

Abstracts of Accepted Papers

9. On the optimal control of partially observed inventory systems. (with A. Bensoussan and S. Sethi). To appear in Comptes Rendus de l'Académie des Sciences.

This paper introduces recent developments in the analysis of inventory systems with partial observations. The states of these systems are typically conditional distributions, which evolve in infinite dimensional spaces over time. Our analysis involves introducing unnormalized probabilities to transform nonlinear state transition equations to linear ones. With the linear equations, the existence of the optimal feedback policies are proved for two models where demand and inventory are partially observed. In a third model where the current inventory is not observed but a past inventory level is fully observed, a sufficient statistic is provided to serve as a state. The last model serves as an example where a partially observed model has a finite dimensional state. In that model, we also establish the optimality of the basestock policies, hence generalizing the corresponding classical models with full information.

8. Pricing and production game under revenue sharing and information update. (with S. Luo). Journal of Revenue and Pricing Management Vol.3, No.4. Fall 2005.

We study the pricing and production games in a vendor managed inventory system (VMI) where the supplier determines the quantity and the retailer determines the price. We illustrate the role of revenue sharing in mitigating the double marginalization in VMI. We also investigate the positive effect of information updates on the performance of VMI. At the same time, we discuss the existence and uniqueness of Nash equilibrium in one period game, subgame perfect Nash equilibrium in a two-period game with perfect information in the second period and the myopic Nash equilibrium in finitely repeated games. Our numerical studies demonstrate the substantial economic value of revenue sharing and information updates.

7. Capacity-driven acceptance of customer orders for a multi-stage batch manufacturing system: models and algorithms (with D. Chen, P. Chen, M. Freimer, P.L. Jackson, V. Melkonian and R.O. Roundy). To appear in IIE Transactions.

An automatic parts manufacturer produces a wide range of parts in a job-shop environment. Many of the manufacturing operations have substantial setups. When a client phones in an order, the manufacturer must decide quickly whether or not it has the capability required to accept the order. We develop a simplified formulation of the order acceptance problem. We formulate the discrete-time version as an integer program. The problem is NP-hard, but in 51 out of 51 test problems the LP relaxation is tight. For larger problems we test several heuristics. Three of the heuristics look promising -simulated annealing, a genetic algorithm and an LP-based heuristic.

6. Optimal Machine Capacity Expansions with nested Limitations under Demand Uncertainty (with R.O. Roundy and S.C. Wood). Naval Research Logistics Vol.51, Iss.2, 2004: 217-241.

This paper studies capacity expansions for a production facility that faces uncertain customer demand for a single product family. The capacity of the facility is modeled in three tiers, as follows. The first tier consists of a set of upper bounds on production that correspond to different resource types (e.g., machine types, categories of manpower, etc). These upper bounds are augmented in increments of fixed size (e.g., by purchasing machines of standard types). There is a second-tier resource that constrains the first-tier bounds (e.g., clean room floor space). The third-tier resource bounds the availability of the second-tier resource (e.g., the total floor space enclosed by the building, land, etc). The second and third-tier resources are expanded at various times in

various amounts. The cost of capacity expansion at each tier has both fixed and proportional elements. The lost sales cost is used as a measure for the level of customer service. The paper presents a polynomial time algorithm (FIFEX) to minimize the total cost by computing optimal expansion times and amounts for all three types of capacity jointly. It accommodates positive lead times for each type. Demand is assumed to be nondecreasing in a “weak” sense.

5. Optimal Capacity Expansion for Multi-Product, Multi-Machine Manufacturing Systems with Stochastic Demand (with R.O. Roundy, F. Zhang and T.Huh). IIE Transactions, Vol.36 No.1, 2004: 23-36.

We consider a discrete-time capacity expansion problem involving multiple product families, multiple machine types, and non-stationary stochastic demand. Capacity expansion decisions are made to strike an optimal balance between investment costs and lost sales costs. Motivated by current practices in the semiconductor and other high-tech industries, we assume that only minimal amounts of finished-goods inventories are held, due to the risk of obsolescence. We assume that when capacity is in short supply, management desires to ensure that a minimal service level for all product families is obtained. Our approach uses a novel assumption that demand can be approximated by a distribution whose support is a collection of rays emanating from a point and contained in real multi-dimensional space. These assumptions allow us to solve the problem as a max-flow, min-cut problem. Computational experiments show that large problems can be solved efficiently.

4. Evaluation of capacity planning practices for the semiconductor industry (with R.O. Roundy). IEEE Transactions on Semiconductor Manufacturing, Vol.15, No.3, 2002: 331-340.

We perform four experiments to evaluate several features of capacity planning approaches used in practice and make concrete suggestions for practitioners. We use an optimal capacity planning technique (FIFEX) with industrial demand and capacity data for realistic conclusions. The first experiment studies the cost effects of allowing for capacity expansion only at special times such as at the beginning of each month or each quarter. We find that these effects are not great if there is at least one capacity expansion opportunity about every 6 months. The second experiment determines the cost implications of capacity planning heuristics. We demonstrate that FIFEX delivers solutions which cost 5-10% less than heuristic solutions. The third experiment examines the sensitivity of costs against the frequency of forecasting and planning. Our experiments indicate that costs can be decreased by 2-7% by doubling forecasting and planning frequency. We suggest that practitioners forecast and plan at least once every quarter. The final experiment compares joint optimization of tool and floor-shell expansions with sequential optimization. 3-9% of costs can be saved by using the joint optimization technique FIFEX instead of sequential optimization. We note that although cost savings are relatively small in percentages, they correspond to tens of million dollars.

3. SeDFAM: Semiconductor demand forecast accuracy model (with R.Roundy). IIE Transactions, Vol.34, No.5, 2002: 449-465.

In the semiconductor industry, many critical decisions are based on demand forecasts. However, these forecasts are subject to random error. In this paper, we lay out a scheme estimating the variance and correlation of forecast errors (without altering given forecasts) and modeling the evolution of forecasts over time. Our scheme allows correlations across time, products and technologies. It also addresses the case of nonstationary errors due to ramps (technology migrations). It can be used to simulate chip demands for production planning / capacity expansion studies.

2. Continuous review inventory models where random lead time depends on lot size and reserved capacity (with J.H. Bookbinder and Y. Gerchak) International Journal of Production Economics, Vol.68, Iss.3, 2000: 217-228.

The processing time of large orders is, in many industries, longer than that of small orders. This renders supply lead times in such settings to be increasing in the order size. Yet that pattern is not reflected in existing inventory control models, especially those allowing for random lead times. This work aims at rectifying this situation. Our setting is an order-quantity/reorder-point model with backordering, where the shortage penalty is incurred per unit per time. The processing time of each unit is random; the processing time of a lot is correlated with its size. For the case where the lead time is proportional to the lot size, we obtain a closed-form solution. That is, unlike the classical (Q, r) model (where lead time is independent of lot size), no iterations are required here. We also analyze a case where the processing time exhibits economies of scale in the lot size. Finally, we consider a situation where a customer can secure shorter processing times by reserving capacity at the supplier's manufacturing facility.

1. Random lead times and expedited orders in (Q, r) inventory systems (with J.H. Bookbinder). European Journal of Operational Research, Vol.115, Iss.2, 1999: 300-313.

This paper considers inventory models of the order-quantity/reorder-point type, or (Q, r) models. In general, the control parameters (Q and r) depend on both the demand process and the replenishment lead time. Although many studies have treated lead time as constant, focusing solely on demand, a (Q, r) model with stochastic lead time could be a building block in Supply Chain Management. Variability in lead time between successive stages is often what disturbs supply chain coordination. In a two-stage system with a constant demand rate, we will concentrate on lead time as a random variable, and develop two probabilistic models. In the first, lead time T is exogenous. Lead time is made endogenous in the second stochastic model through an "expediting factor" τ , the constant of proportionality between random variable \tilde{T} (the expedited lead time) and T (ordinary lead time): $\tilde{T} = \tau T$. For expedited orders ($\tau < 1$), shorter-than-average lead time can be obtained at a cost. Similarly, longer mean lead times result in a rebate for the customer when $\tau > 1$. The second model thus has three decision variables (Q, r, τ) . For each model, we show that the expected cost per unit time is jointly convex in the decision variables and obtain the global minimizer. Numerical examples are given. Sensitivity analyses are conducted with respect to the cost parameters, and suggestions are made for future research.