Cooperation vs. Non-Cooperation: Balancing Inventory and Stockout Risk in Retail Supply Chains Using Fast-Ship

Hao-Wei Chen† • Diwakar Gupta† • Haresh Gurnani‡

†Program in Industrial & Systems Engineering
University of Minnesota, Minneapolis, MN 55455

‡Department of Management, School of Business Administration,
University of Miami, Coral Gables, FL 33124

In this paper, we consider a two-retailer and one-supplier supply chain in which retailers make an extra effort to satisfy demand from customers who experience stockouts — retailers offer to direct-ship out-of-stock items on an expedited basis at no extra cost to customers. This practice is referred to as the fast-ship option. By supporting the fast-ship option, retailers and suppliers can reduce stockout risk without excessively increasing any one party’s inventory risk (excess inventory). These risks are also affected by retailers’ decision to cooperate or not when either one of them runs out of stock, and customers decision to search for the item at the other retailer when their first-choice retailer stocks out. We characterize the supplier’s and the retailers’ optimal decisions and identify situations in which cooperation may not be the retailers’ best strategy. Similarly, having noncooperative retailers may not maximize the supplier’s profit among the two configurations considered. We also identify different regions of parameter values for which the retailers prefer to cooperate.

Keywords: fast-ship option, newsvendor model, Nash equilibrium, retailer cooperation, customer search.
1. Introduction

Matching supply and demand is a perennial problem for retailers that have a short selling season and uncertain demand. A direct consequence of demand uncertainty is that retailers in the vast majority of circumstances end up with either too much or too little inventory, both of which have costly consequences. Too much inventory causes markdowns (often referred to as the inventory risk) and too little inventory leads to stockouts (referred to as the stockout risk). Recognizing that both these types of risks are undesirable, many papers in the Operations Management (OM) literature focus on strategies to help newsvendor-type retailers reduce stockouts while preventing inventory risk from increasing significantly, thereby reducing overall cost of matching supply with demand. These strategies include the use of different types of pooling (e.g. demand, location, and lead time) and the use of reactive capacity. Reactive capacity gives retailers the ability to place a second (and possibly third) order after the selling season begins and a portion of demand is observed, which shifts some inventory risk to suppliers.

Offering the fast-ship option is one of several strategies that uses a second replenishment opportunity to reduce supply-demand mismatch costs (Gupta et al. 2010). Fast-ship option is exercised only after stockout occurs. It allows the retailer to satisfy demand from those customers who are willing to wait a short amount of time to have the item direct-shipped to their addresses. The retailer arranges to have out-of-stock items shipped directly from the supplier at no additional cost to the customers. That is, it absorbs the higher wholesale prices associated with expedited delivery. In this way, the fast-ship option provides a mechanism by which the stockout risk and the inventory risk can be allocated among retailers and suppliers.

In multiple retailer markets, in addition to arranging for fast shipping, retailers can also cooperate with each other in an arrangement by agreeing to transship items among each other and share revenue from such sales. Alternatively, when they do not cooperate, some customers will perform a search on their own before placing fast-ship orders. In an environment in which fast-ship option is available to customers, it is not immediately apparent whether the retailers will benefit more from cooperation (i.e. transshipping) or letting customers search on their own. In this paper, we develop mathematical models to compare the two scenarios — those in which retailers cooperate, which we refer to as the cooperative model (Model C), and those in which they do not, which we refer to as
the noncooperative model (Model N). A schematic of these arrangements is shown in Figure 1.

![Figure 1: Model Illustration](image)

Fast-ship option is used by a number of retailers. For example, Best Buy, an electronic products retailer, offers an Instant Ship option to in-store customers when mobile telephone devices they intend to purchase are not available in store. Apparel retailers like J. Crew and Gap Inc. offer to home-ship items when customers cannot find desired items in the right sizes on the store shelf, absorbing the shipping cost. Office Depot ships ink and toner cartridges free to customers if it stocks out of such items. Similarly some retailers transship and others do not. For example, some auto dealers cooperate with each other and transship vehicles and/or spare parts to meet customer demand (see Jiang and Anupindi 2010, Zhao and Atkins 2009, Shao et al. 2011, and Kwan and Dada 2005) and others do not do so. When the latter happens, some customers who wish to purchase the item may perform a search on their own before backordering. In this paper, we study the attractiveness of the option to cooperate when fast-ship option is also available.

In model C, each retailer places an initial order to satisfy its own demand. In addition, it uses its leftover inventory to fulfill the other retailer’s excess demand. Any excess demand that cannot be fulfilled by inventory sharing between retailers is satisfied by the supplier. Transshipment is assumed to occur instantaneously. Thus, the customers are indifferent between either purchasing an on-the-shelf item or an item that is transshipped. They are also aware that retailers cooperate, which makes it futile to search if their first-choice retailer informs them that the only option to obtain the item is the fast-ship option. When retailers cooperate, the extra revenue (net of transshipment cost) arising from their alliance are allocated between them according to a revenue-sharing rule.

In Model N, retailers do not cooperate. If a retailer stocks out, some of its customers who
experience stockout would want to purchase the item by other means. A fraction of these search on their own and another fraction right away place a fast-ship order. The relative sizes of these fractions will depend on the customer search costs, which are assumed exogenous in our paper. Customers who search the other retailer’s inventory will place a fast-ship order with that retailer if that retailer is also stocked out. Thus, in Model C, retailers preempt customer search by performing the search for the customers, whereas in Model N they let customers decide. In both models, a fraction of customers may remain unserved because they do not wish to wait for future delivery.

We show that a pure strategy Nash equilibrium exists for the retailers’ ordering decisions in Models C and N. We also obtain nonlinear equations whose solutions provide the supplier’s and the retailers’ optimal parameter values in both these configurations. We establish that the supplier will always set the fast-ship wholesale price equal to the retailers revenue per sale in scenarios where it sets the wholesale prices. We then compare the retailers’ and the supplier’s profits under the two models. Most of these comparisons rely on numerical experiments with either high or low customer search costs. In these comparisons, we differentiate between cases in which wholesale prices are determined by market forces and cases in which wholesale prices are determined by the supplier.

When wholesale prices are exogenous and search costs are high, we find that the retailers generally prefer to cooperate, whereas supplier’s profit is higher when they do not cooperate. Retailers’ profits under Model N are higher only if a significantly higher fraction of customers decide to use the fast-ship option under Model N as compared with Model C. In contrast, when search costs are low, then this causes retailers’ profits to be higher when they cooperate, regardless the fraction of customers who place fast-ship orders in Model C or search on their own in Model N. This happens because when almost all customers perform the search under Model N (which is a consequence of low search costs), retailers are not able to capture demand from their loyal customers, i.e. from those who visit their store first and experience a stockout. Hence, they need to stock more up front, which increases their inventory risk and lowers their expected profit. In contrast, in Model C, retailers can pool excess inventories to lower their costs.

With exogenous wholesale prices, it is difficult to identify parameters under which the supplier makes more profit under Model C or Model N because its profit depends on the sizes of the initial and fast-ship orders as well as the margins from these two types of sales. For example, if more customers place fast-ship orders in Model C when both retailers stock out, then this causes the
retailers to order less up front. Consequently, the supplier bears a greater inventory risk (if it produces more up front) or incur a higher production cost (if it lacks supplies to fill fast-ship orders). On the one hand, both these actions could lower its profit. On the other hand, the fast-ship margin could be sufficiently large to make up for the increased costs. That is, detailed comparisons of profit functions are needed to make definite statements about their relative sizes when wholesale prices are exogenous.

The situation is reversed when wholesale prices are determined by the supplier. In such cases, it is easier to anticipate supplier’s preference as compared to retailers’ preferences. The retailers do not benefit from fast-ship orders because they are required to pay their entire revenue for procuring these items from the supplier. The retailers’ profits are driven largely by the initial order quantity, which is in turn determined by the wholesale price for initial orders. Although the supplier chooses this price to maximize its profit under each model, it also needs to consider the retailers’ participation constraint, i.e. retailers’ profits must be non-negative. The retailers’ profit comparisons are more complex because they are dictated by the wholesale price set by the supplier. For the supplier, in contrast, it is possible to anticipate whether greater size of the initial or fast-ship order will matter more. It depends on its margin from either type of sale and the propensity of customers to place fast-ship orders.

The key contribution in this paper is two fold: (1) we obtain solutions to the retailers’ and the supplier’s parameter optimization problems, and (2) we show how the fraction of customer that place fast-ship orders (in both models) and the fraction that perform the search (in the non-cooperative model) affect model preferences of the retailers and the supplier. One might expect the retailers to achieve greater profit under cooperation (because it allows them to pool inventory) and the suppliers to realize greater profit under non-cooperation (because that could increase their ability to extract a greater share of the supply chain profit). Whereas the above intuition is correct in many situations, our analysis shows that such intuition is not uniformly valid. We also identify and explain cases in which the retailers make greater profit by not cooperating and the supplier makes greater profit when retailers cooperate.

The rest of this paper is organized as follows. We review literature in Section 2 and introduce notation and models in Section 3. Section 4 contains the supplier’s and the retailers’ optimal decisions for the two models. In Section 5, we compare our models under two scenarios that arise
in the noncooperative setting. In one scenario, search costs are high, therefore the majority of customers who desire to purchase the item place fast-ship orders without searching. In the second scenario, search cost is low, therefore the majority of customers perform a search first. Section 6 concludes the paper.

2. Literature Review

The models presented in this paper have features that include inventory risk sharing, transshipment, and customer search. The OM literature contains many papers that deal with each of these topics. Retailers in our setting utilize the availability of more than one replenishment opportunity to shift some inventory risk to the supplier. Papers in which the retailers can replenish more than once to reduce inventory/shortage cost can be divided into two groups. In both groups, the retailer places its initial order before the start of the selling season. The first group of papers focus on the effect of either updated demand distribution or improved cost estimates upon getting additional (but incomplete) information after the start of the selling season. After updates are performed, a second order is placed during the selling season to reduce the supply-demand mismatch (see Eppen and Iyer 1997a,b, Gurnani and Tang 1999, and Donohue 2000 for examples of papers belonging to this category). Papers in the second group, which includes our work, assume that the second order is placed at the end of the selling season to satisfy demand from customers who backorder (Cachon 2004, Dong and Zhu 2007, Gupta et al. 2010, and Chen et al. 2011). However, both the model formulation and the purpose of models is quite different in our setting relative to these papers. For example, the supplier in our setting also has a second opportunity to procure additional items to satisfy unmet fast-ship demand albeit at a higher cost. This option is not available to the supplier in Cachon (2004) and Dong and Zhu (2007). Gupta et al. (2010) and Chen et al. (2011) model the second replenishment opportunity for the supplier, but both consider only one retailer.

The dual strategy model described in Netessine and Rudi (2006) is similar to regular and fast-ship option in our setting. When dual strategy is adopted, each retailer uses its stockpile as the primary source of items needed to satisfy demand and drop shipping as a backup source when its stock runs out. However, there are important differences between our work and Netessine and Rudi (2006). First, Netessine and Rudi (2006) assumes that when in-store inventory runs out, all remaining customers agree to receive their items from the drop-ship channel, which corresponds to a
special case of our model. Our model can be applied to many more situations in which an arbitrary fraction of customers do not take advantage of the fast-ship option. Second, the supplier in our model has two replenishment opportunities to effectively manage its cost whereas the supplier in Netessine and Rudi (2006) has a single replenishment opportunity.

In our cooperative model (Model C) the retailers pool excess inventory after satisfying their initial demands, which is related to works that focus on transshipment between retailers. These include papers that study transshipment under a centralized system — for example, Tagaras (1989), Herer and Rashit (1999) and Dong and Rudi (2004), as well as papers in which each retailer makes individually optimal ordering decisions — for example, Anupindi et al. (2001), Rudi et al. (2001), Granot and Sošić (2003), and Sošić (2006). In the latter category, some papers assume that the transshipped item is sold at either a fixed transshipment price (e.g., Rudi et al. 2001, Zhao and Atkins 2009, and Shao et al. 2011), or a proportional transshipment price (e.g., Jiang and Anupindi 2010) and the others assume that the additional profit is divided based on some allocation rules (e.g., Anupindi et al. 2001, Granot and Sošić 2003, and Sošić 2006).

Retailers in our paper make individually optimal decisions and the additional revenue is divided between retailers based on a predetermined allocation rule. This is similar to a two-stage distribution problem in Anupindi et al. (2001). That is, each retailer first places an order to satisfy its own demand. Excess demand at one retailer can be fulfilled by excess inventory at the other retailer and the related profit is distributed among retailers. In addition, Anupindi et al.’s (2001) and our paper both assume that the retailers share all inventory. This is different from the three-stage problem introduced by Granot and Sošić (2003) and Sošić (2006), in which the retailers also determine how much inventory to share. In addition, Granot and Sošić (2003) and Sošić (2006) are concerned with supply chain coordination. The authors identify allocation rules that achieve coordination whereas we focus on how retailers’ decision to cooperate or not affects their’s and supplier’s profits.

Kemahlioglu-Ziya and Bartholdi (2010) also studies a shareable inventory model in which all retailers are served from a common pool of inventory managed by the supplier. The paper introduces a mechanism that allocates the extra profit based on Shapley value among all players, including the supplier. This model feature is different from our setting in which supplier plays a hands-off role in terms of inventory sharing between retailers. Also, unlike our paper, Kemahlioglu-Ziya and Bartholdi (2010) focuses on supply chain coordination rather than comparing cooperation with
non-cooperation.

The non-cooperative model (Model N) is related to papers that model competing newsvendors. In these papers, a fraction of customers who experience stockouts try to get the item from a different store. Depending on the demand model considered, these papers can be divided into two categories. In the first category, retailers’ demands arise either as a result of a deterministic allocation of a random market demand (e.g., Lippman and McCardle 1997 and Caro and Martinez-de Albeniz 2010), or as a result of a random allocation of a deterministic demand (e.g., Nagarajan and Rajagopalan 2009). These papers establish under certain conditions the existence of a unique Nash equilibrium in pure ordering strategies. In the second category, the two retailers face independently distributed demands (see, for example, Parlar 1988 and Avsar and Baykal-Gürsoy 2002). Parlar (1988) studies a single period problem with two independent retailers. Avsar and Baykal-Gürsoy (2002) extends the model to an infinite horizon problem. Our model belongs to the second category discussed above. Two model features make our work substantially different from the above-mentioned papers. First, we include the supplier in our analysis and compare cooperation with non-cooperation from the retailers’ and the supplier’s perspectives. Second, retailers in our model can obtain additional replenishments from the supplier if their local inventory is not enough to met demand from customers that are willing to backorder.

A key contribution of this paper is to develop a framework for comparing the supplier’s and retailer’s profits when retailers either cooperate or not. Similar comparisons are also done in the literature, but with different assumption and under different model settings. For example, in Zhao and Atkins (2009), retailers either agree to transship excess inventory or let unsatisfied customers switch stores. The paper shows that transshipment is beneficial for retailers when either transshipment price is high (e.g., higher than the wholesale price) or relatively fewer customers search for the item on their own. Anupindi and Bassok (1999) studies two models. In the decentralized model, a fraction of unsatisfied customers travel to the other store to buy the item. In the centralized model, the two retailers store their inventory in a common pool. Anupindi and Bassok (1999) shows that a higher \( \alpha \) makes decentralized model more preferable for the supplier. Anupindi et al. (2001) compares a retailer-driven search model (transshipment) and a customer-driven search model (customer search) and shows that the transshipment scheme is better for the retailers (resp. the supplier) when the transshipment price is set by the retailers (resp. the supplier). Also, Shao et al.
(2011) compares supply chain with and without transshipment and shows that the retailers always prefer transshipment whereas the supplier prefers transshipment when the transshipment price is higher than a threshold and wholesale price is exogenous.

We also show that cooperation (resp. non-cooperation) may not be the best option for the retailers (resp. the supplier) but for different reasons. In our models, we observe that the retailers would be better off if they cooperate when more customers search for the item at the other store (which is opposite of the observation in Zhao and Atkins 2009) because retailers often stock more in absence of cooperation, which increases costs and lowers profits. In addition, our setting is very different. Zhao and Atkins (2009) does not include the supplier’s problem. Zhao and Atkins (2009), Anupindi et al. (2001), and Shao et al. (2011) use a transshipment price, whereas we use an allocation rule to divide net revenue from inventory sharing. Finally, because of the existence of the fast-ship option in our models, the profit from the fast-ship order is also an important element of the supplier and the retailers’ total profits. This additional component makes the retailers depend less on transshipment and the supplier rely less on the initial order quantity, complicating players’ preferences. A summary of model features of key papers discussed in this section is provided in Table 1.

3. Notation and Model Formulation

We consider a supply chain with two retailers (denoted by $R_i$, $i \in \{A, B\}$) and a single supplier (denoted by $S$). Retailer $R_i$ faces independent demand $X_i \in \mathbb{R}^+$, sells products at a unit retail price $r$, and pays wholesale prices $w_1$ and $w_2$ for initial orders and fast-ship orders, respectively. We assume that $w_1, w_2 \leq r$ to ensure that both initial and fast-ship orders are profitable for the retailer. In absence of this requirement, the retailer either may not remain a business partner or choose not to offer the fast-ship option to customers when a stockout occurs. If the two retailers decide to cooperate they share revenue (net of shipping cost $\tau_i$ per unit) from the sale of transshipped items.

The supplier has two production\footnote{We say production opportunity or quantity in the remainder of this paper. The word “production” could be easily replaced by “procurement” without affecting our results.} opportunities and it’s production costs are $c_1$ for the initial order and $c_2$ for expedited production where $c_2 \geq c_1$. The supplier pays shipment costs to a third party logistics service provider, which are denoted by $\tau_1$ and $\tau_2$ for initial and fast ship orders, respectively. Because fast-ship orders are shipped on an expedited basis, we assume that $\tau_2 \geq \tau_1$. 

Table 1: Summary of Literature

<table>
<thead>
<tr>
<th>Literature</th>
<th>Centralized</th>
<th>Decentralized</th>
<th>Transshipment</th>
<th>Transshipment Price</th>
<th>Allocation Rule for Transshipment</th>
<th>Customer Searching</th>
<th>Supplier</th>
<th>Wholesale Price</th>
<th>Retail Price</th>
<th>Inventory Risk Sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td>This paper</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>F/D</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Parlar (1988)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tagaras (1989)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lippman and McCardle (1997)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eppen and Iyer (1997a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eppen and Iyer (1997b)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herer and Rashit (1999)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>F</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>Anupindi and Bassok (1999)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>Gurnani and Tang (1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Donohue (2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rudi et al. (2001)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>F/D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avsar and Baykal-Gürsoy (2002)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Granot and Sošić (2003)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dong and Rudi (2004)</td>
<td>✓</td>
<td>✓</td>
<td>F</td>
<td></td>
<td>F/D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cachon (2004)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Sošić (2006)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Netessine and Rudi (2006)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Ulkü et al. (2007)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Dong and Zhu (2007)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Nagarajan and Rajagopalan (2009)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Zhao and Atkins (2009)</td>
<td>✓</td>
<td>✓</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Jiang and Anupindi (2010)</td>
<td>✓</td>
<td>✓</td>
<td>F/P</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>Gupta et al. (2010)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Caro and Martinez-de Albeniz (2010)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Chen et al. (2011)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Shao et al. (2011)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>


Also, the cooperation model in our paper is reasonable when the two retailers are located closer to each other than to the supplier. Therefore, we also assume that $\tau_1 \leq \tau_2$.

We consider two models consisting respectively of cooperative (denoted by superscript C) and non-cooperative retailers (denoted by superscript N). When retailers cooperate, each retailer’s excess demand can be fulfilled by the other retailer’s excess inventory through transshipment. If both retailers run out of stock, then each retailer offers to procure items via the fast-ship mode for its customers. At that point in time, a fraction $\alpha_i^C$ of $R_i$’s unserved customers take advantage of the fast-ship option and $(1 - \alpha_i^C)$ remain unserved.
In contrast, if retailers choose not to cooperate, then upon stocking out each retailer immediately offers to procure items via the fast-ship mode for its customers. Depending on customers’ individual search costs, a fraction \( \alpha_i^N \) of \( R_i \)'s customers take advantage of this offer whereas a fraction \( \beta_i \) choose to search the other retailer’s inventory. Those who search the item at the second store either immediately purchase the item if available on the second retailer’s shelf or place a fast ship order if that retailer also stocks out. Thus, the fraction \( (1 - \alpha_i^N - \beta_i) \) of customers who experience stockout remain unserved.

The wholesale prices could be determined either by market forces, which we refer to as exogenous prices, or determined by the supplier, which we refer to as endogenous prices. When the supplier chooses wholesale prices, they may be different in the two models. In such cases, we add superscript C or N to the wholesale prices to indicate the model setting. Retailers choose order quantities \( (q_i, i \in \{A, B\}) \) after learning \( w_1 \) and \( w_2 \), and finally the supplier decides production quantity or inventory level \( (y + q_A + q_B) \) upon learning each retailer’s initial order size. Notation used in this paper is summarized in Table 2. We next develop expressions for the retailers’ and the supplier’s profit functions, denoted by \( \pi^k_{R_i} \) and \( \pi^k_S \), respectively, where \( k \in \{C, N\} \).

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters:</td>
<td></td>
</tr>
<tr>
<td>( X_i )</td>
<td>Retailer-( i )'s demand; density and distribution functions ( f_i(\cdot) &gt; 0 ) and ( F_i(\cdot) )</td>
</tr>
<tr>
<td>( \alpha_i^k )</td>
<td>Partial backorder rate in Model ( k \in {C, N} ), ( \alpha_i^k \in [0, 1] )</td>
</tr>
<tr>
<td>( \beta_i )</td>
<td>Customer self-searching rate in Model N, ( \beta_i \in [0, 1 - \alpha_i^N] )</td>
</tr>
<tr>
<td>( w_1, w_2 )</td>
<td>( ^\dagger )Wholesale price per unit for regular/fast-ship orders, ( w_1, w_2 \leq r )</td>
</tr>
<tr>
<td>( \tau_1, \tau_2 )</td>
<td>Regular/expedited shipping cost for each item from the supplier, ( \tau_2 \geq \tau_1 )</td>
</tr>
<tr>
<td>( \tau_t )</td>
<td>Transshipment cost between retailers, ( \tau_t \geq 0 )</td>
</tr>
<tr>
<td>( c_1, c_2 )</td>
<td>Supplier’s first/second production costs per unit, ( c_2 \geq c_1 \geq 0 )</td>
</tr>
<tr>
<td>( r )</td>
<td>Retail price per unit, ( r \geq 0 )</td>
</tr>
<tr>
<td>Decision Variables:</td>
<td></td>
</tr>
<tr>
<td>( y )</td>
<td>Additional units produced by S during the first replenishment, ( y \geq 0 )</td>
</tr>
<tr>
<td>( q_i )</td>
<td>Retailer-( i )'s order quantity, ( q_i \geq 0, i \in {A, B} )</td>
</tr>
</tbody>
</table>

\( ^\dagger \)In endogenous cases, \( w_1 \) and \( w_2 \) are decision variables.

**Model C – Cooperating Retailers**

When the two retailers cooperate, transshipment is assumed to occur instantaneously. That is, customers do not differentiate between a transshipped item and one that is purchased from on-
hand inventory. For ease of exposition, we assume that the two retailers equally share net revenue from sales through transshipment. Our model can be easily generalized to a situation where the net revenue is divided differently. Given that the two retailers order \( q_i \) and \( q_j \) respectively, where \((i, j) \in \{(A, B), (B, A)\}\), the expected profit for \( R_i \) can be written as follows.

\[
\pi^C_{R_i}(q_i, q_j) = rE[(q_i \wedge X_i)] - w_1q_i + (r - w_2)\alpha_i^C E[(X_i - q_i)^+ - (q_j - X_j)^+] + \frac{1}{2}(r - \tau_1)E\{(X_i - q_i)^+ \wedge (q_j - X_j)^+\} + \{(X_j - q_j)^+ \wedge (q_i - X_i)^+\}. \tag{1}
\]

In (1), the first two terms denote \( R_i \)'s profit from sales that come from its on-hand inventory. The third term denotes the profit from fast-ship orders and the last term denotes the allocated revenue from transshipment. Each retailer anticipates the other player’s response and chooses its Nash equilibrium quantity. Specifically, \( R_i \) orders \( \hat{q}^C_i = \arg\max_q \pi^C_{R_i}(q_i, \hat{q}^C_j) \), where \( \hat{q}^C_j \) is the equilibrium order quantity of \( R_j \) and \((i, j) \in \{(A, B), (B, A)\}\).

Next, we present the supplier’s profit function. Suppose that the two retailers order \( q_A \) and \( q_B \), respectively. Then, the supplier’s profit function can be written as follows when it produces \( y \) additional units up front.

\[
\pi^C_S(y \mid q_A, q_B) = (w_1 - \tau_1 - c_1)(q_A + q_B) - c_1y + (w_2 - \tau_2)E(\alpha^C_A \zeta^C_A + \alpha^C_B \zeta^C_B) - c_2E(\alpha^C_A \zeta^C_A + \alpha^C_B \zeta^C_B - y)^+, \tag{2}
\]

where \( \zeta^C_A = [(X_A - q_A)^+ - (q_B - X_B)^+] \) and \( \zeta^C_B = [(X_B - q_B)^+ - (q_A - X_A)^+] \). Equation (2) can be explained as follows. The first two terms denote the supplier’s profit from initial orders of the two retailers. The third term is the revenue from the fast-ship sales and the fourth term is the cost of second production, which is required only if the number of fast-ship orders exceed \( y \). When wholesale prices are exogenous, the supplier’s only decision is the optimal production quantity \( \hat{y}^C(q_A, q_B) = \arg\max_y \pi^C_S(y \mid q_A, q_B) \). When the supplier chooses wholesale prices, its decision problem can be written as follows:

\[
(\hat{w}^C_1, \hat{w}^C_2) = \arg\max_{w_1, w_2} \pi^C_S(\hat{y}^C(q^C_A(w_1, w_2), q^C_B(w_1, w_2))) \tag{3}
\]
subject to
\[
\pi^C_{R_i}(q^C_i(w_1, w_2), q^C_j(w_1, w_2)) \geq 0, \quad (i, j) \in \{(A, B), (B, A)\},
\]
\[w_1, w_2 \leq r,
\]

where \(q^C_i(w_1, w_2)\) denotes \(R_i\)’s equilibrium order quantity given \((w_1, w_2)\).

**Model N – Non-Cooperating Retailers**

When the two retailers do not cooperate, customers who experience stockout at their preferred retailer have three options: (1) place a fast-ship order at the store they visit (fraction \(\alpha^N_i\) do so), (2) search the other retailer’s inventory and either purchase an in-stock item or place a fast-ship order with the other retailer (fraction \(\beta_i\) do so), or (3) walk away (fraction \(1 - \alpha^N_i - \beta_i\) do so). Given that the two retailers order \(q_i\) and \(q_j\) respectively, where \((i, j) \in \{(A, B), (B, A)\}\), the expected profit of retailer \(R_i\) can be written as follows:

\[
\pi^N_{R_i}(q_i, q_j) = rE[\min(q_i, X_i)] - w_1q_i + (r - w_2)\alpha^N_i E(X_i - q_i)^+ \\
+ rE(\{\beta_j(X_j - q_j)^+ \wedge (q_i - X_i)^+\}) + (r - w_2)E[\beta_j(X_j - q_j)^+ - (q_i - X_i)^+]^+.\]

(6)

In (6), \(rE[\min(q_i, X_i)] - w_1q_i\) denotes profit from initial sales to \(R_i\)’s customers. The third term denotes the profit from customers who place fast-ship orders at the first store they visit. The last two terms represent the profit derived from satisfying customers traveling from the other store using either on-hand inventory or the fast-ship option. Similar to Model C, \(R_i\) orders \(\hat{q}^N_i = \arg \max_q \pi^N_{R_i}(q_i, \hat{q}^N_j)\), where \(\hat{q}^N_j\) is the equilibrium order quantity of \(R_j\) and \((i, j) \in \{(A, B), (B, A)\}\).

The supplier’s profit function in Model N is as follows.

\[
\pi^N_S(y \mid q_A, q_B) = (w_1 - c_1 - \tau_1)(q_A + q_B) - c_1y + (w_2 - \tau_2)E[\alpha^N_A(X_A - q_A)^+ + \alpha^N_B(X_B - q_B)^+] \\
+ (w_2 - \tau_2)E[\zeta^N_A(\beta_A) + \zeta^N_B(\beta_B)] \\
- c_2E[\alpha^N_A(X_A - q_A)^+ + \alpha^N_B(X_B - q_B)^+] + \zeta^N_A(\beta_A) + \zeta^N_B(\beta_B) - y^+ 
\]

(7)

where \(\zeta^N_A(\beta_A) = [\beta_A(X_A - q_A)^+ - (q_B - X_B)^+]^+\) and \(\zeta^N_B(\beta_B) = [\beta_B(X_B - q_B)^+ - (q_A - X_A)^+]^+\). The supplier decides optimal initial production quantity after observing retailers’ orders. Similar to
Model C, the optimal production quantity is \( \hat{y}^N(\hat{q}_i^N(w_1, w_2), \hat{q}_j^N(w_1, w_2)) = \arg \max_y \pi_S^N(y \mid q_A, q_B). \)

Furthermore, let \( \hat{q}_i^N(w_1, w_2) \) denote the equilibrium order quantity given \((w_1, w_2)\). If the supplier also chooses wholesale prices, then it faces the following optimization problem.

\[
(w_1^N, w_2^N) = \arg \max_{w_1, w_2} \pi_S^N(\hat{y}^N(\hat{q}_i^N(w_1, w_2), \hat{q}_j^N(w_1, w_2)))
\]

subject to

\[
\pi_{R_i}^N(\hat{q}_i^N(w_1, w_2), \hat{q}_j^N(w_1, w_2)) \geq 0, \quad (i, j) \in \{(A, B), (B, A)\},
\]

\[
w_1, w_2 \leq r.
\]

4. Retailers’ and Supplier’s Operational Choices

This section is divided into two parts. In the first part, we focus on retailers’ and supplier’s quantity decisions assuming that wholesale prices are given. Then, in the second part, we comment on how the supplier chooses wholesale prices when prices are endogenous.

4.1 Quantity Decisions

Model C

We first determine the retailers’ equilibrium order quantities \((\hat{q}_A^C, \hat{q}_B^C)\). In order to calculate each retailer’s best response, we show in Proposition 1 that \(\pi_{R_i}^C\) is strictly concave in \(q_i, i \in \{A, B\}\).

**Proposition 1.** For each \((i, j) \in \{(A, B), (B, A)\}\), retailer-i’s profit \(\pi_{R_i}^C(q_i, q_j)\) is strictly concave in \(q_i\) if \(\alpha_i^C \leq \frac{(r-\tau_i)}{2(r-w_2)}\).

Note that the inequality \(\alpha_i^C \leq \frac{(r-\tau_i)}{2(r-w_2)}\) would be true in many practical situations because \(w_2\) would be much greater than \(\tau_i\). As discussed in the Introduction section, we expect \(\tau_i \leq \tau_2\) to hold in the vast majority of cases in practice. Because \(w_2\) includes \(\tau_2\) and should also at least cover production cost \(c_1\), we expect \(\tau_i \ll w_2\) to hold. In fact, we show in Section 4.2 that when wholesale prices are endogenous, the supplier sets \(w_2 = r\), in which case the above inequality is true. For these reasons, we assume hereafter that \(\alpha_i^C \leq \frac{(r-\tau_i)}{2(r-w_2)}\). The result in Proposition 1 implies that a unique best response \(q_i\) exists for each \(q_j\) and the Nash equilibrium solution is obtained by solving
the following two equations, one for each pair of \((i, j)\) where \((i, j) = \{(A, B), (B, A)\}\).

\[
w_1 = r \bar{F}_i(q_i) - \alpha^C_i(r - w_2)[F_i(q_i)F_j(q_j) + \int_0^{q_j} \bar{F}_j(q_i + q_j - x_i)dF_j(x_j)] + \frac{(r - \tau_t)}{2}[-F_i(q_i)F_j(q_j) + \int_0^{q_j} \bar{F}_j(q_i + q_j - x_i)dF_j(x_j)] + \int_0^{q_i} \bar{F}_i(q_i + q_j - x_j)dF_i(x_i)].
\] (11)

Using (11), we can ascertain how retailers’ order quantities change in \(\alpha^C\), as shown in Proposition 2.

**Proposition 2.** When the two retailers are identical, each retailer’s equilibrium order quantity is weakly decreasing in \(\alpha^C\).

The result in Proposition 2 has an intuitive explanation. With a higher \(\alpha^C\), the retailers are more likely to profit from fast-ship orders when stockout occurs. In other words, the overall shortage cost for the retailers is lower. Consequently, retailers choose to stock less up front. Note that when \(w_2 = r\), the equilibrium order quantity is not affected by \(\alpha^C\) because retailers do not earn a profit from fast-ship sales.

We next shift attention to the supplier’s problem. For the supplier, we observe that \(\pi^C_S(y \mid q_1, q_2)\) in Equation (2) is strictly concave in \(y\) for each pair of \((q_1, q_2)\). Hence, the optimal \(y\) can be obtained by setting \(\partial \pi^C_S(y \mid q_1, q_2)/\partial y\) to 0. This yields \(\hat{y}^C = \max(0, \eta^C)\), where \(\eta^C\) is obtained by solving Equation (12) below.

\[
\frac{c_1}{c_2} = \int_{q_A}^{q_B} \frac{\eta^C + \alpha^C_A(x_A - q_A) + q_B}{{\alpha^C_B}} dF_A(x_A) + \bar{F}_A(q_A + \frac{\eta^C}{\alpha^C_A}) \bar{F}_B(q_B) + \int_0^{q_B} \bar{F}_A(q_A + q_B - x_B)dF_B(x_B) + \int_0^{q_A} \bar{F}_B(q_B + q_B + q_A - x_A)dF_A(x_A). \] (12)

In (12), the first two terms capture the probability that stockout occurs at both retailers and the sum of fast-ship orders is greater than \(\eta^C\). The last two terms denote the probability that stockout event occurs at one retailer and the total fast-ship demand received by the supplier is greater than \(\eta^C\). Therefore, if \(\eta^C \geq 0\) then it can be thought of as a production level at which the marginal cost for procuring up front \((c_1)\) is equal to the marginal expected cost of procuring later
(c_2 \times \text{Prob}[S \text{ requires the second replenishment}]). Note that if \eta^C < 0, it implies that the supplier’s profit is decreasing in the extra production level \( y \). Hence, \( \hat{y}^C \) would be set to 0.

**Model N**

Similar to Model C, we begin by obtaining order quantities for the two retailers.

**Proposition 3.** For each \((i,j) \in \{(A, B), (B, A)\}\), \( \pi^N_{R_i}(q_i, q_j) \) is strictly concave in \( q_i \).

The result in Proposition 3 implies that a unique best response \( q_i \) exists for each \( q_j \) and the equilibrium order quantity is obtained by solving Equations (13), one for each pair of \((i, j)\) where \((i, j) = \{(A, B), (B, A)\}\).

\[
\begin{align*}
\frac{c_1}{c_2} = & \int_{q_A}^{q_B} \frac{\eta^N}{\gamma_A} F_B \left( q_B + \frac{\eta^N - \gamma^N_A (x - q_A)}{\gamma_B} \right) dF_A(x) + \bar{F}_A \left( q_A + \frac{\eta^N}{\gamma_A} \right) \bar{F}_B \left( q_B \right) \\
& + \int_{q_A}^{q_B} \frac{\beta_A \eta^N}{\alpha_A} F_A \left( q_A + \frac{\eta^N}{\alpha_A} \right) + F_A \left( q_A - \frac{\beta_A \eta^N}{\alpha_A} \right) \bar{F}_B \left( q_B + \frac{\eta^N}{\alpha_B} \right) \\
& + \int_{q_B - \frac{\beta_B \eta^N}{\alpha_B}}^{q_B} F_A \left( q_B - x + \eta^N \gamma_A \right) + q_A dF_B(x) + \int_{q_A - \frac{\beta_B \eta^N}{\alpha_B}}^{q_A} F_B \left( q_A - x + \eta^N \gamma_B \right) + q_B) dF_A(x),
\end{align*}
\]

(14)
where $\gamma_i = \alpha_i^N + \beta_i$. Similar to Model C, when $\eta^N \geq 0$, $\eta^N$ can be interpreted as the production level at which the marginal cost of procuring up front ($c_1$) is equal to the marginal expected cost of procuring later ($c_2 \times \text{Prob}[S \text{ requires the second replenishment}]$). When $\eta^N < 0$, no additional production is needed.

### 4.2 Wholesale-Price Decisions

The problem of finding supplier’s optimal wholesale prices is a hard problem because the supplier must take into account retailers’ equilibrium response. As noted in the Section 4.1, we do not have a closed-form expression for the equilibrium order quantities. We partially address this difficulty by proving that the supplier facing identical retailers would set $\hat{w}_2^k = r$ in Model $k \in \{C, N\}$ to maximize its expected profit (Proposition 5). This helps reduce the problem of finding optimal wholesale prices to a straightforward line search.

**Proposition 5.** Suppose that the two retailers are identical. When the wholesale prices are chosen optimally by the supplier, $\hat{w}_1^k < \hat{w}_2^k = r$ for Model $k \in \{C, N\}$.

The results shown in Proposition 5 can be explained as follows. When $w_2$ is higher, both retailers would respond by ordering a higher initial order quantity because the benefit from offering the fast-ship option is diminishing. This benefits the supplier because the profit from the initial order and the fast-shop orders both increase. Since $\hat{w}_2^k$ is always set to $r$ in both models, we can focus on obtaining an optimal $w_1$. However, the complexity of the supplier’s profit function precludes further analytical results. The optimal initial wholesale price $\hat{w}_1^k$ can be obtained numerically.

### 5. Insights

In this section, we compare the supplier and retailers’ profits under cooperative and non-cooperative scenarios via numerical examples in which we vary production costs, backorder rates, and customer search rates. With identical retailers, we focus attention on two scenarios: exogenous and supplier-determined wholesale prices. In the first scenario, the wholesale prices are identical in the cooperative and non-cooperative models. In contrast, in the second scenario, the supplier chooses individually optimal prices that are in general different for the two models. Because profits of the two players depend on a number of problem parameters, we isolate two special cases that we believe are of particular interest. These two cases consider high and low customer search costs, as
Recall that in our model, customer choice is explained by their propensity to either try to obtain the item by some other means if their first-choice retailer is stocked out, or walk away without purchasing. In the non-cooperative model, customers who want to purchase the item after experiencing a stockout at one retailer have two choices: they may decide to first search at the other retailer (fraction $\beta$), or right away place a fast-ship order with their first-choice retailer (fraction $\alpha^N$). When search cost is high, we expect the majority of customers who desire to purchase the item to place fast-ship orders without searching, which gives rise to the first of two cases of interest with $\beta = 0$. In contrast, when search cost is low, we expect the majority of customers to search first. The latter gives rise to the second case of interest in which $\alpha^N = 0$. Note that when $\beta = 0$ (respectively, $\alpha^N = 0$), it does not imply that $\alpha^N$ (respectively, $\beta$) equals 1 because in general $\beta + \alpha^N \leq 1$.

Before presenting our results and insights, we point out what one might conclude based on first-pass intuitive arguments. One might expect the retailers to achieve greater profit under cooperation (because it allows them to pool inventory) and the suppliers to realize greater profit under non-cooperation (because that could increase their ability to extract a greater share of the supply chain profit). Whereas the above intuition is correct in many situations, the purpose of the ensuing analysis is to identify and explain cases in which the retailers make greater profit by not cooperating and the supplier makes greater profit when retailers cooperate.

5.1 Exogenous Wholesale Prices and High Search Cost

Consider an example in which $(w_1, w_2)$ are determined by market forces and if retailers do not cooperate, then $\beta = 0$. Throughout this section, we omit retailer index because the two retailers are assumed to be identical. Other parameters are as follows: Gamma distributed demand with $E[X] = 225$ and $\text{Var}(X) = 15^3$, $r = 14$, $w_1 = 10$, $w_2 = 12$, $\tau_1 = 0$, $\tau_2 = 1$, and $\tau_t = 0.8$. With $\beta = 0$, we vary $\alpha^C \in [0, 1]$ and $\alpha^N \in [0, 1]$ for two different sets of production costs $(c_1, c_2) \in \{(5, 6), (9, 11)\}$. The regions in which the two players realize greater profits are shown in Figure 2. These regions are labeled I, II and III with the following distinction: in Region I both retailers and the supplier earn greater profit when retailers do not cooperate, in Region II retailers make more profit by cooperating whereas the supplier makes more profit when the two retailers do not
cooperate, and in Region III both retailers and the supplier make greater profit when retailers do not cooperate. In addition, we hereafter use right arrow “→” in figure legends to denote preferences. For example, $R \rightarrow N$ and $S \rightarrow N$ means that both the supplier and the two retailers prefer non-cooperation model.

![Figure 2: Profit Comparisons – Exogenous $w_1$ ($\beta = 0$)](image)

The dotted line joining coordinates $(0, 0)$ and $(1, 1)$ shows the special case in which $\alpha^C = \alpha^N$. In the cooperative model, some customers who would wish to purchase the item when their first-choice retailer stocks out would have their demand met by transshipment. Therefore, one may argue that the fraction $\alpha^C$ who would place a fast-ship order should be at most $\alpha^N$. That is, the region above the 45-degree line joining coordinates $(0, 0)$ and $(1, 1)$ is of particular interest. Upon focusing attention on this set of parameters, we note that the first-pass intuition holds in many cases because Region II ($R \rightarrow C$ and $S \rightarrow N$) dominates other regions. Also, retailers’ profits are higher under Model N only if $\alpha^N$ is significantly higher than $\alpha^C$. This happens only in Region I in Figures 2(a) and 2(b). This is because a higher $\alpha^N$ implies that a larger portion of excess demand can be satisfied in Model N. With exogenous $w_2$, the supplier does not have the ability to extract all excess profit from the increased demand for fast-ship orders. This is different in Section 5.3 where the supplier chooses $w_2 = r$.

Explaining supplier’s profit function comparisons is more complicated because it depends on the sizes of the initial and fast-ship orders as well as the margins from these two types of sales.
For example, Propositions 2 and 4 show that a higher $\alpha^C$ in Model C, or a higher $\alpha^N$ in model N, causes the retailers to order less up front. This could cause the supplier to bear a greater portion of inventory cost (if it produces more up front) or incur a higher production cost (if it lacks supplies to fill fast-ship orders), and thereby lower its profit. However, the fast-ship margin could be sufficiently large to make up for the increased costs. For example, we see in Figure 2(a) with $(c_1, c_2) = (5, 6)$, the supplier prefers Model C (respectively, N) if $\alpha^C$ (respectively, $\alpha^N$) is high. This happens because its worst-case margin from fast-ship orders (which equals $w_2 - \tau_2 - c_2 = 12 - 1 - 6 = 5$) is as much as the margin from initial sales (which equals $w_1 - \tau_1 - c_1 = 10 - 0 - 5 = 5$). However, the comparison changes in Figure 2(b) with $(c_1, c_2) = (9, 11)$. Now, the supplier prefers model N both when $\alpha^N$ is high and when $\alpha^C$ is high. It prefers Model C only when both $\alpha^C$ and $\alpha^N$ are relatively small — observe the size of Region III in Figures 2(a) and 2(b). We explain these differences next.

In Model N, the increase in fast-ship orders is much greater and the supplier can benefit from it by producing more up front, i.e. increasing $y$. Note that the profit margin if it supplies fast-ship order from inventory (rather than second production) is $w_2 - \tau_2 - c_1 = 12 - 1 - 9 = 2$, whereas the initial sales margin is $w_1 - \tau_1 - c_1 = 10 - 0 - 9 = 1$. That is, the supplier still makes more profit when $\alpha^N$ is high because the drop in initial orders is not so high as to dominate the gain from more profitable fast-ship orders. In contrast, in Model C, fast-ship orders are needed only if both retailers run out. In this case, the decrease in initial order quantity when $\alpha^C$ increases may cause a larger decrease in profit than the gain from higher expected fast-ship orders. Therefore, a higher value of $\alpha^C$ does not cause the supplier to prefer Model C.

More formally, with $\beta = 0$, the fast-ship amounts under the two models can be written as follows

\[ FS^C = \alpha^C (X_A + X_B - q^C_A - q^C_B), \quad \text{and} \]
\[ FS^N = \alpha^N [(X_A - q^N_A)^+ + (X_B - q^N_B)^+]. \]

When $\alpha^C = \alpha^N$ and $q^C_i \geq q^N_i$, it is easy to see that $FS^N \geq FS^C$. In other words, the supplier earns higher profit from the fast-ship order when the two retailers do not cooperate. The initial order quantity decreases as $\alpha^C$ and $\alpha^N$ increase (see Propositions 2 and 4) and a larger portion of the supplier profit comes from the fast-ship orders. Therefore, the supplier prefers Model C only
when \( \alpha_C \) and \( \alpha_N \) are low because in that case, the benefit from higher initial order dominates the benefit from fast-ship orders. In other instances, it prefers Model N because the profit from the fast-ship orders is more significant.

The above discussion begs the question: when can we compare the initial order quantities in the two models? For this purpose, we define

\[
G(q_i, q_j) = -\bar{F}_i(q_i)F_j(q_j) + \int_0^{q_j} F_j(q_i + q_j - x_i) dF_j(x_j) + \int_0^{q_i} F_i(q_i + q_j - x_j) dF_i(x_i). \tag{15}
\]

If \( \beta = 0 \), we can show from (11) and (13) that \( q_C > q_N \) provided \( G(q_N, q_N) \geq 0 \) and \( \alpha_C = \alpha_N \).

In general, whether the order quantity would be higher when retailers cooperate would depend on the demand distribution and problem parameters. This is similar to observations made in Yang and Schrage (2009), in which conditions that result in a higher initial order size are identified when inventory pooling occurs. In our case, \( G(q_i, q_j) > 0 \) is more likely if \( w_1 \) is high relative to \( r \) (because it causes \( q \) to be small). If \( q_C > q_N \) holds, then the supplier earns greater profit from initial orders when the two retailers cooperate.

### 5.2 Exogenous Wholesale Prices and Low Search Cost

Next, we focus attention on an example in which \( \alpha_N = 0 \), i.e. search costs are low. Other parameters remain the same as in Section 5.1. In the experiments reported in this section, we vary \( \alpha_C \in [0, 1] \) and \( \beta \in [0, 1] \) and show regions in which the two players realize greater profit in Figure 3. Note that the region labels are consistent with those used in Section 5.1. Also, the 45-degree lines joining coordinates \((0, 0)\) and \((1, 1)\) in Figures 3(a) and 3(b) denote cases in which \( \alpha_C = \beta \).

We observe that low search cost causes retailers’ profit to be higher when they cooperate, regardless of \( \alpha_C \) and \( \beta \) values. This is a major difference as compared to Section 5.1. It is because with \( \alpha_N = 0 \), retailers are not able to capture demand from their loyal customers, i.e. those who search their store first and experience a stockout. Hence, they need to stock more up front, which increases their inventory cost and lowers their expected profit. In contrast, in Model C, retailers can pool inventories to lower the cost of lost sales even if \( \alpha_C \) is low.

Similar to Section 5.1, the supplier’s profit comparisons are more complex. In Figure 3(a), large values of \( \beta \) makes Model N more attractive to the supplier, whereas large values of \( \alpha \) makes C more desirable. The reason is that the supplier benefits from increased fast-ship orders that are
sold at a higher wholesale price $w_2$ if $\alpha^C$ is high, and from increased initial order size if $\beta$ is high. That said, we find that the size of Region III in which the supplier prefers C reduces a great deal in Figure 3(b) relative to Figure 3(a) because the supplier’s margin from fast-ship orders is much smaller when $(c_1, c_2) = (9, 11)$. In such cases, increase in the size of fast-ship orders is insufficient to make up for the smaller initial order size.

Focusing next on the special case when $\alpha^C = \beta$, we observe that the supplier prefers Model C only when both $\alpha^C$ and $\beta$ are low. This can be explained as follows. Although in general the supplier’s profit is affected by both initial and fast-ship sales, in the special case with $\alpha^N = 0$, the supplier’s relative profit is determined largely by the initial order size. This is because the fast-ship quantities in Models C and N are not significantly different when $\alpha^N = 0$. From (11) and (13), we can show that $q^C > q^N$ when $G(q^N, q^N) \geq 0$ and $\alpha^C = \beta = 0$ whereas $q^C < q^N$ when $\alpha^C = \beta = 1$. Furthermore, $q^C$ is decreasing in $\alpha^C$ and $q^N$ is increasing in $\beta$ (see Propositions 2 and 4). Therefore, there must exist a threshold $\delta \in [0, 1]$ such that $q^C > q^N$ for $\alpha^C = \beta < \delta$ and $q^C < q^N$ for $\alpha^C = \beta > \delta$. Hence, the supplier’s preference changes from model C to N as $\alpha^C$ and $\beta$ increase. We confirmed these arguments by checking that the values of $\alpha^C$ and $\beta$ at which the supplier is indifferent between Models C and N in Figures 3(a) and 3(b) happen at a point at which $q^C = q^N$. Plots of $q^C$ and $q^N$ are not shown in the interest of brevity.

In summary, the supplier’s profit is generally greater when retailers do not cooperate, whereas
the retailers profit is always greater when they do cooperate. However, when the $\alpha^C$ and $\beta$ are small and approximately the same size, both parties make greater profits when retailers cooperate.

### 5.3 Endogenous Wholesale Prices

We continue with the same set of example parameters as in Section 5.1, but select wholesale prices in each model that would be optimal from the supplier’s perspective. Note that $\hat{w}_2^k = r$ in both models (Proposition 5). Therefore, in these examples, we performed a line search to find an optimal $w_1$ for each combination of example parameters. As in earlier examples, two special cases are considered: $\beta = 0$ or $\alpha^N = 0$. Profit comparisons are shown in Figures 4 and 5. In addition to Regions I–III defined in Section 5.1, we introduce a new label – Region IV – to indicate a region in which retailers prefer N, but the supplier prefers C. Such a region did not occur in our examples when wholesale prices were exogenous.

![Preference](image1)

![Optimal Quantities and Wholesale Prices](image2)

**Figure 4: Profit Comparisons – Endogenous $w_1$ ($\beta = 0$)**

Because $\hat{w}_2^k = r$, retailers’ profits are not affected by $\alpha^C$ and $\alpha^N$ values. Their profits are driven largely by the initial order quantity, which is in turn determined by $\hat{w}_1^k$. Although the supplier chooses the wholesale prices to maximize its profit under each model, it also needs to consider the retailers’ participation constraint, i.e. their profit must be non-negative. Still, the supplier is able to extract nearly the entire excess profit in the supply chain.

Consider the supplier’s profit comparisons first. Because $\hat{w}_2^k = r$, the supplier benefits a great deal from the fast-ship orders. Therefore, it prefers Model C (resp. N) for higher $\alpha^N$ or $\beta$ (resp. $\alpha^C$).
Note that the region in which the supplier prefers Model N (region II) is different in Figures 4(a) and Figure 5(a). The main reason for such difference is that the supplier is forced to reduce the wholesale price $w_1$ when $\beta$ is high and $\alpha^N = 0$ (see Figure 5(b)) to ensure retailers’ participation. This significantly reduces the profitability from Model N and makes Model C much more attractive for the supplier for higher $\beta$.

In this instance, the retailers’ profit comparisons are more complex because they are dictated a great deal by the supplier’s choice of $w_1$. For example, if we compare Figures 2(a) and 4(a), we notice that the case in which the retailers prefer Model N does not exist in scenario with endogenous wholesale price. This is because with $\hat{w}_2^{k} = r$, the retailers do not benefit at all from the fast-ship option. Hence, having the ability to pool inventory via transshipment in Model C is more attractive to the retailers. For additional comparisons of retailers’ profits in the ensuing paragraph, we restrict attention to the special cases in which $\alpha^C = \alpha^N$ when $\beta = 0$, or $\alpha^C = \beta$ when $\alpha^N = 0$.

In Figure 4(a), we observe that with $\beta = 0$, if $\alpha^C = \alpha^N$ is small, then the retailers generally prefer Model C. This is because $q^C$ is large and consistently much greater than $q^N$, sufficient to overcome the loss of profit from a higher $\hat{w}_1^C$ relative to $\hat{w}_1^N$ (see Figure 4(b)). However, when $\alpha^N = \alpha^C$ is large, $q^C$ declines sharply, $\hat{w}_1^C$ increases and approaches $\hat{w}_1^N$. Therefore, the retailers marginally prefer non-cooperation. The comparison is more complicated if $\alpha^N = 0$ because of the retailers’ participation constraint. We note in Figure 5(b) that $\hat{w}_1^N$ is decreasing in $\beta$ when
$\beta \in [0.3, 1]$. For small to moderate values of $\beta = \alpha^C$, retailers prefer to cooperate because the effect of lower $\hat{w}_1^C$ and higher $q^C$ dominates, but for high values, retailers, in fact, prefer non-cooperation on account of smaller $\hat{w}_1^N$ and higher $q^N$. The retailers’ profits are small and dictated mostly by the supplier’s choice of the wholesale prices.

6. Conclusions

In decentralized supply chains, the fast-ship option can be used to mitigate both inventory and stockout risks. However, the type of interaction between the retailers would influence the search behavior of consumers which in turn affects risk allocation and supply chain partners’ profits. We investigated a two-retailer supply chain under two different structures. In the cooperation model (denoted as Model C), if a customer encounters an out-of-stock situation, the retailers offer the transshipment option to customers (by incurring a transshipment cost but with no extra cost to the customer) which is assumed to occur instantaneously if the other retailer has excess inventory. As such, the customers do not have to search and are indifferent between the transshipped item or the one that is purchased from on-hand inventory. If, however, the other retailer is also out of stock, a certain fraction of customers use the fast-ship option at the original retailer. In the non-cooperation model (denoted as Model N), retailers do not offer to transship and a fraction of customers who encounter an out-of-stock situation use the fast-ship option at their first-choice retailer, whereas, a certain fraction are willing to search for the item at the other retailer. We identified conditions under which the retailers would prefer to cooperate.

Intuitively, non-cooperation may yield a higher profit for the supplier because the retailers may stock more to avoid losing customers to the other retailer, whereas the retailers may prefer to cooperate and share any excess inventory with each other. Our results indicate that this may not always be true. When wholesale prices are exogenous and search costs are high, we find that our intuition is indeed valid. But when the fast-ship option is significantly more favored by customers under the non-cooperation case, the retailers may prefer not to cooperate. For the supplier, the preference is strongly affected by the initial order quantity from the retailers and its margins from regular and fast-ship orders.

When wholesale prices are optimally chosen by the supplier, we show that the retailers do not benefit from fast-ship sales as the supplier is able to extract all profits from those sales. As such,
the retailers’ preferences depend on the profits derived from the initial sale which is based on the wholesale price for the regular order. For the case of high search costs, our analysis reveals that while the supplier mostly prefers the retailers not to cooperate, in cases when fast-ship fraction of customers are comparable under both cooperation and non-cooperation cases, the supplier may be better off from retailer cooperation.

The contribution of the paper is to develop analytical results for the optimal decisions for the retailers and the supplier under the two models of cooperation and non cooperation. A key insight from the paper is that cooperation between retailers may not be the retailers’ best option under all conditions. Sometimes, it may be more profitable for retailers to let customer perform the search their own, which is contrary to preliminary intuition. Similarly, non-cooperation between retailers is not always the best option for the supplier. Whereas some form of inventory risk allocation is inevitable in supply chains, there is no dominant strategy that works best for all scenarios. Different schemes should be adapted based on product characteristics and target markets. This paper serves to shine light on the combined roles of cooperation and fast-shipping in mitigating inventory and stockout risks.

Appendix

Proof of Proposition 1. From Equation (1), we obtain

\[
\frac{\partial^2 \pi^C_{R_A}(q_A, q_B)}{\partial q_A^2} = -\tau_2 f_A(q_A) - \frac{1}{2}(r - \tau_2)f_A(q_A)F_B(q_B) - \left[\frac{1}{2}(r - \tau_2) - \alpha_C(r - w_2)\right]f_A(q_A)F_B(q_B) \\
+ \int_0^{q_B} f_A(q_A + q_B - x_B)dF_B(x_B) - \frac{1}{2}(r - \tau_2)\left(\int_0^{q_A} f_B(q_A + q_B - x_A)dF_A(x_A)\right).
\]

(16)

Suppose that \((r - \tau_2)/2 \geq \alpha_A(r - w_2)\). One can easily see that \(\frac{\partial^2 \pi^C_{R_A}(q_A, q_B)}{\partial q_A^2}\) in (16) is less than 0. Hence proved.

Proof of Proposition 2. When the two retailers are identical, \(q = q^A = q^B\). Hence, the equilibrium
order quantities $q$ must satisfy (11) and can be rewritten as follows.

\[
\begin{align*}
\hat{w}_1 &= r \hat{F}(q) - \alpha C(r - w_2) [\hat{F}(q)F(q) + \int_0^q \hat{F}(q + x)dF(x)] \\
&\quad + \frac{(r - \tau_1)}{2} [-\hat{F}(q)F(q) + \int_0^q \hat{F}(q + x)dF(x) + \int_0^q \hat{F}(q + x)dF(x)].
\end{align*}
\]

Let $q_L$ and $q_H$ denote the equilibrium order quantity for $\alpha L$ and $\alpha H$, where $\alpha L \leq \alpha H$. Because we observe that the RHS of (17) is decreasing in $\alpha C$ for a fixed $q$, the RHS of (17) with $q_H$ must be higher than that with $q_L$ for a fixed $\alpha C$. By taking derivative of RHS of (17) with respect to $q$, we obtain

\[
- rf(q) - \alpha C(r - w_2) [-f(q)\hat{F}(q) - 2 \int_0^q f(2q + x)dF(x)] + \frac{r - \tau_2}{2} [f(q) - 4 \int_0^q f(2q + x)dF(x)] \\
\leq - rf(q) + \frac{r - \tau_2}{2} [f(q)\hat{F}(q) + f(q) - 2 \int_0^q f(2q + x)dF(x)] \\
\leq - rf(q) + \frac{r - \tau_2}{2} [f(q) + f(q) - 2 \int_0^q f(2q + x)dF(x)] \\
\leq - rf(q) + (r - \tau_2)f(q) \leq 0,
\]

where the first inequality comes from the assumption that $(r - \tau_2)/2 \geq \alpha C(r - w_2)$. In other words, the RHS of (17) is decreasing in $q$ for a fixed $\alpha C$. Hence, $q_H \leq q_L$ must hold.

**Proof of Proposition 3.** From (6), we obtain

\[
\frac{\partial \pi_{RA}^N(q_A, q_B)}{\partial q_A} = -w_1 + (r - \alpha A^N(r - w_2))F_A(q_A) + w_2 \int_0^{q_A} \hat{F}_B \left( \frac{q_A - x_A}{\beta_B} + q_B \right) dF_A(x),
\]

and

\[
\frac{\partial^2 \pi_{RA}^N(q_A, q_B)}{\partial q_A^2} = -(r - \alpha A^N(r - w_2))f_A(q_A) + w_2 \hat{F}_B(q_B)f_B(q_B) \\
- \frac{w_2}{\beta_B} \int_0^{q_A} f_B \left( \frac{q_A - x_A}{\beta_B} + q_B \right) dF_A(x) > 0.
\]

Hence, $\pi_{RA}^N(q_A, q_B)$ is strictly concave in $q_A$. Similar arguments can be applied to $\pi_{RB}^N(q_B, q_A)$. \(\square\)

**Proof of Proposition 4.** When the two retailers are identical, the equilibrium order quantities $q =$
\[ q^A = q^B \] must satisfy (13) and can be rewritten as follows.

\[
w_1 = (r - \alpha^N (r - w_2)) \bar{F}(q) + w_2 \int_0^q \bar{F} \left( \frac{q - x}{\beta} + q \right) dF(x).
\]

(20)

It is easy to check that the RHS of (20) is decreasing in either \( \alpha^N \) or \( q \). Let \( q_L \) and \( q_H \) denote the equilibrium order quantity for \( \alpha^N_L \) and \( \alpha^N_H \), where \( \alpha^N_L \leq \alpha^N_H \). In order to make the equality in (20) hold, \( q_H \leq q_L \) must hold because \( \alpha^N_L \leq \alpha^N_H \).

Similarly, Let \( \hat{q}_L \) and \( \hat{q}_H \) denote the equilibrium order quantity for \( \beta_L \) and \( \beta_H \), where \( \beta_L \leq \beta_H \). Because the RHS of (20) is increasing in \( \beta \), \( \hat{q}_H \geq \hat{q}_L \) must be true so (20) can be satisfied.

\[ \square \]

**Proof of Proposition 5.** We show proof for Model C only as the arguments for Model N are almost the same. This is proved by contradiction. Recall that \( (\hat{w}_C^1, \hat{w}_C^2) \) is the supplier selected wholesale prices. Let \( q^C(w_1, w_2) \) be the corresponding order quantity when the supplier chooses \( (w_1, w_2) \).

Assume that \( \hat{w}_C^2 < r \). In addition, we observe from (11) that the equilibrium \( q^C(w_1, w_2) \) is decreasing (resp. increasing) in \( w_1 \) (resp. \( w_2 \)) when the two retailers are identical. Therefore, when the supplier sets \( w_2^C = r \), there exists a \( w'_1 > \hat{w}_1^C \) such that \( q^C(w'_1, r) = q^C(\hat{w}_2^C, \hat{w}_2^C) \). Note that observing from (2), the supplier’s profit \( \pi_S^C(y \mid q_A, q_B) \) is increasing in both \( w_1 \) and \( w_2 \) for fixed \( y \), \( q_A \) and \( q_B \). In other words, the supplier’s profit with \( (w_1, w_2) = (w'_1, r) \) is higher than that with \( (w_1, w_2) = (\hat{w}_1^C, \hat{w}_2^C) \), which contradicts our assumption that \( \hat{w}_2^C < r \). Hence, \( \hat{w}_2^C = r \) must be true.

We next show that \( \hat{w}_1^C < \hat{w}_2^C \) again by contradiction. We first assume \( \hat{w}_1^C = \hat{w}_2^C \) is true. Then by taking derivative of (2) with respect to \( q \), we observe that the changes in profit is at least

\[ 2[(\hat{w}_1 - \tau_1 - c_1) - (\hat{w}_2 - \tau_2 - c_1)(\bar{F}(q)(1 + F(q)))] > 2[(\hat{w}_1 - \tau_1 - c_1) - \alpha^C(\hat{w}_2 - \tau_2 - c_1)] \geq 0 \]

because \( \tau_2 \geq \tau_1 \) and \( \hat{w}_1^C = \hat{w}_2^C \). In other words, \( \pi_S^C(y \mid q, q) \) when \( \hat{w}_1^C = \hat{w}_2^C \) is increasing in \( q \) while keeping all other parameter fixed. In addition, it is easy to check that \( \pi_S^C(y \mid q, q) \) in increasing in \( w_2 \) for fixed \( q \). This along with the fact that equilibrium \( q \) is increasing in \( w_2 \), there exist a \( w'_2 > \hat{w}_1^C \) such that \( \pi_S^C \) is higher. This contradicts our assumption that \( \hat{w}_1^C = \hat{w}_2^C \). Hence, \( \hat{w}_1^C < \hat{w}_2^C \) must be true.

\[ \square \]

**References**


