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

Andrew M. Davis, Rihuan Huang
Samuel Curtis Johnson Graduate School of Management, Cornell SC Johnson College of Business, Cornell University, Ithaca, NY 14853, adavis@cornell.edu, rh567@cornell.edu

Douglas J. Thomas
Darden School of Business, University of Virginia, Charlottesville VA 22903, ThomasD@darden.virginia.edu

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 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.

Key words: Inventory Sharing, Behavioral Operations, Supply Chain Contracting, Risk Pooling
History: June 17, 2021

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 have 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 automobile (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).

While 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 to a setting without any inventory sharing at all. This is largely due to 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. 2020), 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 to retailers acting independently (Eppen 1979). There is also theoretical work which 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 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. 2020). Although, there are two papers which 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 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 to 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 to 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 to 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 which 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 to 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.

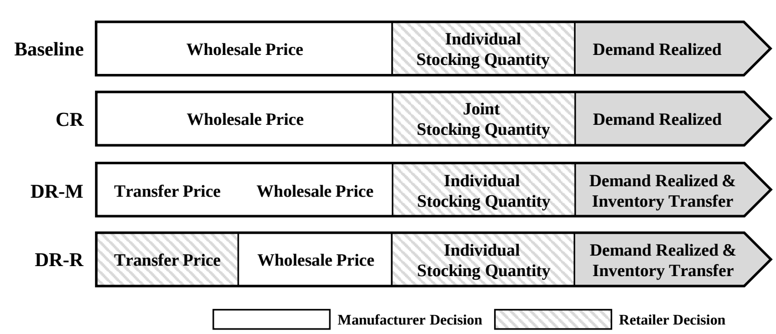
2. Normative Theory and Predictions

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 which 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, while 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.

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



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 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 (CR)

Under the centralized retailer inventory-sharing strategy, CR, 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 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 $\left(q_i^c, q_j^c\right)$ derived from Equation (5)

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

Given $\left(q_i^c, q_j^c\right)$, 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 strategy, 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 over-stocking retailer shares any leftover inventory to an under-stocking 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

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

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)

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

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).

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

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

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

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 0 and p for low-margin products).

2.4. Experimental Predictions and Hypotheses

Due to 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 above: 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. While 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	1041.67	1242.26	1434.44	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%

Note: We assume that a continuous approximation of demand is sufficiently precise for predictions. Also, note (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 (e-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). And, 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. And 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. While 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 (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, thus:

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 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 While 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, please 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 do 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.

While 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 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 can not 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 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. (2020) conduct a survey of 54 firms and find that 35 (65%) rely on automated algorithms which are adjusted by human managers and 10 (19%) rely 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. Please see e-companion EC.5.3 for sample instructions and screenshots.

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

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 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 single retailer. Also, following existing experimental supply chain and economics research, our study includes university student participants (Donohue et al. 2019, Kagel and Roth 2017), 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 which 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.

[3] An earlier version of this paper, which was prior to the COVID-19 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. Due to the pandemic, we ran this treatment synchronously online through Zoom. We provide further details in e-companion EC.5.1 (e.g., we put each individual participant in their own private breakout room with an experimenter, required cameras to be on, etc), and 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 did not have a meaningful effect on our results, we re-ran a session of each of the original three treatments (CR, DR-M, DR-R), yielding 57, 60, and 60 participants in each. Overall, we found similar results with the online implementation and thus include all data in our analysis.

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 fraction of time that retailers came to an agreement over the stocking quantity in CR and the transfer price in DR-R was 95.18% and 97.08%. 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 fraction of time retailers accepted the manufacturer's wholesale price and set a positive stocking quantity was 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.

[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 non-parametric tests, with decisions collapsed within subject.

[6] In e-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.

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

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

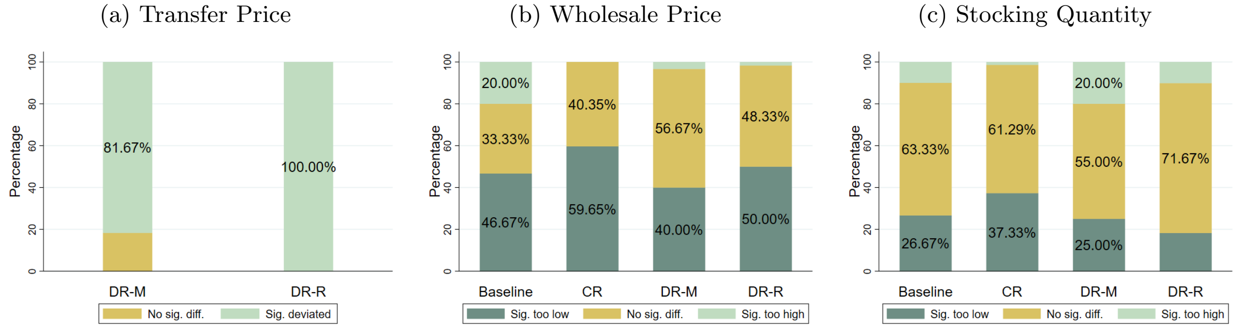
Note: Standard errors, across participants, reported in parentheses. Results for DR-R and CR are conditioning on agreement. Stocking quantity in CR is one-half of the average joint stocking quantity. Predicted wholesale prices in DR-M and DR-R are conditioning on observed transfer prices. Predicted stocking quantities are conditioning on observed wholesale prices (and transfer prices). Unconditional normative predictions, when applicable, reported in square brackets. Significance of regressions comparing observed versus conditional normative predictions given by † $p < 0.005$ (the corrected critical p-value). Transfer prices deviate from the normative predictions and wholesale prices are often set too low. Stocking quantities are set close to predictions or low.

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). While 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 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 each individual participant we conducted a test between their decisions and the conditional predictions. Due to 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 2a, 81.67% in DR-M and 100% in DR-R of participants set transfer prices that significantly deviated from the normative predictions. In Figure 2b, 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 2c, a majority of participants set quantities in a way that is not significantly different from the conditional prediction, but a reasonable percentage understock in Baseline, CR, and DR-M. This coincides with the aggregate results. Thus we have our first result, which largely rejects Hypothesis 1:

Figure 2 Percentages of Participants Deviating in Pricing and Stocking Decisions



Note: Classification of a subject is based on Wilcoxon signed-rank tests between observed decisions and (conditional) optimal decisions (at 10% level). Percentages below 20% are omitted for due to limited space.

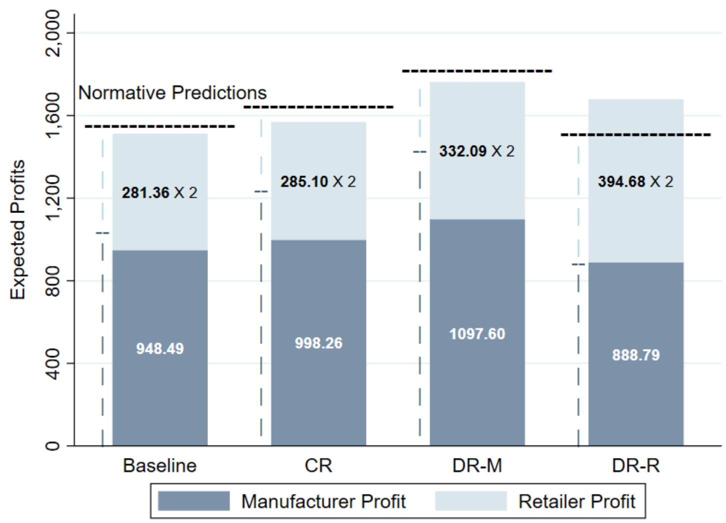
Result 1 **(Contract Terms)**. *Transfer prices in the two decentralized inventory-sharing strategies, DR-M and DR-R, deviate symmetrically from the normative predictions. Wholesale prices are set too low in all three of the inventory-sharing environments, CR, DR-M, and DR-R. Stocking quantities are 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 were 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 rounds all of our main results (including ones highlighted later) continue to hold. Second, retailer stocking quantity deviations do not increase with higher wholesale prices. And 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. While we will compare the observed results to the normative theoretical predictions momentarily, we also include the normative predictions by dashed horizontal and vertical lines.

Figure 3 Average Observed Expected Profits



Note: 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 to Baseline (and DR-M higher than CR).

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 to 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, 1097.60 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 to the Baseline setting and centralized 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 being 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 **(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 to 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 to 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). While 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

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)	1097.60† (22.35)	888.79 (20.68)	1041.67	1242.26	1434.44	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% (0.008)	71.54% (0.011)	80.36%† (0.006)	76.54%† (0.007)	71.27%	75.55%	83.60%	69.26%

Note: Standard errors, across participants, reported in parentheses. Significance of regressions comparing observed versus normative predictions given by † $p < 0.017$ (the corrected critical p-value). There are significant differences between all profits and efficiencies compared to the normative benchmarks except manufacturer profit in DR-R, and efficiency in Baseline and CR.

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 **(Efficiency)**. *Both decentralized retailer inventory-sharing strategies, DR-M and DR-R, achieve a higher efficiency compared to 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 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 have decision authority over the transfer price, depending on the situation.

4.2.1. Connecting Contract Terms with Profits and Efficiency Here we connect our results by summarizing how the the observed contract-term deviations highlighted in Result 1 can largely account for Results 2, 3, and 4 (profit, equity, and efficiency). Beginning with Result 2, which finds that DR-M generates a Pareto improvement over the Baseline and CR conditions, due to 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 to 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 (and 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 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 to 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 ${}^{(\pi(w,q(w,t)|t)-\pi(w^*,q(w^*,t)|t)}\!/_{\pi(t^*,w^*,q(w^*,t^*))}$. 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% to 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 sub-optimal 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 retailers earn more than theory predicts and DR-R becomes quite equitable.

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

	Baseline		CR		DR-M		DR-R	
	Retail.	Manuf.	Retail.	Manuf.	Retail.	Manuf.	Retail.	Manuf.
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%

Note: 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 to 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.

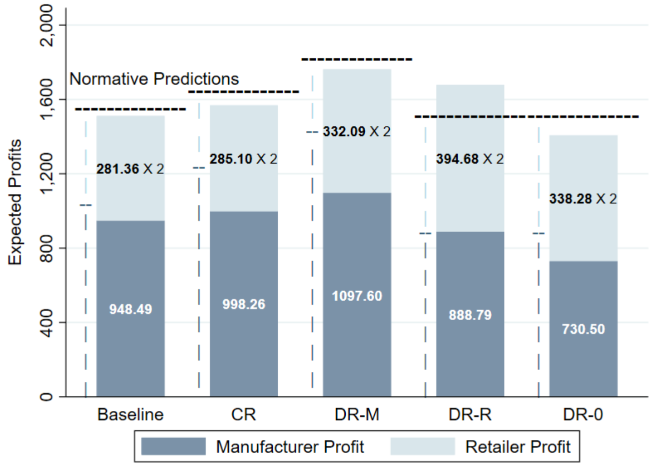
5. Alternative DR Treatment

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 who dictates that all retailers act independently but are required to share product with one another at a transfer price of zero. And 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 to 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 to 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.

In 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

Figure 4 **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]

Note: 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, reported in square brackets). Significance of regressions versus conditional normative predictions given by † $p < 0.005$ (the corrected critical p-value).

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 to when it is exogenously set to zero.*

There are at least two plausible (non-exclusive) 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 (resp. 1), then the number of observations increases to 12.0% (resp. 25.8%). In this case, retailers in DR-R understock by an average of 0.68 units (resp. 1.96 for $t < 1$). This degree of understocking is far less than what we observe in DR-0, which is 4.04. Overall, this analysis suggests that while 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, who 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, below we provide an overview of this behavioral bias and relegate detailed theoretical analyses, estimation procedures, and more results in e-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. 2020). 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, while 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 $\pi^d_{r,i}$ are the expected profit of manufacturer and retailer i 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

$$u_m^F = \pi_m - \lambda_m(\pi_m - \pi^d_{r,i})^+, \quad u^{d,F}_{r,i} = \pi^d_{r,i} - \lambda_r(\pi_m - \pi^d_{r,i})^+, \tag{10}$$

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 e-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 sub-optimal transfer prices on profits is relatively small compared to 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: $\theta \to 0$ is fully rational and $\theta \to \infty$ is fully random decisions (Su 2008).

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	-13620.24	-13449.36	-13566.87	-13390.01
Sum of separate estimations	LL_s	-13469.88	-13313.81	-13404.42	-13248.71

Note: "LL" represents log-likelihoods. Number of observations in each LL is 3370. $\hat{\theta}$ is 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 e-companion EC.3.

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 then sums up the five log-likelihoods (LL_s). Both approaches are depicted in Table 5. A series of likelihood-ratio tests

[7] We also considered the 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 e-companion EC.3.

reveals significant differences across all applicable estimations, such that the full-fairness model with both retailer and manufacturer concerns, $\text{Fair}_{m,r}$, 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 -13449.36 versus -13566.87). This suggests that the inclusion of manufacturer-fairness concerns is especially important to the overall fit.

In the top of Table 6 we provide the transfer price, wholesale price, and quantity predictions using the parameter estimates from the full-fairness model, $\text{Fair}_{m,r}$, 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, 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 17.34 versus 17.35. Last, the predicted quantities are within one unit of the observed quantities, in all five treatments.

Table 6 Fairness Predictions and Observed Values (top) and Parameter Estimates (bottom), by Treatment

		Baseline		CR		DR-M		DR-R		DR-0	
		$\text{Fair}_{m,r}$	Obs.	$\text{Fair}_{m,r}$	Obs.	$\text{Fair}_{m,r}$	Obs.	$\text{Fair}_{m,r}$	Obs.	$\text{Fair}_{m,r}$	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	17.34	17.35
	q	41.45	41.69	39.04	38.37	43.33	42.86	40.42	39.80	31.69	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		$\hat{\lambda}_m$=0.187	
	$\hat{\lambda}_r$	$\hat{\lambda}_r$=0.054		$\hat{\lambda}_r$=0.139		$\hat{\lambda}_r$=0.065		$\hat{\lambda}_r$=0.012		$\hat{\lambda}_r$=0.154	

Note: Number of observations for Baseline, CR, DR-M, DR-R, and DR-0 are 360, 434, 960, 932, and 684 (3370 total). DR-R and CR are conditioning on agreement. "Obs." is observed average values. $\hat{\theta}$ is 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 e-companion EC.3.

The bottom 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 sub-optimal transfer prices. Manufacturer-fairness concerns ($\hat{\lambda}_m$) are generally larger than retailer-fairness concerns ($\hat{\lambda}_r$) in all treatments, but one 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, 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). But, 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 regards 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 due to 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 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]

[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.

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 e-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 to 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, while 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. While 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. 2020) 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. While 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, to sharing inventory at a transfer price that they negotiate with the other retailer, or to outsource 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. While the two cases are the same in theory, it could be interesting to explore behaviorally for future work. Another limitation is that we assume that transfer prices and retailer demand distributions are common knowledge. While 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

We gratefully acknowledge Lawrence Robinson for helpful discussions, seminar participants at Indiana University, the University of Texas at Dallas, Arizona State University, and financial support from Cornell University and the University of Virginia.

References

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

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

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

Beer R, Rios IA, Saban D (2020) Increased transparency in procurement: The role of peer effects. *Management Science* Articles in advance, https://doi.org/10.1287/mnsc.2020.3894.

Bolton GE, Chen Y (2019) Other-regarding behavior: Fairness, reciprocity, and trust. Donohue K, Katok E, Leider S, eds., *The Handbook of Behavioral Operations*, chapter 6, 199–236 (Wiley).

Bolton GE, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. *The American Economic Review* 90(1):166–193.

Bolton GE, Ockenfels A, Thonemann UW (2012) Managers and students as newsvendors. *Management Science* 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.

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

Chen DL, Schonger M, Wickens C (2016) otree – an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9:88–97.

Chen KY, Wu D (2019) Buyer-supplier interactions. Donohue K, Katok E, Leider S, eds., *The Handbook of Behavioral Operations*, chapter 13, 459–487 (Wiley).

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

Davis AM (2019) Biases in individual decision-making. Donohue K, Katok E, Leider S, eds., *The Handbook of Behavioral Operations*, chapter 5, 149–198 (Wiley).

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 Science* 60(11):2666–2683.

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

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

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

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

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

Fehr E, Schmidt KM (1999) A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* 114:817–868.

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

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

Hyndman K, Embrey M (2019) Econometrics for experiments. Donohue K, Katok E, Leider S, eds., *The Handbook of Behavioral Operations*, chapter 2, 35–88 (Wiley).

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

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

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

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

Katok E, Villa S (2021) Centralized or decentralized transfer prices: A behavioral approach for improving supply chain coordination. *Manufacturing & Service Operations Management* Articles in advance, https://doi.org/10.1287/msom.2020.0957.

Kremer M, Siemsen E, Thomas DJ (2016) The sum and its parts: Judgmental hierarchical forecasting. *Management Science* 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*, chapter 7, 237–286 (Wiley).

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

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

Narus JA, Anderson JC (1996) Rethinking distribution: adaptive channels. *Harvard Business Review* 74:112–120.

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

Padmanabhan VP, Rudi N, Tsetlin I (2010) Who prices? manufacturer's versus retailers' control of retail prices. Working paper, INSEAD.

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

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

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

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

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

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

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

Su X (2008) Bounded rationality in newsvendor models. *Manufacturing & Service Operations 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 Science* 56(5):766–784.

Villa S, Castañeda JA (2018) Transshipments in supply chains: A behavioral investigation. *European Journal of Operational Research* 269(2):715–729.

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

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

Zhao H, Xu L, Siemsen E (2020) Inventory sharing and demand-side underweighting. *Manufacturing & Service Operations Management* Articles in advance, https://doi.org/10.1287/msom.2020.0875.

Electronic Companion

EC.1 Supplemental Experimental Prediction Figures

Figure EC.1.1 Manufacturer Expected Profits with Respect to the Transfer Price and the Wholesale Price

(a) Transfer Price (b) Wholesale Price

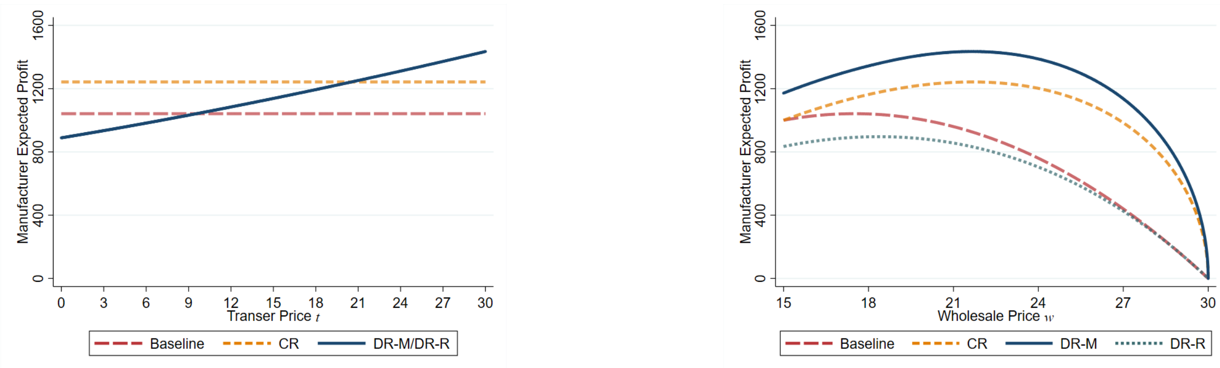


Figure EC.1.2 Retailer Expected Profits with Respect to the Transfer Price and the Wholesale Price

(a) Transfer Price (b) Wholesale Price

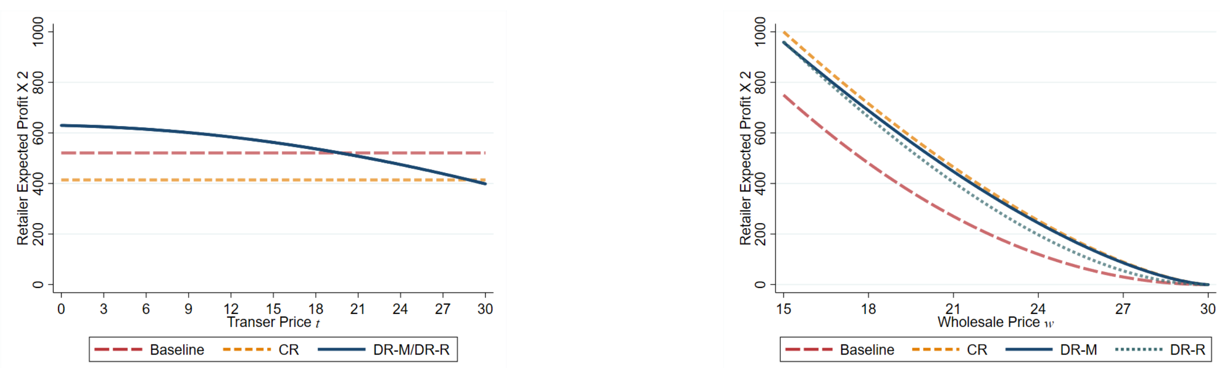
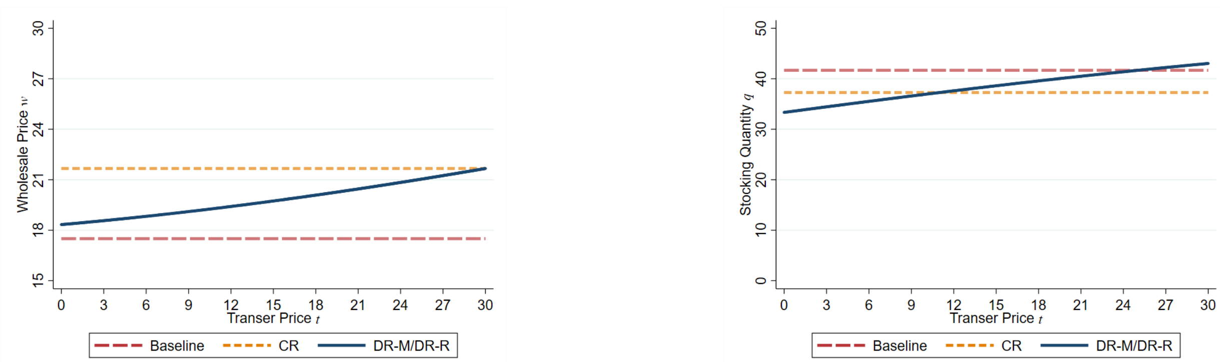


Figure EC.1.3 Wholesale Price and Stocking Quantity with Respect to the Transfer Price

(a) Wholesale Price w (b) Stocking Quantity q



EC.2 Power Analysis

Here we provide a power analysis for all hypothesis test results by tables and formal results (1-5). Due to complexity of conducting power analysis on random effects regression, we show power of t-tests with decisions collapsed within subject (t-tests generally lead to higher p-values in our hypothesis testing). Power levels are reported with critical p-values using Bonferroni corrections, assuming an unadjusted critical p-value of 0.05. The p-values of t-tests are reported in square brackets. Sample sizes are reported in parentheses when applicable.

Table EC.2.1 Power for Table 2 and Result 1

	Baseline (30)	CR (57)	DR-M (60)	DR-R (60)
Transfer Price	-	-	>0.99 [<0.01]	>0.99 [<0.01]
Wholesale Price	0.07 [0.16]	>0.99 [<0.01]	>0.99 [<0.01]	>0.99 [<0.01]
Stocking Quantity	0.14 [0.02]	0.81 [<0.01]	0.16 [0.04]	0.02 [0.28]

Table EC.2.2 Power for Table 3

	Baseline (30)	CR (57)	DR-M (60)	DR-R (60)
Manufacturer Profit	0.95 [<0.01]	>0.99 [<0.01]	>0.99 [<0.01]	0.02 [>0.99]
Retailer Profit	0.64 [0.01]	>0.99 [<0.01]	>0.99 [<0.01]	>0.99 [<0.01]
Supply Chain Efficiency (%)	0.11 [0.23]	0.40 [0.03]	0.66 [<0.01]	>0.99 [<0.01]

Table EC.2.3 Power for Results 2 and 4

	DR-M (60) vs.		Baseline (30) vs.	
	Baseline (30)	CR (57)	CR (57)	DR-R (60)
Manufacturer Profit	0.96 [<0.01]	0.51 [<0.01]		
Retailer Profit	0.96 [<0.01]	0.75 [<0.01]		
Supply Chain Efficiency (%)	>0.99 [<0.01]	0.92 [<0.01]	0.05 [0.37]	0.72 [<0.01]

Table EC.2.4 Power for Result 3

	DR-R (60) vs.		
	Baseline (30)	CR (57)	DR-M (60)
R. Share	>0.99 [<0.01]	>0.99 [<0.01]	>0.99 [<0.01]
M. R. Gap	>0.99 [<0.01]	>0.99 [<0.01]	>0.99 [<0.01]

Note: "R. Share" is retailers' share of supply chain profit. "M. R. Gap" is absolute difference between manufacturer and retailer profits.

Table EC.2.5 Power related to DR-0 and Result 5

	DR-R (60)		Normative Predictions	
	Manufacturer Profit	Retailer Profit	Wholesale Price	Stocking Quantity
DR-0 (57)	>0.99 [<0.01]	0.93 [<0.01]	0.53 [<0.01]	0.77 [<0.01]

EC.3 Behavioral Model Details

Here we present behavioral model details and estimations. We provide information for the decentralized retailer inventory-sharing strategy (DR), and note that the centralized retailer inventory-sharing strategy (CR) follows similar logic.

EC.3.1. Fairness Model

EC.3.1.1. Manufacturer Advantageous Fairness

In DR, a manufacturer with advantageous fairness concerns maximizes its expected utility function

$$u_m^{d,F} = \pi_m - \lambda_m(\pi_m - \pi_{r,i}^d)^+, \tag{EC.3.1}$$

where λ_m is the manufacturer's degree of fairness concerns. Note that in Equation (EC.3.1) the manufacturer compares its profit with a retailer's normative expected profit without taking retailer's fairness concern into account. This is because our decision support system shows the manufacturer a retailer's optimal quantity and expected profit. Moreover, this allows us to separate the two parties' fairness concerns.

Figure EC.3.1 provides a numerical analysis of how manufacturer advantageous concerns affect its decisions in DR-M and DR-R. While Figure EC.3.1a shows that λ_m does not affect the manufacturer's optimal transfer price in DR-M, Figure EC.3.1b shows that a higher λ_m makes the manufacturer choose a lower wholesale price in DR-M and DR-R as it achieves a more equitable outcome. In DR-M, one explanation for this result is that a lower wholesale price reduces inequity while keeping the manufacturer expected profit relatively high (since retailers will set a higher quantity in response). On the other hand, reducing the transfer price will lead to lower quantities and have a more significant detrimental impact on the manufacturer's expected profit. Further, the transfer price has less impact on retailer profit compared to the wholesale price. These effects can be seen in Figures EC.1.1 and EC.1.2 in Section EC.1.

EC.3.1.2. Retailer Disadvantageous Fairness

With fairness concerns, a decentralized retailer i is optimizing its expected utility

$$u_{r,i}^{d,F} = \pi_{r,i}^d - \lambda_r(\pi_m - \pi_{r,i}^d)^+, \tag{EC.3.2}$$

where λ_r is the degree of retailer i's fairness concerns.

The optimal quantity $(q_i^{d,F}, q_j^{d,F})$ at equilibrium is given by

$$\begin{aligned}
&\alpha(q_i) - \beta_i(q_i,q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i,q_j)\left(\frac{p-t}{p}\right) = \frac{p-w}{p} - \frac{\lambda_r}{1+\lambda_r}\frac{w-c}{p} && \text{if } \pi_m - \pi_{r,i}^d > 0,\\
&\alpha(q_i) - \beta_i(q_i,q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i,q_j)\left(\frac{p-t}{p}\right) = \frac{p-w}{p} && \text{if } \pi_m - \pi_{r,i}^d \le 0.
\end{aligned} \tag{EC.3.3}$$

Similarly, retailers i and j in CR are optimizing the joint expected utility function

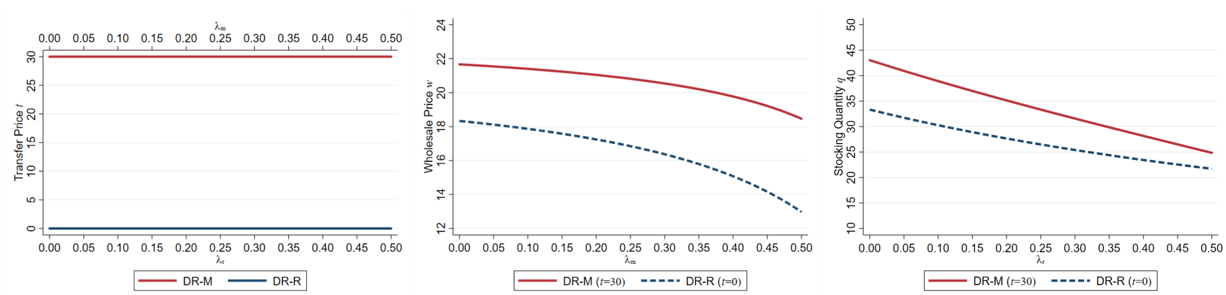
$$u_r^{c,F} = \pi_r^c - 2\lambda(\pi_m - \pi_r^c/2)^+. \tag{EC.3.4}$$

Note that the fairness term in utility function (EC.3.4) considers only half of the joint retailer profit because when making decisions, retailers observe individual profit instead of joint profit. Therefore, the disutility term is at the individual level and multiplied by 2 in the joint utility function. Optimal quantities in CR follow a similar logic as DR.

Additional numerical analyses (not presented here) indicate that λ_r may potentially lead to a lower transfer price, when the production cost is very high. However this is not applicable in our parameter setting (see Figure EC.3.1a) since the normative theory predicates that retailers should set the transfer price at zero. Turning to the quantity, Figure EC.3.1c shows that a higher degree of disadvantageous fairness concerns results in a lower quantity, which leads to a more fair outcome.

Figure EC.3.1 Impact of Fairness Concerns on Transfer Price, Wholesale Price and Stocking Quantity

(a) Transfer Price t w.r.t. λ_m or λ_r (b) Wholesale Price w w.r.t. λ_m (c) Stocking Quantity q w.r.t. λ_r



EC.3.2. ERC model

We also consider the equity, reciprocity, and competition (ERC) model in Bolton and Ockenfels (2000) where manufacturers suffer disutility when earning more than 1/3 of the supply chain profit. Here we provide some technical details and later show it fits slightly worse than fairness formulation presented above. In ERC, retailers decide quantities following the normative theory, i.e., maximizing Equation (4) in CR and Equation (7) in DR. The manufacturer's expected utility function with decentralized retailers is

$$u_m^{d,ERC} = \pi_m - 100\frac{\omega}{3}\left(\frac{\pi_m}{\pi_m + \pi_{r,i}^d + \pi_{r,j}^d} - \frac{1}{3}\right)^2. \tag{EC.3.5}$$

In Equation (EC.3.5), the disutility term reflects concern when the total profit is not evenly distributed among the three parties and ω is the degree of such concern. The number 100 is a scalar given the maximum individual demand in our experiments is 100. The utility function with centralized retailers can be derived by replacing retailer expected profit function in Equation (EC.3.5).

EC.3.3. Random Errors

For random errors in transfer prices, let Ω represent the decision space of t. Let u_m and u_r be the general utility function of the manufacturer and a single retailer. Transfer prices in DR-M and DR-R are chosen with probabilities ρ_m and ρ_r, respectively

$$\rho_m(t,\theta,\cdot)=\frac{e^{u_m(t,\cdot)/\theta}}{\sum\limits_{t\in\Omega} e^{u_m(t,\cdot)/\theta}}, \quad \rho_r(t,\theta,\cdot)=\frac{e^{u_r(t,\cdot)/\theta}}{\sum\limits_{t\in\Omega} e^{u_r(t,\cdot)/\theta}}, \tag{EC.3.6}$$

where θ is the degree of rationality: $\theta \to 0$ means the decision maker is fully rational and $\theta \to \infty$ means they make fully random decisions, i.e., each transfer price will be picked with the same probability. Equation (EC.3.6) can be applied to the normative theory or any behavioral model mentioned above by specifying the form of u_m and u_r, and corresponding parameters.

EC.3.4. Estimation Methodology and Results

We now structurally estimate the parameters of our behavioral model. A heterogeneity analysis of our data shows that participants make sub-optimal decisions for wholesale prices and stocking quantities around both sides of the optimal points. Therefore, for our maximum-likelihood estimation (MLE) we use truncated normal distributions for these two decisions with conditional optimal values as the means and estimate the standard deviations. For the wholesale price the normal distribution is truncated at unit cost $c=5$ and unit selling price $p=30$. For the stocking quantity the distribution is truncated at lower bound 0 and upper bound 100 of demand for decentralized retailers. For centralized retailers in CR, the upper bound is 200. The probability density function of a truncated normal distribution $\varphi(x;\mu,\sigma,a,b)$ is defined by

$$\varphi(x;\mu,\sigma,a,b)=\frac{\phi(\frac{x-\mu}{\sigma})}{\sigma\left(\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)\right)}, \tag{EC.3.7}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function and cumulative distribution function of the standard normal distribution. To be consistent with our analysis of contract terms, non-agreement data in CR and DR-R are excluded from estimations.

Table EC.3.1, EC.3.2 and Table EC.3.3 show MLE estimation results for Baseline, CR, DR-M, DR-R and DR-0. Random errors model is estimated for DR-M and DR-R only as they include transfer price decisions. Compared to the full-fairness model, ERC fits much worse in CR and slightly worse in three DR treatments.

Table EC.3.1 Maximum-likelihood Estimation Results in Baseline

	Fair_m	Fair_r	$\text{Fair}_{m,r}$	ERC
$\hat{\lambda}_m$	0.147 (0.065)		0.147 (0.178)	
$\hat{\lambda}_r$		0.054 (0.017)	0.054 (0.018)	
$\hat{\omega}$				0.098
$\hat{\sigma}_w$	2.514	2.576	2.514	2.514
$\hat{\sigma}_q$	9.522	9.313	9.313	9.522
LL	-1161.32	-1158.40	-1155.47	-1161.32

Note: Number of observations is 360 in Baseline. "LL" represents log-likelihoods. Standard errors in parentheses are derived from bootstrapping 2000 times. Standard errors of ERC are unable to be computed in reasonable time.

Table EC.3.2 Maximum-likelihood Estimation Results in CR and DR-0

	CR				DR-0			
	Fair_m	Fair_r	$\text{Fair}_{m,r}$	ERC	Fair_m	Fair_r	$\text{Fair}_{m,r}$	ERC
$\hat{\lambda}_m$	0.392 (0.014)		0.392 (0.015)		0.186 (0.031)		0.186 (0.052)	
$\hat{\lambda}_r$		0.139 (0.024)	0.139 (0.024)			0.152 (0.023)	0.152 (0.023)	
$\hat{\omega}$				0.351				0.211
$\hat{\sigma}_w$	2.704	3.880	2.704	2.704	3.204	3.357	3.204	3.204
$\hat{\sigma}_q$	9.064	8.148	8.148	13.657	11.991	11.549	11.549	11.991
LL	-1309.21	-1351.23	-1285.74	-1391.61	-2356.95	-2338.75	-2328.43	-2356.95

Note: Number of observations are 434 in CR (non-agreement data excluded) and 684 in DR-0. "LL" represents log-likelihoods. Standard errors in parentheses are derived from bootstrapping 2000 times. Standard errors of ERC are unable to be computed in reasonable time.

Table EC.3.3 Maximum-likelihood Estimation Results in DR-M and DR-R

	DR-M					DR-R				
	Errors	Fair_m	Fair_r	$\text{Fair}_{m,r}$	ERC	Errors	Fair_m	Fair_r	$\text{Fair}_{m,r}$	ERC
$\hat{\theta}$	250.709	125.508	234.290	118.184	239.736	19.306	33.829	20.570	35.010	16.134
	(24.990)	(21.546)	(24.492)	(13.511)		(1.960)	(3.503)	(2.175)	(6.261)	
$\hat{\lambda}_m$		0.382		0.382			0.332		0.332	
		(0.067)		(0.032)			(0.022)		(0.037)	
$\hat{\lambda}_r$			0.065	0.065				0.017	0.012	
			(0.019)	(0.020)				(0.013)	(0.013)	
$\hat{\omega}$					0.354					0.441
$\hat{\sigma}_w$	3.938	3.222	3.938	3.222	3.301	4.156	3.544	4.156	3.544	3.507
$\hat{\sigma}_q$	12.893	12.893	12.748	12.748	12.893	9.988	9.988	9.992	9.988	9.988
LL	-4440.10	-4399.43	-4433.17	-4392.47	-4405.73	-4123.56	-4086.90	-4122.87	-4086.60	-4090.46

Note: Number of observations are 960 in DR-M and 932 in DR-R (non-agreement data excluded). "Errors" means only transfer errors parameter is estimated. "LL" represents log-likelihoods. Standard errors in parentheses are derived from bootstrapping 2000 times. Standard errors of ERC are unable to be computed in reasonable time.

EC.4 Robustness Check: Revenue Sharing with Decentralized Retailers

Here we report results of two revenue-sharing experiments. These serve as a robustness check, allowing us to determine whether the behavioral irregularities in our original experiments extend to such a setting and whether fairness can capture decisions. Due to their favorable performance, we focus on the DR-M and DR-R settings. We assume that the revenue share split is exogenous, ensuring that any new treatments have the same number of decisions as our original experiments.

EC.4.1. Normative Theory

We first briefly summarize the normative theory for a revenue-sharing contract with decentralized retailers. Consider the normative theory outlined previously in Section 2. By backward deduction, with revenue sharing, retailer i's expected profit function is expressed by Equation (EC.4.1)

$$\tilde{\pi}^d_{r,i} = \psi \mathbb{E}\left[p \min(d_i, q_i) + tT_i + (p-t)T_j\right] - wq_i. \tag{EC.4.1}$$

where ψ is the ratio of revenue-sharing. The optimal stocking quantity under Nash equilibrium $(\tilde{q}^d_i, \tilde{q}^d_j)$ is given by Equation (EC.4.2)

$$\alpha(q_i) - \beta_i(q_i, q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i, q_j)\left(\frac{p-t}{p}\right) = \frac{\psi p - w}{\psi p}, \tag{EC.4.2}$$

and we have the following proposition (please see Section EC.4.3 for all proofs).

PROPOSITION EC.1. *There exists a unique Nash equilibrium, $(\tilde{q}^d_i, \tilde{q}^d_j)$, in a revenue-sharing contract with the decentralized retailer inventory-sharing strategy.*

The following lemma shows retailer quantities at equilibrium have similar characteristics under a revenue-sharing contract versus a wholesale price contract:

LEMMA EC.1. *Under a revenue-sharing contract, a decentralized retailer i's optimal inventory level, $\tilde{q}^d_i$, is increasing in the transfer price t.*

Given retailers' quantities, the manufacturer's expected profit function is

$$\tilde{\pi}_m = (w-c)(q_i + q_j) + (1-\psi)\mathbb{E}[p \min(d_i + d_j, q_i + q_j)]. \tag{EC.4.3}$$

Similar to a wholesale price contract, the manufacturer always prefers a higher transfer price.

LEMMA EC.2. *When the manufacturer optimally sets its wholesale price, its expected profit increases as the transfer price t increases.*

It is worth noting that under inventory sharing, one can show that a revenue-sharing contract can coordinate the supply chain when w and ψ are jointly decided by the manufacturer. However, as we fix ψ in our experiments for consistency with the main experiments, the manufacturer's profit-maximizing wholesale price will not perfectly coordinate the supply chain.

EC.4.2. Experimental Design & Results

In this new experiment we refer to DR-M and DR-R under a revenue-sharing contract as DRRS-M and DRRS-R. All experimental protocols are identical to the original design in Section 3. We set the revenue-sharing ratio ψ to 0.7, which generates retailer profit predictions that are close to the normative predictions without revenue sharing. All normative predictions for the revenue-sharing treatments are shown in Table EC.4.1. Each treatment consisted of 42 participants from the same participant pool as our main experiments. Average payment per participant was \$26.94.

Table EC.4.1 Normative Predictions under Revenue Sharing with Decentralized Retailers

Treatment	DRRS-M	DRRS-R
Transfer Price t	30.00	0.00
Wholesale Price w	14.00	10.82
Stocking Quantity q	47.14	40.34
Manufacturer's Profit $\tilde{\pi}_m$	1571.35	1116.90
Retailer's Profit $\tilde{\pi}_r$	183.32	318.80
Supply Chain Efficiency	88.39%	80.02%

Note: (1) transfer prices continue to be predicted at extreme values, (2) wholesale prices are now both predicted to be below the potential anchor points of $(p+c)/2 = 17.5$ and $p/2 = 15$, and (3) stocking quantities are both predicted to be below individual retailer mean demand of 50.

To highlight some of the predictions, first, as in our original experiment the optimal transfer price in DRRS-M is 30 and in DRRS-R is 0. Second, wholesale prices are predicted to be 14.00 and 10.82 (formerly 21.67 and 18.33), allowing us to determine whether wholesale prices are being offered in line with fairness, or, whether they are simply anchored on midpoints $(p+c)/2 = 17.5$ or $p/2 = 15$. Third, optimal stocking quantities are still predicted to be below 50. Fourth, manufacturers in both treatments should continue to earn significantly higher profits than retailers. Fifth, DRRS-M and DRRS-R are predicted to generate higher efficiency than the original DR-M and DR-R treatments. But, as noted above, these efficiencies are below 100% due to self-interested manufacturers: 88.39% and 80.02% (formerly 83.60% and 69.26%).

We provide contract terms in Table EC.4.2. Comparing the observed decisions and normative predictions to one another (significance given in the middle two columns), we see the same patterns as in our original experiment. In particular, transfer prices neglect to be set at extreme values (both $p < 0.004$, the corrected critical p-value) and wholesale prices are set lower than predicted (both $p < 0.004$). This latter finding suggests that wholesale prices being offered more generously appears to be a robust result, as opposed to an anchoring bias. Also, quantities are set rather well, albeit slightly low.

The left-hand side of Table EC.4.3 presents profits and efficiency in our original treatments, whereas the right-hand side depicts the new revenue-sharing treatment data. Beginning with each

Table EC.4.2 Average Contract Prices, Quantities, and Predictions under Revenue Sharing

	Observed Results		Normative Predictions		Behavioral Predictions	
	DRRS-M	DRRS-R	DRRS-M	DRRS-R	DRRS-M	DRRS-R
Transfer Price	19.85 (0.99)	9.50 (0.71)	30.00†	0.00†	19.85	9.25
Wholesale Price	11.91 (0.37)	10.87 (0.29)	12.85† (0.11)	11.74† (0.07)	11.85	10.86
Stocking Quantity	45.64 (1.27)	44.68 (1.05)	47.19 (0.71)	45.30 (0.42)	45.75	44.06

Note: Standard errors, across participants, reported in parentheses. Results for DRRS-R are conditioning on agreement. Normative predictions are conditioning on previous decisions. Behavioral predictions are derived from fitting treatment level data with the full-fairness model. Significance of regressions comparing observed versus normative results given by † $p < 0.004$ (the corrected critical p-value).

party's profits in Table EC.4.3, while manufacturers benefit from revenue sharing, retailers are actually worse off. Regarding efficiency, only DRRS-R achieves a higher efficiency than without revenue sharing, 81.94% versus 76.54% originally ($p < 0.008$, the corrected critical p-value). Combining these observations we have:

Result EC.1. *Under a decentralized retailer inventory-sharing strategy where the retailers set the transfer price, revenue sharing improves supply chain efficiency over wholesale price contracts, but at the expense of retailers.*

Table EC.4.3 Average Profits and Efficiency between the Original Experiment and Revenue-Sharing Experiment

	Original Experiment		Revenue Sharing	
	DR-M	DR-R	DRRS-M	DRRS-R
Manufacturer Profit	1097.60 (22.35)	888.79 (20.68)	1270.99† (37.78)	1150.73† (21.23)
Retailer Profit	332.09 (7.40)	394.68 (7.74)	276.84† (13.76)	322.90† (6.07)
Supply Chain Efficiency (%)	80.36% (0.006)	76.54% (0.007)	83.23% (0.007)	81.94%† (0.007)

Note: Standard errors, across participants, reported in parentheses. Significance of regressions versus original experiment results given by † $p < 0.008$ (the corrected critical p-value). Revenue sharing leads to higher efficiency in DRRS-R, although it comes at the expense of retailer profits.

Returning to Table EC.4.2, the final columns depict treatment-specific predicted decisions from the full-fairness model presented in Section 6. Comparing observed decisions to the fairness predictions, there are no significant differences. The observed transfer price in DRRS-M is the same as the prediction, both 19.85, and is very close in DRRS-R, 9.50 versus 9.25. Wholesale prices are

11.91 versus 11.85 in DRRS-M and 10.87 versus 10.86 in DRRS-R. Regarding stocking quantities, they are 45.64 versus 45.75 in DRRS-M and 44.68 versus 44.06 in DRRS-R.

EC.4.3. Proofs

Proof of Proposition EC.1. To prove there is a unique Nash equilibrium, we have to show the slope of the reaction function is monotonic with an absolute value less than 1. Recall that $f^c(d_i, d_j)$ is the joint probability density function of demands. First define the following marginal probabilities:

$$b_{ij}^1 = \int_0^{q_i} f^c(d_i, q_i + q_j - d_i)\,\mathrm{d}d_i, \qquad b_{ij}^2 = \int_{q_j}^{\infty} f^c(q_i, d_j)\,\mathrm{d}d_j,$$

$$g_{ij}^1 = \int_0^{q_i+q_j} f^c(d_i, q_i + q_j - d_i)\,\mathrm{d}d_i, \qquad g_{ij}^2 = \int_0^{q_j} f^c(q_i, d_j)\,\mathrm{d}d_j.$$

Implicit differentiation of Equation (EC.4.2) leads to

$$\frac{\partial \tilde{q}_i^d}{\partial \tilde{q}_j^d} = -\frac{tb_{ij}^1 + (p-t)g_{ij}^1}{p(b_{ij}^2 + g_{ij}^2) + t(b_{ij}^1 - b_{ij}^2) + (p-t)(g_{ij}^1 - g_{ij}^2)}. \tag{EC.4.4}$$

Equation (EC.4.4) is a special case of Equation (11) in Rudi et al. (2001), which has been shown that the slope of the reaction function is non-positive and less than 1 in absolute value. □

Proof of Lemma EC.1. The proof follows the proof of Lemma 1 in Shao et al. (2011). At equilibrium, the Implicit Function Theorem and the symmetry of retailer i, j lead to

$$\frac{\partial \tilde{q}_i^d}{\partial t} = \frac{(\partial^2 \tilde{\pi}_{r,i}^d / \partial q_i \partial t)[(\partial^2 \tilde{\pi}_{r,i}^d / \partial q_i \partial q_j) - (\partial^2 \tilde{\pi}_{r,i}^d / \partial^2 q_i^2)]}{|H|}, \tag{EC.4.5}$$

where $|H|$ is the positive determinant of the Hessian matrix. For the numerator of Equation (EC.4.5), recall the notations of marginal probabilities in proof of Proposition EC.1, and we have

$$\frac{\partial^2 \tilde{\pi}_{r,i}^d}{\partial q_i \partial t} = \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) > 0, \tag{EC.4.6}$$

$$\frac{\partial^2 \tilde{\pi}_{r,i}^d}{\partial q_i \partial q_j} - \frac{\partial^2 \tilde{\pi}_{r,i}^d}{\partial^2 q_i^2} = pf(q_i) - tb_{ij}^2 - (p-t)g_{ij}^2 > 0, \tag{EC.4.7}$$

where $f(q_i)$ is the probability density function of retailer i's demand. Inequality (EC.4.7) is derived by $p \geq t$, $f(q_i) > b_{ij}^2$ and $f(q_i) > g_{ij}^2$. Therefore, the numerator of Equation (EC.4.5) is greater than 0, which means $\partial \tilde{q}_i^d / \partial t > 0$ at equilibrium. □

Proof of Lemma EC.2 The derivative of manufacturer expected profit function (EC.4.3) with respect to transfer price t is:

$$\frac{\mathrm{d}\tilde{\pi}_m}{\mathrm{d}t} = \frac{\partial \tilde{\pi}_m}{\partial t} + \frac{\partial \tilde{\pi}_m}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial \tilde{\pi}_m}{\partial \tilde{q}^d}\frac{\partial \tilde{q}^d}{\partial t}. \tag{EC.4.8}$$

Note that in Equation (EC.4.8) the subscript of retailer is dropped as quantities are equal at equilibrium. The first term of the right hand side is zero as t is cancelled out at equilibrium. The second term is also zero when w is optimal. Equation (EC.4.8) can be rewritten as

$$\frac{\mathrm{d}\tilde{\pi}_m}{\mathrm{d}t} = 2(w-c)\frac{\partial \tilde{q}^d}{\partial t}. \tag{EC.4.9}$$

From Lemma EC.1 we know that $\partial \tilde{q}^d/\partial t > 0$. Because $w > c$, we have $\mathrm{d}\tilde{\pi}_m/\mathrm{d}t > 0$. $\square$

EC.5 Experimental Details

EC.5.1. Details for Online Experimental Sessions

All online synchronous experimental sessions (one session of each in CR, DR-M, and DR-R, and all sessions in Baseline and DR-0) adhered to the following procedures. A participant logs in to a Zoom session and participates in the study in oTree through an internet browser. All participants are required to play on a desktop or laptop computer with a working camera.

1. Each participant is assigned a random 4-digit label prior to a session. This label acts as a participant's Zoom name to ensure anonymity.
2. Participants arriving at the Zoom session first wait in the waiting room. One participant is admitted at a time and verified through a photo ID.
3. Admitted and verified participants wait in a Zoom breakout room until the session begins.
4. Each participant accesses oTree via a unique link sent only to them.
5. After all participants arrive and are checked in, a researcher reads through the instructions and answers any questions in Zoom.
6. The game begins. To ensure that participants cannot see each other during this time, each participant is assigned to an individual Zoom breakout room with a researcher who will answer any questions. To be clear, this requires the experimenter to log into multiple browsers and computers, in order to be "alone" with each individual in each breakout room.
7. During the session, participants are required to be muted and have their cameras turned on.
8. After the session, participants are paid via PayPal.

EC.5.2. Detailed Results for Online Experimental Sessions

In Table EC.5.1 we provide a summary of decisions between in-person versus virtual, along with regressions comparing the two with corrected p-values.

Table EC.5.1 Average Contract Prices and Quantities In-person versus Online

	In-person			Online		
	CR [42]	DR-M [42]	DR-R [42]	CR [15]	DR-M [18]	DR-R [18]
Transfer Price	-	20.86 (0.83)	$7.58^{\dagger}$ (0.53)	-	18.85 (1.33)	3.21 (0.52)
Wholesale Price	19.20 (0.30)	18.27 (0.36)	16.90 (0.33)	19.03 (0.60)	18.71 (0.52)	16.48 (0.74)
Stocking Quantity	$39.89^{\dagger}$ (0.62)	43.18 (1.35)	40.92 (1.11)	34.12 (1.40)	42.11 (1.02)	37.20 (1.43)

Note: Number of participants reported in square brackets. Standard errors, across participants, reported in parentheses. Significance of regressions with random effects given by † $p < 0.00625$ (the corrected p-value).

EC.5.3. Sample Experiment Instructions in treatment DR-R

Instructions

You are about to participate in a decision making experiment. If you follow these instructions carefully and make good decisions, you will earn money that will be paid to you in cash at the end of the session. Your earnings will depend on your decisions, the decisions of other participants, and chance. Please do not talk with any other participant, and please do not use any resources outside of those given to you for the duration of the experiment.

Game Overview

This is a three-player game consisting of two retailers and one manufacturer. You will be randomly assigned to one of two roles in each round: one of the two retailers or the manufacturer. A retailer independently purchases units of a product from the manufacturer at a wholesale price per unit, and sells units to customers for $30 per unit (all $ are laboratory dollars). The manufacturer produces units at a cost of $5 per unit. For each retailer, customer demand is randomly and independently determined in each round, from 0 to 100, with each integer in that range equally likely.

After demand is known, if one retailer ("sender") has leftover units and the other retailer ("receiver") has excess demand, leftover units will be automatically transferred to the receiver who will pay the sender a transfer price for each transferred unit. The transferred quantity is the lower number between leftover units and excess demand. The sender cannot make extra money from any remaining units after an inventory transfer.

Timeline of the Game

You will play in 12 rounds. Each round has 3 stages. Decisions at each stage are specified as follows.

a) If you are manufacturer:

Stage 1: Wait for retailers to decide an inventory transfer price.

Stage 2: Set a wholesale price.

Stage 3: Wait for retailers to decide their stocking quantities.

b) If you are retailer:

Stage 1: Decide jointly an inventory transfer price with the other retailer through a 2-minute negotiation.

Stage 2: Wait for the manufacturer to decide a wholesale price.

Stage 3: Decide independently your own stocking quantity. Note that if you are not satisfied with the wholesale price set by the manufacturer, you can set your stocking quantity as 0. In this case, you earn $0 and the manufacturer earns $0 from you but may earn profit from the other retailer.

Inventory Transfer Price Negotiation between Retailers

During the retailers' negotiation over the inventory transfer price, you can send offers ranging from \$0 to \$30 as well as receive offers from the other retailer. For a received offer, you can decide whether to accept or decline it. The negotiation ends immediately once any offer is accepted, and the accepted inventory transfer price will apply if an inventory transfer happens later in that round. If an agreement has not been reached when time ends, there will be additional 10 seconds for each retailer to accept the last offer proposed by the other retailer. Note that if both retailers accept the other's final offer, the first accepted offer will be the inventory transfer price in that round. If no agreement is reached after that, the game will continue without any inventory transfer. In this case, each retailer only sells units to its own market.

Decision Support

At each stage, there will be a testing section and decision-making section, such that you can test your decisions before submission. Slide the scroll bar(s) and you will see average profits of all parties. Note that all the average profits are calculated assuming that any following player makes optimal decisions to maximize their average profits. For example, in inventory transfer price testing, average profits are calculated by assuming the manufacturer sets the wholesale price to maximize its average profit, and retailers choose the stocking quantities to maximize their average profits. In stocking quantity testing, initially you can test your own stocking quantity and assume the other retailer stocks optimally responding to your quantity. However, you can also override this and set the other retailer's quantity by unchecking the checkbox. Screenshots of the 3 stages are shown below.

Profit Calculations

The profit equations are as follows:

$$\begin{aligned}
\textit{Retailer profit} &= \$30 \times \textit{Units Sold} - \textit{Wholesale Price} \times \textit{Stocking Quantity} \\
&\quad + (\$30 - \textit{Inventory Transfer Price}) \times \textit{Units Received} \\
&\quad + \textit{Inventory Transfer Price} \times \textit{Units Sent} \\
\textit{Manufacturer profit} &= (\textit{Wholesale Price} - \$5) \times \textit{Units Purchased by Both Retailers}
\end{aligned}$$

"Units Sold" equals the lower number between realized demand and the stocking quantity. "Units Sent" and "Units Received" equals the lower number between leftover units and excess demand, and will be 0 if no transfer happens or no agreement about the inventory transfer price is made.

Results

After 3 stages, demand will be revealed and inventory transfer automatically happens, if applicable. Then you will see all information of that round in the result page, including your profit and other parties' profits.

This concludes one round. In total there will be 12 rounds. At the beginning of each round, you will be randomly re-matched with two other participants and randomly assigned a role. Note also that customer demand in one round is completely independent from customer demand in any other round.

Example

These numbers are simply used to illustrate the sequence of decisions and should not be construed as "good" or "bad" contract terms or stocking decisions.

Decisions:

1. Retailers agree to an inventory transfer price of \$20.00.
2. The manufacturer sets the wholesale price to be \$15.00.
3. Retailer 1 chooses to stock 67 units. Retailer 2 chooses to stock 42 units.

Outcomes:

1. Demand is realized. Demand for Retailer 1 is 52 units. Demand for Retailer 2 is 49 units.
2. Initial sales occur. Retailer 1 sells 52 units to her market at \$30 per unit and has 15 leftover units. Retailer 2 sells 42 units to his market and has 7 units of excess demand.
3. Inventory transfer occurs. Retailer 1 transfers 7 units to Retailer 2 at \$20 per unit.

$$\text{Retailer 1's profit: } \$30 \times 52 + \$20 \times 7 - \$15 \times 67 = \$695$$

$$\text{Retailer 2's profit: } \$30 \times 42 + (\$30 - \$20) \times 7 - \$15 \times 42 = \$700$$

$$\text{Manufacturer's profit: } (\$15 - \$5) \times (67 + 42) = \$1090$$

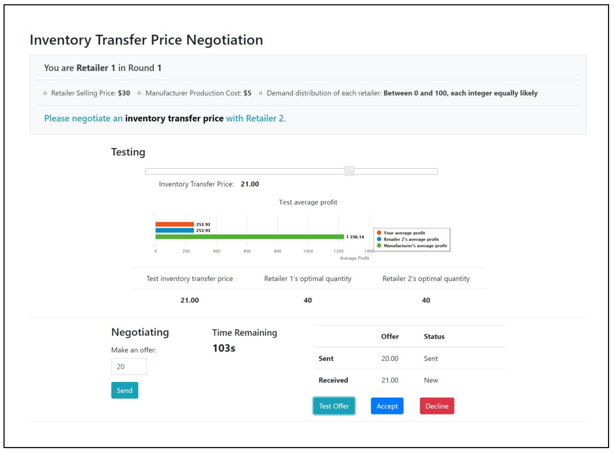
Payment

At the end of the session the actual earnings from the game will be converted to US dollars at the rate of 370 laboratory dollars for \$1 US dollar. These profits will be added to your \$7 show-up fee, displayed on your screen, and paid to you in cash at the end of the session.

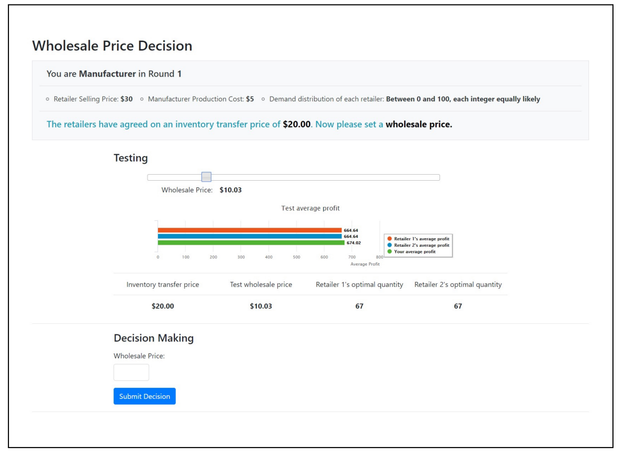
Note: *Instructions above are for in-person sessions. Instructions for online sessions are slightly adapted for online environment (e.g., payment methods).*

Figure EC.5.1 Screenshots of Experiment Interface

(a) Stage 1: Inventory transfer price negotiation



(b) Stage 2: Wholesale price decision



(c) Stage 3: Stocking quantity decision