leveraging exploration) shortens, which is captured by the decrease with t of the multiplicative factor z_t. The summation of a variance and an expectation in the second term of (1) is not a mistake, but rather a consequence of the period length being equal to 1.

Finally, when assessing the performance of the index policy defined above, our primary benchmark is the *greedy policy*, which consists of selecting in each period the N products with the highest immediate expected profit $r_s\mathbb{E}[\gamma_s]$. Note that the greedy policy still involves learning despite its myopic nature, but the impact of assortment decisions on future learning is ignored. As a result, several authors also refer to it as *passive learning*.

4. Conclusions

We have developed a discrete-time DP model for the dynamic assortment problem faced by a fast-fashion retailer refining his estimate of consumer demand for his products over time. The main assumptions made were: (i) independent products, (ii) no lost sales, (iii) constant demand rates, and (iv) immediate assortment implementation. Under these assumptions, we have formulated this problem as a multiarmed bandit with finite horizon and multiple plays per stage. Using the Lagrangian decomposition of weakly coupled DPs, we have derived a closed-form index policy and we have derived an upper bound for the optimal profit-to-go, which allows us to assess the suboptimality gap of the suggested policy.

A simulation study indicates that the index policy always performs at least as well as the greedy policy (or passive learning), and significantly outperforms it in scenarios with diffuse or biased prior demand information. Also, numerical computations of the bound mentioned above suggest that the index policy is close to optimal. In general, the improvement of the suggested index policy on the greedy rule increases with the planning horizon length and the variance of the initial priors.

Although the major assumptions of our model may be particularly strong in some environments, our approach was partly motivated by the belief that the closed-form policy that they allow to derive constitutes a useful starting point for designing heuristics or developing extensions in more complex environments.

References

Berry, D. A., B. Fristedt. 1985. *Bandit Problems, Sequential Allocation of Experiments*. Chapman and Hall, New York.

Bertsimas, D., A. Mersereau. 2004. A learning approach to customized marketing. Working paper, University of Chicago, Chicago, IL.

Fisher, M. L., A. Raman, A. S. McClelland. 2000. Rocket science retailing is almost here—Are you ready? *Harvard Bus. Rev.* **78**(4) 115–124.

Promised Leadtime Contracts and Renegotiation Incentives Under Asymmetric Information

Holly S. Lutze
School of Management, University of Texas at Dallas, Richardson, Texas 75083-0688, hlutze@utdallas.edu

Advisor: Özalp Özer
Department of Management Science and Engineering, Stanford University, Stanford, California 94305-4026

Firms often establish supply chain relationships through contracts that provide rules for interaction. These contracts help align incentives for decision making and establish how partners will share both the benefits of interaction and the risks from uncertain supply or demand. We propose a new multiperiod contract form, the promised leadtime contract. The contract reduces supplier risk from future demand uncertainty, and it eliminates buyer risk from uncertain inventory availability. The supplier agrees to ship buyer orders in full after a promised leadtime, and the buyer pays the supplier for this privilege. The supplier and buyer may each carry inventory, depending on the agreed on promised leadtime and their respec-