Capacity Choice and Allocation: Strategic Behavior and Supply Chain Performance

Gérard P. Cachon • Martin A. Lariviere
The Fuqua School of Business, Duke University, Durham, North Carolina 27708
gpc@mail.duke.edu • larivier@mail.duke.edu

We consider a simple supply chain in which a single supplier sells to several downstream retailers. The supplier has limited capacity, and retailers are privately informed of their optimal stocking levels. If retailer orders exceed available capacity, the supplier allocates capacity using a publicly known allocation mechanism, a mapping from retailer orders to capacity assignments. We show that a broad class of mechanisms are prone to manipulation: Retailers will order more than they need to gain a more favorable allocation. Another class of mechanisms induces the retailers to order exactly their needs, thereby revealing their private information. However, there does not exist a truth-inducing mechanism that maximizes total retailer profits.

We also consider the supplier's capacity choice. We show that a manipulable mechanism may lead the supplier to choose a higher level of capacity than she would under a truth-inducing mechanism. Nevertheless, her choice will appear excessively restrictive relative to the prevailing distribution of orders. Furthermore, switching to a truth-inducing mechanism can lower profits for the supplier, the supply chain, and even her retailers. Hence, truth-telling is not a universally desirable goal.
(*Supply Chain*; *Game Theory*; *Capacity Allocation*; *Bull-Whip Effect*; *Incentive Contracts*)

1. Introduction

Consider a supplier selling to multiple retailers. What should she do when retailer orders exceed her capacity? In many supply chains, the answer is to put the retailers on "allocation," rationing capacity through quantity limits instead of price mechanisms. Going on allocation is a common occurrence in industries in which capacity expansion is costly and time consuming (e.g., steel and paper). It also occurs with popular new products such as initial public offerings of stocks, fashionable toys at Christmas time, or hot automobiles.

Many allocation schemes are used in practice. Makers of nicotine patches based allocations on sales histories and location (Hwang and Valeriano 1992). Many car companies use a "turn-and-earn" scheme, rationing hot models on the basis of past sales. For a time, Acuras were allotted through an unofficial system of price premiums and side payments, leading to several fraud convictions (Henderson 1995); the company now allocates vehicles on the basis of sales and a customer satisfaction index. In some instances, explicit preferences are granted contractually; Frito-Lay, for example, has exclusive access to Procter & Gamble's new fat substitute (Frank 1996). Alternatively, powerful customers may demand priority. (For a toy industry example, see Gruley and Pereira 1996.)

The chosen allocation scheme matters when retailers anticipate different levels of demand. Those expecting high demand would optimally set a high stocking level while the less optimistic would choose lower levels. Ideally, when the sum of needs exceeds available capacity, stock would be allocated among the retailers to maximize their total profits. Although

0025-1909/99/4508/1091$05.00

collectively rational, such an arrangement requires that each retailer necessarily receives less than his optimal amount. Despite grumbling from individual retailers, the supplier could allocate stock to maximize total profits if she knew each retailer's ideal stocking level. But an omniscient supplier is unlikely. Instead, a supplier can only use her prior beliefs regarding retailer needs and submitted orders to construct an allocation mechanism. Even if her beliefs are accurate in expectation, uncertainty opens the door to strategic manipulation; retailers may game the system, distorting their orders to receive larger allotments.

In such a setting, this paper addresses three primary issues:

- Which allocation mechanisms are "manipulable" and induce retailers to misrepresent their needs? Which mechanisms are "truth-inducing" and lead to the truthful reporting of retailer information?
- Does the entire supply chain, the supplier, or the retailers benefit from restricting the supply chain to truth-inducing mechanisms?
- How does the chosen allocation mechanism influence how much capacity the supplier elects to build?

Truth-inducing mechanisms are an appealing goal. If retailers ordered exactly their needs, the supplier could allocate more to those with the larger market. Conversely, manipulable mechanisms might precipitate an avalanche of orders, preventing the supplier from determining who truly needs the most stock. Some with high expected demand may receive too little and others with low expected demand may receive too much. In the end, the system serves all retailers poorly. Lee et al. (1997) demonstrate that allocating capacity in proportion to orders induces strategic behavior. We show that a larger class of mechanisms is manipulable. In particular, the Pareto allocation mechanism, which would maximize total retailer profits under full information, is always manipulable under asymmetric information. In other words, if the supplier does not know the retailers' needs with certainty, she cannot guarantee that capacity will be allocated so as to maximize the retailers' profits. Truth-inducing mechanisms exist, and a numerical study demonstrates that they can yield total retailer profits that are reasonably close to the maximum.

Our second question asks whether truth telling helps the supply chain. At first glance the answer seems trivial. More information must be better, so of course the retailers should tell the truth. This persuasive intuition is the cornerstone of many supply chain management innovations: JIT; Efficient Consumer Response (Gary 1993); Quick Response (Abernathy et al. 1995); Accurate Response (Fisher et al. 1994, Fisher and Raman 1996). However, in addition to not maximizing retailer profits, truth-inducing mechanisms fail to maximize the supplier's profits. Once the supplier has sunk the cost of building capacity, her profits are maximized when capacity is fully utilized. In some instances, the retailers' total truthful needs will be less than the supplier's capacity, and the supplier's profits would be higher if the retailers increased their orders. Hence, supply chain profits can increase when a truth-inducing mechanism is replaced by a manipulable mechanism that creates order inflation.

The allocation mechanism is, of course, irrelevant if capacity is never a binding constraint. Our third question thus considers how the chosen allocation mechanism influences the supplier's capacity choice. We study this question using a game of asymmetric information. We characterize the set of equilibria and prove an unexpected result: when the supplier uses a manipulable mechanism we call linear allocation, the expected sum of retailers' orders *declines* as she invests in greater capacity. Paradoxically, if she chooses a small capacity, the supplier will face a large market; if she chooses a large capacity, she will face a small market. The supplier consequently has an incentive to restrict capacity, which is consistent with the casual observation that some industries (e.g., automobiles, toys) experience perennial capacity shortages. Alternatively, it argues against claiming supplier mismanagement or ineptness as explanations for these shortages. While this perverse incentive leads to extremely tight capacity relative to the prevailing distribution of orders, the supplier may in fact choose greater capacity than she would under a truth-inducing mechanism. Furthermore, not only can the use of linear allocation increase supply chain profits, it can increase both the supplier's and the retailers' profits when capacity is expensive.

Manipulable mechanisms can thus benefit a supply chain for two reasons. For a given capacity level, they reduce the expected amount of idle capacity, increasing supplier profits. Since they reduce the probability of idle capacity, the supplier may build more capacity than she would under a truth-inducing mechanism, benefiting all players. Further, in our numerical study, manipulable mechanisms do not lead to terrible allocations; in equilibrium, the retailer with the highest need inflates his order the most and so receives the highest allocation. Hence, higher capacity utilization does not come at the expense of efficient capacity assignments.

The next section outlines the Allocation Game and relates the model to the extensive literature on capacity management. Section 3 describes allocation mechanisms and shows that a broad class of mechanisms is open to manipulation. Also, the conditions required for a mechanism to induce truth-telling are outlined and Bayesian equilibria in the Allocation Game are characterized. Section 4 describes the Capacity Game in which the supplier chooses her allocation mechanism and how much capacity to build. Section 5 presents a numerical example. Section 6 concludes and discusses future research.

2. The Allocation Game

We consider a one-period setting in which a single supplier sells to $N \geq 2$ retailers. The retailers enjoy local monopolies and are not in direct competition. The individual retailers have private information regarding the optimal amount of inventory they should stock. For example, each may face a newsvendor problem and be privately informed of some parameter of his market's demand distribution. Heterogeneity among retailer expectations can occur for a variety of reasons including geographic locations, promotion plans, product selection, and pricing strategies.

We assume that before the period begins: (1) the supplier has chosen a finite capacity K and consequently can produce no more than K units during the period; and (2) the supplier has announced publicly the allocation mechanism she will use if total retailer orders exceeds available capacity. While both capacity and the allocation mechanism are taken as given in the Allocation Game, they are endogenous decisions in the Capacity Game.

During the period, the following sequence of events occurs: (1) each retailer learns his (and only his) private information; (2) retailers simultaneously submit their orders, which the supplier fills according to the posted allocation mechanism; and (3) retailers experience consumer demand.

Retailers submit orders independently, and orders are the only communication between the retailers and the supplier. No retailer can credibly announce his information to other players, and no side contracts between the supplier and any retailer are allowed. In short, a retailer can only influence his allotment through his order. The supplier charges a constant wholesale price w per unit, and a retailer must accept and pay for any allocation up to his full order. The supplier cannot deliver to a retailer more than he has ordered. Finally, stock allotted to one retailer cannot be diverted to another.

This model reflects an industry in which (1) capacity is sufficiently expensive that K may be less than total retailer orders; (2) short term changes in demand can be substantial, but short term modifications to capacity are infeasible; (3) and spot markets impose prohibitively high transaction costs. Industries generally meeting these characteristics include toys, automobiles, and personal computers.

2.1. Literature Review

The literature on the management of capacity is extensive, so we only provide a description of major classes of models and how they relate to this work.

In queueing models customers are always served, but as the arrival rate approaches the service rate of the facility (i.e., its capacity), the waiting time for service increases. One generally seeks pricing schemes that induce individual customers to implement system-optimal solutions (e.g., Bell and Stidham 1983, Dewan and Mendelson 1990, Mendelson and Whang 1990, Lederer and Li 1994). In the peak-load pricing problem, customers do not wait to be served but can choose when to request service, albeit higher prices are charged in peak demand periods (e.g., Oren et al. 1983, Gale and Holmes 1993). Our work differs from both areas in important ways. We develop a one-period model, so there is no notion of customers' waiting or shifting consumption. For instance, all retailers want delivery of popular toys at the start of the Christmas season, and few would accept delivery in January. Also, we assume a constant wholesale price; a retailer is unable to signal his priority by accepting higher priced service.

A constant wholesale price is a reasonable assumption for our setting because at the time the retailers submit their orders no one knows how constrained capacity will be. The socially optimal wholesale price would depend on the degree to which capacity is constrained. To implement efficient pricing, the supplier would have to post a schedule contingent on all possible realizations of orders, and each retailer would have to submit a listing of orders for all possible prices he could be charged. (See Porteus and Whang 1991 for a model that does implement a menu of contracts.) Clearly, transaction costs would be overwhelming, and we feel that this is a primary reason for the prevalence of constant wholesale pricing in the markets we consider. (See Peck 1996 for another explanation.) The queuing literature avoids this complication by analyzing the long run performance of the system. Although correct on average, the posted price at any given time may be too high or too low. See Harris and Raviv (1981) and Maskin and Riley (1989) for models in which a capacity constrained supplier chooses a selling mechanism that can price discriminate among retailers.

Several researchers (e.g., Topkis 1968, Kaplan 1969, Nahmias and Demmy 1981, Ha 1997) have studied systems with sequentially arriving customers of observable priority classes. Here, all orders arrive at once, and the supplier has no way of prioritizing orders.

The impact of contract terms on ordering policies and supply chain performance has received substantial attention (e.g., Eppen and Iyer 1997, Bassok and Anupindi 1998, Tsay and Lovejoy 1998, Donohue 1996). Anand and Mendelson (1997) examine system performance under alternative information and management structures when the supply chain acts as a team. In work closer to ours, Lee et al. (1997) outline a one period model with identical retailers served by a supplier experiencing stochastic capacity shocks. When capacity is insufficient, it is allocated in proportion to orders. They show that competition for restricted capacity may lead retailers to inflate their orders.

3. Allocation Game Analysis

This section studies the Allocation Game. Profit functions and allocation mechanisms are formally defined. Several allocation mechanisms are identified, including the Pareto allocation mechanism. Two equilibrium concepts are applied, and equilibria are identified.

3.1. Profit Functions and Allocation Mechanisms

Define $\pi_i(a_i, \theta_i)$ as retailer i's profits when his private information (or type) is θ_i and he receives an allocation of a_i, where θ_i is drawn from a closed interval Θ_i. $\pi_i(a_i, \theta_i)$ is twice differentiable with respect to a_i and θ_i with

$$\frac{\partial^2 \pi_i(a_i, \theta_i)}{\partial a_i^2} < 0 \quad \text{and} \quad \frac{\partial^2 \pi_i(a_i, \theta_i)}{\partial a_i \partial \theta_i} > 0.$$

For a fixed θ_i, we assume $\pi_i(a_i, \theta_i)$ is maximized at a finite allocation level $a_i^*(\theta_i)$. Let θ be the vector of types, let a be a vector of allocations, and let $a^*(\theta) = \{a_1^*(\theta_1), \ldots, a_N^*(\theta_N)\}$ be the vector of optimal allocations. From prior assumptions, $a_i^*(\theta_i)$ is strictly increasing in θ_i. We assume there exists at least one realization of types such that the total of optimal allocations exceeds available capacity. The local monopoly assumption means that a given retailer's profits are independent of the other retailers' allocations and types.

The set of feasible allocations $\mathcal{A}$ is a subset of $\Re^N$ such that for all $a \in \mathcal{A}$, $a \geq 0$ and $\Sigma_{i=1}^N a_i \leq K$. Let m be the vector of retailer orders, let m_i be retailer i's order, and let m_{-i} be the vector of the other retailers' orders. We assume that $m_i \leq \bar{M}$ for all i where $\bar{M}$ is a large constant. (This is a technical restriction needed for Theorem 8.) An *allocation mechanism* is a function g that assigns a feasible allocation to each vector of orders, $g(m) \in \mathcal{A}$. Let $g_i(m)$ be retailer i's allocation. The supplier can never allocate to a retailer more than the retailer orders, i.e., $g_i(m) \leq m_i$. This assumption has both intuitive appeal and technical consequences.

See Laffont (1988) for a more detailed treatment on mechanism design.

Perhaps the most intuitive allocation mechanism is *proportional allocation*, for which:

$$g_i(m) = \min\left\{ m_i,\ Km_i \Big/ \sum_{j=1}^{N} m_j \right\}.$$

When capacity binds, proportional allocation gives to each retailer the same fraction of his order. An alternative scheme, *linear allocation*, awards each retailer his order minus a common deduction. To be specific, index the retailers in decreasing order of their order quantities, i.e., $\{m_1 \geq m_2 \geq \cdots \geq m_N\}$. Retailer i is allocated $g_i(m, \tilde{n})$, where

$$g_i(m, \tilde{n}) = \begin{cases} m_i - \dfrac{1}{\tilde{n}} \max\left(0, \displaystyle\sum_{j=1}^{\tilde{n}} m_j - K\right), & i \leq \tilde{n}, \\ 0, & i > \tilde{n}, \end{cases}$$

and $\tilde{n}$ is the largest integer less than or equal to N such that $g_{\tilde{n}}(m, \tilde{n}) \geq 0$. Linear allocation is cumbersome to define, but we demonstrate later that it has useful analytical properties.

An allocation mechanism is *efficient* if $\sum_{i=1}^{N} m_i \leq K$ implies $g(m) = m$, and $\sum_{i=1}^{N} m_i > K$ implies $\sum_{i=1}^{N} g_i(m) = K$. Thus, efficient allocation mechanisms never waste capacity. An allocation mechanism is *increasing* if for all i, $\hat{m}_i > m_i$ and m_{-i}, $g_i(\hat{m}_i, m_{-i}) \geq g_i(m_i, m_{-i})$. (Decreasing mechanisms do exist. For instance, serve the retailers from the smallest order to greatest order.) Under an increasing mechanism, a retailer never receives less by ordering more, so ordering $a_i^*(\theta_i)$ dominates any smaller order. Given these obvious virtues, we consider only efficient and increasing allocation mechanisms and restrict our attention to orders $m_i \geq a_i^*(\theta_i)$.

A mechanism is *individually responsive* (IR) if, for all i, $0 < g_i(m) < K$ implies

$$g_i(\hat{m}_i, m_{-i}) > g_i(m_i, m_{-i}), \qquad \hat{m}_i > m_i.$$

Under an IR allocation mechanism, if a retailer is receiving a positive allocation and orders more, the retailer gets more unless he has already been allocated all of capacity. We will show that individual responsiveness delineates allocation mechanisms that induce truthful ordering from those that induce order inflation. Both proportional and linear allocation are efficient, increasing, and individually responsive.

3.2. Pareto Allocation Mechanism

The *Pareto allocation mechanism*, denoted g^*, is the mechanism that maximizes the sum of retailer profits assuming all retailers truthfully submit their optimal orders, $a^*(\theta)$. It can be interpreted as maximizing supply chain profits subject to no retailer ever receiving more than he truthfully desires at the prevailing wholesale price. As such, it is worth determining whether it can be implemented in practice, i.e., will retailers indeed order truthfully when the supplier implements the Pareto mechanism? Even if the Pareto mechanism cannot be implemented, it is a benchmark against which other schemes can be compared. Before tackling these questions, we develop some general properties of Pareto mechanisms.

LEMMA 1. *The following are properties of the Pareto allocation mechanism* g^*:

(i) $\partial\pi_i(g_i^*(a^*(\theta)), \theta_i)/\partial a_i = \partial\pi_j(g_j^*(a^*(\theta)), \theta_j)/\partial a_j$ $\forall i, j$ *such that* $g_i^*(a^*(\theta)) > 0$ *and* $g_j^*(a^*(\theta)) > 0$;

(ii) g^* *is increasing, efficient and individually responsive.*

The proof, and all others, appears in the Appendix.

The Pareto mechanism balances the needs of markets of different sizes by eliminating all profitable trades between retailers. No retailer could find a partner willing to sell at a price he is willing to pay. As a consequence, the Pareto mechanism must recognize the smallest changes in every retailer's marginal valuation of stock and thus must be IR.

Further characterizations of Pareto mechanisms require additional structure. Suppose all retailers face symmetric problems so that for any possible type $\hat{\theta}$, $\pi_i'(a, \hat{\theta}) = \pi_j'(a, \hat{\theta})$, for $i \neq j$ and $a \geq 0$. In a symmetric problem, retailers who draw the same type always place the same value on a marginal unit of stock. The following shows that for many symmetric problems, either proportional or linear allocation is the Pareto mechanism.

LEMMA 2. *Assume that all retailers face symmetric problems.*

(i) *Suppose there exists a function* $\tau(\theta_i)$ *such that* $\tau(\bar{\theta}) = 1$ *for a fixed type* $\bar{\theta}$ *and*:

$$\pi_i'(a, \theta_i) = \pi_i'(a/\tau(\theta_i), \bar{\theta}). \tag{1}$$

Then proportional allocation is the Pareto mechanism.

(ii) *Suppose there exists a function* $\tau(\theta_i)$ *such that* $\tau(\bar{\theta}) = 0$ *for a fixed type* $\bar{\theta}$ *and*:

$$\pi_i'(a, \theta_i) = \pi_i'(a - \tau(\theta_i), \bar{\theta}). \tag{2}$$

Then linear allocation is the Pareto mechanism.

In the first part of the lemma, the retailers' problems differ by a scale parameter while those in the second part differ by a shift parameter. Both cover a variety of problems. For example, suppose the retailers face newsvendor problems with demand distributions taken from the same family F and let θ_i represent a parameter of retailer i's demand distribution. If we have that $F(x \mid \theta_i) = F(x/\tau(\theta_i) \mid \bar{\theta})$ (e.g., an exponential distribution with mean θ_i), the first part of Lemma 2 holds and proportional allocation is the Pareto mechanism. If, however, $F(x \mid \theta) = F(x - \tau(\theta) \mid \bar{\theta})$ (e.g., a normal distribution with mean θ_i and a constant standard deviation), the second part of the lemma holds and linear allocation is the Pareto mechanism. Note that linear allocation may assign an allocation of zero; maximizing retailer profits may require denying stock to particularly small markets.

3.3. Dominant Strategy Equilibria

Retailer profits are maximized if they order $a^*(\theta)$ and the supplier employs the Pareto mechanism. But will retailers truthfully request their optimal allocation or will they inflate their order? We study this question with two equilibrium concepts: dominant strategy equilibria and Bayesian equilibria. In a dominant strategy equilibrium, each retailer has an order that maximizes his profits regardless of the orders of the other retailers. Discussion of Bayesian equilibria is deferred to §3.5. In all cases, we consider only pure strategy equilibria.

Let $x_i(\theta_i)$ be a function mapping from Θ_i to $[a_i^*(\theta), \bar{M}]$. It thus defines a strategy for player i, dictating an order for each possible type. Similarly, $x_{-i}(\theta_{-i})$ denotes the vector of orders submitted by all retailers but retailer i. The functions $x^*(\theta) = \{x_1^*(\theta_1), \ldots, x_N^*(\theta_N)\}$ form a dominant equilibrium if for all i and θ,

$$\pi_i(g_i(x_i^*(\theta_i), m_{-i})) \geq \pi_i(g_i(m_i, m_{-i})) \quad \forall\, m_i \in [a_i^*(\theta), \bar{M}],\ \forall\, m_{-i}.$$

We are particularly interested in allocation mechanism under which the retailers order their optimal allocations in a dominant equilibrium, i.e., $x^*(\theta) = a^*(\theta)$. In those cases we can confidently expect each retailer to order his true needs since that is an optimal strategy no matter how the other retailers behave, even if they are irrational. As the following theorem states, such an equilibrium does not exist for a broad class of allocation mechanisms.

THEOREM 3. *All retailers truthfully reporting their optimal allocations,* $a^*(\theta)$, *is not a dominant equilibrium under an individually responsive allocation mechanism.*

For $a^*(\theta)$ to be a dominant equilibrium a retailer must be willing to order his optimal allocation even when he knows for certain that capacity will be insufficient. However, when capacity binds, some retailer must be below his optimal allocation. By the IR property, he can increase his ultimate allocation, and his profits, by inflating his order. Thus, truth-telling cannot be a dominant equilibrium. If the supplier implements an IR mechanism such as linear allocation, she must accept that the retailers will game the system. In particular, the Pareto allocation mechanism is subject to manipulation.

THEOREM 4. *All retailers truthfully reporting their optimal allocations,* $a^*(\theta)$, *is not a dominant equilibrium under a Pareto allocation mechanism.*

The result follows immediately from Theorem 3 and Lemma 1. A well-intentioned supplier may want to implement the Pareto allocation mechanism to maximize the retailers' profits, but she is thwarted by the retailers' self-interested behavior.

This finding is similar to results in the social choice literature, especially the celebrated impossibility theorem of Gibbard and Satterthwaite (Gibbard 1973, Satterthwaite 1975). They show that any voting scheme satisfying certain properties is open to manipulation. However, our model differs from most work in social choice theory along two important dimensions. First, most of the social choice literature considers the provision of purely public goods as opposed to the division of a finite resource. (Sprumont (1991) and Moulin (1993) are exceptions.) Second, that work generally assumes a mechanism designer who maximizes some measure of social welfare while we consider a self-interested supplier.

3.4. Nothing but the Truth

While we have shown that $a^*(\theta)$ cannot be part of a dominant equilibrium for many mechanisms, we have not established if there exists allocation schemes free from strategic manipulation. From Theorem 3, being non-IR is necessary. A stronger property provides a sufficient condition.

THEOREM 5. *For allocation mechanism $g(m)$, suppose that for all i and (m_i, m_{-i}) such that $g_i(m_i, m_{-i}) < a_i^*(\theta_i)$ there does not exist a $\hat{m}_i$ such that $g_i(\hat{m}_i, m_{-i}) > g_i(m_i, m_{-i})$. Then all retailers truthfully reporting their optimal allocations, $a^*(\theta)$, is a dominant equilibrium.*

The theorem requires that a retailer can never raise his allocation (by increasing his order) when his allocation is less than $a_i^*(\theta_i)$. With non-IR mechanisms a higher order does not guarantee a retailer a higher allocation in all states of the world, but it may raise a retailer's allocation in some states of the world when the retailer is allocated less than $a_i^*(\theta_i)$. Under the conditions of Theorem 5, ordering more than $a_i^*(\theta_i)$ can increase one's allocation only when one would have received $a_i^*(\theta_i)$, making extra stock undesirable.

Several mechanisms satisfy the requirements of Theorem 5. The simplest is *lexicographic allocation*. Retailers are ranked in some manner independent of their order sizes (say, alphabetically) and allocated stock in accordance with that ranking. Retailer i receives the minimum of his order and the as yet unallocated capacity. The scheme induces truth telling because whenever those ranked above i have claimed all of capacity, he has no means by which to increase his allocation. Note that $a^*(\theta)$ remains a dominant equilibrium even if no retailer knows the value of K. Since there is nothing special about any one ranking of the retailers, truth telling is a dominant equilibrium for all $N!$ permutations of the retailers. Consequently, it is also a dominant equilibrium for lotteries over the orderings as long as the chance of selecting a particular permutation is independent of the orders submitted (e.g., filling orders from the largest to smallest is not allowed).

Uniform allocation is also a truth-inducing mechanism (Sprumont 1991). As with linear allocation, index the retailers in decreasing order of their order quantity, i.e., $\{m_1 \geq m_2 \geq \cdots \geq m_N\}$. Retailer i is allocated $g_i(m, \hat{n})$, where

$$g_i(m, \hat{n}) = \begin{cases} \dfrac{1}{\hat{n}}\left(K - \displaystyle\sum_{j=\hat{n}+1}^{N} m_j\right), & i \leq \hat{n}, \\ m_i, & i > \hat{n}, \end{cases}$$

and $\hat{n}$ is the largest integer less than or equal to N such that $g_{\hat{n}}(m, \hat{n}) \leq m_{\hat{n}}$. Uniform allocation always favors small retailers. Despite this inequity, it may result in higher supply chain profits than lexicographic allocation. Lexicographic allocation guarantees a wide spread in marginal valuations of stock as some receive their full orders and others get nothing. Uniform allocation narrows this gap, moving the system closer to the Pareto-mechanism standard of equal marginal valuations. Further, in symmetric problems, retailers drawing the same type receive the same allocation, eliminating profitable trades among them.

3.5. Bayesian Equilibria

In a dominant equilibrium, a retailer must have a single action that is a best response to all realizations of orders, regardless of their likelihood. A Bayesian equilibrium only requires that each retailer maximizes his profits in expectation, assuming other retailers follow the same Bayesian equilibrium. A formal definition requires additional notation. Let $\mu(\theta)$ be the joint distribution of types, and let $\mu_i(\theta_{-i})$ be the joint distribution of the types of all retailers but i conditional on i's type, θ_i. We assume $\mu(\theta)$ (and hence $\mu_i(\theta_{-i})$) is common knowledge. Let $\Pi_i(x_i(\theta_i), x_{-i}(\theta_{-i}))$ equal retailer i's expected payoff given the allocation mechanism g and his type θ_i:

$$\Pi_i(x_i(\theta_i), x_{-i}(\theta_{-i})) = \int_{\Theta_{-i}} \pi_i(g_i(x_i(\theta_i), x_{-i}(\theta_{-i})))\, d\mu_i(\theta_{-i}),$$

where $\Theta_{-i} = \Pi_{j \neq i} \Theta_j$. It holds that $x^*(\theta) = \{x_1^*(\theta_1), \ldots, x_N^*(\theta_N)\}$ is a Bayesian equilibrium if for $i = 1, \ldots, N$ and all θ,

$$\Pi_i(x_i^*(\theta_i), x_{-i}^*(\theta_{-i})) \geq \Pi_i(m_i, x_{-i}^*(\theta_{-i})) \quad \forall\, m_i \in [a_i^*(\theta_i), \bar{M}].$$

Despite imposing less stringent constraints on a retailer's optimal policy, truth telling is not a Bayesian equilibrium under an IR mechanism.

THEOREM 6. *All retailers truthfully reporting their optimal allocations, $a^*(\theta)$, is not a Bayesian equilibrium under an individually responsive allocation mechanism.*

The intuition is the same as in Theorem 3: If the mechanism guarantees one a larger allocation by ordering more, a rational retailer will exploit that guarantee.

3.6. Bayesian Equilibrium with Relaxed Linear Allocation

We have shown that under an IR mechanism truth telling is not a Bayesian equilibrium. We now ask whether any Bayesian equilibrium exists under an IR mechanism. We would like to examine this question in as general a setting as possible, placing no additional restrictions on the form of retailer's profits $\pi_i(a_i, \theta_i)$ or the joint distribution of types θ, $\mu(\theta)$. We therefore turn to the theory of supermodular games.

From Milgrom and Roberts (1990), the Allocation Game is a supermodular game if for all i, $\Pi_i(x_i(\theta_i), x_{-i}(\theta_{-i}))$ has increasing differences in $(x_i(\theta_i), x_{-i}(\theta_{-i}))$. (Additional conditions are confirmed in the proof of Theorem 8.) Increasing differences requires that for all $\hat{m}_i > m_i$ and all $\hat{x}_{-i}(\theta_{-i}) \geq x_{-i}(\theta_{-i})$,

$$\Pi_i(\hat{m}_i, \hat{x}_{-i}(\theta_{-i})) - \Pi_i(m_i, \hat{x}_{-i}(\theta_{-i})) \geq \Pi_i(\hat{m}_i, x_{-i}(\theta_{-i})) - \Pi_i(m_i, x_{-i}(\theta_{-i})). \tag{3}$$

Roughly speaking, when retailer i increases his order from m_i to $\hat{m}_i$ (holding his type θ_i constant), the change in expected profits is larger when the other retailers have placed larger orders. This property implies that when retailer j increases his order when he is type θ_j, retailer i will want to increase his own order regardless of this type.

Whether (3) holds depends on the allocation mechanism the supplier chooses.

THEOREM 7. *$\Pi_i(x_i(\theta_i), x_{-i}(\theta_{-i}))$ has increasing differences in $(x_i(\theta_i), x_{-i}(\theta_{-i}))$ when the allocation mechanism satisfies the following: For all $\hat{m}_i > m_i$ and $\hat{m}_{-i} > m_{-i}$,*

$$g_i(\hat{m}_i, \hat{m}_{-i}) - g_i(m_i, \hat{m}_{-i}) \leq g_i(\hat{m}_i, m_{-i}) - g_i(m_i, m_{-i}), \tag{4}$$

and when $g_i(m_i, m_{-i}) < m_i$,

$$g_i(\hat{m}_i, \hat{m}_{-i}) - g_i(m_i, \hat{m}_{-i}) = g_i(\hat{m}_i, m_{-i}) - g_i(m_i, m_{-i}). \tag{5}$$

The first condition is reasonable: The change in retailer i's allocation as he increases his order from m_i to $\hat{m}_i$ does not increase as others order more. The second condition requires that the change in retailer i's allocation as he increases his order from m_i to $\hat{m}_i$ be constant for any set of orders from the other retailers as long as retailer i receives less than his order in all cases. This is a strong condition; one would expect when capacity is binding that the change in retailer i's allocation would vary with the orders of others. Indeed, this condition does not hold for proportional or linear allocation. Linear allocation fails because capacity assignments must be nonnegative. Suppose retailer j orders a moderate amount, and retailer i receives a positive allocation with $\hat{m}_i$; he must enjoy a gain in his allocation when moving from m_i to $\hat{m}_i$. Now suppose retailer j orders a very large amount and retailer i receives a zero allocation at $\hat{m}_i$. He then gains no increase in allocation when moving from m_i to $\hat{m}_i$. See Cachon and Lariviere (1998) for a simpler model that provides some results under proportional and linear allocation.

Relaxed linear allocation does satisfy the second condition of Theorem 7. Under relaxed linear allocation:

$$g_i(m) = \min\left\{ m_i,\ m_i - \frac{1}{N}\left(\sum_{j=1}^{N} m_j - K\right)\right\}.$$

While the relaxed and original versions of linear allocation yield the same allocations whenever every retailer is given a positive amount, the relaxed mechanism can produce infeasible allocations, i.e., $g_i(m) < 0$ is possible. Despite this complication in implementation, §5 presents numerical evidence that the relaxed version provides an excellent approximation to the true mechanism. Further, the relaxed version allows additional analysis.

THEOREM 8. *Under relaxed linear allocation, the set of pure strategy Bayesian equilibrium is nonempty and possesses upper and lower equilibria,* $\bar{x}(\theta)$ *and* $\underline{x}(\theta)$.

An equilibrium is called an upper equilibrium, $\bar{x}(\theta)$, if there does not exist another Nash equilibrium, $x^*(\theta)$, such that for any i and θ, $\bar{x}_i(\theta) < x_i^*(\theta)$. An equilibrium is called a lower equilibrium, $\underline{x}(\theta)$, if there does not exist another Nash equilibrium, $x^*(\theta)$, such that for any i and θ, $\underline{x}_i(\theta) > x_i^*(\theta)$. When $\bar{x}(\theta) = \underline{x}(\theta)$, there is clearly a unique equilibrium. See Milgrom and Roberts (1990) for a method to evaluate $\bar{x}(\theta)$ and $\underline{x}(\theta)$.

In addition to existence, the supermodular property also facilitates comparative statics. In particular, we are interested in how changes in capacity impact the retailers' orders.

THEOREM 9. *Under relaxed linear allocation, increasing the supplier's capacity lowers each retailer's equilibrium order.*

Theorem 9 implies that the distribution of total retailer orders stochastically declines as capacity increases. Hence, the supplier's capacity choice will affect her expected sales.

4. The Capacity Game

The Allocation Game is the second stage of the Capacity Game. In the first stage, the supplier chooses her allocation mechanism and capacity, anticipating how the retailers will subsequently behave. We assume that capacity costs the supplier c per unit to acquire but that she incurs no additional costs to convert raw capacity into finished goods. Both the mechanism and capacity announcements are credible: The supplier cannot renege on the allocation scheme after seeing the retailers' orders, and each retailer can independently verify that the supplier has exactly K units of capacity available.

For the chosen allocation mechanism g and each capacity level K, assume there exists a Bayesian equilibrium to the Allocation Game. We are assured this holds when the allocation mechanism is either truth-inducing or relaxed linear allocation. Let $\Phi(y \mid K)$ be the distribution function of total demand at the supplier when the supplier chooses capacity K (i.e., the sum of equilibrium retailer orders to the supplier given K and g). Assume $\Phi(y \mid K)$ is differentiable in both y and K. The supplier's expected profits given her chosen mechanism and capacity are

$$\pi_s(K) = w\left[K(1 - \Phi(K \mid K)) + \int_0^K y \, d\Phi(y \mid K)\right] - cK.$$

The supplier's problem is simplest under a truth-inducing mechanism because equilibrium retailer orders are independent of K. The distribution of total demand at the supplier consequently does not vary with K, and the supplier's capacity choice is a standard newsvendor problem. Hence, the supplier's profit-maximizing capacity with a truth-inducing allocation mechanism, K_t, solves

$$\Phi(K_t \mid K_t) = (w - c)/w. \tag{6}$$

Alternatively, the supplier could choose a manipulable mechanism. Here, $\Phi(y \mid K)$ depends on K, since equilibrium orders are not independent of capacity. For example, according to Theorem 9, $x^*(\theta)$ is nonincreasing in K under relaxed linear allocation, making $\Phi(y \mid K)$ nondecreasing in K. Letting K_l equal the capacity that maximizes the supplier's expected profit for a specified manipulable mechanism, the following suggests that the supplier will choose an apparently restrictive capacity given the distribution of orders she faces.

THEOREM 10. *Letting* K_l *equal the supplier's profit maximizing capacity choice given an allocation mechanism for which* $\Phi(y \mid K)$ *is nondecreasing in* K, *it holds that* $\Phi(K_l \mid K_l) \le w - c/w$.

Hence, a supplier observing $\Phi(y \mid K_l)$ would choose a higher capacity than K_l if she (incorrectly) assumed that $\Phi(y \mid K_l)$ were independent of the chosen capacity. Thus, a profit-maximizing supplier may choose a capacity level that leads to chronic and persistent shortages, even if capacity is relatively cheap. While K_l may appear too low relative to the order distribution it induces, it may be greater than K_t. That is, the supplier may choose a higher capacity than she would under a truth-inducing allocation mechanism.

It should be noted that Theorem 10 assumes $\Phi(y \mid K)$ is differentiable in y and K. This assumption does not necessarily hold in equilibrium, so the theorem should be seen as an approximation. However, the numerical study below confirms that it is an excellent one.

To summarize the Capacity Game, the supplier first evaluates her profits given a truth-inducing mechanism and an optimal capacity choice. (Note that both her profits and capacity choice are independent of the truth-inducing scheme used.) These profits are compared to her profits under all manipulable mechanisms under consideration. The supplier then chooses the mechanism and matching capacity that offer the highest expected profit.

5. Numerical Study

To complement our analytical findings, we present a numerical study of the Capacity Game. We consider a setting in which each of five retailers faces linear demand:

$$Q(P) = \theta - P,$$

where P is the price the retailer charges consumers, θ is the retailer's private information, and Q is the quantity the retailer sells. The retailers face symmetric problems and the setting fulfills the requirements of the second part of Lemma 2 (with $\tau(\theta) = (\theta - \bar{\theta})/2$) so linear allocation is the Pareto mechanism. To facilitate computation of equilibria, we assume retailers are independently assigned one of five types, i.e., $\theta \in \{4, 5, 6, 7, 8\}$. (The results of the previous section hold under discrete types with appropriate modifications.) The probabilities that a retailer is assigned these types are {0.05, 0.25, 0.40, 0.25, 0.05}. As the retailers are symmetric, let the subscript denote the retailer's type.

It is straightforward to verify that $a_\theta^* = \max\{0, (\theta - w)/2\}$ and that given an allocation a, a retailer's profits are

$$\pi_\theta(a) = (\theta - a)a - wa,$$

where it is assumed the retailer must bring his full allocation to market. (Allowing the retailer to withhold stock has no qualitative impact.) We consider $w \in \{1, 1.5, \ldots, 4\}$ with $c \in \{0.1w, 0.3w, 0.5w, 0.7w, 0.9w\}$ for a total of 35 problems. We set $\bar{M} = 10a_8^*$.

5.1. Results

Table 1 reports the results for the truth-inducing allocation mechanisms relative to system-optimal performance and the Pareto mechanism.[1] The capacity, K_t, maximizes the supplier's profits assuming a truth-inducing allocation mechanism while K_o would maximize the performance of the integrated system. Note that the decentralized system can provide significantly less capacity than the centralized one. Each value of K_t is a solution to a standard newsvendor problem since retailer orders are independent of the chosen capacity (because a truth-inducing mechanism is implemented). Consequently, the probability that the sum of orders to the manufacturer is less than capacity is constant for a given c to w ratio. This probability is significantly higher than the corresponding probability for the centralized system.

The Pareto mechanism gives an upper bound on expected retailer profits for a given capacity level. One sees that both uniform and lexicographic allocation achieve remarkably close to what the Pareto mechanism would achieve when capacity is relatively inexpensive. However, when capacity becomes expensive (relative to the wholesale price) the supplier builds less capacity. As a result, the allocation rule is implemented more frequently and the performance of the two rules declines. Over all scenarios, uniform allocation provides better performance than lexicographic allocation.

[1] That is, the performance of an integrated system controled by one decision maker. System-optimal actions were found by implementing the Pareto mechanism with $w = 0$ (since there is a zero marginal cost of production) and assuming the retailers order $\theta/2$. The optimal capacity was found by searching over the range $[0, 5\theta_8/2]$.

Table 1 Truth-Inducing Allocation Mechanisms

		Supplier		Expected Retailer Profits (% of Pareto)		System Optimal	
c/w	w	K_t/K_o	P(orders $< K_t$)	Uniform Allocation	Lexicographic Allocation	P(system optimal sales $< K_o$)	Decentralization Penalty*
0.9	4.0	58%	4.8%	93.1%	80.6%	0.0%	21.9%
0.9	3.5	67%	4.8%	95.3%	84.6%	0.0%	14.8%
0.9	3.0	73%	4.8%	96.7%	86.9%	0.0%	10.3%
0.9	2.5	77%	4.8%	97.5%	89.1%	0.0%	7.4%
0.9	2.0	81%	4.8%	98.1%	90.3%	0.0%	5.5%
0.9	1.5	85%	4.8%	98.5%	91.7%	0.0%	4.1%
0.9	1.0	85%	4.8%	98.8%	92.6%	1.6%	3.1%
0.7	4.0	56%	24.0%	96.5%	90.9%	0.0%	25.3%
0.7	3.5	58%	24.0%	97.6%	93.1%	0.0%	15.9%
0.7	3.0	72%	24.0%	98.3%	94.2%	0.0%	11.6%
0.7	2.5	79%	24.0%	98.7%	95.4%	0.0%	7.7%
0.7	2.0	83%	24.0%	99.0%	96.1%	0.0%	5.1%
0.7	1.5	86%	24.0%	99.2%	96.8%	0.4%	3.2%
0.7	1.0	89%	24.0%	99.4%	97.3%	4.8%	1.9%
0.5	4.0	50%	40.7%	98.0%	94.8%	0.0%	31.9%
0.5	3.5	60%	40.7%	98.6%	96.1%	0.0%	22.5%
0.5	3.0	65%	40.7%	99.0%	96.8%	0.0%	15.5%
0.5	2.5	73%	40.7%	99.3%	97.5%	0.1%	10.3%
0.5	2.0	80%	40.7%	99.4%	97.9%	0.4%	6.5%
0.5	1.5	87%	40.7%	99.5%	98.3%	1.6%	3.8%
0.5	1.0	89%	40.7%	99.6%	98.6%	11.8%	1.9%
0.3	4.0	46%	59.3%	99.2%	97.4%	0.1%	37.5%
0.3	3.5	54%	59.3%	99.4%	98.1%	0.4%	27.4%
0.3	3.0	62%	59.3%	99.6%	98.5%	1.6%	19.4%
0.3	2.5	71%	59.3%	99.7%	98.8%	1.6%	13.0%
0.3	2.0	78%	59.3%	99.8%	99.1%	4.8%	8.2%
0.3	1.5	84%	59.3%	99.8%	99.3%	11.8%	4.6%
0.3	1.0	90%	59.3%	99.8%	99.4%	24.0%	2.1%
0.1	4.0	46%	88.2%	99.9%	99.6%	11.8%	41.7%
0.1	3.5	53%	88.2%	99.9%	99.7%	24.0%	31.5%
0.1	3.0	62%	88.2%	99.9%	99.8%	24.0%	22.8%
0.1	2.5	71%	88.2%	99.9%	99.9%	24.0%	15.7%
0.1	2.0	77%	88.2%	100.0%	99.9%	40.7%	9.9%
0.1	1.5	85%	88.2%	100.0%	99.9%	40.7%	5.5%
0.1	1.0	90%	88.2%	100.0%	99.9%	59.3%	2.5%

* Decrease in supply chain profits under uniform allocation, as % of optimal supply chain profits

- The truth-inducing mechanisms we consider perform well when capacity is relatively cheap. Efficiency losses become significant when capacity is expensive.

The *decentralization penalty* relates the performance of a decentralized supply chain under uniform allocation to the centralized supply chain's optimal profits. This penalty can be substantial (up to 42%) and is driven by two observations made above: The supplier in the decentralized system builds too little capacity and her capacity is too often idle.

Table 2 presents results when relaxed linear allocation is implemented, but the capacities are those that

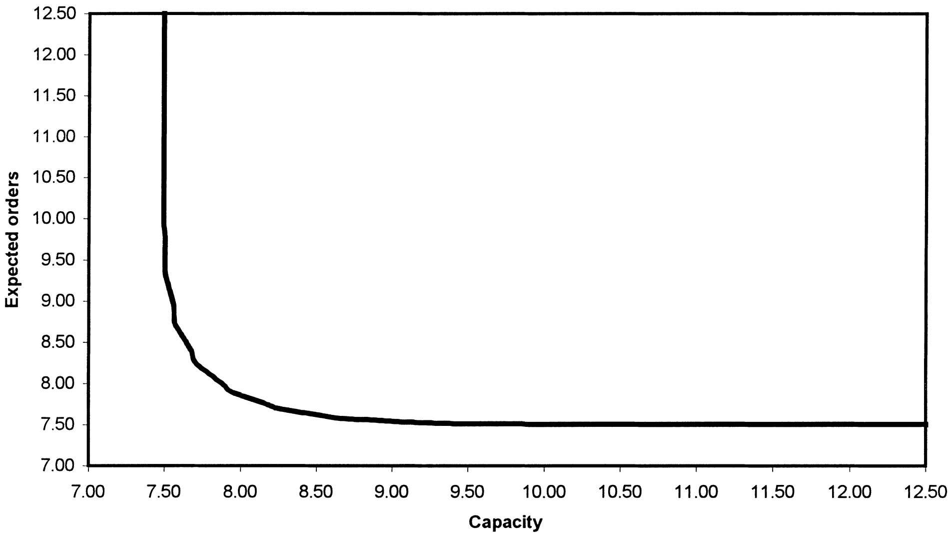
Table 2 Supply Chain Performance Under Linear Allocation with Capacity of K_t

			Change in profits from uniform to linear			
c/w	w	$P(\text{orders} < K_t)$	Chain	Supplier	Retailer	Decen. Penalty
0.9	4.0	0.0%	2.4%	10.8%	0.3%	20.0%
0.9	3.5	0.0%	1.4%	7.8%	0.2%	13.5%
0.9	3.0	0.0%	0.9%	6.1%	0.1%	9.5%
0.9	2.5	0.0%	0.6%	5.0%	0.1%	6.9%
0.9	2.0	0.0%	0.4%	4.2%	0.1%	5.1%
0.9	1.5	0.0%	0.2%	3.6%	0.1%	3.8%
0.9	1.0	0.0%	0.2%	3.2%	0.0%	2.9%
0.7	4.0	0.0%	9.0%	18.8%	1.1%	18.6%
0.7	3.5	0.0%	5.9%	14.1%	0.7%	10.9%
0.7	3.0	0.0%	4.0%	11.3%	0.5%	8.1%
0.7	2.5	0.0%	2.8%	9.4%	0.4%	5.2%
0.7	2.0	0.0%	1.9%	8.1%	0.3%	3.3%
0.7	1.5	0.0%	1.3%	7.1%	0.2%	2.0%
0.7	1.0	0.0%	0.8%	6.3%	0.2%	1.1%
0.5	4.0	4.8%	10.9%	19.0%	−0.5%	24.5%
0.5	3.5	4.8%	7.4%	14.6%	−0.3%	16.8%
0.5	3.0	4.8%	5.2%	11.9%	−0.2%	11.1%
0.5	2.5	4.8%	3.7%	10.0%	−0.2%	7.0%
0.5	2.0	4.8%	2.5%	8.6%	−0.1%	4.1%
0.5	1.5	4.8%	1.7%	7.6%	−0.1%	2.2%
0.5	1.0	4.8%	1.0%	6.8%	−0.1%	1.0%
0.3	4.0	59.3%	3.8%	5.4%	0.5%	35.1%
0.3	3.5	59.3%	2.7%	4.3%	0.4%	25.4%
0.3	3.0	59.3%	2.0%	3.5%	0.3%	17.7%
0.3	2.5	59.3%	1.5%	3.0%	0.2%	11.7%
0.3	2.0	59.3%	1.1%	2.6%	0.1%	7.2%
0.3	1.5	59.3%	0.8%	2.3%	0.1%	3.9%
0.3	1.0	59.3%	0.5%	2.0%	0.1%	1.6%
0.1	4.0	88.2%	0.5%	0.6%	0.1%	41.4%
0.1	3.5	88.2%	0.4%	0.5%	0.1%	31.2%
0.1	3.0	88.2%	0.3%	0.4%	0.1%	22.6%
0.1	2.5	88.2%	0.2%	0.4%	0.0%	15.5%
0.1	2.0	88.2%	0.2%	0.3%	0.0%	9.8%
0.1	1.5	88.2%	0.1%	0.3%	0.0%	5.4%
0.1	1.0	88.2%	0.1%	0.2%	0.0%	2.4%

maximize the supplier's profits under a truth-inducing mechanism. While these capacities generally do not maximize the supplier's profits given relaxed linear allocation, we provide these data to highlight the impact of only switching from a truth-inducing allocation mechanism to a manipulable allocation mechanism. In each scenario the data are from the lower equilibrium, which has the least order inflation of all possible equilibria. All of these equilibria are also equilibria under linear allocation, which indicates that the relaxed version of the linear rule provides an excellent approximation of the feasible version. Further, no retailer receives a negative allocation in any of these equilibria, so all allocations are implementable. The remaining data we report are also for the relaxed version.

The table indicates that switching from uniform allocation to linear allocation generally raises supplier

Graph 1 Expected Orders with Linear Allocation (w = 3.0, c = 1.5, Lower Equilibrium)



profits. Further, since retailers' profits change little, total supply chain profits tend to increase, reducing the decentralization penalty. There is a simple intuition for this result. The marginal change in the supplier's profits when a retailer increases his order is w. At $a_i^*(\theta_i)$, there is a zero marginal change in the retailer's profits when his allocation is increased (because $\pi_i(a_i, \theta_i)$ is strictly concave in a_i). Hence, inducing the retailers to increase their orders above $a^*(\theta)$ has little impact on their profits but a significant impact on the supplier's profits, benefiting the supply chain. In short, reducing idle capacity in this example is more important than a perfect allocation of capacity among the retailers.

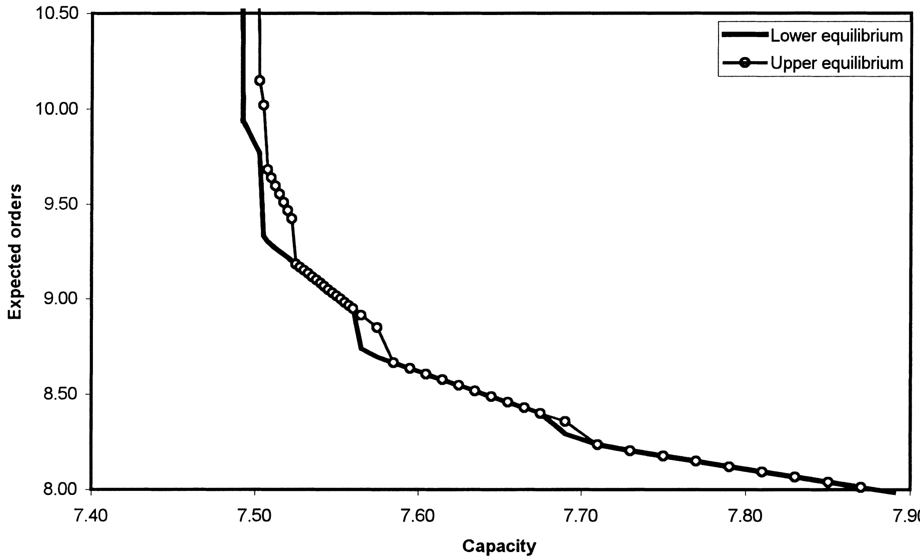
Graph 1 presents the supplier's expected retailer orders from the lower equilibrium as a function of her chosen capacity for one scenario. As suggested by Theorem 9, the supplier's expected orders decline as she builds more capacity, but the theorem offers little insight into the rate of decline. Graph 1 indicates that the orders the supplier expects are relatively insensitive to her capacity decision when capacity is ample, but as capacity is reduced to a critical level, retailer orders rise dramatically. The phenomenon is driven by a "cascade of expectations." Retailer i raises his order to gain a better allotment, which lowers the expected allocation of others. They respond by inflating their orders, inducing retailer i to raise his further. For all other scenarios studied, the analogous graph displays the same pattern.

- As capacity is reduced, the pattern of retailer orders under linear allocation suggests an "all-or-nothing" phenomenon: either the supplier receives a moderate level of orders, or she receives an avalanche.

Graph 2 presents both the lower and upper equilibria from the same scenario displayed in Graph 1. The scale has been reduced so that the equilibria are distinguishable. For moderately tight capacity we observed in all scenarios that the lower and upper equilibria were identical, indicating that there is a unique equilibrium. Multiple equilibria are only observed with tight capacity, but even in these cases the gap between the lower and upper equilibria is small.

Table 3 displays the supplier's optimal capacity choice assuming she implements linear allocation, K_l. In all scenarios the supplier chooses the maximum capacity that guarantees her 100% utilization. This demonstrates the very strong incentive to restrict capacity under a manipulable mechanism. Not surprisingly, the supplier always gains by implementing a manipulable mechanism because she faces a lower risk of idle capacity. However, the retailers can also be better off, because the supplier may build more capacity, especially when the wholesale price is high.

Graph 2 Expected Orders with Linear Allocation (w = 3.0, c = 1.5)



• The supply chain generally suffers from adopting truth-inducing mechanisms when the supplier jointly chooses her mechanism and capacity.

Table 3 indicates that the decentralization penalty can still be substantial under linear allocation when capacity is cheap (although it is always lower then when capacity is fixed at K_t). This occurs because the supplier chooses to restrict capacity far below its optimal level. However, when capacity is expensive, the supplier's capacity choice begins to approach the optimal capacity, so the competition penalty is reduced substantially. For example, when $w = 4$ and $c = 3.6$, the competition penalty with uniform allocation is 22% but it is only 3% with linear allocation. In this case, capacity increases from 58% of the system optimal level to 83% of that level.

6. Conclusion

We investigate a model with one supplier and N independent retailers. The retailers enjoy local monopolies in the consumer market but compete via their orders for scarce supplier capacity. We show that some allocation mechanisms induce the retailers to place their optimal order (thereby revealing their private information), while others lead the retailers to inflate their orders in an effort to gain a better allotment of stock. If the retailers were to report their private information truthfully, a benevolent dictator could allocate scarce inventory to maximize retailer profits. We show that this is not possible because the allocation mechanism that maximizes retailer profits provides incentives for them to misrepresent their needs. While there exist reasonable allocation mechanisms that induce truth telling, we find that in a broad sample of scenarios the supply chain is better off under an allocation mechanism that induces retailers to inflate their orders.

We do not wish to conclude that truth telling harms a supply chain. Instead, we conclude that truth telling provides some advantages to the supply chain that should be weighed against the costs of inducing truth telling. In our setting, three factors influence the performance of the system: How much capacity is built, how that capacity is utilized, and how the resulting output is split among retailers. The truthful revelation of information improves performance along the third dimension but can hinder results along the other two. In the economics literature, it has long been known that decentralized supply chains are less profitable than integrated enterprises because of "double marginalization" (Spengler 1950). Each retailer orders too little stock, relative to the system-wide optimal

Table 3 Supply Chain Performance Under Linear Allocation with Capacity of K_l

			Change in profits from uniform to linear			
c/w	w	K_l/K_o	Chain	Supplier	Retailer	Decen. penalty
0.9	4.0	83%	23.9%	58.0%	15.4%	3.2%
0.9	3.5	88%	15.1%	41.6%	9.9%	1.9%
0.9	3.0	91%	10.2%	32.4%	6.9%	1.2%
0.9	2.5	93%	7.2%	26.5%	5.1%	0.7%
0.9	2.0	95%	5.3%	22.5%	3.9%	0.5%
0.9	1.5	98%	3.9%	19.5%	3.1%	0.3%
0.9	1.0	96%	3.0%	17.2%	2.5%	0.2%
0.7	4.0	62%	15.6%	31.7%	2.8%	13.6%
0.7	3.5	62%	10.3%	23.8%	1.8%	7.2%
0.7	3.0	77%	7.0%	19.1%	1.3%	5.4%
0.7	2.5	83%	4.9%	15.9%	1.0%	3.2%
0.7	2.0	87%	3.4%	13.7%	0.8%	1.9%
0.7	1.5	90%	2.3%	12.0%	0.6%	1.0%
0.7	1.0	93%	1.5%	10.6%	0.5%	0.4%
0.5	4.0	50%	11.3%	19.8%	−0.8%	24.3%
0.5	3.5	59%	7.7%	15.2%	−0.5%	16.6%
0.5	3.0	65%	5.4%	12.4%	−0.4%	10.9%
0.5	2.5	73%	3.8%	10.4%	−0.3%	6.9%
0.5	2.0	80%	2.6%	9.0%	−0.2%	4.1%
0.5	1.5	86%	1.7%	7.9%	−0.2%	2.2%
0.5	1.0	89%	1.0%	7.1%	−0.1%	1.0%
0.3	4.0	42%	6.7%	11.4%	−2.9%	33.3%
0.3	3.5	50%	4.7%	8.9%	−2.0%	24.0%
0.3	3.0	58%	3.3%	7.3%	−1.4%	16.7%
0.3	2.5	67%	2.3%	6.2%	−1.1%	11.0%
0.3	2.0	74%	1.6%	5.4%	−0.8%	6.8%
0.3	1.5	80%	1.0%	4.7%	−0.7%	3.7%
0.3	1.0	86%	0.4%	4.2%	−0.5%	1.7%
0.1	4.0	36%	1.9%	4.1%	−4.3%	40.6%
0.1	3.5	43%	1.3%	3.2%	−2.9%	30.6%
0.1	3.0	52%	0.8%	2.7%	−2.1%	22.2%
0.1	2.5	60%	0.5%	2.3%	−1.6%	15.3%
0.1	2.0	67%	0.2%	2.0%	−1.2%	9.7%
0.1	1.5	75%	0.0%	1.8%	−1.0%	5.5%
0.1	1.0	81%	−0.2%	1.6%	−0.8%	2.7%

order, because his marginal cost is higher than the supply chain's marginal cost. In addition, the supplier chooses capacity that is too low, again relative to the supply chain optimal capacity, because her marginal benefit for each additional unit of capacity is less than the supply chain's.

In the cases we consider, capacity choice and utilization are the dominant concerns. Implementing an individually responsive allocation mechanism, such as linear allocation, mitigates both of these problems. Retailers increase their orders to compete for capacity, and the supplier may build more capacity. The supply chain approaches the outcome of an integrated firm. While linear allocation cannot guarantee the best allocation of capacity, it nevertheless achieves a reasonably good allocation since retailers with greater need receive more stock. Of course, different settings may change the balance between the factors driving performance. For example, the less elastic retail demand is, the less retailer orders will change with the wholesale price. Double marginalization will not pose as significant a problem, and assuring high capacity utilization may become less important than directing allocations efficiently. A truth-inducing mechanism may then be preferable.

Although manipulable mechanisms may lead to a significant distortion of orders, we emphasize that retailers in our model act rationally. It is tempting to label inflated orders as "sub-optimal" or "irrational," but such characterizations are likely incorrect. Similarly, the rational supplier may consistently appear to serve only a tiny fraction of the potential market, but it is the irrational supplier who discovers that, paradoxically, her market collapses when she builds more capacity.

Our model demonstrates that a seemingly destructive behavior like dramatic order inflation can actually improve the supply chain performance. We believe our model is a good representation of many industries, but additional work is needed to determine the robustness of our results. In particular, future research should explore models with pricing schedules, multiple products, and multiple time periods. The latter expands the class of feasible allocation mechanisms (e.g., allocations based on past sales), and potentially introduces the issue of cyclical demand. Overall, we feel there are rich opportunities to explore other supply chain models that integrate traditional operations decisions (like capacity planning) with the strategic decisions of independent firms seeking to maximize their own welfare.[2]

[2] The authors would like to thank Jim Anton, Hervé Moulin, Sridhar Tayur, Paul Zipkin, as well as the seminar participants at Carnegie-Mellon University, Duke University, INSEAD, Stanford University, The Wharton School, and the 1996 MSOM Conference at the Amos Tuck School, Dartmouth College. The helpful comments of the Associate Editor and reviewers are graciously acknowledged.

Appendix.

LEMMA 1. *The Pareto allocation is found from the following program*:

$$\operatorname*{Max}_{a \in \mathcal{A}} \sum_{i=1}^{N} \pi_i(a_i, \theta_i).$$

The Kuhn-Tucker necessary conditions, $\partial \pi_i(a_i, \theta_i)/\partial a_i = \lambda \geq 0\ \forall i$ such that $a_i > 0$, prove (i). For (ii), increasing and efficient is obvious. If $\sum_{i=1}^{N} a_i^*(\theta_i) < K$, $\lambda = 0$ and $a_i = a_i^*(\theta_i)\ \forall i$ is optimal and IR. If orders exceed capacity, $\lambda > 0$ and $a_i < a_i^*(\theta_i)\ \forall i$. Suppose i currently receives a positive quantity, and consider the situation if he had a higher type, $\hat{\theta}_i > \theta_i$. He orders $a_i^*(\hat{\theta}_i) > a_i^*(\theta_i)$ while other orders remain unchanged. λ must increase and a_j falls for $i \neq j$. As g^* is efficient, a_i must rise, and g^* is IR. □

LEMMA 2. *Note that (1) implies that $a_i^*(\theta_i) = \tau(\theta_i) a_i^*(\bar{\theta})$. Additionally, using Lemma 1, we must have that $g_i^*(a^*(\theta)) = \tau(\theta_i) g_i^*(a^*(\bar{\theta}))$. Assuming capacity availability binds, we have that $g^*(a^*(\bar{\theta})) = K/\sum_i^N \tau(\theta_i)$, making proportional allocation optimal. For (ii), (2) implies that $a_i^*(\theta_i) = a_i^*(\bar{\theta}) + \tau(\theta_i)$ and that $g_i^*(a^*(\theta)) = g_i^*(a^*(\bar{\theta})) + \tau(\theta_i)$. Assuming that when capacity binds the base type $\bar{\theta}$ receives stock, it must be that*

$$\sum_{i=1}^{\tilde{n}} g_i^*(a^*(\theta)) = \tilde{n} g_i^*(a^*(\bar{\theta})) + \sum_{i=1}^{\tilde{n}} \tau(\theta_i) = K,$$

which yields $g_i^*(a^*(\bar{\theta})) = (K - \sum_{i=1}^{\tilde{n}} \tau(\theta_i))/\tilde{n}$. Linear allocation is then the Pareto mechanism. □

THEOREM 3. *By assumption $\sum_{j=1}^{N} a_j^*(\theta_j) > K$ for some θ. Hence, $g_i(a_i^*(\theta_i), m_{-i}) < a_i^*(\theta_i)$ for some i. For $a^*(\theta)$ to be a dominant equilibria, i must always maximize his profits by ordering $a_i^*(\theta_i)$. We show that this is not so. If $g_i(m_i, m_{-i})$ is continuous in m_i at $m_i = a_i^*(\theta_i)$, there exists an $\epsilon > 0$ such that $g_i(a_i^*(\theta_i), m_{-i}) < g_i(a_i^*(\theta_i) + \epsilon, m_{-i}) \leq a_i^*(\theta_i)$. As $\pi_i(a_i)$ is increasing for $a_i < a_i^*(\theta_i)$, truth-telling cannot be optimal. If $g_i(m_i, m_{-i})$ is discontinuous in m_i at $m_i = a_i^*(\theta_i)$, consider a decreasing sequence m_i^n such that $\lim_{n\to\infty} m_i^n = a_i^*(\theta_i)$. As $g_i(m_i^n, m_{-i}) \leq m_i^n$ for all n, $\lim_{n\to\infty} g_i(m_i^n, m_{-i}) \leq a_i^*(\theta_i)$. For strict inequality, an $\epsilon > 0$ exists as before. For equality, $g_i(a_i^*(\theta_i) + \delta, m_{-i}) > a_i^*(\theta_i)$ for $\delta > 0$, although one may get arbitrarily close to $a_i^*(\theta_i)$. By concavity of $\pi_i(a_i)$, there exists a δ^* such that $\pi_i(g_i(a_i^*(\theta_i) + \delta, m_{-i})) > \pi_i(g_i(a_i^*(\theta_i), m_{-i}))$ for all $\delta \in (0, \delta^*)$; ordering $a_i^*(\theta_i)$ is not optimal.* □

THEOREM 5. *If $g_i(a_i^*(\theta_i), m_{-i}) = a_i^*(\theta_i)$, i prefers truth telling. If $g_i(a_i^*(\theta_i), m_{-i}) < a_i^*(\theta_i)$, $g_i(a_i^*(\theta_i), m_{-i}) \geq g_i(m_i, m_{-i})$ for any $m_i > a_i^*(\theta_i)$, making truth-telling optimal.* □

THEOREM 6. *Assume i orders $a_i^*(\theta_i) + \epsilon$ for some $\epsilon > 0$. From the definition of individually responsive, $g_i(a_i^*(\theta_i) + \epsilon, x_{-i}(\theta_{-i})) > g_i(a_i^*(\theta_i), x_{-i}(\theta_{-i}))$. Since a retailer never is allocated more than the retailer orders, $g_i(a_i^*(\theta_i) + \epsilon, x_{-i}(\theta_{-i})) \leq a_i^*(\theta_i) + \epsilon$. Hence, there exists an ϵ such that*

$$\int_{\Theta_{-i}} (\pi_i(g_i(a_i^*(\theta_i) + \epsilon, x_{-i}(\theta_{-i}))) - \pi_i(g_i(a_i^*(\theta_i), x_{-i}(\theta_{-i})))) \, d\mu_i(\theta_{-i}) > 0.$$

The above holds $\forall \theta_{-i}$ such that $g_i(a_i^*(\theta_i) + \epsilon, x_{-i}(\theta_{-i})) \leq a_i^*(\theta_i)$. When $g_i(a_i^*(\theta_i) + \epsilon, x_{-i}(\theta_{-i})) > a_i^*(\theta_i)$, it holds that

$$\lim_{\epsilon\to 0} \pi_i(g_i(a_i^*(\theta_i) + \epsilon, x_{-i}(\theta_{-i}))) - \pi_i(g_i(a_i^*(\theta_i), x_{-i}(\theta_{-i})) \geq 0,$$

since $\lim_{\epsilon \to 0} g_i(a_i^*(\theta_i) + \epsilon, x_{-i}(\theta_{-i})) \le a_i^*(\theta_i)$. $\square$

THEOREM 7. *It is easy to show that if each function in a set has increasing differences in its arguments, then any convex combination of the functions also has increasing differences in its arguments. Therefore, to show (3) it is sufficient to demonstrate that the players' profits functions have increasing differences in* (m_i, m_{-i}) *for each realization of* θ. *Define* $\kappa = \sum_{j \ne i}^{N} m_j - K$. *If a retailer's profit function has increasing differences in* (m_i, κ), *then it has increasing differences in* (m_i, m_{-i}). *Note that* $\kappa > 0$ *is possible. For convenience define* $g_i(m_i, \kappa)$ *as retailer i's allocation. Since* $\mu(\theta)$ *is independent of the retailers' orders, (3) holds if for each realization of* θ,

$$\pi_i(g_i(\hat{m}_i, \hat{\kappa})) - \pi_i(g_i(m_i, \hat{\kappa})) \ge \pi_i(g_i(\hat{m}_i, \kappa)) - \pi_i(g_i(m_i, \kappa)), \quad (7)$$

where $\hat{\kappa} > \kappa$. *Define* $z_i(a, \delta) = \pi_i(a + \delta) - \pi_i(a)$, *where* $\delta \ge 0$ *is assumed. Since* π_i *is strictly concave, it is easy to show the following:*

(i) $z_i(a, \delta)$ *is decreasing in* a,
(ii) if $z_i(a, \delta) \ge 0$, *then* $z(a, \beta\delta) \ge 0$ *for* $\beta \in [0, 1]$,
(iii) $z_i(a, \delta) < 0$, *then* $z(a, \beta\delta) \ge z(a, \delta)$ *for* $\beta \in [0, 1]$.

Define $\hat{\delta} = g_i(\hat{m}_i, \kappa) - g_i(m_i, \kappa)$. *From (4),*

$$g_i(\hat{m}_i, \hat{\kappa}) - g_i(m_i, \hat{\kappa}) = \beta\hat{\delta} \quad (8)$$

for some $\beta \in [0, 1]$. *Rewrite (7) as*

$$z_i(g_i(m_i, \hat{\kappa}), \beta\hat{\delta}) \ge z_i(g_i(m_i, \kappa), \hat{\delta}). \quad (9)$$

Several cases are considered to confirm (9). Say $g_i(m_i, \kappa) < m_i$. *From (i) and* $g_i(m_i, \hat{\kappa}) < g_i(m_i, \kappa)$,

$$z_i(g_i(m_i, \hat{\kappa}), \beta\hat{\delta}) \ge z_i(g_i(m_i, \kappa), \beta\hat{\delta}). \quad (10)$$

From (5), $\beta = 1$, *so (10) and (9) are the same. Say* $m_i = g_i(m_i, \kappa)$. *Since* $m_i \ge a_i^*(\theta_i)$, $\hat{\kappa} > \kappa$ *and* $\pi_i'(a \mid a \ge a_i^*(\theta_i)) < 0$,

$$0 > z_i(g_i(m_i, \kappa), \hat{\delta}). \quad (11)$$

Suppose $z_i(g_i(m_i, \hat{\kappa}), \hat{\delta}) \ge 0$. *From (ii), this implies* $z_i(g_i(m_i, \hat{\kappa}), \beta\hat{\delta}) \ge 0$, *which when combined with (11) confirms (9). Now suppose*

$$z_i(g_i(m_i, \hat{\kappa}), \hat{\delta}) \le 0. \quad (12)$$

From (i) and $g_i(m_i, \hat{\kappa}) \le g_i(m_i, \kappa)$,

$$z_i(g_i(m_i, \hat{\kappa}), \hat{\delta}) \ge z_i(g_i(m_i, \kappa), \hat{\delta}). \quad (13)$$

From (iii) and (12), $z_i(g_i(m_i, \hat{\kappa}), \beta\hat{\delta}) \ge z_i(g_i(m_i, \hat{\kappa}), \hat{\delta})$, *which when combined with (13) confirms (9).* $\square$

THEOREM 8. *From Milgrom and Roberts (1990) this game is supermodular if (1) each player's strategy is bounded and a sublattice, (2) each player's payoff function has increasing differences in* (m_i, m_{-i}), *and (3)* π_i *is supermodular in* m_i. *Since each player has a bounded single dimensional strategy space, the first and third conditions are trivial to confirm. Theorem 7 confirms the second condition. Existence of upper and lower equilibria follows immediately (see Milgrom and Roberts 1990).* $\square$

THEOREM 9. *Theorem 7 confirms that each player's profit has increasing differences in* (m_i, κ), *so each player's profit has increasing differences in* $(m_i, -K)$. *Hence, as K increases, each retailer will lower his order quantity, no matter his type.* $\square$

THEOREM 10. *The supplier's profits may be written as:*

$$\pi_s(K) = (w - c)K - w \int_0^K \Phi(y \mid K)\, dy.$$

Assuming that $\Phi(y \mid K)$ *is differentiable with respect to* K, *first order necessary conditions are:*

$$(w - c) - w\Phi(K \mid K) - w \int_0^K \frac{\partial}{\partial K} \Phi(y \mid K)\, dy = 0,$$

and K_l *is implicitly defined by:*

$$\Phi(K_l \mid K_l) = \frac{w - c}{w} - \int_0^{K_l} \frac{\partial}{\partial K} \Phi(y \mid K_l)\, dy.$$

The result then follows from $\Phi(y \mid K)$ *being nondecreasing in* K. $\square$

References

Abernathy, F. H., J. T. Dunlop, J. Hammond, D. Weil. 1995. The information-integrated channel: A study of the U.S. apparel industry in transition. *Brookings Papers on Microeconomic Activity* 175–247.

Anand, K., H. Mendelson. 1997. Information and organization for horizontal multimarket coordination. *Management Sci.* **43** 1609–1627.

Bassok, Y., R. Anupindi. 1998. Analysis of supply contracts with commitments and flexibility. Working Paper, Northwestern University, Evanston, IL.

Bell, C. E., S. Stidman. 1983. Individual versus social optimization in the allocation of customers to alternative servers. *Management Sci.* **29** 831–839.

Cachon, G., M. Lariviere. 1998. An equilibrium analysis of linear and proportional allocation of scarce capacity. Forthcoming in *IIE Trans.*

Dewan, S., H. Mendelson. 1990. User delay costs and internal pricing for a service facility. *Management Sci.* **36** 1502–1517.

Donohue, K. L. 1996. Supply contracts for fashion goods: Optimizing channel profits. Working paper, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Eppen, G. D., A. V. Iyer. 1997. Backup agreements in fashion buying—The value of upstream flexibility. *Management Sci.* **43** 1469–1484.

Fisher, M. L., J. H. Hammond, W. R. Obermeyer, A. Raman. 1994. Making supply meet demand in an uncertain world. *Harvard Bus. Rev.* **72**(3) 83–89.

——, A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* **44** 87–99.

Frank, R. 1996. Frito-Lay puts up more than chips in deal for Olestra. *Wall Street Journal* **May 31** A–3.

Gale, I. L., T. J. Holmes. 1993. Advance-purchase discount and monopoly allocation of capacity. *Amer. Econom. Rev.* **83** 135–146.

Gary, M. 1993. A $30 billion windfall? *Progressive Grocer* **72** 7.

Gibbard, A. 1973. Manipulation of voting schemes: A general result. *Econometrica* **41** 587–601.

Gruley, B., J. Pereira. 1996. FTC says Toys R Us competes unfairly. *Wall Street Journal* **May 23** A–3.

Ha, A. Y. 1997. Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Sci.* **43** 1093–1103.

Harris, R., A. Raviv. 1981. A theory of monopoly pricing schemes with demand uncertainty. *Amer. Econom. Rev.* **71** 347–365.

Henderson, A. B. 1995. U.S. jury convicts ex-Honda officials of taking bribes. *Wall Street Journal* **June 2** B–5.

Hwang, S. L., L. Valeriano. 1992. Marketers and consumers get the jitters over severe shortages of nicotine patches. *Wall Street J.* **May 22** B–1.

Kaplan, A. 1969. Stock rationing. *Management Sci.* **15** 260–267.

Laffont, J. J. 1988. *Fundamentals of Public Economics*. The MIT Press, Cambridge, MA.

Lederer, P. J., L. Li. 1994. Pricing, production, scheduling, and delivery-time competition. *Oper. Res* **45** 407–420.

Lee, H., V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Sci.* **43** 546–558.

Maskin, E., J. Riley. 1989. Optimal multi-unit auctions. Frank Hahn ed., *The Economics of Missing Markets, Information, and Games*. Oxford University Press, New York.

Mendelson, H., S. Whang. 1990. Optimal incentive-compatible priority pricing for the $M/M/1$ queue. *Oper. Res.* **38** 870–883.

Milgrom, P., J. Roberts. 1990. Rationalizability, learning and equilibrium in games with strategic complementarities. *Econometrica* **58** 1255–1278.

Moulin, H. 1993. On the fair and coalitions-strategy proof allocation of private goods. K. Binmore, A. Kirman, and P. Tani (Eds.), *Frontiers of Game Theory*. The MIT Press, Cambridge, MA.

Nahmias, S., W. S. Demmy. 1981. Operating characteristics of an inventory systems with rationing. *Management Sci.* **27** 1236–1244.

Oren, S., S. Smith, R. Wilson. 1983. Competitive nonlinear tariffs. *J. Econom. Theory* **29** 49–71.

Peck, J. 1996. Demand uncertainty, incomplete markets, and the optimality of rationing. *J. Econom. Theory* **70** 342–363.

Porteus, E., S. Whang. 1991. On manufacturing/marketing incentives. *Management Sci.* **37** 1166–1181.

Satterthwaite, M. 1975. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *J. Econom. Theory* **10** 187–217.

Spengler, J. 1950. Vertical integration and antitrust policy. *J. Political Econom.* 347–352.

Sprumont, Y. 1991. The division problem with single-peaked preferences: A characterization of the uniform allocation rule. *Econometrica* **59** 509–519.

Topkis, D. 1968. Optimal ordering and rationing policies in a nonstationary dynamic inventory model with n demand classes. *Management Sci.* **15** 160–176.

Tsay, A., W. Lovejoy. 1998. Quantity flexibility contracts and supply chain performance. Forthcoming in *Manufacturing and Service Oper. Management*.

Accepted by Hau Lee; received December 12, 1996. This paper has been with the authors 6 months for 2 revisions.