Centralizing Inventory in Supply Chains by Using Shapley Value to Allocate the Profits

Eda Kemahlıoğlu-Ziya
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3490, eda_kemahliogluziya@unc.edu

John J. Bartholdi, III
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, john.bartholdi@gatech.edu

How should the excess profit because of inventory pooling be shared amongst firms at different levels along the supply chain? Suppose each of several retailers observes local demand for a common item and places an order at the supplier, which is immediately filled if the supplier has the item in stock. The supplier can fill retailer orders either from their reserved inventories or from a shareable pool of inventory. Using terminology from cooperative game theory, we say that the supplier and the retailers whose orders are filled from the common pool have formed an inventory-pooling coalition and study the use of Shapley value to allocate the expected excess profit because of pooling.

We find that under Shapley value allocations the retailers have incentive to join the inventory-pooling coalition, and the supplier carries the level of inventory that is optimal for the coalition. Shapley value allocations might not lie within the core of the game, but the grand coalition of all players is stable in the farsighted sense. And, although the supplier’s share of the expected excess profit is largest when all the retailers participate in the inventory-pooling coalition, the allocations to the retailers may diminish as the coalition grows. Colluding against the supplier (by merging and forming larger retailers) may seem like an appealing strategy for the retailers to increase their share of the total supply chain profit, but we find that the total expected after-pooling profits of retailers may instead go down because of collusion.

Key words: supply chain management; incentives and contracting

History: Received: September 22, 2007; accepted: May 10, 2010. Published online in Articles in Advance October 13, 2010.

1. A Problem of Supply Chain Coordination

We recently observed a supply chain for consumer electronics in which an upstream contract manufacturer maintained two physically separate inventories of the same part for two competing customers. Everyone recognized that savings were possible if inventory was pooled, but it was not clear to them how to share any savings or what would be the risks of exposing the inventory on which they relied to the demand of a competitor.

Similar problems can be observed in other contexts of supplier–buyer relationships in the electronics industry. Because of highly uncertain demand and shortened product life cycles, the procurement trend in this industry is to delay ownership until after the demand for the part is known. Hence, vendor-managed pull systems are frequently utilized, in which the downstream customer pays for the component only when it is actually used (Anderle 1997, McKeefry 1998). There are different ways pull systems can be implemented. For example, TTI, Inc., an electronics part distributor, offers several inventory management alternatives including “bonded inventory” and “consigned inventory” (TTI, Inc., http://www.ttiinc.com/page/services_scm, last accessed September 12, 2010). Bonded inventory is kept at the TTI warehouse but is electronically segregated and reserved for the customer. Consigned inventory is owned and managed by TTI at the customer’s facility until the customer uses it. TTI cannot benefit from risk pooling under either policy because demand is not filled from a common pool, and so the inventory carrying cost is higher than necessary—but supply is secured.

Steve Vecchiarelli, vice-president of supply chain solutions for Digi-Key Corp, says that buyers prefer to reserve their inventory because “Companies believe that they want the comfort of knowing that they have the supplier’s inventory in their warehouses, assuring supply while mitigating costs” (Atkinson 2007, p. 17). However, when inventory is managed this way, suppliers’ costs increase and are passed to the downstream buyers in the form of lower service levels. When Digi-Key procures from suppliers, they prefer a form of vendor-managed inventory where
replenishments are made from the suppliers’ centralized stores. Vecchiarelli’s reasoning is that “these centralized stores allow suppliers more control without under-performing pockets of inventory impacting supplier costs,” and result in higher service levels at the buyers (Atkinson 2007, pp. 17–18). In their study of replenishment models in the high-tech industry, Eitelwein and Wallenburg (2007) also highlight the potential benefits of keeping inventory in centralized locations and note that to align the incentives in such arrangements, new pricing models that include profit sharing are required. This paper analyzes one such profit-sharing mechanism.

We study these issues in a two-stage supply chain with one upstream player (called the supplier) and multiple downstream players (called the retailers) who face stochastic demand. The supplier carries all the inventory in the supply chain, and the retailers carry none. We consider a single-period problem (which is repeated indefinitely) in which the supplier sets the inventory level and then the retailers observe their local demand and fulfill it to the extent the supplier has stock. The supplier can manage inventory in one of several ways. One policy, which we call the reserved inventory policy, is to hold some inventory reserved for each retailer subject to his service-level constraint, as in the bonded inventory or consignment inventory programs. Another policy, which we call the shareable inventory policy, is to serve all retailers from a common pool of inventory as in the vendor-managed inventory program adopted by Digi-Key. We assume that the status quo in the supply chain is that the supplier manages the inventories of all retailers under the reserved inventory policy, but prefers that they switch to a shareable inventory policy.

It is well known that, given the same level of inventory, the total supply chain profit under the shareable inventory policy will be higher than that under the reserved inventory policy because of the pooling effect. However, retailer profits may be lower than those under the reserved inventory policy because they may be allocated fewer units of inventory. One contribution of this paper is to propose a mechanism that uses solution concepts from cooperative game theory (specifically the Shapley value) to allocate the extra profit because of pooling among the retailers and the supplier. The mechanism ensures that the retailers’ expected profits under the shareable inventory policy are at least as high as those under the reserved inventory policy. In addition, under the proposed allocation mechanism, the supplier sets the inventory level for the shareable inventory policy in a way that maximizes the expected total supply chain profit, that is, the mechanism coordinates the supply chain.

Using terminology from cooperative game theory, the supplier and the set of retailers who switch to the shareable inventory policy are said to form a coalition, and we are interested in the stability of the grand coalition of all the retailers and the supplier. We find that the core of the game that ensues from the allocation of expected excess profit is not empty, and we provide sufficient conditions under which the Shapley value allocations are in the core. We further show that if we use a farsighted notion of stability, namely the largest consistent set, the grand coalition is always stable under Shapley value allocations. This result implies that all players prefer the grand coalition to any smaller coalition, which is important because the total supply chain profit is highest when all players switch to the shareable inventory policy.

We discover that Shapley value allocations give incentives to the players for actions that can change the total supply chain profits and/or profits of the other players under shareable inventories. One such action is investing in reducing the demand variability faced by the supply chain, which increases the total supply chain profit. Demand variability can be reduced in a number of different ways, such as improving forecasting accuracy, smoothing product consumption, etc. We find that the retailers have incentive to take actions to reduce demand variability only when the retail markup is high, whereas the supplier has incentive when the markup is low.

We also investigate whether a subset of retailers will have incentive to collude against the supplier by acting as a larger retailer to increase their expected profits under the shareable inventory policy. Collusion changes the profits of all players, but not the total supply chain profit. Surprisingly, there are circumstances in which the retailers’ expected profits go down if they choose to collude.

2. Literature Review
This paper is broadly related to the literature on inventory pooling, and Eppen (1979) was the first to analyze its benefits. Others extend his results to more general settings (Cherikh 2000 provides a review of the more recent literature). However, these are all models of vertically integrated companies, and so there is no question of how to allocate savings or costs. When inventory or capacity is pooled in a decentralized supply chain, there must be mechanisms to manage both the distribution of goods in case of shortages as well as the fair allocation of costs or profits among supply chain partners. Gerchak and Gupta (1991), Robinson (1993), and Hartman and Dror (1996, 2003) were among the first to study the fair cost allocation problem and Robinson (1993) and Hartman and Dror (1996, 2003) use concepts from cooperative game theory. Müller et al. (2002) show that the core of the cost allocation game is nonempty
for general demand distributions under the assumption that the retailers have identical holding and penalty costs. This work has been extended in several directions (including consideration of profit allocation) by allowing nonidentical selling and wholesale prices and transportation costs (Slikker et al. 2005), multiple sources of inventory (Özen et al. 2008), demand forecast updates (Özen et al. 2006b), and lower bounds on the amount of inventory that retailers must receive (Özen et al. 2006a); and the authors show that the cores of the respective games are nonempty. However, Hartman and Dror (2005) find that the core may be empty if the shortage and inventory costs are not identical.

A stream of related papers analyzes the distribution of virtually centralized goods as a transshipment problem among different players at the same echelon. Anupindi et al. (2001) consider a two-stage model, where in the first stage inventory levels are set non-cooperatively, and in the second stage excess inventories are transshipped cooperatively. Suakkaphong and Dror (2009) relax some assumptions of Anupindi et al. (2001) (e.g., participants are rational and have complete information, local demand is satisfied before transshipment), whereas Granot and Sošić (2003) extend the model of Anupindi et al. and add an intermediate stage where the players explicitly decide on the amount of excess inventory to transship. Rudi et al. (2001) analyze a similar problem, but with only two retailers. These papers propose various profit-allocation schemes and analyze their stability properties. One of the schemes studied by Granot and Sošić (2003) is profit allocation using the Shapley value; Sošić (2006) extends this analysis to show that such a strategy, though not necessarily stable in the myopic sense, is stable when dynamic player deviations are allowed. These papers are closely related to our work in that they consider decentralized systems, but they differ from ours in that they concentrate on profit allocations within a single echelon.

Anupindi and Bassok (1999) are among the first to study the incentives of inventory centralization in a multiechelon setting. They consider a two-level supply chain with a single manufacturer and two retailers, but unlike our model, the inventory decisions are made by the retailers, who bear all the risk. They show that the manufacturer may not always benefit from inventory pooling because total sales may drop. Dong and Rudi (2004) extend the model of Anupindi et al. (2001) to a two-echelon supply chain and explore whether transshipments, which are beneficial for the retailers, are also beneficial for the upstream manufacturer. Again, unlike our model, the retailers hold inventory and make the transshipment decisions. Dong and Rudi (2004) assume that demand is normally distributed, and Zhang (2005) extends their results to general distributions. Zhao et al. (2005) consider a two-retailer distribution system where the retailers simultaneously decide on their stocking and rationing levels. After establishing the equilibrium policies for the retailers, the authors analyze the effects of these policies on the supplier through a numerical study.

The transshipment problem arises when the retailers own inventory. When the supplier decides on the inventory level and carries all the inventory in the supply chain, as is the case in our paper, new issues arise, such as how to induce the supplier to choose the optimal inventory level and how to allocate inventory among retailers in case of shortages. There is a rich literature on capacity allocation. To our knowledge, Cachon and Lariviere (1999) and Deshpande and Schwarz (2005) are the only two that study these two issues simultaneously in a single-period setting. In our paper, we use a modified version of the linear allocation rule proposed by Cachon and Lariviere (1999) to allocate inventory. However, our paper is different from this body of work in that the supplier allocates inventory after retailer demand has been observed.

Similar to ours, some recent papers refrain from modelling inventory/capacity allocation explicitly; rather, they analyze the cooperative bargaining over the use of restricted capacity after demand realization. Plambeck and Taylor (2005) study who (a contract manufacturer or a joint venture between original equipment manufacturers) should own shared capacity, given that after demand realization, profits will be allocated to the parties with respect to their bargaining confidence indices. Plambeck and Taylor (2007a, b) interpret ex post bargaining over excess, shareable capacity as renegotiation of the supply contract and compare different contracts for their ability to induce first-best investment given that a renegotiation step with cooperative bargaining will follow demand realization. Bartholdi and Kemahlıo˘glu-Ziya (2005) study inventory centralization in a two-retailer model and determine with what type of players (differentiated by their minimum service level requirements) it is preferable to pool inventory if the supply chain profit is allocated using the Shapley value. The present paper extends the analysis to an N-retailer model and explores new issues, including myopic and farsighted stability, how the allocations depend on parameters such as demand variance, and how to allocate profits for a given realization of demand such that, in expectation, all players receive their Shapley value shares of excess profit from inventory pooling.

3. Savings Because of Inventory Pooling

Consider a supply chain where a single supplier holds a product at her expense. Each of N retailers observes
local demand, places an order with the supplier, and receives the item immediately, provided the supplier has the item in stock. Both the retailers and the supplier maximize their single-period expected profits. Each retailer’s expected profit is a function only of the local expected sales because they carry no inventory. The retailers may affect the stocking level of the supplier by requiring service levels \(ρ_i\), the minimum acceptable probability of no stockout at retailer \(i\).

Let \(D_i\) denote the random demand (and \(d_i\), its realization) at retailer \(i\) with cumulative distribution function (cdf) \(F_i(\cdot)\), which is assumed to be strictly increasing, differentiable, with finite mean, and with probability distribution function (pdf) \(f_i(\cdot)\). We assume that \(D_i\) for \(i \in \{1 \ldots N\}\) are independent random variables. Let \(p\) be the wholesale price the supplier charges to the retailers, \(w\) the procurement cost per unit, \(h\) the holding cost per unit, and \(p_M\) the markup on wholesale price that the retailers charge. Assume that the demand distributions and the cost and revenue parameters are common knowledge to all of the supply chain players (a common assumption in the related literature, including Anupindi and Bassok 1999, Granot and Sošić 2003, Plambeck and Taylor 2007b, among others). Let \(x\) denote the stock level.

To manage inventory, the supplier might reserve independent supplies for each retailer (reserved inventory policy), or she might pool the inventory and draw from a shared cache when orders arrive (shareable inventory policy). We will compare the expected optimal profits of the supplier and the retailers under the reserved and the shareable inventory management policies. However, first, to benchmark the results, we look at the first-best solution, where the supplier and the retailers are vertically integrated and the inventory-level decision is made centrally.

### 3.1. The Vertically Integrated Supply Chain

If the supply chain is vertically integrated—that is, if both the supplier and the retailers are owned by the same company—then a single centralized authority can set inventory levels to maximize expected profits. We let subscript \(c\) denote the variables under this centralized decision system. The stock-level decision in the centralized supply chain is made with respect to the total demand across all retailers, which we denote by \(D_c\), the cdf of which is \(F_c(\cdot) = F_1(\cdot) \cdots F_N(\cdot)\). The centralized supply chain seeks to maximize expected profit and the objective function is

\[
\max_{x_c} \left( p E[\min\{x_c, D_c\}] - h(x_c - E[\min\{x_c, D_c\}]) - wx_c \right)
\]

for which the optimal stocking level is \(F_c^{-1}(p + p_M - w) / (p + p_M + h)\), where \((p + p_M - w) / (p + p_M + h)\) is the critical fractile. The corresponding expected profit level is the maximum attainable by the supply chain and provides a benchmark for supply chain performance.

### 3.2. Savings in a Decentralized Supply Chain

#### 3.2.1. Reserved Inventories

When inventories must be managed separately, the supplier sets the inventory level for each of the retailers independently to maximize her expected profit subject to the retailer’s service-level constraint:

\[
\max_{x_i} \left( p E[\min\{x_i, D_i\}] - h(x_i - E[\min\{x_i, D_i\}]) - wx_i \right)
\]

subject to \(F_i(x_i) \geq \rho_i\).

This is the classical newsvendor problem; the optimal stocking level \(x_i^*\) is \(F_i^{-1}(\max\{(\rho_i, (p - w)/(p + h))\})\) for retailer \(i\), and the total inventory carried by the supplier is \(\sum_{i=1}^{N} F_i^{-1}(\max\{(\rho_i, (p - w)/(p + h))\})\). We consider reserved inventories the base case, providing a level of total expected profit to which other inventory strategies will be compared.

#### 3.2.2. Shareable Inventories

As the bearer of excess inventory cost, the supplier will benefit from a switch to shareable inventories from reserved inventories because of the risk-pooling effect. In this case, she makes the inventory decision based on the total demand, the cdf of which is the same as in the case of the centralized supply chain, \(F_c(\cdot) = F_1(\cdot) \cdots F_N(\cdot)\). To emphasize that the centralized supply chain and the supplier who decides to pool inventory both face the same demand stream, we will continue to use the subscript \(c\) to denote the demand in this case. To denote that the supplier chooses the inventory level that maximizes her expected profit, we use the subscript \(S\) on \(x\). An all-powerful supplier pools inventory and sets her stock level to maximize her own expected profit as follows:

\[
\max_{x_S} \left( p E[\min\{x_S, D_s\}] - h(x_S - E[\min\{x_S, D_s\}]) - wx_S \right)
\]

The optimal inventory level in this case is \(F_c^{-1}(p - w) / (p + h)\).

This solution leaves money on the table. First, the optimum inventory level under this policy is less than the optimum inventory level for the vertically integrated supply chain. Therefore, the total supply chain profit is not maximized. Second, as observed by Anupindi and Bassok (1999), expected sales, and, therefore, profits, at the retailers may decrease compared to the reserved inventory management policy. This happens when an independent, all-powerful supplier holds less inventory than she would when all inventories are managed independently. In such a case, the retailers may try to prevent their loss through service-level
(or, equivalently, minimum inventory-level) requirements on the supplier, but service-level contracts are an imperfect tool to manage pooled inventories in decentralized supply chains. For example, not all service-level contracts can guarantee the retailers their expected profits under the reserved inventory management policy, and even those that do may create a free-rider problem (Bartholdi and Kemahlioglu-Ziya 2005). In addition, such constraints on the supplier may move total supply chain profits even further away from the maximum level it can attain (because the supplier is asked to carry more inventory than is optimal for the supply chain under pooling). An ideal procedure to manage shareable inventories would maximize the expected excess profit from pooling, transparently share them in a way deemed fair, and discourage manipulation. It is to the search for such a protocol that this paper contributes.

4. Sharing the Savings due to Centralized Inventory

We assume that the status quo in the supply chain is the case in which the supplier follows the reserved inventory policy and the players have agreed on the minimum service levels and the amount of inventory stocked for each retailer (§3.2.1). The supplier proposes switching to a shareable inventory policy where the participating retailers will be replenished from a common pool (the size of which is set by the supplier), and the savings from inventory pooling will be allocated among the players according to the procedure we propose later in this section. The retailers decide whether to switch based on the amount of inventory they believe the supplier will carry, how inventory will be allocated in case of a shortage, and what their expected profit allocations will be; and the supplier sets her inventory level based on how many retailers switch and what her consequent expected profit will be under the savings-sharing procedure.

Consider retailer $i$ who faces demand realization $d_i$ and for whom the supplier carries $x_i^*$ units of inventory under the reserved inventory policy. Under the reserved inventory policy, retailer $i$’s profit is $p_M \min(x_i^*, d_i)$. Under the shareable inventory policy, retailer $i$ may be allocated less inventory than $x_i^*$ and, unless retailer $i$ can be guaranteed to make at least $p_M \min(x_i^*, d_i)$ for any demand realization $d_i$, he will not switch to shareable inventories. This is similar to imposing expectation damages as the remedy for breach of contract (Plambeck and Taylor 2007a, b). Under expectation damages, if the supplier fails to deliver the agreed-upon quantity, she must pay the retailer the equivalent of the profit the retailer would have made if the quantity had been delivered. In our context, the retailer expects to receive $x_i^*$ units of inventory regardless of the inventory management policy of the supplier, and the supplier compensates the retailer if she fails to deliver. Let $\phi_i(x, x_i^*, d_i) \geq 0$ denote the extra profit from pooling allocated to retailer $i$ under the shareable inventory policy when the supplier carries $x$ units of inventory. Then the total profit of retailer $i$ under shareable inventories is $p_M \min(x_i^*, d_i) + \phi_i(x, x_i^*, d_i)$, and that of the supplier is

$$\begin{aligned}
(p + p_M) \min \left\{ x, \sum_{i=1}^N d_i \right\} - h \max \left\{ 0, x - \sum_{i=1}^N d_i \right\} - wx \\
- \sum_{i=1}^N (p_M \min(x_i^*, d_i) + \phi_i(x, x_i^*, d_i)).
\end{aligned}$$

Note that for given total demand $\sum_{i=1}^N d_i$, the supplier’s profit under the shareable inventory policy may be less than what her profit would have been under the reserved inventory management policy. For example, this may happen when $\sum_{i=1}^N x_i^* \geq \sum_{i=1}^N d_i > x$, and the loss can be interpreted as the penalty the supplier incurs for carrying fewer units of inventory than she did under the reserved inventory management policy.

We can now write the players’ expected total profits under the shareable inventory policy. For retailer $i$ we have

$$\Pi(R_i) = p_M E\left[ \min(x_i^*, D_i) \right] + E[\phi_i(x, x_i^*, D_i)],$$

and in the remainder of the paper we let $\pi_{R_i} = p_M E[\min(x_i^*, D_i)]$ and $\phi_{R_i} = E[\phi_i(x, x_i^*, D_i)]$. Assuming that all the retailers agree to switch to the shareable inventory policy, the expected total profit for the supplier is

$$\Pi(S) = E\left[ (p + p_M) \min \left\{ x, \sum_{i=1}^N D_i \right\} - h \max \left\{ 0, x - \sum_{i=1}^N D_i \right\} - wx \right] - \sum_{i=1}^N (\pi_{R_i} + \phi_{R_i}).$$

Let $\pi_{S_i}$ denote the supplier’s expected profit due to retailer $i$ under the reserved inventory management policy. Then we can rewrite $\Pi(S)$ as

$$\sum_{i=1}^N \pi_{S_i} + E\left[ (p + p_M) \min \left\{ x, \sum_{i=1}^N D_i \right\} - h \max \left\{ 0, x - \sum_{i=1}^N D_i \right\} - wx \right]$$

$$- \sum_{i=1}^N (\pi_{S_i} + \pi_{R_i}) - \sum_{i=1}^N \phi_{R_i},$$

where the term denoted by $a$ is the expected total profit of the supply chain under the shareable inventory policy given the total inventory level is $x$, the term denoted by $b$ is the expected total profit of the...
supply chain under the reserved inventory policy, and \( c \) is the total expected additional profit allocated to the retailers for switching inventory policies. Therefore, the sum of the last three terms is the expected profit the supplier gains when all the retailers switch policies, which we denote by \( \phi_S \). Then the supplier’s expected total profit is simply

\[
\Pi(S) = \sum_{i=1}^{N} \pi_S + \phi_S. \tag{2}
\]

Our goal in this paper is to find the extra profit from pooling allocated to retailer \( i \) \( (\varphi_i(x, x_i^*, d_i) \geq 0) \) for all \( i \in [1 \ldots N] \) such that \( \phi_S > 0 \) \( (\phi_R, \geq 0 \) is guaranteed by \( \varphi_i(x, x_i^*, d_i) \geq 0 \). In addition, given \( \phi_S \), we want the supplier to carry the supply-chain-optimal level of inventory under the shareable inventory policy.

Provided that \( \varphi_i(x, x_i^*, d_i) \geq 0 \) for all \( i \in [1 \ldots N] \), the players’ decisions (for the retailers, whether to switch to shareable inventories, and for the supplier, how much inventory to carry) are made based on their expected profit allocations \( \phi_R \) and \( \phi_S \). Therefore, in the rest of the paper we first propose a way of calculating \( \phi_R \) and \( \phi_S \) such that all the players prefer to switch to the shareable inventory policy and the supplier carries the optimal amount of inventory. Then we propose a way of physically allocating the common inventory at the supplier and calculating \( \varphi_i(x, x_i^*, d_i) \geq 0 \) for a given demand realization such that when the retailers are compensated according to this mechanism, in expectation retailer \( i \) receives \( \phi_R \), and the supplier receives \( \phi_S \).

In deriving \( \phi_R \) and \( \phi_S \), we utilize concepts from cooperative game theory. A cooperative game consists of a set of players \( K \) and a characteristic function \( v(J) \) that specifies the maximum value that can be realized by coalition \( J \subseteq K \). The coalition consisting of all the players in \( K \) is called the grand coalition. An allocation \( \phi \) is a vector in which \( \phi_i \) is the payoff to player \( i \).

In our supply chain game, the players are the supplier and the \( N \) retailers; and the characteristic function \( v(J) \) represents the expected increase in profit when the subset \( J \) of players collaborate to pool inventory.

We seek a mechanism by which to allocate the expected excess profit due to inventory pooling among the participants of the inventory-pooling coalition. Ideally, such a mechanism would be fair, transparent, resistant to strategic manipulation, would leave no money on the table, and the allocations would be computable in polynomial time. We first focus on the issue of fairness. There are many ways to formalize notions of fairness. Here are three that seem reasonable.

- All the savings generated by a coalition will be fully allocated, and a player who does not contribute to the savings receives an allocation of zero (efficiency and null player).
- Identical players receive identical allocations (symmetry).
- For any two games \( v \) and \( w \), the total allocation to a player in the sum of the two games \( v + w \) is equal to the sum of his allocations received in \( v \) and \( w \) separately (additivity).

The reader may recognize that Shapley value is the unique allocation scheme that jointly satisfies these three properties (Shapley 1953, Myerson 1991). The Shapley value assigns to each player their expected marginal contribution as a coalition member, and is computed as follows:

\[
\phi_i(v) = \sum_{J \subseteq K \setminus \{i\}} \frac{|J|!(|K| - |J| - 1)!}{|K|!} \left( v(J \cup \{i\}) - v(J) \right). \tag{3}
\]

### 4.1. Allocating Savings by Shapley Value

Let \( \mathcal{N} \) denote the set of retailers (indexed by \( i \), \( i \in [1 \ldots N] \)) and \( S \) the supplier. Coalitions are formed when a subset of players agree to follow the shareable inventory management policy and pool inventory. For each coalition, \( J \subseteq \mathcal{N} \cup \{S\} \), let the value \( v(J) \) of the coalition be the expected additional profit that coalition \( J \) creates by pooling inventory, where \( v(\emptyset) = 0 \). We will use subscripts to represent the elements of set \( J \): The first part of the subscript denotes the retailers in the coalition and the subscript \( S \) will be included if the supplier is also part of the coalition. Thus, if \( J \) consists of retailers 1 and 2 and the supplier, \( v(J) = v_{1,2,S} \). We propose a profit-sharing mechanism in which each player’s share of the increase in expected profit due to pooling is equal to his Shapley value. By definition, Shapley value \( \phi(v) \) allocates to each player a weighted average of his marginal contribution \((v(J \cup \{i\}) - v(J))\) to coalitions of different sizes.

To compute \( v(J) \) for coalition \( J \), we assume that:

- Retailers, who do not agree to pool inventory, continue to obtain their stock from the supplier.
- Retailers do not purchase in bulk from the supplier nor centralize inventory without the participation of the supplier. (This is equivalent to assuming that the retailers cannot communicate without the facilitation of the supplier and/or do not have the physical means to manage centralized inventories. Similarly, Plambeck and Taylor (2007a) do not allow the downstream players to pool capacity without the participation of the upstream player.)

Recall that \( x_i^* \) is the optimal inventory level for retailer \( i \) (whose demand is \( D_i \)) under the reserved inventory management policy. Let \( D_J = \sum_{i \in J \setminus \{S\}} D_i \) denote the total demand coalition \( J \) faces. If \( x \) is the total amount of inventory the supplier carries, the amount of inventory the coalition can efficiently allocate is \( x^J = x - \sum_{i \in J \setminus \{S\}} x_i^* \) because the retailers in
\( \mathcal{N}(J | S) \) are not members of the coalition and the supplier is obligated to carry \( x^*_i \) units of inventory for each such retailer \( i \). Then the value function for coalition \( J | J | \geq 2 \), that is, the additional profit generated by a coalition \( J \) that consists of \( J \) retailers and the supplier, is

\[
v(J) = \left( p + p_M \right) E[\min[x^l, D_j]] - wx^l - h(x^l - E[\min[x^l, D_j]])
- \sum_{i \in [J]} \left( p + p_M \right) E[\min[x^*_i, D_i]] - wx^*_i
- h(x^*_i - E[\min[x^*_i, D_i]])
\]

(4)

The first term in (4) is the maximum expected profit coalition \( J \) can attain by efficiently allocating inventory among \( J - 1 \) retailers after demand realization. The second term is the expected profit coalition \( J \) can attain under the reserved inventory management policy. Because coalitions \( J \) consisting only of retailers cannot generate any additional profits, \( v(J) = 0 \) for \( J \subseteq \mathcal{N} \). Similarly, the supplier cannot unilaterally decide to centralize inventory; therefore, a coalition consisting only of the supplier cannot generate any additional profit and \( v(S) = 0 \).

To simplify the notation, we use \( \pi(J) \) to denote the maximum expected profit the members of coalition \( J \) can attain by efficiently allocating inventory. Recall that \( \pi_R \) denotes the expected profit of retailer \( i \) before pooling, and \( \pi_S \) denotes the expected profit of the supplier from managing retailer \( i \)'s inventory under the reserved inventory management policy. Then we can rewrite (4) as

\[
v(J) = \pi(J) - \sum_{i \in J | S} (\pi_S + \pi_R),
\]

which we use to compute \( v(J | i) - v(J) \) for each \( i \in \mathcal{N} \cup \{i\} \) and \( J \subseteq \mathcal{N} \) and \( J \subseteq \mathcal{N} \cup \{i\} \).

Next we derive the expressions for the Shapley value allocations to the retailers and the supplier. We use \( \phi_R \) and \( \phi_S \) to denote the Shapley value allocations to retailer \( i \) and the supplier, respectively.

**Proposition 1.** In the inventory-pooling game among \( N \) retailers and the supplier, the Shapley value allocations to retailer \( i \) and to the supplier are

\[
\phi_R = \sum_{J \subseteq \mathcal{N} \cup \{i\}: |S| \leq J, |J| \geq 1} \frac{(|J|!(N - |J|)!(N + 1)!)}{\pi(J | i)}
- \sum_{J \subseteq \mathcal{N} \cup \{i\}: |S| \leq J, |J| \geq 2} \frac{(|J|!(N - |J|)!(N + 1)!)}{\pi(J)}
- \frac{1}{2} (\pi_S + \pi_R),
\]

(5)

\[
\phi_S = \sum_{J \subseteq \mathcal{N} \cup \{i\}: |S| \leq J, |J| \geq 1} \frac{(|J|!(N - |J|)!(N + 1)!)}{\pi(J | i)}
- \frac{1}{2} (\pi_S + \pi_R).
\]

(6)

Expressions (5) and (6) provide the means through which expected excess profit is allocated to the players. When all \( N \) retailers participate in pooling, the total expected profits of retailer \( i \) and the supplier are given by (1) and (2), where \( \phi_R \) and \( \phi_S \) are given by (5) and (6), respectively.

5. **Properties of the Shapley Value Allocations**

In a decentralized supply chain with inventory pooling at the supplier, the supplier bears the cost of too much or too little inventory and will tend to understock (§3.2.2). Now, if the supplier’s share of the extra profit due to inventory pooling is her Shapley value (6), how much inventory will the supplier carry? With Theorem 1 we show that under Shapley value allocations, the supplier will carry the level of inventory that maximizes the expected profit of each pooling coalition and, hence, if the grand coalition is formed, she will carry the level of inventory that maximizes the expected profit of the supply chain, that is, the level of inventory that is optimal for the supply chain.

**Theorem 1 (Effectiveness).** Allocation by Shapley value induces the supplier to make the optimal inventory decision for the pooling coalition.

From (6) we observe that the supplier’s Shapley value allocation is a linear function of total supply chain profit and, hence, if the supplier knows that her share of the extra profit due to pooling will be her Shapley value, she will carry the supply-chain-optimal level of inventory. Given Theorem 1, we next show that savings due to pooling in a large coalition are at least as large as the sum of savings achieved in subsets of the large coalition.

**Proposition 2.** The characteristic function of our inventory-pooling game is superadditive; that is, for any disjoint coalitions \( J \) and \( K \), \( v(J \cup K) \geq v(J) + v(K) \).

Superadditivity assures us that adding one more member to a coalition increases the total worth of the coalition by at least the stand alone worth of the new member. From Proposition 2, we conclude that when the expected excess profit due to pooling is allocated with respect to Shapley value, each player makes as much profit as they would under reserved inventories.

**Corollary 1 (No Disincentive to Pool Inventory).** Whether supplier or retailer, each player’s Shapley value allocation is nonnegative.
5.1. Stability of Shapley Value Allocations

An allocation \( \phi \) is said to be in the core of the characteristic function \( v \) if and only if

\[
\sum_{i \in K} \phi_i = v(K) \quad (7)
\]

\[
\sum_{i \in J} \phi_i \geq v(J) \quad \text{for all } J \subseteq K, \quad (8)
\]

where \( K \) is the grand coalition. If an allocation \( \phi \) is not in the core, then there is some subset of players who can do better among themselves than as members of a larger coalition. We next establish that our inventory-pooling game has a nonempty core by showing an allocation that always lies within the core. Consider an allocation that rewards all the extra profits from inventory pooling to the supplier; the retailers’ allocations are zero. Call this the Strong Supplier allocations.

**Theorem 2 (Nonempty Core).** Strong Supplier allocations are in the core.

Strong Supplier allocations are individually rational in the sense that no one has a disincentive to participate. Moreover, they are effective in that they induce the supplier to carry the optimal level of inventory for the coalition, so when the grand coalition is formed, the supplier carries the supply-chain-optimal level of inventory:

**Proposition 3.** Strong Supplier allocations coordinate the supply chain.

If Shapley value allocations are in the core, then it is in the best interest of everyone to pool inventory together. If Shapley value allocations are not in the core, then they are not stable, in the sense that the collaboration may break up into subsets. This would leave money on the table by failing to extract all possible savings from pooling inventory. Shapley value allocations are not in general in the core except, as proved by Shapley (1971), when the game is convex.

**Definition 1.** A game with a set of players \( K \) and characteristic function \( v \) is convex if

\[
v(R \cup \{i\}) - v(R) \geq v(P \cup \{i\}) - v(P) \quad (9)
\]

for all \( i \in K \) and all \( P, R \subseteq K \setminus \{i\} \) with \( P \subset R \).

Özen et al. (2005) study newsvendor games, which arise when \( N \) retailers centralize their inventories and the total supply chain profit is allocated among them after demand realization. They prove that newsvendor games are convex when retailer demands lie within certain classes of independent, symmetric, unimodal distributions. The supplier is not part of these newsvendor games, which simplifies the analysis. However, we are able to show that the convexity of newsvendor games implies convexity of our game and vice versa (in Lemma 1 in the online appendix), which enables us to conclude,

**Proposition 4.** The Shapley value allocation is always in the core when the demand at each retailer

- is normally distributed or,
- has a symmetric, unimodal distribution and the critical fractile is 0.5.

Proposition 4 holds for any number of retailers. If there are only two retailers, the Shapley value is always in the core (Bartholdi and Kemahlıo˘glu-Ziya 2005). However, inventory-pooling games are not always convex (because of the Özen et al. 2005 result that newsvendor games are not convex in general and Lemma 1 in the online appendix), and so Shapley value allocations can fail to be in the core.

Proposition 4 proves that the Shapley value allocations are in the core for specific demand distributions. Our next result states that in our inventory-pooling game, the Shapley value allocations are always in the core (regardless of demand distribution) if all \( N \) retailers are identical.

**Proposition 5.** The Shapley value allocation is in the core when all the retailers are identical.

5.1.1. Is the Core the Right Measure of Stability? Section 5.1 establishes that Strong Supplier allocations are always in the core, whereas Shapley value allocations may not be. However, this does not necessarily imply that Strong Supplier allocations are superior to Shapley value allocations as a means to divide the expected excess profit among supply chain parties. For one thing, Strong Supplier allocations may not be attainable. It is hard to see how a decentralized supply chain could migrate to Strong Supplier allocations unless forced by the supplier. It is natural for the retailers to assert a claim on the excess profits that their participation is responsible for generating. Similar issues arise in some of the core allocations (the dual allocation) proposed by Anupindi et al. (2001), in which players contributing a scarce resource divide the total worth of the coalition among themselves, whereas players contributing resources, of which there are a surplus, receive an allocation of zero.

Although Shapley value allocations may not be in the core, this need not be a telling objection. Membership in the core provides a simple kind of stability; however, as in the case of Strong Supplier allocations, it does not guarantee that the allocation is achievable in practice. It can also fail to prevent strategic manipulation. For example, Young (1985) shows that one-point allocations that are always in the core when the core is nonempty (e.g., the nucleolus) may allow players to increase their allocations by decreasing their contributions to the game (the technical term is that they...
may not be coalitionally monotonic), whereas a player’s Shapley value allocation increases when his contribution to the game increases. If it is important that the players have incentives to increase the total worth of the coalition they are in (as is the case in our context) then coalitionally monotonic allocation rules such as the Shapley value are more suitable (Maschler 1992).

5.1.2. Farsighted Stability. The concept of the core has other shortcomings as a stability measure, such as its myopic nature. From this myopic point of view of stability, the grand coalition is considered unstable when there exists a subset of players who benefit from a one-step deviation from the grand coalition. In other words, the new coalition formed by these players is assumed to be final. However, in many realistic situations one such deviation may trigger subsequent deviations by other players, and at the end of these deviations, the players who initiated the deviations may arrive back at the grand coalition or, worse, at a coalition where they receive less than they did under the grand coalition. Such a possibility could deter the initial deviation, and the grand coalition, which appeared unstable from a myopic view, could deter the initial deviation, and the grand coalition is farsightedly stable.

Theorem 3. The grand coalition is in the LCS and is farsightedly stable (discussion and example adapted from Chwe 1994).

Definition 2. \(X_1\) is directly dominated by \(X_2\) (denoted by \(X_1 \prec X_2\)) if there exists a coalition \(C\) such that \(X_1 \rightarrow_C X_2\) and \(X_1 \prec_C X_2\).

As an example, let \(X_1\) be the grand coalition. If there exists another coalition structure such that \(X_1 \prec X_2\), the allocation \(u^{X_1}\) cannot be in the core.

Definition 3. \(X_m\) is indirectly dominated by \(X_{m+n}\) (denoted by \(X_m \prec X_{m+n}\)) if there exist coalition structures \(X_m, X_{m+1}, \ldots, X_{m+n}\) and coalitions \(C_m, C_{m+1}, \ldots, C_{m+n-1}\) such that \(X_i \rightarrow_C X_{i+1}\) and \(X_i \prec_C X_{m+n}\) for \(i = m, m+1, \ldots, m+n-1\).

Following from the above example, say that a coalition \(C_1\) strictly prefers \(X_2\) to \(X_2\) but cannot move from \(X_2\) to \(X_1\) in one step. From a myopic point of view of stability, \(C_1\) is stuck at \(X_2\). However, if \(C_1\) is able to move to \(X_3\) and then another coalition, \(C_2\), is able to move to \(X_4\) and \(X_3 \ll C_4\), this implies \(X_3 \ll C_4\). Now consider the coalition that has incentive to move from \(X_1\) to \(X_2\); if at least one member does not strongly prefer \(X_4\) to \(X_2\), the possibility of eventually ending up in \(X_4\) deters their move in the first place. Hence, even though \(X_1\) is not myopically stable, it may be stable in the farsighted sense (discussion and example adapted from Chwe 1994).

The example demonstrates that a deviation from a coalition structure is deterred if there exists a stable ending outcome that the deviating coalition does not prefer, and an outcome is stable if all deviations from it can be deterred. This implies that the set of stable outcomes must be consistent and Chwe (1994) calls a set \(Y\) consistent if the following condition holds: \(X \in Y\) if and only if for all \(\hat{X}, C\) such that \(X \rightarrow C \hat{X}\) there exists \(B \in Y\) where \(\hat{X} = B\) or \(\hat{X} \ll B\) such that \(X \ll C\). Chwe proves the existence and the uniqueness of the largest consistent set; that is, the consistent set that contains all consistent sets.

LCS is a weak predictor of coalitional dynamics: it eliminates outcomes that cannot possibly be stable, but does not determine which outcome will be the prevailing one. Our next theorem shows that the grand coalition is in the LCS and is farsightedly stable. Therefore, if the status quo in the supply chain is the grand coalition (i.e., all retailers and the supplier agree to pool inventory and share excess profits with respect to their Shapley values), deviations are unlikely.

Theorem 3. Under Shapley value allocations, the grand coalition is a farsightedly stable outcome.
The theorem does not rule out other farsightedly stable outcomes. However, similar to Sošić (2006), one may argue that the grand coalition is the most likely stable outcome because it maximizes the total supply chain profits. Therefore, for the rest of the paper, we will assume that the grand coalition is formed under Shapley value allocations.

5.2. Effects of Coalition Size on Supplier and Retailer Profits

Although the stability of the grand coalition implies that no player can achieve a higher allocation by breaking away, it does not necessarily imply that individual players’ allocations are maximized under the grand coalition. As more retailers in the supply chain join an inventory-pooling coalition, it seems natural to expect that the savings, and therefore the allocations, will grow as well. Indeed, this is true for the supplier.

**Theorem 4.** The Shapley value allocation to the supplier is nondecreasing in the number of retailers in the coalition.

As the next example demonstrates, the profit allocations to the retailers may diminish as the coalition grows. This might happen if the game is not convex, as is the case with Example 1.

**Example 1.** Consider three retailers. Retailer 1’s demand is 0 with probability 0.5 and 1 with probability 0.5; retailer 2’s demand is 0 with probability 0.4 and 1 with probability 0.6; retailer 3’s demand is 0 with probability 0.01 and 1 with probability 0.99. Let \( p = 1 \), \( p_M = 0.2 \), \( w = 0.6 \), and \( h = 0 \). Prior to inventory pooling, the optimal order quantities for retailers 1 and 2 are zero and that for retailer 3 is one. Retailer 3’s before-pooling expected profit is 0.198 and that of the supplier is 0.39. The optimal order quantities and expected profits for all possible coalitions (coalitions that do not include the supplier are left out because they do not generate value) are:

<table>
<thead>
<tr>
<th>Retailer(s) in coalition</th>
<th>Order quantity</th>
<th>Expected profit in coalition</th>
<th>Retailer(s) in coalition</th>
<th>Order quantity</th>
<th>Expected profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1 and 3</td>
<td>1</td>
<td>0.5940</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.120</td>
<td>2 and 3</td>
<td>2</td>
<td>0.7080</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.588</td>
<td>1, 2, and 3</td>
<td>2</td>
<td>0.9516</td>
</tr>
<tr>
<td>1 and 2</td>
<td></td>
<td></td>
<td></td>
<td>1 and 2</td>
<td>0.360</td>
</tr>
</tbody>
</table>

The Shapley value allocation to retailer 3 is 0.002 when she is in a coalition with retailer 1 and the supplier, but her allocation is only 0.0014 when she is in the grand coalition.

Sprumont (1990, Corollary 2) shows that for convex games, the Shapley value allocations to all players increase as the coalition size grows. This result, coupled with Proposition 4, means that when demand at the retailers is normally distributed or has a unimodal, symmetric distribution and the critical fractile is 0.5, the Shapley value allocations to all players grow as the coalition grows. Hence, for these special types of games, membership in the grand coalition maximizes not only the total allocation to any subset of players, but also the player’s individual allocations.

6. Implementation of Shapley Value Allocations

We have shown that if the parties share the expected gain from inventory centralization in such a way that each receives her Shapley value in expectation, the retailers will agree to switch to a shareable inventory policy and the supplier will carry the supply-chain-optimal amount of inventory. In this section, we propose a mechanism that rations inventory among retailers when total demand exceeds total supply and charges them in such a way that all players’ shares of the expected gain from pooling are equal to their Shapley values. In addition, the mechanism satisfies the individual rationality condition for the retailers for each demand realization vector \( \bar{d} \); in other words, for any demand realization \( d_i \) and inventory allocation \( a_i \), retailer \( i \) makes at least as much profit as he would have made given his inventory level \( x_i^* \) under the reserved inventory management policy.

We first describe the modified linear allocation rule, which we use to ration inventory to the retailers when total demand exceeds total supply. The modified linear allocation rule builds on the linear allocation rule proposed by Cachon and Lariviere (1999). (Any other rationing rule that allocates inventory efficiently could be used instead. Efficiency is required to guarantee that the supplier orders the first-best quantity.) Recall that \( x \) is the total inventory level at the supplier. The supplier first ranks the retailers in decreasing order of \( d_i \) (the linear allocation rule ranks retailers based on the orders they placed) and allocates each retailer \( a_i \) where

\[
a_i = \begin{cases} 
  d_i - \frac{1}{n} \max \left\{ 0, \sum_{j=1}^{\hat{n}} d_j - x \right\} & i \leq \hat{n}, \\
  0 & i > \hat{n},
\end{cases}
\]  

(10)

and \( \hat{n} \) is the largest integer less than or equal to the total number of retailers such that \( a_n \geq 0 \). Although this allocation mechanism may maximize the total profit of the retailers, it may assign zero inventory, resulting in zero profit, to retailers with low demand (Cachon and Lariviere 1999). Therefore, implementing this allocation by itself may not satisfy the individual rationality constraints of the retailers and will not
induce inventory centralization. Instead, consider the following mechanism:

- Prior to demand realization, each retailer $i$ reserves $x_i^*$ units of inventory at a unit price of $p$.
- After demand realization, the supplier allocates inventory $a_i$ to each retailer $i$ according to the modified linear allocation rule.
- The supplier bills or refunds retailer $i$ based on how large $a_i$ is when compared to $x_i^*$.

The third step, the billing process, requires more explanation. Given that $N$ retailers switch to the shareable inventory policy and the total inventory to be shared is $x$, let $\pi_i^1(x, \sum_{j=1}^N d_j)$ denote the supply chain’s total profit for demand realization $\sum_{j=1}^N d_j$, and let $\Pi_i^0(x)$ denote its expectation. Define $\alpha_i = \phi_i/(\Pi_i^1(x) - \Pi_i^0(x)\max(0, x - x_i^*))$ as the ratio of retailer $i$’s Shapley value allocation to the expected contribution retailer $i$ makes to the coalition of the remaining $(N - 1)$ retailers (the inventory available to the coalition of $N - 1$ retailers is $\max(0, x - x_i^*)$ because we account for the unlikely event that $x_i^* > x$).

Let $z_i$ be the unit surcharge or refund charged or paid per unit of inventory allocated to retailer $i$ and let $\varphi_i(x, x_i^*, d_i) = \alpha_i(\pi_i^0(x, \sum_{j \neq i} d_j) - \pi_i^1(x, \sum_{j \neq i} d_j))$ be the extra profit allocated to retailer $i$ under the shareable inventory policy. Then $z_i$ is determined such that the following equality holds for every retailer for whom $a_i \neq x_i^*$:

$$(p + p_M) a_i - p x_i^* - z_i |a_i - x_i^*| = p_M \min\{x_i^*, d_i\} + \varphi_i(x, x_i^*, d_i).$$  \hspace{1cm} (11)

If, for a retailer $i$, $d_i = x_i^*$, then $z_i$ satisfies

$$(p + p_M) x_i^* - (p + z_i) x_i^* = p_M \min\{x_i^*, d_i\} + \varphi_i(x, x_i^*, d_i).$$  \hspace{1cm} (12)

**Theorem 5.** If inventory at the supplier is allocated to the retailers using the modified linear allocation rule and the price adjustment (surcharge or refund) is calculated according to Equations (11) and (12), then all players’ expected after-pooling profits are equal to their Shapley value allocations plus their before-pooling profits, and the supplier carries the supply-chain-optimal level of inventory.

In the mechanism proposed in Theorem 5, for each demand realization, each retailer makes the profit he would have made if he had $x_i^*$ units to sell plus a fraction of his contribution to a coalition of $(N - 1)$ retailers. Others (e.g., Anupindi et al. 2001) have proposed profit allocation mechanisms in which the allocation to each player is determined in part by some fraction of the total supply chain profit for a given demand realization. However, they do not specify how these fractions should be determined. We propose a mechanism where each player receives a fraction that is a function of her Shapley value, which can be defended on notions of fairness.

### 7. Effects of Demand Variance and of Retailer Markup

The main objective of inventory pooling is to mitigate the negative effects of demand variance on supply chain profits. Therefore, it is worthwhile to understand how the expected after-pooling profits of the players change with demand variance and whether they will have incentive to reduce demand variability under the shareable inventory management policy. Demand variability can be reduced by improving forecasting accuracy, but supply chain parties can affect demand variability in other ways, too. For example, a retailer may inadvertently increase his demand variance by encouraging his customers to buy in quantities that are different from their demands, or selling the same product with many different packaging options. If the retailer’s expected after-pooling profit is decreasing in demand variance, he has incentive to reduce demand variability and discontinue such practices. Or if the supplier’s expected after-pooling profit is decreasing in demand variance, she may provide incentives (e.g., discounts) to the retailers to discontinue such practices. Our goal is to understand when the incentives for variability reduction exist under Shapley value allocations.

Because a player’s incentive to decrease demand variance depends on how variance affects her expected total profit rather than her share of expected extra profits due to pooling, we explore the relationship between demand variance and the expected total profits. As we discover, there is an intricate relationship among demand variance, retailer markup, and expected after-pooling profits.

Consider two supply chains. Let demand of retailer $i$ in supply chain 1 be denoted by $X_i$, and distributed with $F_i(\cdot)$, and that in supply chain 2 be denoted by $Y_i$ and distributed with $G_i(\cdot)$. For all $i$, assume that $X_i, Y_i$ are nonnegative, $E[X_i] = E[Y_i] < \infty$, and $X_i \leq_{ew} Y_i$ where $\leq_{ew}$ indicates $Y_i$ is larger than $X_i$ in excess wealth order (also known as right spread order, Shaked and Shanthikumar 2007). (We provide a summary of this and other stochastic orders in the online appendix.) Further assume that for each $i$, $F_i(\cdot)$, and $G_i(\cdot)$ cross exactly once and let $z_i$ denote the crossing point. Note that this is not a restrictive assumption; for example, if the two random variables belong to the same member of the location-scale class (e.g., gamma, Weibull, uniform, etc.) or the generalized location-scale class (e.g., lognormal) distribution class, the assumption holds (Zhang 2005, Song 1994).

The normal distribution, which is frequently invoked as a demand distribution, also belongs to the location-scale distribution class. However, because the normally distributed random variables may take on negative values, we treat this case separately.
Excess wealth order is a *variability* order; in particular, of two random variables the one that is larger in the excess wealth order has a higher variance. Using the properties of excess wealth order, one can show that the demand of any coalition in supply chain 2 is more variable than the demand of the same coalition in supply chain 1. Under the assumptions we make, this also implies that the total expected profit is lower in supply chain 2 (we formally prove this claim as part of the proof of Proposition 6). Because there is less to share in supply chain 2, does that imply that all players’ expected profits will be lower? Our next result answers this question for the supplier.

**Proposition 6.** For each retailer *i*, assume that $G_i \geq \text{ess sup} \ F_i$, $E[X_i] = E[Y_i] < \infty$, and $F_i(\cdot)$ and $G_i(\cdot)$ cross exactly once. Then the following hold:

(i) If $p_M \leq p + h$ and $\max(p, (p-w)/(p+h)) \geq \max z_i$, the supplier’s expected after-pooling profit is higher in the supply chain where retailer *i*’s demand is distributed with $F_i(\cdot)$ for $i \in \{1 \ldots N\}$.

(ii) There exist a threshold markup level, $p_M^*$, such that for $p_M > p_M^*$ the supplier’s expected after-pooling profit is higher in the supply chain where retailer *i*’s demand is distributed with $G_i(\cdot)$ for $i \in \{1 \ldots N\}$.

Since the total supply chain profit is lower when retailer *i*’s demand is distributed with $G_i(\cdot)$ and is more variable, intuition suggests that the supplier’s expected after-pooling profit would be lower, too. This is indeed the case when retailer markup $p_M$ is low, but not if retailer markup is high enough. Recall that the expected total after-pooling profit of the supplier is equal to her expected *before-pooling profit* plus her *Shapley value share* of the expected increase in supply chain profit because of pooling. When demand variance increases, the supplier’s expected before-pooling profit goes down, and the magnitude of this decrease is independent of $p_M$. At the same time, when demand variance increases, the expected increase in supply chain profit (because of pooling) goes up due to the higher level of pooling benefit. This increase is even more pronounced for products that have high retailer markups because stocking out on these products costs the supply chain more and when demand can be better matched with supply through pooling, the expected increase in supply chain profit is higher. When $p_M$ is large enough, the decrease in supplier’s expected before-pooling profit is more than offset by the increase in her share of the expected gain in supply chain profit. This implies that under Shapley value allocations, for products with high markups, the supplier prefers that the retailers face higher demand variance.

To make sharper conclusions, we make additional assumptions. Unless otherwise stated, in the remainder of the paper we assume that the retailers’ demands are independent and normally distributed. In addition, we follow Yang and Schrage (2009) in making the reasonable assumption that the probability of no stockout at the retailers is at least 50% before pooling. Our next result says that, everything else being equal, the retailer has incentive to lower the demand variance of a product on which he makes a higher margin.

**Proposition 7.** When retailers’ demands are independent, identical, and normally distributed with mean $\mu$ and standard deviation $\sigma$, and they each impose a service level requirement of $\rho$, there exist two levels of markup, $p_M^1 \leq p_M^2$, such that:

- For markup levels less than or equal to $p_M^1$, the supplier’s expected after-pooling profit is decreasing and the retailer’s expected after-pooling profit is increasing in demand variance.
- For markup levels greater than or equal to $p_M^2$, the retailer’s expected after-pooling profit is decreasing and the supplier’s expected after-pooling profit is increasing in demand variance.

Proposition 7 confirms our insights regarding the supplier’s expected profit and also states that the retailer’s expected after-pooling profit is increasing in demand variance for low levels of retailer markup and decreasing for high levels of markup. Like the supplier, the expected after-pooling profit of the retailer is equal to his expected before-pooling profit plus his Shapley value share of the expected profit increase because of pooling. When demand variance increases, the retailer’s expected before-pooling profit goes down, but unlike the case of the supplier, the magnitude of this decrease is increasing in $p_M$. For products with low $p_M$, this decrease is offset by the increase in the total additional profit because of pooling, but for products with high $p_M$, the decrease is so large that it cannot be offset, and for these types of products, the retailer’s expected after-pooling profit is decreasing in demand variance.

Proposition 7 by Proposition 7 the retailers’ and the supplier’s incentives to reduce demand variance can be aligned only for products with moderate values of retailer markup. For these types of products, the increase in the expected pooling benefit as demand variance increases is not high enough to offset the decrease in either the supplier’s or the retailer’s expected before-pooling profits.

### 8. Colluding Retailers and Strategic Bargaining

Do retailers have an incentive to collude when the expected savings from inventory pooling are distributed according to Shapley value? How does collusion affect noncolluding retailers? Up to this point, we have assumed that the supplier proposes a switch
to the shareable inventory management policy where the excess profit from pooling will be shared such that in expectation each player’s share of the excess is her Shapley value. Then the retailers decide whether to participate in the pooling coalition. However, once the supplier proposes the switch, the retailers may first want to collude among themselves and act as a single retailer with a demand function corresponding to the sum of their demands. The collusion of the retailers could arise from a buying alliance, or a group of retailers sharing the same parent company where purchasing is centralized, or, as in the electronics industry, a number of OEMs who have delegated component procurement to their common contract manufacturer.

We assume that once a subset of the retailers collude to act as a single retailer in negotiating their after-pooling savings allocation, the supplier treats them as a single retailer (the aggregate retailer) and determines the stock level accordingly. We continue to allow the retailers to impose service-level constraints under the reserved inventory management policy, but assume that either the constraints are not binding (the supplier finds it optimal to carry more inventory) or all the retailers impose the same service-level constraint. The latter is usually the case in practice where, for example, customers in the electronics industry require their suppliers to carry two-weeks’ demand in inventory (Barnes et al. 2000). In the absence of this assumption, it is not clear what the service level requirement of the aggregate retailer would be.

In what follows, we define the optimal collusion as the set of colluding retailers that generates the highest expected total profit for its members. Because collusion among the retailers does not change the total expected profit of the supply chain, an increase in the expected total profit of the colluding retailers implies that their share of the total profit goes up, and we are interested in identifying when collusion leads to such an increase. As long as their total expected profit goes up when compared to the no-collusion case, the colluding retailers can divide the additional profit among themselves such that in expectation each receives more than he would without the collusion. Because our aim is to understand how the total expected profits of the colluding retailers change vis-a-vis those of the supplier and the noncolluding retailers, we abstract away from the details of dividing the profit among the colluding retailers.

Because of the pooling effect, the demand from the aggregate retailer has a lower coefficient of variation (and the aggregate retailer’s expected before-pooling profit is higher), and so one would expect the total after-pooling profits of the colluding retailers to be higher, too. We would also expect the collusion to adversely affect the expected profits of the supplier and any nonparticipating retailers. Surprisingly, these intuitions do not always hold. We first analyze two-retailer and three-retailer supply chains and determine when it is beneficial for the retailers to collude before cooperating with the supplier. It seems difficult to determine analytically the optimal collusion structure for the general N-retailer case, but numerical tests suggest that intuitions from the two and three-retailer cases extend.

8.1. Collusion in a Two-Retailer Supply Chain

Intuition suggests that the retailers will always increase their expected after-pooling profits if they collude. However, our result shows that the retailers may in fact be the worse for it.

Proposition 8. Collusion reduces the expected after-pooling profits of the retailers when both of the following conditions hold:

• $p_M \leq p + h$, and
• the optimal inventory level under inventory pooling is less than the total optimal inventory level without pooling.

The retailer’s and the supply chain’s margins on a product sold are $p_M$ and $p + p_M - w$, respectively, and for higher levels of $p_M$, the retailer’s share of the margin on the product is higher. In addition, everything else being equal, the higher $p_M$ is, the higher the retailer’s expected before-pooling profit. Therefore, when $p_M$ is low and the retailers’ expected before-pooling profits are low, one might think it is advantageous for them to collude. However, our result indicates that if $p_M$ is low enough, then the retailers are worse off when they collude! When retailers collude, their expected before-pooling profits increase. The expected total before-pooling profit of the supply chain increases, too, because the effective demand variability faced by the aggregate retailer is lower. Hence, the expected gain from pooling goes down. The retailers’ expected after-pooling profits go down because of collusion when the increase in their expected before-pooling profits cannot make up for the drop in their share of the expected gain from pooling. This is more likely to happen when $p_M$ is low.

The second part of the sufficient condition requires that the optimal inventory level goes down as a result of inventory pooling. If the retailers are to collude in the parameter region where pooling reduces the optimal inventory level, the supplier benefits in the sense that the level of inventory she needs to carry goes down (because of the reduction in demand variance). This increases the supplier’s base profit in the Shapley value allocation scheme and leaves the retailers worse off. As Gerchak and Mossman (1992) and Yang and Schrage (2009) have shown, the optimal
inventory level under pooling is not always lower than that before pooling, but when demand is normally distributed (and under our assumption that the critical fractile before pooling is greater than a half), the optimal inventory level under pooling is always lower and we can state that collusion reduces the profit allocations of the retailers if $p_M \leq p + h$.

8.2. Collusion in a Three-Retailer Supply Chain

We can completely characterize the optimal collusion structure for a supply chain with three retailers when demand is normally distributed. Our first result shows that at low levels of retailer markup, retailer collusion is disadvantageous, but as retailer markup increases, the size of the optimal collusion grows to encompass all retailers.

**Proposition 9.** When demand is normally distributed there exist two levels of markup, $p_M^1 \leq p_M^2$, such that for $p_M^1 < p_M < p_M^2$, collusion reduces the expected after-pooling profits of the retailers. For $p_M^2 < p_M < p_M^3$, collusion of the retailers with the two lowest standard deviations, and for $p_M^3 < p_M$, collusion of all three retailers maximizes the expected after-pooling profits of the colluding retailers.

Proposition 9 states that retailers prefer to increase collusion size as the retailer markup increases. Similar to Proposition 8, the intuition for this result is that the increase in the expected before-pooling profits of the colluding retailers makes up for the drop in their share of the expected excess profit from pooling only when $p_M$ is high. Another implication of Proposition 9 is that, of the three possible two-retailer collusions, only the one between the retailers with the two lowest standard deviations is part of the optimal collusion structure. Collusion among retailers with higher demand variability would be profitable only at higher $p_M$ values because the expected before-pooling profit of such a collusion is low because of high demand variability. However, our result shows that at such higher levels of $p_M$, collusion of all three retailers dominates any two-retailer collusion.

Our next result shows that the supplier, on the other hand, prefers the largest retailer collusion up to a threshold level of retailer markup, and above that threshold, prefers that the retailers do not collude at all. The supplier’s expected after-pooling profit is her expected before-pooling profit plus her share of the expected excess profit from pooling. When retailers collude, the effective demand variability observed by the supplier goes down and her expected before-pooling profit goes up. The supplier’s expected before-pooling profit is indeed highest when all retailers collude because the demand variability is minimized. However, retailer collusion simultaneously reduces the expected gain from pooling, negatively impacting the supplier’s expected after-pooling profit. When $p_M$ is low, the increase in the supplier’s expected before-pooling profit offsets the decrease in her share of the expected excess profit from pooling, but when $p_M$ is high, the retailers claim such a large share of the excess profit from pooling that the supplier’s expected after-pooling profit goes down.

**Proposition 10.** Under normally distributed demand and up to a threshold level of $p_M$, the supplier’s expected profit is highest when all three retailers collude. Above that threshold, the supplier’s expected profit is highest when the retailers do not collude at all.

Propositions 9 and 10 indicate an anticipated conflict of interest between the retailers and the supplier: when $p_M$ is low, no collusion is optimal for the retailers, whereas three-retailer collusion is optimal for the supplier. So, is it possible for the supplier to benefit from optimal retailer collusion? To answer this question we provide Figure 1, which depicts the ordering of preferences for both types of players with respect to different levels of $p_M$ (the different threshold levels of $p_M$ that divide the regions follow from the proofs of Propositions 9 and 10). For example, an ordering of $1 > 2 > 3$ for a player means that the expected profit is highest under no collusion and lowest under three-retailer collusion. When we refer to a two-retailer collusion, we mean the collusion between the retailers with the two lowest standard deviations. The shaded region on the supplier’s line of preferences designates the only region where the supplier benefits from optimal retailer collusion.

Finally, another interesting question is what happens to the expected after-pooling profit of the noncolluding retailer in the region where two-retailer collusion is optimal. As expected, the noncolluding retailer’s expected after-pooling profit is always lower than that prior to collusion.

**Proposition 11.** Under normally distributed demand, when two retailers collude, the noncolluding retailer’s expected after-pooling profit is lower than his expected after-pooling profit prior to the collusion of the other retailers.

To summarize, collusion will not always benefit the colluding retailers. However, if collusion size and...
structure is chosen optimally with respect to the cost and revenue parameters, colluding retailers may increase their profit shares at the expense of the non-colluding retailer and/or the supplier.

8.3. Collusion in an N-Retailer Supply Chain
For \( N > 3 \), it is difficult to analytically derive the optimal collusion size or structure. Even when we consider only collusion structures where a subset of the retailers collude and the rest stand alone, the total number of alternatives to be compared is \( 1 + \sum_{k=2}^{N} \binom{N}{k} \). If we allow the retailers to further break into groups of 2, 3, etc., the number of alternative collusion structures is prohibitively large for analytical comparison. Therefore, our aim here is to test if the results from §8.2 can be used to provide insights for the general \( N \)-retailer case.

Our analysis of the two- and three-retailer supply chains suggests that the relationship between the retailer markup \( p_M \) and the supplier’s cost and revenue parameters determine the optimal collusion structure. Demand variances of the potentially colluding retailers also impact the profitability of collusion. Proposition 9 states that for a three-retailer supply chain only the collusion of the two retailers with the lowest standard deviations of demand can be optimal. Therefore, for a 10-retailer supply chain, we order the retailers in increasing order of standard deviation. We start with the collusion of the two retailers with the lowest two standard deviations, add the retailers on the list one by one to the collusion, and test whether larger collusion becomes more profitable as \( p_M \) increases. We assume that the mean demand at all retailers is 10 and the standard deviations change between 1 and 1.9 in increments of 0.1. By keeping the mean constant, we keep increasing the coefficient of variation as standard deviation increases. We let \( p = 2 \), \( w = 1 \), and \( h = 0.1 \). In Figure 2, we plot the expected profit gain achieved by adding the next retailer to the list of retailers ordered by standard deviation of demand to the collusion. We compute the expected total profit of a \( k \)-retailer collusion and that of a \( (k - 1) \)-retailer collusion plus the expected profit of the \( k \)th retailer and plot the difference in expected total profits. For a \( k \)-retailer collusion to be preferable to a \( (k - 1) \)-retailer collusion, this difference must be positive. We observe that for low values of \( p_M \), even the collusion of the two retailers with the lowest standard deviations is not profitable. As \( p_M \) increases, the marginal gain from adding the next retailer increases. For high enough \( p_M \), adding the retailer with the highest standard deviation of demand results in the highest marginal gain. Our observations from Figure 2 are consistent with Proposition 9 in that collusion of a large number of retailers is more profitable when \( p_M \) is high.

8.3.1. Effects of Collusion on the Supplier. As Figure 3 illustrates, for low levels of retailer markup, the supplier prefers the largest-possible collusion. As retailer markup increases, the supplier’s expected after-pooling profit first goes down with increasing collusion size, reaching a minimum, and then goes up as collusion size increases, but overall the supplier prefers the no-collusion scenario. At even higher levels of retailer markup, the supplier’s expected profit is decreasing in collusion size and the supplier prefers the no-collusion scenario. Again, our observations are consistent with Proposition 10. Collusion of retailers reduces the effective demand variance the supplier faces from the retailers and increases her expected before-pooling profit. On the other end, collusion also reduces the risk-pooling benefit due to shareable inventories and, hence, reducing the supplier’s share of such benefit. Depending on the level of \( p_M \), either one of these effects dominates and the supplier either prefers the largest collusion or no collusion.

Figure 2  Expected Marginal Profit Gain from Adding the Retailer with the Next-Highest Standard Deviation to the Collusion

![Graph showing expected marginal profit gain](image-url)
9. Conclusions

There are many properties one might want in a mechanism to allocate excess profits due to inventory pooling along a supply chain. We have suggested some, and explored how well Shapley value serves as an allocation mechanism. It is attractive in some ways, especially in that it coordinates the supply chain and does so fairly (at least according to some formalizations of fairness). The problems we find with it do not seem severe: Shapley value allocations might not be in the core, but this need not be a strong objection because the concept of core has its own deficiencies (Myerson 1991); and, in any event, the grand coalition is farsightedly stable in the inventory-pooling game under Shapley value allocations.

We studied Shapley value allocations in the context of inventory pooling, but our results extend to problems of capacity centralization where capacity must be built in advance to be shared by multiple customers. Such problems arise in the biopharmaceutical and semiconductor industries where the lead time to build manufacturing capacity is so long that customer demand may change significantly in the interim. All can benefit from reallocation of capacity after demand realization, but cooperation depends on how the benefits will be shared. Cooperative game theory is a natural way to model profit sharing in this context. Previous work on capacity pooling, such as Plambeck and Taylor (2005, 2007a, b), followed this approach, but mostly refrained from specifying a bargaining outcome. This has merits in their context, but specifying an outcome, the Shapley value, allows us to make precise observations about the stability of allocations and how their values depend on demand variance and other descriptors of the supply chain.

Our paper also contributes to the literature on supply chain coordinating contracts. All such contracts implicitly divide the total profit among different players. However, as Cachon (2003) observes, most of the literature on supply chain coordination presents one contract parameter in terms of another, rather than defining the contract in terms of each player’s share of the total profit. Cachon prefers the latter approach, as do we, because it is more general and will work even under circumstances in which the former will not. In this paper, through the Shapley value mechanism, we propose a way of computing each player’s “fair” share of the supply chain profit by taking into account their profit levels before cooperation was proposed and their contributions to the inventory-pooling coalitions. In addition, because the shares are recalculated when supply chain circumstances (e.g., demand variance, number of retailers, retailer collusion) change, we can obtains insights on how each player’s share of supply chain profits changes as supply chain settings change.

There are many directions in which our analysis may be extended. One interesting avenue is to study supply chains where players have asymmetric demand information. It also seems worthwhile to explore richer behavior in the formation of coalitions, such as allowing for multiple suppliers, or allowing groups of colluding retailers to negotiate reduced prices from a supplier.

Finally, Shapley value is but one way to allocate savings in a supply chain. There are many others and any choice will require compromises. One thing seems clear: participants will be more favorably disposed to an allocation mechanism that is “fair,” verifiable, and whose behavior has been analyzed in detail. As Maschler (1992, p. 616) argues “The real question is not whether a particular solution is good or bad, but rather: In what circumstances should it be recommended?” This paper is a contribution to this project.

Electronic Companion

An electronic companion to this paper is available on the Manufacturing & Service Operations Management website (http://msom.pubs.informs.org/ecompanion.html).
Acknowledgments
The authors thank Harvey Wagner, Gérard Cachon, Stephen Graves, the associate editor, and the three anonymous referees for comments and suggestions that helped improve the paper.

References