Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Retailer Inventory Sharing in Two-Tier Supply Chains: An Experimental Investigation

Andrew M. Davis, Rihuan Huang, Douglas J. Thomas

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–18
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Retailer Inventory Sharing in Two-Tier Supply Chains: An Experimental Investigation

Andrew M. Davis,[a] Rihuan Huang,[a,*] Douglas J. Thomas[b]

[a] Samuel Curtis Johnson Graduate School of Management, SC Johnson College of Business, Cornell University, Ithaca, New York 14853; [b] Darden School of Business, University of Virginia, Charlottesville, Virginia 22903
*Corresponding author

Contact: adavis@cornell.edu, https://orcid.org/0000-0002-1689-2299 (AMD); rh567@cornell.edu, https://orcid.org/0000-0002-6790-6814 (RH); thomasd@darden.virginia.edu, https://orcid.org/0000-0002-2026-1780 (DJT)

Received: November 18, 2019
Revised: January 3, 2021; June 17, 2021; September 9, 2021
Accepted: September 21, 2021
Published Online in Articles in Advance: February 28, 2022

https://doi.org/10.1287/mnsc.2022.4323

Abstract. When multiple retailers hold inventory to satisfy random demand, retailer inventory-sharing strategies can potentially reduce the supply-demand mismatch and increase overall supply chain performance. In this paper, we experimentally investigate alternative inventory-sharing strategies in a two-tier supply chain with an upstream manufacturer and two downstream retailers. In one setting, retailers act as if they are centralized and use a single quantity to fulfill joint demand. In the other, retailers are decentralized and face separate demands, but they can transfer inventory after demands are realized. In this latter decentralized scenario, we also consider whether the manufacturer or retailers have decision authority over the inventory transfer price. One key result is that when the retailers are decentralized and the manufacturer sets the transfer price, both retailers and the manufacturer earn higher profits than in the centralized retailer strategy, which runs counter to theory. We also find that when retailers are decentralized and set their own transfer price, the most equitable distribution of profits is achieved. In an effort to account for these results, we find that a model of fairness captures decisions well. Overall, by investigating how different inventory-sharing strategies affect the distribution of profits in a two-tier supply chain, our results provide guidance to firms considering how, if at all, they should enter such arrangements.

History: Accepted by Jay Swaminathan, operations management.
Funding: The authors acknowledge financial support from Cornell University and the University of Virginia.
Supplemental Material: The data files and electronic companion are available at https://doi.org/10.1287/mnsc.2022.4323.

Keywords: inventory sharing • behavioral operations • supply chain contracting • risk pooling

1. Introduction

Managing random demand is a challenging problem for supply chains. Even in a two-tier supply chain between a single upstream manufacturer and a single downstream retailer, companies must resort to relatively complicated solutions to address it, such as coordinating contracts. However, in a two-tier supply chain with *multiple* retailers, companies have another lever at their disposal for managing the supply-demand mismatch problem: sharing inventory across retailer locations.

Retailer inventory-sharing strategies vary in two key dimensions: whether the retailers operate in a centralized or decentralized manner and when retailers are decentralized, whether the retailers or manufacturer has decision authority over the inventory transfer price. First, retailers may opt to act as if they are effectively centralized and use a single inventory to satisfy joint demand (e.g., Eppen 1979), or they may act independently in a decentralized manner, initially serving their own demand and then transferring inventory at a per unit transfer price (e.g., Robinson 1990). This centralized/decentralized attribute may be fixed given the supply chain setting, but frequently, it may be an organizational choice. Large firms with multiple retail outlets often act in a centralized way, but many choose to grant store managers significant autonomy such that retail outlets effectively operate as if they are decentralized (e.g., DeHoratius and Raman 2007, Van Donselaar et al. 2010). Decentralized inventory sharing has been observed in several industries, including automobiles (Zhao et al. 2005), steel (Robinson 1990), commodities (Park et al. 2016), and fashion (Dong and Rudi 2004). Conversely, franchise networks, such as several of the largest convenience store chains, are technically independent retailers but could coordinate in a centralized manner through the franchisor. Second, once in possession of inventory, retailers may have the right to transfer or resell to other retailers at terms of their choosing, whereas a manufacturer with a strong brand and control over pricing and distribution may be able to dictate terms of inventory transfer between independent authorized retailers (e.g., Padmanabhan et al. 2010).

Although inventory sharing is a common way to increase supply chain efficiency, it may not be the case that all members of the supply chain benefit from a particular strategy. In a two-tier supply chain, theory predicts that certain retailer inventory-sharing strategies may actually lead to lower retailer profits compared with a setting without any inventory sharing at all. This is largely because of an upstream manufacturer having the ability to set the wholesale price for downstream retailers and in the decentralized retailer inventory-sharing strategy, different parties having decision authority over the inventory transfer price (e.g., Shao et al. 2011).

In this paper, we investigate how the potential benefits of inventory sharing are distributed depending on the centralized/decentralized retailer structure and the decision authority over the transfer price (manufacturer or retailers). We consider a two-tier supply chain where an upstream manufacturer endogenously proposes a wholesale price to two downstream retailers. Because supply chain contract decisions are made by managers (e.g., DeHoratius and Raman 2007, Van Donselaar et al. 2010, Zhao et al. 2021), we employ a behavioral approach, complementing existing theoretical research to understand how, if at all, firms should adopt an inventory-sharing strategy. Specifically, we address the following research questions. First, how do inventory-sharing strategies under centralized and decentralized retailer settings compare in terms of distribution of profits? Second, in the decentralized retailer setting, how are profits distributed when the manufacturer, versus the retailers, has decision authority over the transfer price?

We believe our study is the first to address these research questions from a behavioral standpoint. To briefly outline how our work relates to the existing theoretical literature, it is well established that a single joint-retailer inventory, under exogenous prices, can increase retailer profits compared with retailers acting independently (Eppen 1979). There is also theoretical work that looks at inventory sharing with decentralized retailers, where retailers can share inventory in a recourse stage at a transfer price per unit, often referred to as transshipment (Robinson 1990, Rudi et al. 2001, Dong and Rudi 2004). One especially relevant paper in this realm is by Shao et al. (2011), who investigate a two-tier supply chain with endogenous wholesale prices and compare the centralized and decentralized inventory-sharing strategies, thus serving as a useful theoretical foundation for our experimental investigation.

Turning to the behavioral literature, in a centralized retailer inventory-sharing case, Ho et al. (2010) examine stocking quantity decisions in a one-tier supply chain with exogenous prices. For the decentralized retailer inventory-sharing setting, the behavioral literature generally focuses on order quantities under both exogenous wholesale and exogenous transfer prices (Bostian et al. 2012, Villa and Castañeda 2018, Zhao et al. 2021), whereas there are two papers that examine decentralized retailer inventory sharing with endogenous transfer prices set by retailers. First, Li and Chen (2020) experimentally test the model of Rudi et al. (2001) where both stocking quantities and transfer prices are set by retailers. Second, Katok and Villa (2021) evaluate a setting where each retailer decides its own stocking quantity and where retailers are allowed to negotiate the transfer price. In contrast to these papers, we consider endogenous wholesale prices and vary who sets the transfer price. Overall, our paper extends the behavioral literature by directly comparing alternative retailer inventory-sharing strategies with one another by investigating a two-tier supply chain where a manufacturer endogenously proposes a wholesale price to downstream retailers and by varying which party has authority over the transfer price (in the decentralized environment).

We begin our study by leveraging the existing theoretical literature and outlining the normative theory. This includes details for optimal quantities, wholesale prices, transfer prices, and expected profits. We then develop a set of hypotheses and conduct a controlled human subjects experiment to test these predictions. Our main experiment consists of a no inventory-sharing Baseline treatment plus three inventory-sharing treatments: (1) centralized retailer inventory sharing, (2) decentralized retailer inventory sharing where the manufacturer sets the transfer price, and (3) decentralized retailer inventory sharing where the retailers set the transfer price.

A key experimental result is that the decentralized strategy, when the manufacturer sets the transfer price, generates a win-win outcome compared with both the no inventory-sharing and centralized retailer inventory-sharing strategies: both the manufacturer and retailers earn significantly higher expected profits. Further, the decentralized retailer inventory-sharing strategy, when the retailers set the transfer price, leads to the most equitable outcome in terms of distribution of profits. We also find that these two decentralized retailer inventory-sharing strategies yield favorable supply chain efficiency. In sum, the two decentralized retailer inventory-sharing strategies achieve the highest manufacturer profit, retailer profit, equity, and efficiency.

To determine the driver of these results, we further analyze contract decisions and show that the observed contract-term deviations can account for our profit results: transfer prices are not set at the theoretically predicted extreme values, wholesale prices are set lower than the theoretical predictions in all inventory-sharing environments, and quantities are set well or slightly low relative to the theoretical predictions. To dig deeper into what may account for such deviations, we develop a behavioral model of fairness and find that it can capture the observed transfer prices, wholesale prices, and quantities well.

Our study provides insights for both practitioners and researchers. Regarding the former, we first demonstrate that retailers earn a profit that is higher (or at least as high) under all retailer inventory-sharing strategies compared with no inventory sharing. Second, when choosing among the different retailer inventory-sharing strategies, our results indicate that retailers prefer decentralized inventory-sharing strategies to a centralized one. Turning to research, a majority of behavioral supply chain studies explore either a two-tier setting without retailer inventory sharing or a one-tier setting with retailer inventory sharing (focusing on quantity decisions). We extend this literature by investigating different retailer inventory-sharing strategies, including no sharing, in a two-tier supply chain with endogenous wholesale prices. Our work also complements the theoretical work that examines coordination mechanisms for decentralized firms. Notably, past research has shown that certain incentive structures can lead to favorable results for decentralized settings as compared with a centralized one (Celikbas et al. 1999). Our study finds that behavioral tendencies, notably fairness, can lead to similar results, where firm outcomes are better in decentralized settings.

2. Normative Theory and Predictions

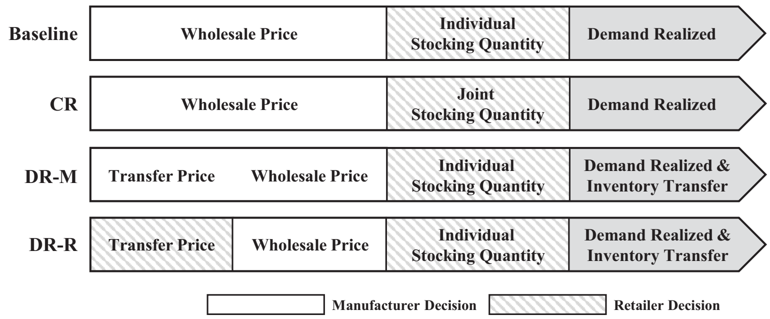
We study a system of one upstream manufacturer and two symmetric downstream retailers. The manufacturer produces a single product at unit cost c and sells it to the two symmetric retailers, indexed by i and j, at wholesale price w. Each retailer decides a quantity, q_i and q_j, purchased from the manufacturer, and sells to its local market with random demand at selling price p. Demands d_i and d_j are independent and follow an identical distribution. Salvage cost is normalized to zero.

For this system, we consider three settings that differ in the inventory-sharing strategy and who has decision authority over the transfer price (if relevant): (1) a no inventory-sharing "Baseline" setting, (2) a *centralized retailer* inventory-sharing strategy (referred to as CR) where retailers make decisions as if they are a single centralized location and set a joint inventory quantity (there is no transfer price), and (3) a *decentralized retailer* inventory-sharing strategy, where retailers share inventory at a transfer price t per unit after demand occurs. In this latter decentralized setting, we also consider who has the decision authority over the transfer price. In one case, the manufacturer sets the transfer price t (referred to as DR-M for *decentralized retailers—manufacturer sets the transfer price*). This setting mimics an environment where a retailer, rather than having to reach out to other individual retailers to see if any have excess units or unmet demand, relies on the upstream manufacturer to help facilitate any sharing. Examples of such "dealer inventory-sharing systems" include Caterpillar, John Deere, and General Motors (Zhao et al. 2005). In the second case, the retailers negotiate and set the transfer price t (referred to as DR-R for *decentralized retailers—retailers set the transfer price*), where retailers take responsibility over their own inventory sharing.

For the Baseline and CR strategies, there are two stages. In stage 1, the manufacturer sets the wholesale price, and in stage 2, the two retailers decide an individual or a joint quantity based on the setting. After quantities are determined, demand is realized, and any inventory sharing automatically occurs. Under the decentralized inventory-sharing strategies, DR-M and DR-R, we follow Shao et al. (2011) and assume that the transfer price is set before the wholesale price, which is common in practice. For example, Narus and Anderson (1996) note that remuneration is decided in advance when firms agree on sharing resources. Also, although there is no difference in predictions in DR-M if we reverse the order of price decisions, if we allow for retailers to set the transfer price after the manufacturer's wholesale price in DR-R, then the normative profit predictions are identical to those in CR. Therefore, the decision sequence in each round of our one-shot environment for DR-M and DR-R consists of three stages. In stage 1, a transfer price is set. In stage 2, the manufacturer sets the wholesale price. In stage 3, the two retailers set stocking quantities. Demands are then realized, and inventory sharing takes place. Figure 1 illustrates these decisions in all four settings.

We note that these four strategies constitute a wide variety of options for retailers and manufacturers. For instance, chain retailers may choose among (a) no inventory sharing (Baseline), (b) a centralized inventory-sharing strategy (CR), or (c) a decentralized inventory-sharing strategy where they set their own transfer price (DR-R). As another example, competing retailers may choose among (a) no inventory sharing (Baseline), (b) a decentralized inventory-sharing strategy where they act independently but agree to share units at a transfer price per unit (DR-R), or (c) a decentralized inventory-sharing

Figure 1. Decision and Event Sequence for Each Inventory-Sharing Strategy



strategy where they yield control of any inventory sharing to an upstream manufacturer, who may have improved visibility into retailers' inventory levels (DR-M). Last, a powerful manufacturer may choose between (a) coordinating any inventory sharing itself as the upstream party and thus, setting the transfer price (DR-M) or (b) allowing the retailers flexibility to set their own terms and manage it themselves (Baseline, CR, or DR-R).

From a theoretical perspective, the main difference between the alternative inventory-sharing strategies is the retailer expected profit function and thus, optimal quantities. The manufacturer's profit function remains the same across all settings:

$$\pi_m = (w - c)(q_i + q_j). \tag{1}$$

In the following discussion, we refer to Equation (1) when showing the manufacturer's profit-maximizing decisions. Unless otherwise noted, we use retailer i when the discussion involves only one retailer, and all results apply to retailer j by symmetry. In the following subsections, we use superscripts b, c, and d for Baseline, centralized retailers (CR), and decentralized retailers (DR-M and DR-R), respectively. Next, we show expected profit functions and optimal decisions by backward induction for each setting.

2.1. No Inventory Sharing—Baseline

When there is no inventory sharing, each retailer faces a standard newsvendor problem. Retailer i decides its stocking quantity q_i to maximize the following expected profit function:

$$\pi^b_{r,i} = \mathbb{E}[p \min(d_i, q_i)] - wq_i. \tag{2}$$

Let $\alpha(\cdot)$ and $f(\cdot)$ denote the cumulative distribution function (CDF) and the probability density function (PDF) of single retailer demand, respectively. The optimal retailer quantity q^b_i and the optimal manufacturer wholesale price must satisfy

$$\alpha(q) = \frac{p - w}{p}, \qquad w = pq^b_i f(q^b_i) + c. \tag{3}$$

2.2. Inventory Sharing—Centralized Retailers

Under the CR inventory-sharing strategy, retailers share a single stocking quantity to satisfy combined demand and maximize their joint expected profit.

Although the centralized retailers will set a single joint stocking quantity, in the following equations we assume that q_i and q_j are each one-half of this joint quantity. The retailer's joint expected profit function is

$$\pi^c_r = \mathbb{E}[p \min(d_i + d_j, q_i + q_j)] - w(q_i + q_j). \tag{4}$$

Let $\alpha^c(q_i, q_j)$ be the probability $\Pr(d_i + d_j < q_i + q_j)$ (i.e., CDF of the joint demand distribution) and $f^c(d_i, d_j)$ be the corresponding PDF. Solving Equation (4) gives the optimal stocking quantity (q^c_i, q^c_j) derived from Equation (5):

$$\alpha^c(q_i, q_j) = \frac{p - w}{p}. \tag{5}$$

Given (q^c_i, q^c_j), the manufacturer's optimal wholesale price is derived from Equation (6):

$$w = p(q_i + q_j) f^c(q_i, q_j) + c. \tag{6}$$

2.3. Inventory Sharing—Decentralized Retailers (DR-M and DR-R)

Under the decentralized retailer inventory-sharing strategies, DR-M and DR-R, each retailer acts independently initially, setting its own stocking quantity to maximize its own expected profit. After the demand is known, an overstocking retailer shares any leftover inventory to an understocking retailer at transfer price t per unit, if possible. For simplicity, t is assumed to be in $[0, p]$. Let $T_i = \min((q_i - d_i)^+, (d_j - q_j)^+)$ be the transferred quantities from i to j (i.e., the minimum of i's leftovers and j's excess demand). Similarly, we define $T_j = \min((q_j - d_j)^+, (d_i - q_i)^+)$ as the transferred quantities from j to i. As with past studies, we assume that the transportation cost of shared units is zero.

In DR-M and DR-R, retailer i's expected profit function is given by

$$\pi^d_{r,i} = \mathbb{E}\big[p \min(d_i, q_i) + tT_i + (p - t)T_j\big] - wq_i. \tag{7}$$

In this setting, Rudi et al. (2001) show that a unique Nash equilibrium exists.[1] The equilibrium stocking quantity $\left(q_i^d, q_j^d\right)$ satisfies Equation (8):

$$\alpha(q_i) - \beta_i(q_i, q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i, q_j)\left(\frac{p-t}{p}\right) = \frac{p-w}{p}, \qquad (8)$$

where $\beta_i(q_i, q_j) = \partial T_i/\partial q_i = \Pr(q_i + q_j - d_j < d_i < q_i)$ and $\gamma_i(q_i, q_j) = -\partial T_j/\partial q_i = \Pr(q_i < d_i < q_i + q_j - d_j)$. Intuitively, $\beta_i(q_i, q_j)$ is the probability of transferring from retailer i to j, and $\gamma_i(q_i, q_j)$ is the probability of transferring from retailer j to i.

The manufacturer's profit-maximizing wholesale price is derived from Equation (9):

$$w = q_i^d[p\alpha(q_i^d) - t\beta_i(q_i^d, q_j^d) + (p-t)\gamma_i(q_i^d, q_j^d)] + c. \qquad (9)$$

Given these quantities and wholesale prices, we next split the two cases, DR-M and DR-R, to discuss optimal transfer prices.

2.3.1. Transfer Price Set by the Manufacturer (DR-M).

When the manufacturer sets the transfer price, DR-M, Shao et al. (2011) show that the stocking quantity and wholesale price are monotonically increasing in the transfer price. Therefore, the manufacturer prefers a higher transfer price and sets $t = p$.

2.3.2. Transfer Price Set by the Retailers (DR-R).

In contrast with DR-M, when retailers set t, DR-R, they set a relatively low transfer price. Specifically, retailers set $t = 0$ for high-margin products across the two-tier supply chain (it may lie between zero and p for low-margin products).

2.4. Experimental Predictions and Hypotheses

Because of the strategic interaction between the manufacturer and retailers and that pricing and quantity decisions are often established by human managers in practice, we test this theory using a behavioral approach. To this end, we conduct a controlled between-subjects experiment with four treatments, each of which corresponds to one of the theoretical settings: Baseline, CR, DR-M, and DR-R. We provide details about our experimental methodology in the next section and here, outline specific parameters, experimental predictions, and hypotheses.

In all four treatments, we use a retail selling price $p = 30$ and a manufacturer unit production cost $c = 5$. Each retailer faces an integer demand drawn from a uniform distribution between 0 and 100. Although our selling price and unit production cost parameters appear to be for a relatively high-margin product, recall that this is across the entire supply chain (the retailer's normative critical fractile is actually less than 50% in all treatments) (see quantity predictions in Table 1). Importantly, unlike existing behavioral papers on inventory sharing, in all treatments the retailer's critical fractile may vary through w being endogenously set by the manufacturer.

Table 1. Normative Theoretical Predictions in the Experiment

	Baseline	CR	DR-M	DR-R
Transfer price t	—	—	30.00	0.00
Wholesale price w	17.50	21.67	21.67	18.33
Stocking quantity q	41.67	37.27	43.03	33.33
Manufacturer profit π_m	1,041.67	1,242.26	1,434.44	888.89
Retailer profit π_r	260.42	207.04	199.23	314.81
Supply chain efficiency, %	71.27	75.55	83.60	69.26

Notes. We assume that a continuous approximation of demand is sufficiently precise for predictions. Also, note that (1) transfer prices are predicted to be extreme values in both DR-M and DR-R, (2) wholesale prices are predicted to be equal to or above the potential anchor points of $(p + c)/2 = 17.5$ and $p/2 = 15$, and (3) quantities are predicted to be below 50.

Table 1 illustrates the experimental predictions for contract terms, profits, and supply chain efficiency (we provide detailed plots in Electronic Companion EC.1). First, beginning with the contract terms in Table 1, the predicted transfer price is always extreme, 30 in DR-M and 0 in DR-R. Intuitively, in DR-M, a manufacturer will set $t = p$ to incentivize retailers to order a higher quantity, which earns them a higher profit (i.e., a higher transfer price allows a retailer to earn a higher price on units sent but also requires them to pay more for units received). Additionally, in DR-R, retailers will set $t = 0$, leading to lower stocking quantities (i.e., units sent are less valuable, and units received are more affordable). Second, regarding wholesale prices, the predicted values are always equal to or higher than two potential anchor points, $(p + c)/2 = 17.5$ and $p/2 = 15$, which is useful when comparing across treatments. Third, stocking quantities are all predicted to be below 50. Our first experimental hypothesis revolves around these contract-term point predictions.

Hypothesis 1 (Contract Terms). *Transfer prices, wholesale prices, and stocking quantities will be set such that they coincide with the normative theoretical predictions.*

Turning to the manufacturer and retailer profits in Table 1, for manufacturers, it is unsurprising to see that they prefer to have decision authority over the transfer price, leading to the highest profits in DR-M. Overall, the manufacturer's preferred order among the four treatments is predicted as DR-M > CR > Baseline > DR-R. For retailers, they too prefer to have decision authority over the transfer price, yielding the highest profits in DR-R. Interestingly, retailers earn the second-highest profit in Baseline. At first, this may seem counterintuitive, as inventory sharing should be beneficial for retailers. Although this is true with exogenous wholesale prices, it does not necessarily hold in a two-tier supply chain with endogenous wholesale prices. In short, because retailers benefit from inventory sharing, manufacturers are able to charge a higher wholesale price in equilibrium (e.g., 21.67 in CR and DR-M versus 17.50 in Baseline). Across the four treatments, the retailers' preferred order is predicted as DR-R > Baseline > CR > DR-M, and we have Hypothesis 2.

Hypothesis 2 (Profits). *Manufacturer profit in the four treatments, from highest to lowest, will be DR-M > CR > Baseline > DR-R. Retailer profit in the four treatments, from highest to lowest, will be DR-R > Baseline > CR > DR-M.*

Continuing with profits in Table 1, it is noteworthy that the manufacturer always earns significantly more than the retailer. Among those four treatments, DR-R is the most equitable.

Hypothesis 3 (Equity). *The manufacturer will always earn a higher profit than the retailer, but DR-R will yield the most equitable distribution of profits between the manufacturer and retailer.*

Last, we develop a hypothesis for supply chain efficiency, which is calculated as the sum of the manufacturer's and retailers' expected profits divided by the first-best fully integrated supply chain benchmark. In the last row of Table 1, one can observe that DR-M is predicted to achieve the highest efficiency, followed by CR, Baseline, and DR-R:

Hypothesis 4 (Efficiency). *Supply chain efficiency in the four treatments, from highest to lowest, will be DR-M > CR > Baseline > DR-R.*

2.4.1. Behavioral Discussion. Although our experimental hypotheses rely on the normative theory, we would be remiss if we did not highlight that certain behavioral biases may impact decisions (for a summary of biases in operations management in individual decisions, other-regarding behavior, and strategic interactions, see Bolton and Chen 2019, Davis 2019, and Leider 2019). For instance, participants may be susceptible to bounded rationality (e.g., Su 2008) and set transfer prices in a way that does not perfectly coincide with the extreme normative predictions of 30 in DR-M and 0 in DR-R. As another example, in settings where a proposer can make a one-shot offer to a responder and the proposer is predicted to earn a disproportionately high split of overall profits, experimental studies have found evidence of fairness (Roth 1995). In our experiment, this suggests that the manufacturer may, at least qualitatively, set wholesale prices below the normative predictions. Last, existing newsvendor experiments indicate that a pull-to-center bias may push stocking quantities higher than the normative predictions in our setting (Schweitzer and Cachon 2000), but at the same time, certain supply chain experiments also find evidence of an understocking bias (e.g., Davis et al. 2014), so a directional deviation for quantities is unclear.

Although it is important to recognize how certain behavioral factors may influence decisions qualitatively, the resulting profit and efficiency implications are difficult to predict, as they rely on the magnitude of any observed deviations. For example, consider DR-R. Suppose that the observed transfer price is set above its normative prediction and that the observed wholesale price is set below its normative prediction. The deviation in the transfer price increases manufacturer profit, but the deviation in the wholesale price decreases manufacturer profit (and the reverse is true for retailer profit). As a consequence, depending on the magnitude of these two competing effects, it is unclear as to what the observed profits will be, relative to the normative prediction (and relative to the other treatments). Fortunately, by utilizing a controlled experiment, we cannot only test the normative theory and our hypotheses, but we can also identify how any potential deviations impact contract terms, profits, and efficiency.

3. Experimental Methodology

All four experimental treatments, Baseline, CR, DR-M, and DR-R, follow the normative theory and decision sequence outlined in Figure 1 in Section 2. In particular, in the Baseline treatment, each round begins with the manufacturer setting a wholesale price (there is no transfer price). After the wholesale price decision, each retailer then independently sets its own stocking quantity. Demand for each retailer is then realized, and profits are earned.

The CR treatment differs from the Baseline condition in that, after the wholesale price is set, the two retailers set a joint stocking quantity. Specifically, for up to two minutes, either retailer can send a quantity offer to the other retailer. The receiver can either accept or reject the offer. No other communication is allowed. If a quantity is agreed upon, then it becomes the joint quantity for that round. If there is no agreement after two minutes, then there is an additional 10 seconds for each retailer to consider the last offer proposed by the other retailer. If retailers still fail to reach an agreement after the extra 10 seconds, then all three players earn an outside option profit of zero. We opted for this process in an attempt to mimic a scenario where input from both retailers is incorporated (rather than one party unilaterally setting a quantity, which the other may not agree to).[2] After the joint stocking quantity is set, demand is realized, and profits are earned.

In DR-M, each round begins with the manufacturer deciding the transfer price and wholesale price. After this, each retailer determines its own stocking quantity. Demand is then realized, and inventory sharing automatically occurs, if applicable. The DR-R treatment is similar except that each round begins with the two retailers jointly setting the transfer price through a two-minute process. This process is identical to the quantity negotiation in CR. If retailers fail to reach an agreement after time expires (including the extra 10 seconds), the round continues without any inventory sharing after demand is realized. After the transfer price decision, the manufacturer sets the wholesale price, and then, the retailers set their stocking quantities. Finally, demand is realized, and inventory sharing automatically occurs, if applicable.

We provide a decision support tool for both roles to minimize complexity and to create a more realistic environment. For instance, there is empirical evidence that managers often use computers as an input for final decisions in supply chain settings. Zhao et al. (2021) conducted a survey of 54 firms and found that 35 (65%) relied on automated algorithms, which are adjusted by human managers, and 10 (19%) relied on averages between human decisions and system orders. With this in mind, our decision support tool consists of slide bars and a dynamic figure of expected profits. Specifically, participants can test their decisions by sliding the bar(s), and the expected profits of all three parties will be depicted on the figure. For the transfer price and wholesale price decisions, expected profits are calculated assuming that subsequent decisions are made optimally (which participants are aware of). For stocking decisions in Baseline and CR, there is one bar for the quantity. In DR-M and DR-R, each retailer has two slide bars: one for their own and one for the other retailer's stocking quantity (by checking a box, they can use the tool so that the other retailer's quantity is the best response, or they can manually set the other retailer's quantity). In all treatments, the test quantity scroll bar is initially set at the optimal quantity. See Electronic Companion EC.5.3 for sample instructions and screenshots.

One might note that our decision support tool for the retailer stocking quantity decisions is relatively strong. To provide justification for this, past studies have investigated how stocking quantities are set under various inventory-sharing strategies, whereas our paper differs in studying how contract terms are set in a two-tier setting. By providing decision support, we give the normative theory a fair chance of being confirmed. Also, by simplifying the stocking quantity decision, we mitigate concerns about the stocking quantity being set slightly differently (i.e., jointly) in CR. Last, if we were to automate the quantity decision, then retailers would only make decisions in DR-R (for the transfer price) but not in Baseline, CR, and DR-M. This would lead to unfair comparisons across treatments and may also overlook any other-regarding preferences. In sum, we opted for human retailers to set stocking quantities with strong decision support.

Turning to sample sizes, our Baseline, CR, DR-M, and DR-R treatments consist of 30, 57, 60, and 60 participants.[3] We included larger samples sizes for the three inventory-sharing treatments because they have not been investigated before in the laboratory, whereas the Baseline treatment is closely related to existing research between a single manufacturer and a single retailer. Also, following existing experimental supply chain and economics research, our study includes university student participants (Kagel and Roth 2017, Donohue et al. 2019) who were recruited from a large university where cash was the only incentive offered. Several studies have shown that students make similar decisions as managers in operational settings, such as inventory management and forecasting (e.g., Katok et al. 2008, Bolton et al. 2012, Kremer et al. 2016). Nevertheless, we have not seen a supply chain contracting experiment that compares contracting decisions between students and managers, so we recognize this as a limitation.

Our experiment was implemented through oTree (Chen et al. 2016). Each session consisted of 12 rounds. In each round, participants were randomly assigned a role and matched with two other participants. This means that both roles (and trios) were randomly determined each round, similar to Ozer et al. (2011) and Davis and Leider (2018). Before a session started, a researcher read through the instructions out loud and answered any questions. Participants were then required to answer several multiple-choice comprehension questions about the game. Participants received cash based on profits from all rounds in the game plus a $7 show-up fee. Average earnings were roughly $25 across all treatments. Each session lasted for 70 minutes on average.

4. Results

In this section, we present our experimental results. Following our hypotheses, we take a bottom-up approach and begin with contract terms in Section 4.1. We then proceed to investigate profits and efficiency in Section 4.2. In this subsection, we also discuss how any deviations in contract terms, relative to theory, account for observed profits and efficiency. For our analysis, the rate of agreements/acceptances by retailers was high and similar across treatments. Specifically, the fractions of time that retailers came to an agreement over the stocking quantity in CR and the transfer price in DR-R were 95.18% and 97.08%, respectively. These near-100% agreement rates are not particularly surprising as they represent a joint decision rather than a zero-sum negotiation.[4] In addition, the fractions of time retailers accepted the manufacturer's wholesale price and set a positive stocking quantity were also high and similar across all four treatments: 98.33% in Baseline, 97.70% in CR, 97.50% in DR-M, and 98.75% in DR-R. Thus, unless otherwise stated, we include all data in our analysis.

Given the panel structure of our data, unless otherwise noted, we use regression analysis with random effects for all hypothesis tests (Hyndman and Embrey 2019).[5] To provide a specific example, consider our fourth hypothesis on efficiency. Because one manufacturer and two retailers are randomly matched together as a trio and earn the same efficiency (in a round), we only include those observations for the role of the manufacturer for this test. We follow a similar approach any time there is a risk of double or triple counting observations (e.g., efficiency, retailer profit and the joint quantity in CR, the transfer price in DR-R, etc.). Also, because each experimental hypothesis consists of a family of multiple comparisons (List et al. 2019), we adjust the critical p-values using Bonferroni corrections, assuming an unadjusted critical p-value of 0.05.[6]

4.1. Contract Terms

Average observed transfer prices, wholesale prices, and stocking quantities for all four treatments are summarized in the left-hand side of Table 2. Beginning with transfer prices for the decentralized retailer inventory-sharing strategies, DR-M and DR-R, one can see that observed transfer prices deviate from their extreme predictions, which is inconsistent with Hypothesis 1. In DR-M, manufacturers set the transfer price lower than the normative prediction, 20.26 versus 30 ($p < 0.005$, the corrected critical p-value), whereas in DR-R, retailers set the transfer price higher than the normative prediction, 6.27 versus 0 ($p < 0.005$). These two observations suggest that any bias influencing transfer prices may be present for both manufacturers in DR-M and retailers in DR-R.

To investigate wholesale price and stocking quantity decisions, the right-hand side of Table 2 presents the normative predictions conditioned on any previous decisions. For instance, predicted wholesale prices in DR-M and DR-R are conditioned on transfer prices, and all stocking quantities are conditioned on wholesale prices (and transfer prices, if applicable). Although all tests are between the observed data and these conditional predictions, we also report the unconditional normative predictions in square brackets.

Wholesale prices, in all three inventory-sharing treatments (CR, DR-M, DR-R), are set significantly lower than the conditional predictions and contradict Hypothesis 1: 19.15 versus 21.67 in CR, 18.40 versus 20.41 in DR-M, and 16.78 versus 18.89 in DR-R (all $p < 0.005$). Wholesale prices are also set low in the Baseline condition, 16.94 versus 17.50, but the difference is not significant. As for stocking quantities, there is only a significant difference between observed decisions and conditional predictions in CR, 38.37 versus 42.22 ($p < 0.005$). Across the other treatments, if anything, there may be a slight understocking bias relative to the conditional predictions (41.69 versus 43.66 in Baseline, 42.86 versus 44.81 in DR-M, and 39.80 versus 40.58 in DR-R), which is somewhat surprising given the level of decision support that we provided to retailers. Combined with the fact that wholesale prices are set too low in all inventory-sharing treatments, this indicates that a behavioral bias may be

Table 2. Average Contract Prices, Quantities, and Normative Theoretical Predictions

	Observed results				Normative rredictions			
	Baseline	CR	DR-M	DR-R	Baseline	CR	DR-M	DR-R
Transfer price	—	—	20.26* (0.71)	6.27* (0.48)	—	—	30.00	0.00
Wholesale price	16.94 (0.39)	19.15* (0.27)	18.40* (0.30)	16.78* (0.32)	17.50	21.67	20.41 (0.09) [21.67]	18.89 (0.04) [18.33]
Stocking quantity	41.69 (0.94)	38.37* (0.67)	42.86 (0.99)	39.80 (0.91)	43.66 (0.53) [41.67]	42.22 (0.30) [37.27]	44.81 (0.35) [43.03]	40.58 (0.43) [33.33]

Notes. Standard errors, across participants, are reported in parentheses. Results for DR-R and CR are conditioning on agreement. Stocking quantity in CR is one-half of the average joint stocking quantity. Predicted wholesale prices in DR-M and DR-R are conditioning on observed transfer prices. Predicted stocking quantities are conditioning on observed wholesale prices (and transfer prices). Unconditional normative predictions, when applicable, are reported in square brackets. Transfer prices deviate from the normative predictions, and wholesale prices are often set too low. Stocking quantities are set close to predictions or low.

*Significance of regressions comparing observed vs. conditional normative predictions given by $p < 0.005$ (the corrected critical p-value).

affecting wholesale price and to a lesser extent, quantity decisions.

We also conducted a heterogeneity analysis by classifying those participants who did or did not make a particular decision optimally. This is instructive in determining whether any of our aggregate results are driven by a small group of individuals. For individual participants, we conducted a test between their decisions and the conditional predictions. Because of the limited number of observations, we opt for Wilcoxon signed-rank tests and count a significant deviation if there is a difference at the 10% level. Beginning in Figure 2(a), 81.67% in DR-M and 100% in DR-R of participants set transfer prices that significantly deviated from the normative predictions. In Figure 2(b), between 40% and 59.65% of participants set wholesale prices that were significantly too low. These first two figures indicate that the average transfer price and wholesale price deviations do not appear to be driven by a small subset of participants. Last, in Figure 2(c), a majority of participants set quantities in a way that is not significantly different from the conditional prediction, but a reasonable percentage understocks in Baseline, CR, and DR-M. This coincides with the aggregate results. Thus, we have our first result, which largely rejects Hypothesis 1.

Result 1 (Contract Terms). Transfer prices in the two decentralized inventory-sharing strategies, DR-M and DR-R, deviate symmetrically from the normative predictions. Wholesale prices are set too low in all three of the inventory-sharing environments, CR, DR-M, and DR-R. Stocking quantities are set reasonably well or slightly low in all settings.

As a final comment regarding contract terms, we also investigated dynamics: for instance, whether there were any experience effects across rounds, whether the magnitude of stocking quantity deviations was correlated with wholesale prices, whether decisions differed after playing a particular role, and more. These analyses yielded three insights. First, there were some moderate learning effects in early rounds, but if we exclude these round, all of our main results (including ones highlighted later) continue to hold. Second, retailer stocking quantity deviations do not increase with higher wholesale prices. Third, decisions do not appear to significantly differ after a participant plays a particular role. This last statement, from a methodological standpoint, indicates that having participants change roles during the experiment did not influence decisions and from a practical standpoint, suggests that managers need not worry about behavior varying (for better or worse) if they are in the unique position to rotate personnel across different roles and responsibilities.

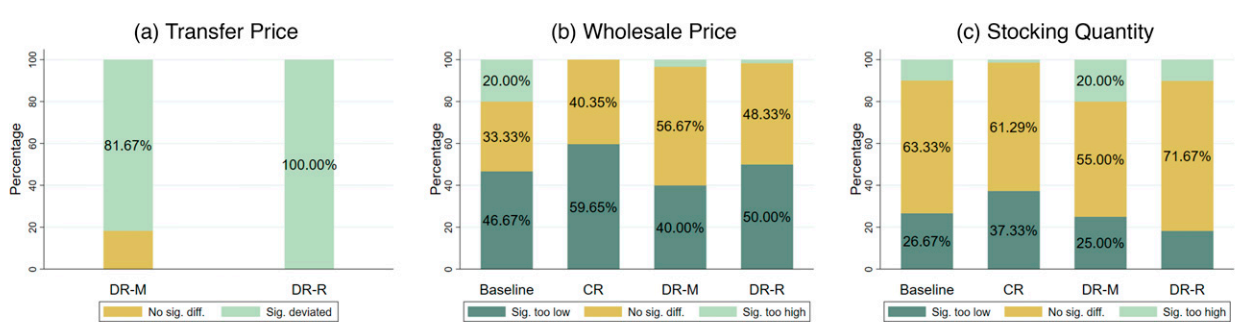
4.2. Profits and Efficiency

Figure 3 depicts the manufacturer expected profit (bottom darker portion), each of the retailer's expected profits (top light portion), and supply chain expected profit (overall height) in all four treatments. Although we will compare the observed results with the normative theoretical predictions momentarily, we also include the normative predictions by dashed horizontal and vertical lines.

Beginning with the manufacturer and retailer profits in Figure 3, a key result is that the DR-M treatment provides a win-win outcome compared with both the Baseline and CR conditions: retailers earn significantly higher profits, 332.09 versus 281.36 and 285.10 (both $p < 0.004$, the corrected critical p-value), and manufacturers earn significantly higher profits as well, 1,097.60 versus 948.49 and 998.26 (both $p < 0.004$). This leads to our second result.

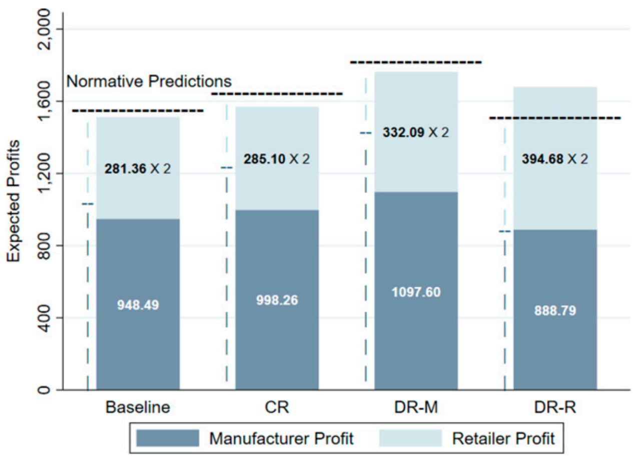
Result 2 (Profits). A decentralized retailer inventory-sharing strategy where the manufacturer sets the transfer price, DR-M, achieves a win-win outcome compared with the Baseline setting and centralized

Figure 2. (Color online) Percentages of Participants Deviating in Pricing and Stocking Decisions



Notes. (a) Transfer price. (b) Wholesale price. (c) Stocking quantity. Classification of a subject is based on Wilcoxon signed-rank tests between observed decisions and (conditional) optimal decisions (at the 10% level). Percentages below 20% are omitted because of limited space.

Figure 3. (Color online) Average Observed Expected Profits



Notes. Dashed lines represent normative predictions (horizontal for supply chain and vertical for distribution). DR-M yields a Pareto improvement over Baseline and CR. DR-R provides the most equitable distribution of profits. Total supply chain profit is higher in both DR-M and DR-R compared with Baseline (and DR-M is higher than CR).

retailer inventory-sharing strategy, CR, in that both the manufacturer and retailers earn significantly higher profits.

To provide further context around this result and Hypothesis 2 (profits), theory predicts manufacturers to earn the highest profit in DR-M (when they have authority over the transfer price). In this sense, the second hypothesis is validated for manufacturers. However, a different picture emerges for retailers. Whereas Hypothesis 2 predicts that retailers should prefer an order of DR-R > Baseline > CR > DR-M, the data reject this order, and we instead observe DR-R > DR-M > CR ≈ Baseline. The primary difference is the favorable performance of DR-M over both Baseline and CR for retailers. A managerial implication of this is that, if retailers are contracting with a powerful manufacturer, enabling inventory sharing while ceding decision authority of the transfer price to the manufacturer will still lead to a relatively high profit. We will discuss this more in Section 8.

Proceeding with Hypothesis 3 (equity), another result we can glean from Figure 3 is that manufacturers do indeed earn more than retailers, and DR-R generates the most equitable distribution of profits between the manufacturer and retailers, among the four treatments. This is true in terms of both the percentage profit split and the absolute difference in profits. For instance, the percentage of total supply chain profits that are earned by each retailer in DR-R is 23.5% in DR-R ($394.68/(2 \times 394.68 + 888.79)$), whereas in the other treatments, this percentage is between 18.2% and 18.8%. A proportions test indicates that the retailer's share in DR-R is indeed significantly higher than the other treatments (all three $p < 0.008$, the corrected critical p-value). Similarly, the absolute difference in manufacturer profit and average retailer profit is only 494.11 in DR-R, yet the differences are between 667.13 and 765.51 in the other three treatments (all three $p < 0.008$). Therefore Hypothesis 3 is supported, and we have Result 3.

Result 3 (Equity). Manufacturers always earn more than retailers, but a decentralized retailer inventory-sharing strategy where the retailers set the transfer price, DR-R, achieves the most equitable outcome in terms of distribution of expected profits between the manufacturer and retailers.

Turning to how profits compare with the normative predictions, the left-hand side of Table 3 depicts the observed average profits along with hypothesis tests versus the normative predictions (the latter of which are illustrated in the right-hand side of the table). Regarding the distribution of manufacturer and retailer profits compared with theory, we see a consistent pattern across all treatments: manufacturers earn significantly less (or the same) than the normative predictions, and retailers earn significantly more than the normative predictions. In short, outcomes are more equitable than theory predicts (even in DR-R, which should already lead to the most equitable payoffs). This observation will be relevant when we explore behavioral biases in Section 6.

Comparing efficiency across treatments (Hypothesis 4), in Table 3, DR-M and DR-R achieve the highest efficiency (DR-R should, in theory, have the lowest efficiency). Although the difference between DR-M and DR-R is not significant, DR-M is significantly higher than the Baseline and CR conditions (both $p < 0.008$, the corrected critical p-value), and DR-R is significantly higher than Baseline ($p < 0.008$). Interestingly, there is no statistical difference between Baseline and CR, where moving from the Baseline to CR should, in theory, increase efficiency. Another notable observation from this table is that the supply chain profit in DR-R is higher than the normative prediction, which we will explore in detail later. Until then, Hypothesis 4 is rejected, and we have Result 4.

Result 4 (Efficiency). Both decentralized retailer inventory-sharing strategies, DR-M and DR-R, achieve a higher efficiency compared with the Baseline setting (and DR-M is higher than the centralized retailer-sharing strategy, CR). Further, there is no significant increase in efficiency when moving from the Baseline setting to a CR strategy.

Overall, a combination of the experimental results around profits, equity, and efficiency provides evidence that both the Baseline setting, where two retailers neglect to share inventory, and the centralized retailer

Table 3. Average Observed Profits, Efficiency, and Normative Theoretical Predictions

	Observed results				Normative predictions			
	Baseline	CR	DR-M	DR-R	Baseline	CR	DR-M	DR-R
Manufacturer profit	948.49* (21.96)	998.26* (25.29)	1,097.60* (22.35)	888.79 (20.68)	1,041.67	1,242.26	1,434.44	888.89
Retailer profit	281.36* (7.72)	285.10* (7.17)	332.09* (7.40)	394.68* (7.74)	260.42	207.04	199.23	314.81
Supply chain efficiency, %	68.93 (0.019)	71.54 (0.018)	80.36* (0.011)	76.54* (0.013)	71.27	75.55	83.60	69.26

Notes. Standard errors, across participants, are reported in parentheses. There are significant differences between all profits and efficiencies compared with the normative benchmarks except manufacturer profit in DR-R and efficiency in Baseline and CR.

*Significance of regressions comparing observed vs. normative predictions given by $p < 0.017$ (the corrected critical p-value).

inventory-sharing strategy, CR, fail to perform best across (a) retailer profit, (b) manufacturer profit, (c) equity of profits, and (d) supply chain efficiency. Thus, if a firm is using any of these metrics as their primary criteria for how to share inventory, they should consider a decentralized retailer inventory-sharing strategy, where the retailers or manufacturer has decision authority over the transfer price depending on the situation.

4.3. Connecting Contract Terms with Profits and Efficiency.

Here, we connect our results by summarizing how the observed contract-term deviations highlighted in Result 1 can largely account for Results 2–4 (profit, equity, and efficiency). Beginning with Result 2, which finds that DR-M generates a Pareto improvement over the Baseline and CR conditions because of observed wholesale prices being lower than predicted in all treatments and transfer prices being set too low in DR-M (which benefits the retailer), the manufacturer earns less and the retailers earn more than the normative theory predicts. This combined with the fact that DR-M is predicted to generate the highest manufacturer profit and the highest total supply chain profit allows manufacturers to effectively redistribute some of their profits to retailers and make both parties better off compared with Baseline and CR (but not DR-R, as it is predicted to generate the highest retailer profit).

For Result 3, DR-R generates the most equitable distribution of profits (more equitable than theory predicts) for two reasons. First, DR-R is predicted to provide the most equitable outcome between parties. Second, manufacturers still offer more generous wholesale prices than theory predicts. A combination of the theoretical prediction and this experimental deviation results in the most equitable split of profits, even more so than theory predicts. Regarding Result 4 (that DR-M and DR-R achieved the highest supply chain efficiency), theory predicts DR-M to earn the highest efficiency. Because quantities are set similarly in all treatments (i.e., close to or slightly below the conditional predictions) and both the unconditional and conditional quantity predictions are the highest in DR-M, its observed efficiency is highest. Regarding DR-R, it even outperforms its theoretical prediction (observed efficiency of 76.54% versus prediction of 69.26%) because manufacturers offer lower wholesale prices than optimal and retailers set transfer prices higher than optimal. Both of these effects drive quantities higher (observed quantities of 39.80 versus unconditional prediction of 33.33), and hence, there is a higher efficiency than theory predicts.

We provide a summary of these effects in Table 4, which illustrates the percentage impact of a particular price and quantity deviation on manufacturer and retailer profits, relative to theory. Each effect is calculated comparing the observed profit with the conditional optimal profit for that decision divided by the normative prediction. For instance, the wholesale price impact in DR-M, for each respective party, is $(\pi(w, q(w,t) \mid t) - \pi(w^*, q(w^*,t) \mid t)/\pi(t^*, w^*, q(w^*, t^*))$. Beginning at the bottom of this table, we observe that, unsurprisingly, slightly lower than optimal quantities only have a relatively small impact on both parties' profits (third row). Turning to wholesale prices (second row), we observe that manufacturers decrease their own earnings but greatly increase retailer profits. For instance, manufacturers give up between 3.97% and 10.93% in profits through lower wholesale prices, but this increases retailer profits by 13.23%–52.38%. Therefore, manufacturers set wholesale prices in a way that is more generous than theory predicts.

In reviewing the transfer price deviation effects (first row) in Table 4, the party with decision authority sets the transfer price in a way that hurts themselves and helps the other party in a significant way. For example, in DR-M, manufacturers give up 13.46% of their profits by setting suboptimal transfer prices, but this translates into an increase of 26.90% in retailer profits. Ultimately, this helps contribute to the redistribution of wealth from the manufacturer to the retailer and the "win-win" outcome over Baseline and CR. Last, in DR-R, the transfer price deviation by retailers only slightly decreases their own profits, by 3.49%, but the impact of the wholesale price deviation by the manufacturer is much larger, increasing retailer profits by 34.14%, such that the net result is that

Table 4. Manufacturer and Retailer Profit Implications from Price and Quantity Deviations

	Baseline		CR		DR-M		DR-R	
	Retailer	Manufacturer	Retailer	Manufacturer	Retailer	Manufacturer	Retailer	Manufacturer
Transfer price, %	—	—	—	—	26.90	−13.46	−3.49	11.82
Wholesale price, %	13.23	−3.97	52.38	−6.28	47.75	−6.44	34.14	−10.93
Stocking quantity, %	−5.19	−4.97	−7.87	−9.10	−7.96	−3.59	−5.25	−0.93
Total, %	8.04	−8.95	44.50	−15.37	66.69	−23.48	25.39	−0.04

Notes. The percentage impact of a particular price and quantity deviation on manufacturer and retailer profits, relative to theory. Each percentage effect is calculated comparing the observed profit with the conditional optimal profit for that decision divided by the normative prediction. For instance, the wholesale price impact in DR-M, for each respective party, is $(\pi(w,q(w,t)\,|\,t)-\pi(w^*,q(w^*,t)\,|\,t)/\pi(t^*,w^*,q(w^*,t^*))$. Results for DR-R and CR are conditioning on agreement.

retailers earn more than theory predicts and DR-R becomes quite equitable.

5. Alternative DR Treatment

One may note that the DR-M and DR-R strategies differ from Baseline and CR not only in how inventory is shared but also, by including a term, the transfer price, that is endogenously set. Therefore, before turning to any behavioral biases that may be driving decisions, here we briefly investigate a decentralized retailer inventory-sharing strategy where the transfer price is exogenously set to zero (DR-0). In addition to providing a sharper comparison among certain treatments, investigating this DR-0 inventory-sharing strategy is useful for two reasons. First, in practice, this represents a chain retailer that dictates that all retailers act independently but are required to share product with one another at a transfer price of zero. Second, retailers are disadvantaged in all of the scenarios we have considered thus far, but among them, DR-R performs best in terms of retailer profits. By exogenously setting the transfer price to zero, this should reduce the profit of manufacturers and help retailers earn a higher profit, relative to the outcomes of DR-R (note that the normative predictions in this new variant are the same as DR-R).

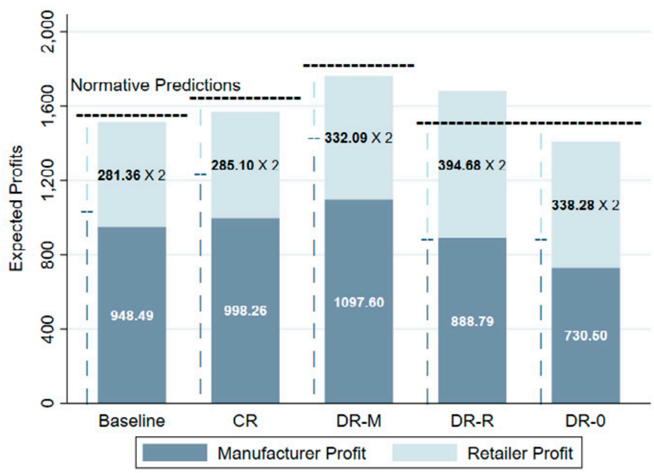
We ran the DR-0 treatment with 57 additional participants. In the left side of Figure 4, we illustrate the profit results for our original four treatments plus the DR-0 treatment (far-right column of the figure). Comparing DR-0 with Baseline and CR, in both cases we observe that retailers do earn a higher profit in DR-0 (and manufacturers earn less), which supports theory ($p < 0.005$, the corrected critical p-value). Also, comparing DR-0 with DR-R, we see that manufacturers do indeed earn a lower profit in DR-0, which supports theory (730.50 versus 888.89, $p < 0.005$). However, despite the theoretical advantage of DR-0 over DR-R for retailers, DR-0 actually achieves a retailer profit that is significantly less than DR-R experimentally (338.28 versus 394.68, $p < 0.005$). As a consequence, DR-R Pareto dominates DR-0, counter to theory.

On the right side of Figure 4, we report average observed wholesale prices and quantities for DR-0, along with the normative predictions. Consistent with the original three inventory-sharing treatments, we find that wholesale prices are again set too low relative to the normative benchmark ($p < 0.005$). However, unlike DR-R, where quantities were set close to optimal, in DR-0 retailers significantly understock relative to the conditional prediction: 31.59 versus 35.63 ($p < 0.005$). This directly accounts for the poor profit performance of DR-0 for both parties: manufacturers offer a more generous wholesale price, thus reducing their own profits, but retailers neglect to capitalize on this and instead, significantly understock, hurting both parties' (and the supply chain's) profits. We, therefore, supplement our previous findings with the following fifth result.

Result 5 (DR-0). A decentralized retailer inventory-sharing strategy with an exogenous transfer price of zero contributes to significant understocking in quantities. As a result, DR-R achieves a win-win outcome over DR-0: both parties earn a higher profit when the transfer price is endogenously set by retailers compared with when it is exogenously set to zero.

There are at least two plausible (nonexclusive) factors that may drive the observed understocking bias in DR-0. First, retailers may understock when the transfer price is zero, regardless of how the transfer price is set. Second, retailers may understock when transfer prices are set exogenously. To investigate this, we compare observations in DR-R where the retailers set a transfer price equal to zero, so that the only difference between DR-R and DR-0 is whether the transfer price is set endogenously or exogenously. However, we observe that retailers set a transfer price equal to zero only 2.6% of the time in DR-R, making it difficult to draw conclusions. If we expand this analysis to include any transfer price less than 0.5 (1), then the number of observations increases to 12.0% (25.8%). In this case, retailers in DR-R understock by an average of 0.68 units (1.96 for $t < 1$). This degree of understocking is far less than what we observe in

Figure 4. (Color online) Average Observed Expected Profits in the Main Experiment Plus DR-0 and Price and Quantity Decisions in DR-0



	DR-0	
	Observed	Predictions
Wholesale Price	17.35* (0.33)	18.33
Stocking Quantity	31.59* (0.96)	35.63 (0.56) [33.33]

Notes. For price and quantity decisions, standard errors, across participants, are reported in parentheses. Predicted stocking quantities are conditioned on observed wholesale prices (unconditional normative predictions, when applicable, are reported in square brackets).
*Significance of regressions vs. conditional normative predictions given by $p < 0.005$ (the corrected critical p-value).

DR-0, which is 4.04. Overall, this analysis suggests that although retailers appear to dislike transfer prices of zero (evidenced by the low frequency of observations with $t = 0$ in DR-R), the excessive understocking in DR-0 may be primarily driven by the fact that such low transfer prices are exogenously set in DR-0. We will explore this further when we fit our behavioral model in the next section.

6. Behavioral Model

Thus far, we have found a number of profit and supply chain efficiency differences across alternative inventory-sharing strategies. We have also observed how these differences can be attributed to deviations in price and quantity decisions relative to the normative theory. In this section, we investigate a plausible behavioral bias that can account for such contract-term deviations.

In determining which bias to investigate, our previous experimental analysis is instructive. A key finding was that the observed distribution of profits is more equitable than predicted. In particular, we found that manufacturers, which are predicted to earn a significantly larger share of the total profits, offered wholesale prices that are more generous than theory predicts, leading to a more equitable distribution of profits than theory predicts. This indicates that fairness may be influencing decisions (Fehr and Schmidt 1999, Bolton and Ockenfels 2000). For brevity, we provide an overview of this behavioral bias and relegate detailed theoretical analyses, estimation procedures, and more results to Electronic Companion EC.3.

Fairness has been examined in a number of supply chain studies (e.g., Cui et al. 2007, Kalkanci et al. 2014, Beer et al. 2021). We follow a similar approach and assume that a retailer (the responder) suffers disutility when their expected profit is less than the manufacturer's profit (we do not consider disutility when the retailer earns more than the manufacturer, which rarely occurred in our data). In addition, although a majority of studies on fairness typically focus on the responding party, there is empirical evidence that proposers in favorable positions often share more of their earnings than theory predicts. This is seen in experimental dictator games, where a proposer who has authority over how much of a surplus to share with a responder routinely offers around 25% of the surplus as opposed to the equilibrium of zero (e.g., Forsythe et al. 1994, Andreoni et al. 2010). Therefore, we also consider that the manufacturer suffers disutility when their expected profit is greater than the retailer's. Recall that π_m and $\pi_{r,i}^d$ are the expected profits of manufacturer and retailer i, respectively, under the normative theory in the decentralized setting (the centralized case follows similar logic). Under fairness concerns, the expected utility functions for manufacturer's utility and the decentralized retailer i's utility are given by

$$u_m^F = \pi_m - \lambda_m(\pi_m - \pi_{r,i}^d)^+, \quad u_{r,i}^{d,F} = \pi_{r,i}^d - \lambda_r(\pi_m - \pi_{r,i}^d)^+, \tag{10}$$

where λ_m represents the manufacturer's degree of fairness over advantageous inequality and λ_r represents retailer i's degree of fairness concerns over disadvantageous inequality. This formulation is directly related to the well-established fairness model of Fehr and Schmidt (1999).[7]

Qualitatively, fairness can account for lower wholesale prices by manufacturers, relative to the normative theory. The intuition is straightforward in that fairness-minded manufacturers prefer to offer lower wholesale prices, which equalizes profits. We also find that fairness, according to the model, can account for retailers stocking less than the normative predictions (see Electronic Companion EC.3), as it leads to more equitable profits, but we expect this effect to be small in our estimations given that retailers set quantities well or only slightly low in certain treatments. Regarding transfer prices, we find that fairness cannot account for the deviation of transfer prices in DR-M or DR-R. Because transfer prices were set too low by manufacturers in DR-M and too high by retailers in DR-R, a symmetric bias may be influencing decisions. Further, the impact of suboptimal transfer prices on profits is relatively small compared with the deviations in wholesale prices (see Table 4). Combining these two observations, we model transfer price deviations with random errors so that they follow a multinomial logit distribution. Transfer prices yielding a higher expected utility are chosen with a higher probability and vice versa, where θ is the degree of rationality; $\theta \rightarrow 0$ is fully rational, and $\theta \rightarrow \infty$ is fully random decisions (Su 2008).

We fit the model to the data using maximum likelihood estimation. Table 5 provides the results for a model with only transfer price errors (transfer errors), each type of fairness separately ($Fair_m$ and $Fair_r$), and the "full" model of fairness with both types of fairness ($Fair_{m,r}$). For each model, we fit the data two ways. The first constrains the parameters to be the same across all five treatments (LL_c). The second fits a set of parameters separately for each treatment and then sums up the five log likelihoods (LL_s). Both approaches are depicted in Table 5. A series of likelihood ratio tests reveals significant differences across all applicable estimations, such that the full fairness model with both retailer and manufacturer concerns, $Fair_{m,r}$, provides the most favorable fit for both procedures. Aside from this main takeaway, in evaluating the two nested fairness models, $Fair_m$ and $Fair_r$, we see that the log likelihood is considerably better in $Fair_m$ (e.g., LL_c of −13,449.36 versus −13,566.87). This suggests that the inclusion of manufacturer fairness concerns is especially important to the overall fit.

In the upper part of Table 6, we provide the transfer price, wholesale price, and quantity predictions using the parameter estimates from the full fairness model, $Fair_{m,r}$, for each treatment. In an effort to evaluate these predictions, we also include the average observed transfer price, wholesale price, and quantities from the data. Beginning with transfer prices in DR-M, the fairness model predicts a price of 20.22 versus 20.26 observed, and in DR-R, the prediction is 6.32 versus 6.27 observed. Wholesale prices are also quite accurate; going from left to right in Table 6 (Baseline, CR, DR-M, DR-R, and DR-0), the fairness model predictions versus observed are 16.94 versus 16.94, 19.13 versus 19.15, 18.35 versus 18.40, 16.82 versus 16.78, and 17.35 versus 17.35, respectively. Last, the predicted quantities are within one unit of the observed quantities in all five treatments.

The lower part of Table 6 shows the parameter estimates by treatment. To begin, the random errors in transfer prices ($\hat{\theta}$) are lower in DR-R than DR-M, which agrees with the analysis in Table 4 showing that retailers give up little profit by setting marginally suboptimal transfer prices. Manufacturer fairness concerns ($\hat{\lambda}_m$) are generally larger than retailer fairness concerns ($\hat{\lambda}_r$) in all treatments, but one must recognize that manufacturers always earn a disproportionately larger split of the overall supply chain profit (i.e., average manufacturer utility is still higher than average retailer utility).

A closer look at the fairness estimates for a given role, across treatments, yields additional insights. Starting with the manufacturer, the estimates are highest in CR and DR-M. This is consistent with the notion that the predicted distribution of profits is largest in these two treatments. Continuing with DR-R and DR-0, the estimates are lower than DR-M and CR, which is natural considering that they should lead to the most equitable distribution of profits. However,

Table 5. Aggregate Behavioral Model Estimation Results

Description	Parameter	Transfer errors	$Fair_m$	$Fair_r$	$Fair_{m,r}$
Transfer price errors	$\hat{\theta}$	193.383	92.563	165.504	85.814
Manufacturer fairness	$\hat{\lambda}_m$	—	0.334		0.333
Retailer fairness	$\hat{\lambda}_r$	—	—	0.086	0.089
Constrained estimation	LL_c	−13,620.24	−13,449.36	−13,566.87	−13,390.01
Sum of separate estimations	LL_s	−13,469.88	−13,313.81	−13,404.42	−13,248.71

Notes. LL represents log likelihoods. The number of observations in each LL is 3,370. $\hat{\theta}$ is the estimated parameter for transfer price errors. $\hat{\lambda}_m$ and $\hat{\lambda}_r$ are estimated fairness parameters for the manufacturer and retailers, respectively. Full details provided in Electronic Companion EC.3.

Table 6. Fairness Predictions, Observed Values, and Parameter Estimates by Treatment

	Baseline		CR		DR-M		DR-R		DR-0	
	$Fair_{m,r}$	Obs.	$Fair_{m,r}$	Obs.	$Fair_{m,r}$	Obs.	$Fair_{m,r}$	Obs.	$Fair_{m,r}$	Obs.
Fairness predictions										
t	—	—	—	—	20.22	20.26	6.32	6.27	—	—
w	16.94	16.94	19.13	19.15	18.35	18.40	16.82	16.78	17.35	17.35
q	41.49	41.69	39.04	38.37	43.33	42.86	40.42	39.80	31.71	31.59
Parameter estimates										
$\hat{\theta}$	—		—		$\hat{\theta}$ = 118.184		$\hat{\theta}$ = 35.010		—	
$\hat{\lambda}_m$	$\hat{\lambda}_m$ = 0.147		$\hat{\lambda}_m$ = 0.392		$\hat{\lambda}_m$ = 0.382		$\hat{\lambda}_m$ = 0.332		$\hat{\lambda}_m$ = 0.186	
$\hat{\lambda}_r$	$\hat{\lambda}_r$ = 0.054		$\hat{\lambda}_r$ = 0.139		$\hat{\lambda}_r$ = 0.065		$\hat{\lambda}_r$ = 0.012		$\hat{\lambda}_r$ = 0.152	

Notes. Numbers of observations for Baseline, CR, DR-M, DR-R, and DR-0 are 360, 434, 960, 932, and 684 (3,370 total). DR-R and CR are conditioning on agreement. Obs. indicates observed average values. $\hat{\theta}$ is the estimated parameter for transfer price errors. $\hat{\lambda}_m$ and $\hat{\lambda}_r$ are estimated fairness parameters for the manufacturer and retailers, respectively. Full details are provided in Electronic Companion EC.3.

whereas theory predicts that DR-R and DR-0 should be identical, we observe that manufacturers have higher fairness concerns in DR-R. This is intuitive if one recalls that, in our DR-R data, retailers first set a noisy transfer price higher than the normative prediction of zero. As shown in our analysis in Table 4, this translates into a higher profit for the manufacturer and hence, a larger (conditional) predicted difference in profits. Therefore, manufacturer fairness concerns should indeed be higher in DR-R relative to DR-0. Last, the manufacturer's fairness estimate is lowest in Baseline. One possible explanation for this is that the normative theory, under inventory sharing, predicts that a manufacturer should set a higher wholesale price than in the Baseline setting (i.e., a manufacturer should recognize that retailers have a risk-pooling benefit under inventory sharing and set a higher price). However, our results suggest that a manufacturer does not follow this prescription under retailer inventory sharing and instead, offers a wholesale price that is significantly below what theory predicts, leading to fairness estimates that are higher in the inventory-sharing conditions than the Baseline setting.

A similar story emerges with regard to retailer fairness concerns across treatments. For instance, retailer fairness estimates are relatively higher in CR and DR-M, which is expected given that theory predicts a large difference in profits. Comparing the two directly, the higher estimate in CR is likely because of two retailers jointly setting a quantity: if one of them has fairness concerns and the other does not, this could lead to lower quantities and thus, a higher fairness estimate. It is also noteworthy that the retailer's estimate in Baseline, which is roughly average across the treatments, is nearly identical to that found in past supply chain experiments with a single retailer (e.g., Davis et al. 2014 estimate the value to be 0.050 versus 0.054 in our study).[8] Last, another interesting observation is that DR-0 has the highest retailer fairness estimate (and far higher than DR-R). Recall that, at the end of Section 5, we conducted an analysis indicating that retailers did not excessively understock in DR-R for transfer prices that were equal to or close to zero. This suggests that retailers react negatively to being required to share inventory at an exogenous transfer price of zero, resulting in them understocking. As a consequence, procedural fairness may be manifesting itself in the higher retailer fairness estimate for DR-0.[9]

7. Robustness Check

As a robustness check of our results, we also investigate an alternative type of contract with inventory sharing rather than a simple wholesale price contract. In particular, we run two additional treatments, which consider revenue sharing between the retailers and the manufacturer. Our objective is to determine whether the behavioral deviations observed in our main experiment hold for a different contract under inventory sharing, which has not been explored before. We provide a high-level summary of the results here and share more details in Electronic Companion EC.4.

Given their favorable performance, we consider the two decentralized inventory-sharing strategies with endogenous transfer prices, DR-M and DR-R, but add an exogenous revenue share for the manufacturer of 30%. Each treatment includes 42 participants and uses the same protocols as our main experiment. For results, all of the price and quantity deviations observed in our main experiment are found in this revenue-sharing contract: (a) transfer prices are not set at the extreme predictions, (b) wholesale prices are set significantly too low in both treatments (and notably, less than the potential anchors of 17.5 and 15), and (c) quantities are set close to theory but slightly low. Turning to the fairness model, we generate the following behavioral predictions relative to the new data (with normative predictions in parentheses): (a) transfer price predictions are 19.85 versus 19.85 observed in DR-M (normative 30) and 9.25 versus 9.50 observed in

DR-R (normative 0), (b) wholesale price predictions are 11.85 versus 11.91 observed in DR-M (normative 12.85) and 10.86 versus 10.87 in DR-R (normative 11.74), and (c) stocking quantity predictions are 45.75 versus 45.64 observed in DR-M (normative 47.19) and 44.06 versus 44.68 observed in DR-R (normative 45.30). Overall, in this alternative contract setting, we observe similar deviations in decisions as in our main experiment and similar quality predictions from our fairness model, providing further support for fairness as an explanation for the observed deviations.

8. Discussion

Here, we summarize key managerial implications from our study and how our study contributes to the existing literature.

8.1. Managerial Implications

From a managerial perspective, it is unlikely that firms will have the ability to choose among all four of the different inventory-sharing strategies we explore. However, at a minimum, we posit that retailers are able to consider at least one (or more) of these strategies. Thus, the first key question for retailers is if they should adopt an inventory-sharing strategy. Our results suggest that the answer is yes; observed retailer profits are higher (or at least as high) under all inventory-sharing strategies compared with the Baseline environment (see Table 3). Given this observation, the next question becomes *how* retailers should adopt inventory-sharing strategies. First, consider the choice between centralized and decentralized inventory sharing. Our results indicate that retailers prefer decentralized inventory-sharing strategies to a centralized one. Second, regarding the decision authority over the transfer price, our results are consistent with theory and indicate that retailers prefer to negotiate the transfer price rather than have it determined by another party, such as the manufacturer.

Turning to the manufacturer, our results are consistent with theory and indicate that manufacturers prefer serving decentralized retailers when they can set the transfer price, although it is worth noting that there is a large gap between the normative manufacturer profit prediction and the observed manufacturer profit (Table 3). So, although manufacturers still prefer this setting to all others, they are less effective than they could be at using wholesale and transfer pricing power to extract higher profits. Our results are also consistent with normative theory in that manufacturers are worst off when retailers are decentralized and have authority over the transfer price. These results suggest that manufacturers should exert effort to try to control the terms of inventory transfer when possible.

8.2. Contribution to Literature

From a research perspective, our study contributes to the behavioral literature on supply chain contracting and inventory sharing. Although a majority of initial supply chain contracting experiments considered a single manufacturer and single retailer (e.g., Kalkanci et al. 2011, Becker-Peth et al. 2013, Zhang et al. 2016; see Chen and Wu 2019 for a summary), more recently, a number of important studies have extended such a setting and allowed for multiple retailers and inventory sharing. Given that they are the first to evaluate such a setting experimentally, these works consider a one-tier supply chain and focus on order quantities (e.g., Ho et al. 2010, Zhao et al. 2021) and transfer prices (e.g., Li and Chen 2020, Katok and Villa 2021).

Similar to the approach taken by theoretical research on inventory sharing, after there is an established behavioral literature on a one-tier supply chain context, it is natural to move to a behavioral work in a two-tier setting. This allows the field to capture a wider range of supply chain structures in practice. Further, by shifting to a two-tier setting with a strategic interaction, many of the theoretical predictions can differ from those of a one-tier setting, including the potential benefits of inventory sharing (it also allows for a rich analysis of endogenous wholesale prices and distribution of profits). To this end, we build on the existing literature by investigating different retailer inventory-sharing strategies, including no sharing, in a two-tier supply chain.

In many ways, our work complements the existing theoretical research that studies coordination mechanisms for decentralized units. For instance, Celikbas et al. (1999) and Balasubramanian and Bhardwaj (2004) investigate coordination issues between a "revenue-maximizing" marketing department and a "cost-minimizing" manufacturing department within an organization. They find that allowing the two departments to operate in a decentralized manner, with correctly set penalty terms, can achieve outcomes that are equal to or even better than when the two departments are centralized. Our study complements such research in finding that similar outcomes can be achieved by behavioral tendencies. Of course, we would be remiss to say that we are the first supply chain experiment to find evidence of fairness. However, by conducting a novel experiment on a two-tier supply chain with alternative inventory-sharing strategies, we are able to dig deeper and identify the specific impacts that such behavioral biases have on contracts and profits.

9. Conclusion

We investigate how different inventory-sharing strategies affect the distribution of profits in a two-tier supply chain. Our results provide guidance to firms considering how, if at all, they should enter such arrangements. In particular, we examine two important dimensions: (1) whether retailers should adopt a centralized or decentralized inventory-sharing strategy (or not share inventory at all) and (2) when decentralized, which party should have decision authority over the transfer price.

We consider four conditions in our study: a no inventory-sharing setting, a centralized retailer inventory-sharing strategy, and two decentralized retailer inventory-sharing strategies (one where the manufacturer has authority over the transfer price and one where the retailers have authority). We also run a fifth variant with decentralized retailers where the transfer price is exogenously set to zero. Although it may not be feasible for all retailers to choose among all of these scenarios in practice, a vast majority should have the ability to select among multiple options. For instance, even competing retailers (likely the most restricted setting in terms of options) may choose among not sharing inventory, sharing inventory at a transfer price that they negotiate with the other retailer, or outsourcing any inventory-sharing responsibilities to the upstream manufacturer.

To summarize our experimental results, we find evidence that decentralized retailer inventory-sharing strategies perform well depending on the metric of interest. In particular, one important result around profits is that a decentralized retailer inventory-sharing strategy, where the manufacturer sets the transfer price (DR-M), leads to a win-win outcome over both the no inventory-sharing strategy and the centralized retailer inventory-sharing strategy: both the manufacturer and retailers earn significantly higher expected profits. Another key insight is that the decentralized retailer strategy, when the retailers set the transfer price (DR-R), leads to the most equitable outcomes. Last, we observe that both decentralized retailer inventory-sharing strategies with endogenous transfer prices (DR-R and DR-M) generate the highest supply chain efficiency relative to the other strategies (Baseline, CR, and DR-0).

Our analysis of contract terms demonstrates that contact-term decisions deviate from the normative theory in systematic ways, which can account for the profit and efficiency differences we observe (in a follow-up robustness experiment, we also find that these deviations persist in an inventory-sharing setting with a revenue-sharing contract). In an effort to account for these deviations, we find that a model of fairness, which includes both manufacturer and retailer fairness concerns, can organize the data well.

In terms of limitations, in the decentralized inventory-sharing strategy where the retailers have decision authority over the transfer price, we assume that the transfer price is set before the wholesale price. We opted for this not only because it is observed in practice but because if the sequence is reversed, then the profit predictions are identical to those in the centralized retailer inventory-sharing strategy. Although the two cases are the same in theory, it could be interesting to explore them behaviorally for future work. Another limitation is that we assume that transfer prices and retailer demand distributions are common knowledge. Although beyond the scope of this study, investigating how private information affects outcomes with inventory sharing could lead to new insights. Finally, the decentralized retailer inventory-sharing strategy where the retailers set the transfer price is unique in that each party has control over a price. Future work could examine similar contracts, where different prices are set by different parties, in a dedicated study.

Acknowledgments

The authors acknowledge Lawrence Robinson for helpful discussions and seminar participants at Indiana University, the University of Texas at Dallas, and Arizona State University. The authors also thank Jay Swaminathan, the associate editor, and three reviewers for their comments.

Endnotes

[1] See Rudi et al. (2001) for a general solution with asymmetric retailers.

[2] We note that allowing one party to set the quantity for multiple retailers is an interesting avenue for future research.

[3] An earlier version of this paper, which was prior to the coronavirus disease 2019 pandemic, did not include the Baseline treatment and included 42 participants in each of the three treatments (CR, DR-M, and DR-R). During the review process, we added the Baseline condition. Because of the pandemic, we ran this treatment synchronously online through Zoom. We provide further details in Electronic Companion EC.5.1 (e.g., we put each individual participant in his or her own private breakout room with an experimenter, required cameras to be on, etc.); here, note that these online sessions followed the same protocols and used the same participant pool as our original in-person experiments. Further, to determine whether switching to online methods had a meaningful effect on our results, we reran a session of each of the original three treatments (CR, DR-M, DR-R), yielding 57, 60, and 60 participants, respectively, in each. Overall, we found similar results with the online implementation and thus, include all data in our analysis.

[4] We also did not observe any sort of deadline effect (i.e., a mass of agreements at the end of the two-minute time period), which is frequently observed in more adversarial negotiation experiments (Roth et al. 1988).

[5] We obtain similar results if we use t tests or nonparametric tests, with decisions collapsed within subject.

[6] In Electronic Companion EC.2, we also include a power analysis (based on t tests with decisions collapsed within subject), which indicates high power (>90%) for most all of our significant results.

[7] We also considered the equity, reciprocity and competition (ERC) model of Bolton and Ockenfels (2000), where manufacturers suffer disutility when earning more than 1/3 of the supply chain profit. It provides a good, albeit slightly worse, fit as the formulation presented. Details are provided in Electronic Companion EC.3.

[8] Despite observing generous wholesale prices by proposers, they do not consider manufacturer fairness concerns.

[9] An experiment explicitly designed to examine the effects of procedural fairness in supply chain contracts is an exciting opportunity for future work.

References

Andreoni J, Harbaugh WT, Vesterlund L (2010) Altruism in experiments. Durlauf S, Blume L, eds. *Behavioural and Experimental Economics* (Palgrave Macmillan, London), 6–13.

Balasubramanian S, Bhardwaj P (2004) When not all conflict is bad: Manufacturing-marketing conflict and strategic incentive design. *Management Sci.* 50(4):489–502.

Becker-Peth M, Katok E, Thonemann UW (2013) Designing buyback contracts for irrational but predictable newsvendors. *Management Sci.* 59(8):1800–1816.

Beer R, Rios IA, Saban D (2021) Increased transparency in procurement: The role of peer effects. *Management Sci.* 67(12):7511–7534.

Bolton GE, Chen Y (2019) Other-regarding behavior: Fairness, reciprocity, and trust. Donohue K, Katok E, Leider S, eds. *The Handbook of Behavioral Operations* (Wiley, Hoboken, NJ), 199–236.

Bolton GE, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. *Amer. Econom. Rev.* 90(1):166–193.

Bolton GE, Ockenfels A, Thonemann UW (2012) Managers and students as newsvendors. *Management Sci.* 58(12):2225–2233.

Bostian AJA, Holt CA, Jain S, Ramdas K (2012) Is transshipment a behaviorally-robust risk-pooling strategy? Working paper, London Business School, London.

Celikbas M, Shanthikumar JG, Swaminathan JM (1999) Coordinating production quantities and demand forecasts through penalty schemes. *IIE Trans.* 31(9):851–864.

Chen KY, Wu D (2019) Buyer-supplier interactions. Donohue K, Katok E, Leider S, eds. *The Handbook of Behavioral Operations* (Wiley, Hoboken, NJ), 459–487.

Chen DL, Schonger M, Wickens C (2016) oTree—an open-source platform for laboratory, online, and field experiments. *J. Behav. Experiment. Finance* 9:88–97.

Cui TH, Raju JS, Zhang ZJ (2007) Fairness and channel coordination. *Management Sci.* 53(8):1303–1314.

Davis AM (2019) Biases in individual decision-making. Donohue K, Katok E, Leider S, eds. *The Handbook of Behavioral Operations* (Wiley, Hoboken, NJ), 149–198.

Davis AM, Leider S (2018) Contracts and capacity investment in supply chains. *Manufacturing Service Oper. Management* 20(3):403–421.

Davis AM, Katok E, Santamaría N (2014) Push, pull, or both? A behavioral study of how the allocation of inventory risk affects channel efficiency. *Management Sci.* 60(11):2666–2683.

DeHoratius N, Raman A (2007) Store manager incentive design and retail performance: An exploratory investigation. *Manufacturing Service Oper. Management* 9(4):518–534.

Dong L, Rudi N (2004) Who benefits from transshipment? Exogenous vs. endogenous wholesale prices. *Management Sci.* 50(5):645–657.

Donohue K, Katok E, Leider S, eds. (2019) *The Handbook of Behavioral Operations* (Wiley, Hoboken, NJ).

Eppen GD (1979) Note: Effects of centralization on expected costs in a multi-location newsboy problem. *Management Sci.* 25(5):498–501.

Fehr E, Schmidt KM (1999) A theory of fairness, competition and cooperation. *Quart. J. Econom.* 114(3):817–868.

Forsythe R, Horowitz JL, Savin N, Sefton M (1994) Fairness in simple bargaining experiments. *Games Econom. Behav.* 6(3):347–369.

Ho TH, Lim N, Cui TH (2010) Reference dependence in multilocation newsvendor models: A structural analysis. *Management Sci.* 56(11):1891–1910.

Hyndman K, Embrey M (2019) Econometrics for experiments. Donohue K, Katok E, Leider S, eds. *The Handbook of Behavioral Operations* (Wiley, Hoboken, NJ), 35–88.

Kagel JH, Roth AE, eds. (2017) *Handbook of Experimental Economics*, vol. 2 (Princeton University Press, Princeton, NJ).

Kalkanci B, Chen KY, Erhun F (2011) Contract complexity and performance under asymmetric demand information: An experimental evaluation. *Management Sci.* 57(4):689–704.

Kalkanci B, Chen KY, Erhun F (2014) Complexity as a contract design factor: A human-to-human experimental study. *Production Oper. Management* 23(2):269–284.

Katok E, Villa S (2021) Centralized or decentralized transfer prices: A behavioral approach for improving supply chain coordination. *Manufacturing Service Oper. Management* 24(1):143–158.

Katok E, Thomas D, Davis A (2008) Service-level agreements as coordination mechanisms: The effect of review periods. *Manufacturing Service Oper. Management* 10(4):609–624.

Kremer M, Siemsen E, Thomas DJ (2016) The sum and its parts: Judgmental hierarchical forecasting. *Management Sci.* 62(9): 2745–2764.

Leider S (2019) Behavioral analysis of strategic interactions: Game theory, bargaining, and agency. Donohue K, Katok E, Leider S, eds. *The Handbook of Behavioral Operations* (Wiley, Hoboken, NJ), 237–286.

Li S, Chen KY (2020) The commitment conundrum of inventory sharing. *Production Oper. Management* 29(2):353–370.

List JA, Shaikh AM, Xu Y (2019) Multiple hypothesis testing in experimental economics. *Experiment. Econom.* 22:773–793.

Narus JA, Anderson JC (1996) Rethinking distribution: Adaptive channels. *Harvard Bus. Rev.* 74:112–120.

Ozer O, Zheng Y, Chen KY (2011) Trust in forecast information sharing. *Management Sci.* 57(6):1111–1137.

Padmanabhan VP, Rudi N, Tsetlin I (2010) Who prices? Manufacturer's vs. retailers' control of retail prices. Working paper, INSEAD, Fontainebleau, France.

Park SJ, Lai G, Seshadri S (2016) Inventory sharing in the presence of commodity markets. *Production Oper. Management* 25(7): 1245–1260.

Robinson LW (1990) Optimal and approximate policies in multiperiod, multilocation inventory models with transshipments. *Oper. Res.* 38(2):278–295.

Roth AE (1995) Bargaining experiments. Kagel J, Roth AE, eds. *Handbook of Experimental Economics* (Princeton University Press, Princeton, NJ), 253–348.

Roth AE, Murnighan J, Schoumaker F (1988) The deadline effect in bargaining: Some experimental evidence. *Amer. Econom. Rev.* 78(4):806–823.

Rudi N, Kapur S, Pyke DF (2001) A two-location inventory model with transshipment and local decision making. *Management Sci.* 47(12):1668–1680.

Schweitzer ME, Cachon GP (2000) Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Sci.* 46(3):404–420.

Shao J, Krishnan H, McCormick ST (2011) Incentives for transshipment in a supply chain with decentralized retailers. *Manufacturing Service Oper. Management* 13(3):361–372.

Su X (2008) Bounded rationality in newsvendor models. *Manufacturing Service Oper. Management* 10(4):566–589.

Van Donselaar KH, Gaur V, Van Woensel T, Broekmeulen RA, Fransoo JC (2010) Ordering behavior in retail stores and implications for automated replenishment. *Management Sci.* 56(5): 766–784.

Villa S, Castañeda JA (2018) Transshipments in supply chains: A behavioral investigation. *Eur. J. Oper. Res.* 269(2):715–729.

Zhang Y, Donohue K, Cui TH (2016) Contract preferences and performance for the loss averse supplier: Buyback vs. revenue sharing. *Management Sci.* 62(6):1734–1754.

Zhao H, Deshpande V, Ryan JK (2005) Inventory sharing and rationing in decentralized dealer networks. *Management Sci.* 51(4): 531–547.

Zhao H, Xu L, Siemsen E (2021) Inventory sharing and demand-side underweighting. *Manufacturing Service Oper. Management* 23(5):1217–1236.