Managing Multilocation Demand in 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

Demand aggregation is a common method for managing random demand in a multilocation setting. In this paper, we experimentally investigate two alternative demand aggregation strategies in a 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 set the inventory transfer price. One key result is that the decentralized retailer demand aggregation strategy, when the manufacturer sets the transfer price, produces a win-win outcome over the centralized retailer strategy: both the manufacturer and retailers earn higher profits. We also find that the decentralized retailer strategy, when the retailers set the transfer price, outperforms the normative theory and leads to the most equitable distribution of profits. In an effort to account for these results we develop a simple behavioral model and show that it captures decisions well. We then test this model in an out-of-sample experiment, which considers revenue-sharing contracts with demand aggregation, and find that it accurately predicts decisions.

Key words: Demand Aggregation, Risk Pooling, Behavioral Operations, Supply Chain Contracting

History: November 18, 2019

1. Introduction

Managing random demand is a challenging problem for supply chains. Even in a two-stage supply chain between a single upstream manufacturer and single downstream retailer, companies must resort to relatively complicated solutions to address it, such as adopting coordinating contracts (Cachon [2003]). However, in a two-stage supply chain with multiple retailers, companies have another lever at their disposal for managing the supply-demand mismatch problem: demand aggregation.

Demand aggregation strategies may differ as to how they handle random demand. In one case, retailers may opt to act as if they are effectively centralized. They then use a single inventory to satisfy joint demand and maximize combined retailer expected profits (e.g., Eppen [1979]). This
centralized retailer strategy represents a traditional method of risk pooling in operations management where companies may utilize, for example, a single distribution center to fulfill multiple retailers’ demands. In another demand aggregation strategy, retailers are decentralized and act independently when setting their initial stocking quantities. However, after each retailer faces its own demand, they then have the ability to transfer inventory between each other at a transfer price per unit (e.g., Robinson 1990). This decentralized retailer demand aggregation strategy is used in various industries, such as automobile (Zhao et al. 2005), steel (Robinson 1990), commodities (Park et al. 2016), and fashion (Dong and Rudi 2004).

While demand aggregation can benefit the supply chain by increasing overall service levels and improving channel coordination, it may not be the case that all members of the supply chain benefit from a particular strategy. In a two-stage supply chain, this is largely due to an upstream manufacturer having the ability to set the wholesale price for downstream retailers and, in the decentralized retailer aggregation strategy, different parties having decision rights over the inventory transfer price. For example, regarding the former, much of the existing literature focuses on retailers without considering an upstream manufacturer who endogenously sets the wholesale price (Rudi et al. 2001, Hu et al. 2007, Li and Chen 2019, etc). For the latter, Shao et al. (2011) show that under the decentralized retailer aggregation strategy, decision rights over the transfer price can significantly alter the distribution of expected profits in the supply chain.

In this paper, we investigate these two common demand aggregation strategies for managing multilocation retailer demand. In both cases we consider a two-stage supply chain where an upstream manufacturer endogenously proposes a wholesale price to two downstream retailers. Further, because supply chain contract decisions are made by managers, we employ a behavioral approach and address the following research questions. First, how do the centralized retailer and decentralized retailer demand aggregation strategies compare in terms of supply chain efficiency and distribution of profits? Second, under the decentralized retailer strategy, how are supply chain outcomes affected when the manufacturer, versus the retailers, sets the transfer price? Third, if contract decisions deviate from the normative theoretical predictions, is there a behavioral model that can better capture such decisions? Lastly, after identifying which demand aggregation performs best in our initial experiment, how else can we further improve supply chain performance?

We begin our study by outlining the normative theory for our setting. These include theoretical derivations for optimal quantities, wholesale prices, and transfer prices, and resulting expected profits. We then conduct a controlled human-subjects experiment to test these predictions and a set of behavioral hypotheses. Our experiment consists of three treatments: (1) a centralized retailer aggregation setting, (2) a decentralized retailer aggregation setting where the manufacturer sets
the transfer price, and (3) a decentralized retailer aggregation setting where the retailers set the transfer price.

Our experimental results indicate that the two decentralized retailer demand aggregation strategies yield higher supply chain efficiency compared to the classic centralized retailer setting. More importantly, we find that the decentralized case, when the manufacturer sets the transfer price, generates a win-win outcome compared to centralized retailer strategy: both the manufacturer and retailers earn higher expected profits. Further, in the decentralized retailer setting when the retailers set the transfer price, expected profits are higher than the normative theory for both parties and also leads to the most equitable outcome in terms of distribution of profits. Interestingly, a combination of these results provides initial evidence that the classic risk pooling strategy of centralizing retailers fails to generate the highest supply chain efficiency, highest manufacturer profit, highest retailer profit, or the most equitable split of profits.

To determine the driver of these results, we proceed to analyze contract decisions and observe that wholesale prices are set too low in all three demand aggregation environments. We also find that transfer prices are not set at the theoretically predicted extreme values. As a result of these deviations, we develop a behavioral model of fairness-minded retailers and bounded rationality and find that it captures contract decisions well.

Lastly, given the favorable performance of both decentralized retailer demand aggregation strategies, we attempt to answer our final research question by investigating these settings with a revenue-sharing contract. To our knowledge, there is no existing work on coordinating contracts in a decentralized retailer demand aggregation environment. Therefore, we first review the theoretical details of such a setting. We then conduct an additional experiment and observe that revenue-sharing contracts can achieve even higher levels of supply chain efficiency, although, these gains come at the expense of retailers. We also use this new experiment to conduct an out-of-sample test of our behavioral model. In particular, using the estimated behavioral parameters from our initial experiment we generate a set of predictions for our revenue-sharing treatments and show that it accurately predicts contract decisions in this alternative environment.

2. Related Literature
Demand aggregation and risk pooling have been investigated both theoretically and behaviorally in the operations management literature. Here we review this literature beginning with the theoretical research on the centralized retailer and decentralized retailer aggregation strategies, and then proceed to detail related behavioral studies. Lastly, because our paper focuses on contract decisions in a two-stage supply chain, we also highlight the behavioral literature on supply chain contracting.

In terms of theoretical work on the centralized retailer demand aggregation strategy, it is well-known that a single joint inventory between retailers can increase retailer profits compared to a
setting where each retailer acts independently (Eppen 1979). Moreover, the centralized retailer strategy has also been shown to benefit the manufacturer in certain cases (Anupindi and Bassok 1999; Ozer 2003). There is also a rich theoretical literature on the decentralized retailer aggregation strategy through transshipment. In particular, Robinson (1990) investigates the optimal ordering policy in a multi-period setting and shows that transshipment is beneficial to retailers even when there is a transshipment cost. Rudi et al. (2001) compare ordering decisions by retailers under both centralized and decentralized retailer strategies in an asymmetric two-location system. They also consider how the transfer price affects both retailers and find that a coordinating transfer price exists. However, Hu et al. (2007) provide a counterexample and show that the coordinating transfer price only exists conditionally.

Additional theoretical research has extended this stream of literature by evaluating a two-stage supply chain with endogenous wholesale prices, which is the setting we study. For instance, Dong and Rudi (2004) consider the decentralized retailer strategy with endogenous wholesale prices and find that manufacturers generally benefit from transshipment because retailers are less sensitive to wholesale prices under transshipment, which also decreases retailer profits. A more general version of their model is then proposed by Zhang (2005). A theoretical paper especially relevant to our study is Shao et al. (2011). They extend the work of Rudi et al. (2001) by investigating a two-stage supply chain with endogenous wholesale prices and directly compare the centralized and decentralized retailer strategies. For a more extensive literature review on decentralized retailers with transshipment, please see Paterson et al. (2011).

Because supply chain decisions often involve human managers, recent papers have begun to investigate demand aggregation from a behavioral standpoint. Regarding the centralized retailer case, Ho et al. (2010) conduct a human-subjects experiment and investigate stocking quantity decisions. Similar to past newsvendor experiments with a single retailer, they observe a pull-to-center effect, which can negate the risk-pooling benefit when demands are highly correlated. In terms of how wholesale prices are set, we are unaware of any behavioral studies that evaluate contracting in a two-stage supply chain with centralized retailer demand aggregation.

There are also a select number of behavioral papers which investigate the decentralized retailer strategy of transshipment. Complementing our study on contract decisions, these papers focus exclusively on retailer decisions, often order quantities. For instance, Bostian et al. (2012) investigate stocking quantity decisions by decentralized retailers and determine whether inventory transshipment is a behaviorally robust risk pooling strategy. In their experiment both wholesale prices and transfer prices are exogenous. Their results show that average profits are lower than the normative theoretical predictions but higher than the predictions without transshipment. Villa and
Castañeda (2018) focus on quantity decisions with transshipment, also under exogenous wholesale and transfer prices, but compare different retailer interactions. They find that face-to-face interactions can sometimes improve profits. Zhao et al. (2018) examine stocking quantity decisions under decentralized retailers with transshipment, but allow retailers to decide whether to request/fulfill inventory after demand is realized. They find that retailers suffer from a demand side under-weighting bias, i.e., they underestimate the potential profit gains from transshipment and tend to understock.

There are two behavioral papers which examine the decentralized retailer demand aggregation strategy and allow for endogenous transfer prices by retailers. In particular, continuing within a single-stage supply chain, Li and Chen (2019) experimentally test the two-retailer model of Rudi et al. (2001) where both the stocking quantity and transfer price decisions are set by retailers. They develop a 2 × 2 between-subjects experimental design which varies whether transfer prices are set before or after demand is realized and whether transshipment is voluntary or automatic. Similarly, Villa and Katok (2018) consider a setting where each retailer decides its own stocking quantity, but retailers are allowed to negotiate the transfer price. They also manipulate the critical fractile through three different exogenous selling prices. They find that the likelihood of coming to an agreement is higher when transfer prices are set before stocking quantity decisions are made and that transfer prices do not differ across different critical fractiles.

As mentioned previously, our paper differs from the existing behavioral literature by comparing both demand aggregation strategies to one another. In addition, we do so with human-decision makers in a two-stage supply chain where a manufacturer endogenously proposes a wholesale price to downstream retailers. Shifting to a two-stage supply chain where each party is interested in maximizing its own profits is not only a significant difference, but it also means that behavioral studies on supply chain contracting are related to our study. With regards to this literature, Katok and Wu (2009) is one of the first papers to explore how contract terms are set. They compare a wholesale price, buyback, and revenue-sharing contract with automated retailers and observe that the buyback and revenue-sharing contracts increase profits, but not as much as the theory predicts. Kalkanci et al. (2011, 2014) investigate how contract complexity, through increasing the number of wholesale prices with quantity breakpoints, affects decisions in a supply chain experiment. Davis et al. (2014) study wholesale price contracts, while varying where the inventory risk resides in the supply chain, and find that wholesale prices are set slightly more generously than the normative theory predicts. Becker-Peth et al. (2013) examine how to design buyback contracts for irrational newsvendors, and Zhang et al. (2016) compare buyback versus revenue-sharing contracts for suppliers when they can choose between contract types. For a more comprehensive summary of the behavioral supply chain contracting literature please see Chen and Wu (2019).
3. Normative Theory and Behavioral Hypotheses

3.1. Normative Theory

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 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.

We consider three settings which differ in two dimensions: the demand aggregation strategy and transfer pricing power. Specifically, the first dimension varies whether retailers aggregate demand by acting as if they are centralized or decentralized. When retailers are centralized, they share a single stocking quantity to satisfy combined demand and maximize their joint expected profit (i.e., there is no transfer price). We refer to this case as CR hereafter for centralized retailers. When retailers are decentralized, each retailer acts independently, sets its own stocking quantity, and maximizes its own expected profit. After the demand is known, an over-stocking retailer transfers any leftovers to an under-stocking retailer at transfer price $t$ per unit, if possible. For simplicity, $t$ is assumed to be in $[0,p]$. The transferred quantity from retailer $i$ to $j$, $T_i$, is the minimum of $i$’s leftovers and $j$’s excess demand, $T_i = \min((q_i - d_i)^+, (d_j - q_j)^+)$. As with past studies, we assume that the transportation cost of transferred units is zero.

The second dimension we vary is transfer pricing power. Here, we consider two cases: in one, the manufacturer sets the transfer price $t$, referred to as DR-M hereafter 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 transfers. 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 hereafter for decentralized retailers - retailers set the transfer price. In this setting retailers take responsibility over their own inventory transfers.

Following Shao et al. (2011), in a supply chain with decentralized retailers, DR-M and DR-R, we assume that the transfer price is set before the wholesale price and stocking quantity. For example, Narus and Anderson (1996) note that appropriate remuneration is decided in advance when firms agree on sharing resources and capabilities. 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 given the transfer price. In stage 3, the two retailers set stocking quantities. Then demands are realized and inventory transfer happens automatically, if possible. As for CR, there are only two stages. In stage 1, the manufacturer sets
the wholesale price and in stage 2 the two retailers decide a joint quantity. After this demand is realized. Figure 1 summarizes the decision and event sequence in all three settings.

We derive optimal decisions by backward induction starting with symmetric retailer $i$ and $j$’s stocking quantity decisions $(q_i, q_j)$. In the decentralized retailer case, retailer $i$’s expected profit function is given by

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

(1)

In this decentralized retailer setting, for both DR-M and DR-R, Rudi et al. (2001) show that a unique Nash equilibrium exists for asymmetric retailers, which applies to our symmetric case. The equilibrium stocking quantity $(q_{d,i}^*, q_{d,j}^*)$ satisfies Equation (2)

$$
\alpha_i(q_i) - \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) = p - w,
$$

(2)

where $\alpha_i(q_i) = \Pr(d_i < q_i)$, $\beta_i(q_i, q_j) = \Pr(q_i + q_j - d_j < d_i < q_i)$ and $\gamma_i(q_i, q_j) = \Pr(q_i < d_i < q_i + q_j - d_j)$. In words, $\beta_i(q_i, q_j)$ 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 $i$. See Rudi et al. (2001) for a more general solution with a fixed transfer cost and asymmetric retailers.

For centralized retailers, CR, the joint expected profit function is

$$
\pi_{c,r} = E[p \min(d_i + d_j, q_i + q_j)] - w(q_i + q_j),
$$

(3)

and the optimal stocking quantity $(q_{c,i}^*, q_{c,j}^*)$ is derived from Equation (4)

$$
\alpha_i(q_i) - \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) = p - w.
$$

(4)

To be clear, for the centralized retailer case CR, the retailers will set a single joint stocking quantity. In the above equations we simply assume that $q_i$ and $q_j$ are each one half of this joint quantity.

In all three demand aggregation cases, the manufacturer solves $w$ to maximize its profit function $\pi_m = (w - c)(q_i + q_j)$. For the transfer price $t$ with decentralized retailers, Shao et al. (2011) show
that when the manufacturer sets the transfer price, it always sets $t = p$. When retailers set $t$, however, they prefer a relatively low transfer price. Specifically, retailers will set $t = 0$ for high-margin products (across the two-stage supply chain), although it may lie between 0 and $p$ for low-margin products.

Lastly, note that we always assume that the transfer price is set before the wholesale price. While this is observed in practice, we admit that the sequence of these decisions may be reversed. However, as we discuss in the conclusion, if we allow retailers to set the transfer price after observing the manufacturer’s wholesale price, then the profit predictions in DR-R coincide with those of the centralized retailer case CR. Therefore, we do not consider this scenario in our experimental design.

### 3.2. Behavioral Hypotheses

While the normative theory presented above provides a number of testable benchmarks, past experimental results indicate that behavioral biases may lead to potential deviations. Therefore, based on prior behavioral work we now hypothesize specific deviations from the normative predictions and discuss the impact of these deviations on profits. Our three hypotheses revolve around contract price and quantity decisions and are presented in reverse chronological decision order since any deviation in quantity decisions will influence pricing decisions. After stating these hypotheses, we then provide a set of profit implications, assuming these hypotheses are confirmed.

Beginning with stocking quantity decisions, although a pull-to-center bias or understocking bias may be observed (Villa and Katok 2018, Zhao et al. 2018), we intend to mitigate these effects by providing extensive decision support to experimental participants, which will be further explained in the next section. This is due to the fact that our study aims to investigate contract parameters, whereas past studies have explored stocking quantities. Hence we have Hyp-q:

**Hypothesis-q.** Stocking quantities will not be significantly different from the normative predictions in all three settings (CR, DR-M, DR-R).

For wholesale prices, past experimental papers have shown that when one party has the ability to make a one-shot ultimatum wholesale price offer to another party, they may set it slightly more favorably than the normative theory predicts (e.g., Davis et al. 2014). This is especially true when the first mover is predicted to earn a disproportionately high split of overall profits (e.g., Roth et al. 1991). Therefore, we hypothesize that the manufacturer in all three treatments will set wholesale prices below the normative predictions, formalized in Hyp-w:

**Hypothesis-w.** Wholesale prices will be set significantly lower than the normative predictions in all three settings (CR, DR-M, DR-R).
Turning to transfer prices, recall that the normative theory predicts that the manufacturer in DR-M should always set it to the highest value \( t = p \), and retailers in DR-R should set \( t = 0 \) for high-margin products. Given that high-margin products, across entire supply chains, are common in practice we focus on this setting. As a result, the normative predictions of transfer prices in both decentralized settings are extreme. However, the behavioral supply chain literature suggests that these values are unlikely to be observed with human decision makers. For example, past experiments find human decision makers exhibit aversion to extreme predictions even if the predicted strategy is optimal (e.g., Gurnani et al. 2014). In addition, there may be other treatment-specific biases causing similar deviations. For example, in DR-M, although a rational manufacturer will always set the transfer price as high as possible, the positive relationship between the transfer price and the manufacturer’s profit may not be recognized by human decision makers. It is also possible that human decision makers may have other-regarding preferences that drive the transfer price down from \( t = p \), particularly since predicted profits are significantly higher for the manufacturer. In DR-R, the transfer price can be regarded as both a cost and revenue for retailers: a lower transfer price is attractive as a lower cost of obtaining transferred units, but a higher transfer price means more potential profit from sending transferred units. Facing this trade-off, retailers may end up with a greater transfer price.

Hypothesis. In DR-M, the transfer price set by the manufacturer will be significantly lower than the selling price \( p \). In DR-R, the transfer price set by retailers will be significantly greater than 0.

Assuming our behavioral hypotheses are supported, we can determine how they affect efficiency and profits relative to the normative predictions. Also, because these are simply consequences of observed decisions, we do not formalize them into a set of hypotheses and instead present a table of profit implications. To this end, Table summarizes how supply chain efficiency, manufacturer expected profit, and retailer expected profit, will differ with respect to the normative predictions assuming all three of our hypotheses are supported.

To provide more details around Table in CR, the lower wholesale price predicted by Hyp-w will hurt the manufacturer and benefit retailers. As a result, the manufacturer (retailers) will earn less (more) than the normative prediction. Further, since a lower wholesale price pushes the stocking quantity higher towards the first-best quantity (for a fully-integrated supply chain), we expect that supply chain efficiency will increase in CR relative to the normative benchmark.

1 Moreover, the transfer price is jointly set by two retailers in DR-R. Past experimental supply chain papers on bargaining have shown evidence that the midpoint of the contracting space is a focal point (e.g., Davis and Leider 2018, Davis and Hyndman 2019). Because neither retailer will benefit from inventory transfer if they do not come to an agreement, the midpoint \( (p/2) \) may also be an outcome that is more likely to be accepted, giving us more reason to expect that the transfer price in DR-R will be greater than zero.
Table 1  Directional Efficiency and Profit Implications Relative to the Normative Predictions, Given the Behavioral Hypotheses

<table>
<thead>
<tr>
<th></th>
<th>CR</th>
<th>DR-M</th>
<th>DR-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply Chain Efficiency</td>
<td>↗</td>
<td>Depends</td>
<td>↗</td>
</tr>
<tr>
<td>Manufacturer Expected Profit</td>
<td>↘</td>
<td>↘</td>
<td>Depends</td>
</tr>
<tr>
<td>Retailer Expected Profit</td>
<td>↗</td>
<td>↗</td>
<td>Depends</td>
</tr>
</tbody>
</table>

Note: This table shows how observed efficiency and profits should compare to the normative theoretical predictions, assuming all three behavioral hypotheses are supported: quantities are set optimally, wholesale prices are set too low, and transfer prices deviate from the extreme predictions (i.e., less than \( p \) in DR-M and more than 0 in DR-R). In certain cases a clear directional prediction is unavailable due to competing forces.

In DR-M, the manufacturer will suffer from both a lower wholesale price (Hyp-w) and lower transfer price (Hyp-t). This will translate into lower profits for the manufacturer and higher profits for retailers, compared to the normative predictions. Regarding supply chain efficiency, because the lower transfer price reduces stocking quantities but a lower wholesale price increases stocking quantities, the direction of change in efficiency will depend on magnitude of these two deviations.

In DR-R, the manufacturer’s profit decreases from a lower wholesale price (Hyp-w) and increases from a higher transfer price (Hyp-t). The reverse can be said for the retailer’s profit. Because these two effects impact each party’s profit in opposite ways, a clear prediction is not possible with respect to profits. However, supply chain efficiency should be higher because a lower wholesale price and a higher transfer price should both contribute to higher stocking quantities.

4. Experimental Design

Aside from testing our normative predictions and behavioral hypotheses, one of the benefits of conducting a controlled-laboratory experiment is that we can shed light on those profit implication predictions where two forces might be running counter to each other. For example, in DR-M, supply chain efficiency may be higher or lower compared to the normative prediction, depending on the magnitude of the deviations in wholesale prices and transfer prices. In addition, while Table 1 pertains to how we expect efficiency and profits will compare to the normative predictions, determining how each treatment will compare directly to one another will also depend on the size of any observed deviations. Hence experimental data are necessary to help tease out these results.

Our experiment consists of a between-subjects design and three treatments: CR, DR-M and DR-R. All treatments follow the normative theory and decision sequence outlined in subsection 3.1. For CR the transfer price does not apply such that each round begins with the manufacturer setting a wholesale price. After the wholesale price is set retailers set a joint stocking quantity. In

\(^2\) To be clear, the manufacturer proposes a wholesale price as opposed to setting it. We informed retailers that if they wished to reject the manufacturer’s wholesale price offer, to set a stocking quantity of zero.
order to allow a clean comparison between the CR and the DR-M/DR-R treatments (where both retailers make quantity decisions), in CR the two retailers set a joint stocking quantity together through a two-minute negotiation. During this time, 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. A negotiation ends once an offer is accepted and the joint quantity is set for that round. If there is no agreement when time expires 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. After a potential agreement, 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. Following stocking quantities, demand is realized and inventory transfers automatically occur, if applicable. The DR-R treatment is similar to DR-M except each round begins with the two retailers jointly setting the transfer price through a negotiation. This negotiation lasts for up to two minutes and is similar 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 transfer after demand is realized. After the negotiation, the manufacturer sets the wholesale price and then the retailers set their stocking quantities. Finally, demand is realized and inventory transfers automatically occur, if applicable.

In an effort to reduce complexity, we provide an extensive decision support tool for all roles in all stages. The 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 immediately updated on the figure. For the transfer price and wholesale price decisions, expected profits are calculated by assuming players in following stages make optimal decisions (which participants are aware of). For stocking decisions with centralized retailers in CR, there is one bar for the joint quantity, which is initially set at the optimal quantity. In DR-M and DR-R, each retailer has two slide bars for their own and the other retailer’s stocking quantities. A retailer then has two ways to use the tool in DR-M and DR-R, by checking/unchecking a box. The first is that the retailer only tests its own quantity and the other retailer’s quantity will be automatically set as the best response to the tested quantity. The second is that the retailer can manually set the other retailer’s quantity. At the beginning of this decision, the test slide bars are set at the optimal quantities. Please see the online Appendix EC.B for sample instructions and screenshots of our experimental interface.

One might note that our decision support tool for the retailer stocking quantity decisions is relatively strong. This is due to a number of reasons. First, Zhao et al. (2018) conduct interviews with over two dozen managers and find that across all industries explored, inventory decisions are
supported by software where managers have the ability to override the system. Second, past studies have investigated how stocking quantities are set under alternative demand aggregation strategies, whereas our paper focuses more on how contract terms are set. By providing such decision support we give the normative theory a fair chance of being confirmed and avoid any biases in stocking quantity decisions. Third, by simplifying the stocking quantity decision we alleviate any concerns about the stocking quantity being set in two different ways between CR and DR-M/DR-R. Lastly, if we instead automate the quantity decision, this would mean that retailers do not make any decisions in CR and DR-M, but they do in DR-R. This would not only lead to unfair comparisons across treatments, but, it may also overlook any other-regarding preferences. In sum, we opted for human retailers to set stocking quantities with heavy decision support in an effort to not only replicate a realistic environment but to also generate useful results.

For our experimental parameters 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 manufacturer, retailer, and end demand. In other words, unlike related behavioral papers, the retailer’s critical ratio varies because $w$ is endogenously set by the manufacturer.

As discussed above, when retailers are decentralized and the manufacturer sets the transfer price (DR-M), it will always set $t=p$ to incentivize retailers to order a higher quantity, which brings the manufacturer more 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, hence a higher quantity). However,
for our experimental parameters, when retailers are decentralized and set the transfer price (DR-R), \( t=0.50 \) will maximize retailer expected profits. Figure 2 shows how the manufacturer’s and retailer’s expected profits vary with respect to the transfer price.

It is also worth noting that when the transfer price is low, such as in DR-R, retailers will want to set lower stocking quantities (i.e., transfers sent are less valuable and transfers received are more affordable). Because this decreases the manufacturer’s expected profit, the manufacturer then prefers to set a lower wholesale price to incentivize retailers to stock more. This can be seen in Table 2, which presents the normative theoretical predictions for our experimental parameters. Supply chain efficiency is calculated by comparing the sum of the manufacturer and retailers’ expected profits relative to the first-best, fully-integrated supply chain benchmark.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Normative Theoretical Predictions in the Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR</td>
</tr>
<tr>
<td>Supply Chain Efficiency</td>
<td>75.55%</td>
</tr>
<tr>
<td>Manufacturer Expected Profit ( \pi_m )</td>
<td>1242.26</td>
</tr>
<tr>
<td>Retailer Expected Profit ( \pi_r )</td>
<td>207.04</td>
</tr>
<tr>
<td>Transfer Price ( t )</td>
<td>N/A</td>
</tr>
<tr>
<td>Wholesale Price ( w )</td>
<td>21.67</td>
</tr>
<tr>
<td>Stocking Quantity ( q )</td>
<td>37</td>
</tr>
</tbody>
</table>

Note: Transfer price \( t \) is n/a in CR because retailers are centralized. Quantity \( q \) is rounded to an integer value. Also, (1) transfer prices are predicted to be extreme values in both DR-M and DR-R, (2) wholesale prices are all predicted to be above the potential anchor points of \( (p + c)/2 = 17.5 \) and \( p/2 = 15 \), and (3) stocking quantities are all predicted to be below individual retailer mean demand of 50.

A few comments are in order regarding the predicted decisions in Table 2 and how they compare across treatments. First, in the two decentralized retailer conditions, DR-M and DR-R, the predicted transfer price is always extreme, 30 or 0.50. Therefore, neither transfer price prediction is likely to have an unfair advantage of being confirmed by having an interior solution. Second, regarding wholesale prices, the predicted values are always less than two potential anchor points: \( (p + c)/2 = 17.5 \) and \( p/2 = 15 \). In particular, past supply chain bargaining experiments have shown evidence of superficial fairness, where retailer-manufacturer pairs arbitrarily set a wholesale price that is roughly in the middle of the contracting space. Because all predicted wholesale prices are above these anchor points, this potential bias should not account for differences across the three treatments. And third, all stocking quantities are predicted to be below 50, hence any ordering bias, if present, should also neglect to account for differences across treatments.

The optimal transfer price in DR-R is 0.00 in theory, but because quantities are restricted to integers in our experiment, the retailer expected-profit maximizing transfer price is actually 0.50.
Our experiment was implemented through oTree (Chen et al. 2016), an open-source platform for behavioral research. In total 126 participants, including undergraduate and graduate students, were recruited from a large northeast university through an online recruitment system. Each treatment had 42 participants across three sessions. Each session consists of 12 rounds. In each round, participants were randomly assigned a role and matched with two other participants. Before a session started, a researcher read through the instructions and answered any questions. Participants were then required to answer several multiple-choice questions about the game in order to make sure that they fully understood the task. Participants received cash based on earned profits from all rounds in the experiment plus a $7 show-up fee. Average earnings were $25.03 across all three treatments. Each session lasted for 70 minutes on average.

5. Experimental Results
In this section we present our experimental results. We take a top-down approach and begin with results relating to profits in Section 5.1 both between treatments and compared to the normative predictions. To find out what drives these observed results and to formally test our behavioral hypotheses, we then analyze contract decisions in Section 5.2.

In terms of our analysis, there was some moderate learning in the early rounds. We found that dropping the first 25% of decisions removes such learning effects for both manufacturers and retailers. Also, the rate of agreements/acceptances by retailers was high across all 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 94.97% and 96.03%. These near-100% agreement rates are not particularly surprising as they each represent a joint decision rather than a zero-sum negotiation. In addition, the fraction of time retailers accepted the manufacturer’s proposed wholesale price and set a positive stocking quantity was high across all three treatments: 96.83% in CR, 96.83% in DR-M, and 100% in DR-R. Due to these high agreement/acceptance rates across all treatments, we include all data in our analysis (outside of the first 25% of decisions mentioned earlier). Lastly, all hypothesis tests are two-sided t-tests with the individual participant as an independent observation and all regressions are run with random effects.

5.1. Profits
Figure 3 depicts the supply chain expected profit (overall height), manufacturer expected profit (bottom darker portion) and each of the retailer’s expected profits (top light portion), in all three treatments. While we will compare the observed results to the normative theoretical predictions momentarily, we also include the normative predictions for supply chain profits by dashed lines. As

---

4 For example, we did not observe any sort of deadline effect where agreements were made as time expired.
one can note in Figure 3 the two decentralized retailer strategies achieve the highest supply chain profits/efficiency: DR-M and DR-R. While the difference in supply chain profits is not different between DR-M and DR-R, both are significantly higher than the CR treatment (both $p < 0.01$). It is also interesting to note that the supply chain profit in DR-R is higher than the normative prediction ($p < 0.01$). This is consistent with our directional expectations assuming our behavioral hypotheses are supported, which we will explore later. Overall, we have our first main result.

Result 1. Both decentralized retailer demand aggregation strategies, DR-M and DR-R, achieve a higher supply chain profit and efficiency compared to the centralized retailer strategy, CR.

![Figure 3](image)

Note: Total supply chain profit is higher in both DR-M and DR-R compared to CR. DR-M yields a Pareto improvement over CR: both the manufacturer and retailers earn higher profits. DR-R provides the most equitable distribution of profits and generates total supply chain profits that are higher than predicted.

Turning to the distribution of profits in Figure 3 we find that DR-M provides a win-win outcome compared to CR: retailers earn significantly higher profits, 302.53 versus 267.84 ($p < 0.05$), and manufacturers earn higher profits, 1112.52 versus 1031.86 ($p < 0.10$). This leads to our second main result.

Result 2. A demand aggregation strategy with decentralized retailers where the manufacturer sets the transfer price, DR-M, achieves a win-win outcome compared to the standard centralized retailer case, CR: both the manufacturer and retailers earn higher profits.

Another important result we can delineate from Figure 3 is that DR-R generates the most equitable distribution of profits between the manufacturer and retailers, among the three treatments. This leads to our third result:
RESULT 3. A demand aggregation strategy with decentralized retailers 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.

Interestingly, a combination of these three results provides initial evidence that the classic centralized retailer demand aggregation strategy (CR) neglects to perform best in terms of supply chain efficiency, manufacturer profit, retailer profit, and distribution of profits.

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

<table>
<thead>
<tr>
<th></th>
<th>Observed Results</th>
<th>Normative Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR</td>
<td>DR-M</td>
</tr>
<tr>
<td>Supply Chain Efficiency (%)</td>
<td>71.72%(^1)</td>
<td>78.89%(^1)</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Manufacturer Profit</td>
<td>1031.86(^1)</td>
<td>1112.52(^1)</td>
</tr>
<tr>
<td></td>
<td>(34.20)</td>
<td>(30.73)</td>
</tr>
<tr>
<td>Retailer Profit</td>
<td>267.84(^1)</td>
<td>302.53(^1)</td>
</tr>
<tr>
<td></td>
<td>(9.01)</td>
<td>(9.73)</td>
</tr>
</tbody>
</table>

Note: Standard errors of means reported in parentheses across subjects. Significance of \( t \)-tests versus normative predictions given by \(^1\) \( p < 0.01 \), \(^\dagger\) \( p < 0.05 \) and \( ^\ast \) \( p < 0.10 \). There are significant differences between all observed efficiencies and profits compared to the normative benchmarks. DR-R outperforms the normative predictions for both the manufacturer and retailers.

Turning to how efficiency and profits compare to the normative predictions, the left-hand side of Table 3 depicts the observed average expected profits and hypothesis tests versus the normative predictions, the latter of which are illustrated in the right-hand side of the table. As one can see all efficiencies and profits are significantly different from the normative predictions. We can also use this table to determine how efficiency and profits compare to our directional profit predictions assuming that our experimental hypotheses are supported, which were presented in Table 1. In short, most of the directional predictions are confirmed and statistically significant at the 5% level or lower. For example, in CR and DR-M, manufacturers do indeed earn less than the normative predictions and retailers earn more than the normative predictions, which could indicate that other-regarding preferences are present. More importantly, we can also tease out the three cases where there was not a clear prediction due to competing forces: DR-M efficiency, DR-R manufacturer profit, and DR-R retailer profit. In the DR-M efficiency case, it appears that efficiency is lower than the normative prediction. This could imply that the effect of (potentially) lower transfer prices more than offsets any supply chain efficiency benefits from (potentially) lower wholesale prices. In DR-R, it was unclear how both manufacturer and retailer profits would compare to the normative predictions. Interestingly, in our data we observe that both parties earn significantly
higher profits compared to the predictions ($p < 0.01$ for retailers and $p < 0.10$ for the manufacturer).

This brings us to our fourth main experimental result, which we will analyze in more detail in the next subsection:

**Result 4.** A demand aggregation strategy with decentralized retailers where the retailers set the transfer price, DR-R, achieves a win-win outcome compared to the normative theoretical predictions: both the manufacturer and retailers earn higher profits.

### 5.2. Contract Pricing and Quantity Decisions

In order to shed light on the drivers of our profit results we now investigate contract pricing and quantity decisions. To ensure a fair comparison across treatments we condition on both quantity agreement and transfer price agreement. Average observed transfer prices, wholesale prices, and stocking quantities for all three treatments are summarized in the left-hand side of Table 4.

Beginning with transfer prices for the decentralized retailer strategies, DR-M and DR-R, one can see that observed decisions deviate by nearly the same amount in both treatments compared to the normative predictions. In particular, in DR-M manufacturers set the transfer price roughly 8 units lower than the normative prediction, 21.85 versus 30 ($p < 0.01$), whereas in DR-R, the average transfer price is roughly 7 units higher than the normative prediction, 7.27 versus 0.50 ($p < 0.01$). These two observations suggest that any bias influencing transfer prices may be present for both manufacturers in DR-M and retailers in DR-R. Nevertheless, Hyp-t is supported.

| Table 4 Average Contract Prices, Quantities, and Normative Theoretical Predictions |
|---------------------------------|----------------|----------------|----------------|----------------|----------------|
|                                | Observed Results | Normative Predictions |
|                                | CR   | DR-M | DR-R | CR   | DR-M | DR-R |
| Transfer Price                 |      |      |      |      |      |      |
| -                              |      |      |      |      |      |      |
| 21.85‡ (0.90)                  |      |      |      |      |      |      |
| 7.27‡ (0.69)                   |      |      |      |      |      |      |
| Wholesale Price                |      |      |      |      |      |      |
| 19.54‡ (0.29)                  |      |      |      |      |      |      |
| 18.69‡ (0.41)                  |      |      |      |      |      |      |
| 17.06‡ (0.32)                  |      |      |      |      |      |      |
| Stocking Quantity              |      |      |      |      |      |      |
| 38.43‡ (0.74)                  |      |      |      |      |      |      |
| 42.34‡ (1.60)                  |      |      |      |      |      |      |
| 40.18‡ (1.28)                  |      |      |      |      |      |      |
| 41.54 (0.32)                   |      |      |      |      |      |      |
| 44.41 (0.48)                   |      |      |      |      |      |      |
| 40.44 (0.54)                   |      |      |      |      |      |      |

Note: Standard errors of means reported in parentheses across subjects. 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). Significance of t-tests versus conditional normative predictions given by ‡ $p < 0.01$. Transfer prices deviate from the normative predictions and wholesale prices are set universally too low. Stocking quantities are set relatively well, albeit slightly low.

To investigate whether manufacturers and retailers make optimal wholesale price and stocking quantity decisions, the right part of Table 4 presents the normative predictions when conditioning
on any previous decisions. For instance, predicted wholesale prices in DR-M/DR-R are conditioned on transfer prices, and all stocking quantities are conditioned on given wholesale prices (and transfer prices, if applicable). Beginning with wholesale prices, in all three treatments they are lower than the predictions (all $p < 0.01$), which supports Hyp-w and is consistent with other-regarding preferences. As for stocking quantities, there is no significant difference between average decisions and predictions in DR-M and DR-R, although they do appear to be slightly low. This indicates that the decision support provided was somewhat effective in helping subjects stock optimally. However, stocking quantities in CR are lower than the normative prediction. As a result, Hyp-q is partially supported.

5.2.1. Explanation of Main Profit Results
Recall that one of our first key experimental results was that DR-M provides a Pareto improvement over CR. Now that we have reviewed the contract and quantity decisions we can explain this result in more detail. First, according to the normative predictions DR-M should lead to the highest supply chain efficiency. In our data we do observe average stocking quantities that are relatively close to the original normative predictions without any conditioning: 42.34 versus 43.00. Thus, observed efficiency is still rather high in DR-M. Second, observed wholesale prices are set lower than predicted, which contributes to the manufacturer earning less and the retailers earning more than the normative theory predicts. As a consequence of these decisions, DR-M generates a higher efficiency but also a win-win outcome over the CR treatment through a redistribution of profits from the manufacturer to the retailer.

We can also account for our prior results pertaining to DR-R and how it generates profits which are higher than the normative predictions for both the manufacturer and retailers. Specifically, a heterogeneity analysis shows that manufacturers in DR-R who benefit from a relatively higher transfer price tend to be reciprocal to retailers and set lower wholesale prices. This is detailed in Figure 3 which shows wholesale price responses with respect to transfer price decisions in both DR-M and DR-R. The solid black line depicts the normative theoretical relationship between the two prices in both treatments: wholesale prices should increase in observed transfer prices. Indeed, we find this theoretical relationship is confirmed in the DR-M condition (illustrated by the orange dots and fitted line). However, in the DR-R treatment the reverse is observed: higher transfer prices by retailers lead to lower wholesale prices by manufacturers (illustrated by the blue dots and fitted line). To summarize, in DR-R, retailers set transfer prices too high and manufacturers set wholesale prices too low in return. As a result, manufacturers gain more from higher stocking quantities than the cost of lower wholesale prices, and retailers gain more from lower wholesale prices than the cost of higher transfer prices. Eventually, both parties are better off relative to the normative predictions.

5 This explains why the observed efficiency is slightly below the normative benchmark, contrary to our behavioral predictions.
Figure 4 Wholesale Price Responses to Observed Transfer Prices in DR-M and DR-R

Note: Wholesale prices should increase in observed transfer prices in both DR-M and DR-R. Yet, in DR-R, where higher transfer prices are observed relative to the normative theory, manufacturers reciprocate with lower wholesale prices. As a result efficiency increases and both parties earn higher profits, relative to the normative theory.

6. Behavioral Model

Thus far we have observed a number of results which deviate with respect to the normative benchmarks. Here we attempt to address our third research question and develop a behavioral model that is more useful in organizing the data. Specifically, we posit that fairness-minded retailers who are averse to disadvantageous inequality may be effective in capturing decisions. The impetus for this relates to a number of observations. First, manufacturers were predicted to earn a disproportionately large split of total supply chain profits. Second, we saw in our data that manufacturers set wholesale prices universally lower than the normative predictions, which reduces their own profits and provides the retailer with higher profits. And third, retailers erred slightly on the side of understocking relative to the normative benchmarks. Therefore, let decentralized retailer $i$’s utility be given by Equation (5)

$$u_{r,i}^d = \pi_{r,i}^d - \lambda (\pi_m - \pi_{r,i}^d)^+, \quad (5)$$

where $\lambda$ represents retailer $i$’s degree of fairness concerns over disadvantageous inequality. To be clear, unlike the full fairness model that considers two-sided fairness (e.g., Fehr and Schmidt 1999, Cui et al. 2007), or peer-induced fairness (e.g., Beer et al. 2019), we only consider disutility when the manufacturer earns more than each retailer. In our setting, the scenario where the retailer earns more than the manufacturer only happens when the manufacturer sets a very low wholesale price, which we do not observe.
The stocking quantity \((\tilde{q}_i^d, \tilde{q}_j^d)\) for fairness-minded decentralized retailers (DR-M and DR-R) must satisfy Equation (6)

\[
\alpha_i(q_i) - \beta_i(q_i, q_j) \left(\frac{t}{p}\right) + \gamma_i(q_i, q_j) \left(\frac{p-t}{p}\right) = \frac{p-w}{p} - \frac{\lambda}{1+\lambda} \frac{w-c}{p}
\]

if \(\pi_m - \pi_{r,i}^d > 0\),

\[
\alpha_i(q_i) - \beta_i(q_i, q_j) \left(\frac{t}{p}\right) + \gamma_i(q_i, q_j) \left(\frac{p-t}{p}\right) = \frac{p-w}{p}
\]

if \(\pi_m - \pi_{r,i}^d \leq 0\).

Given this behavior by retailers, the manufacturer then sets a wholesale price \(w\) to maximize \(\pi_m = (w - c)(\tilde{q}_i^d + \tilde{q}_j^d)\), where \(\tilde{w}\) represents the optimal wholesale price assuming retailers have fairness concerns. The transfer price \(t\) will be set by the manufacturer or retailers depending on the specific environment. The quantity \((\tilde{q}_i^d, \tilde{q}_j^d)\) and wholesale price \(\tilde{w}\) in the centralized retailer case CR follow similar logic.

While we provide more detailed theoretical results in Appendix B, this model predicts that both stocking quantities and wholesale prices should be lower than the normative predictions. Both of these deviations are consistent with our data. The intuition is that under fairness concerns retailers tend to set lower stocking quantities and the manufacturer prefers to set a lower wholesale price to induce retailers to order a quantity closer to the original normative benchmark. However, this bias cannot account for the observed lower transfer prices in DR-M. The fact that transfer prices were set too low by manufacturers in DR-M and too high by retailers in DR-R suggests that a symmetric bias, unlike fairness, may be at play in both of the decentralized retailer settings. Therefore, to accomplish our goal of developing a model that is useful in capturing the data we assume that transfer prices are set in accordance with a certain degree of random errors. In other words, transfer prices are based on a bounded rationality (BR) model where the transfer price is chosen following a multinomial logit distribution: transfer prices yielding a higher expected utility will be chosen with a higher probability than transfer prices yielding lower expected utility (Su 2008). Let \(\Omega\) represent the decision space of \(t\). Transfer prices in DR-M and DR-R are chosen with probabilities \(\theta^m\) and \(\theta^r\), respectively

\[
\theta^m(t, \beta, \lambda) = \frac{e^{\pi_m(t, \lambda)/\beta}}{\sum_{t \in \Omega} e^{\pi_m(t, \lambda)/\beta}}, \quad \theta^r(t, \beta, \lambda) = \frac{e^{\pi_{r,i}^d(t, \lambda)/\beta}}{\sum_{t \in \Omega} e^{\pi_{r,i}^d(t, \lambda)/\beta}}.
\]

Where \(\beta\) is the degree of rationality: \(\beta \rightarrow 0\) means the decision maker is fully rational and \(\beta \rightarrow \infty\) means they make fully random decisions, i.e., each transfer price will be picked with the same probability. Note that \(\Omega\) in Equation (7) is discrete because in our experiments choices of prices are limited to two digits. To summarize, in our behavioral model retailers’ fairness concerns push down both wholesale prices and stocking quantities in all three treatments, while transfer prices are set according to a bounded rationality model.
6.1. Estimation

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 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 truncated normal distribution

$$
\phi(x; \mu, \sigma, a, b) = \frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)},
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function and cumulative distribution function of the standard normal distribution.

Table 5 shows the estimation results. For each treatment, the left column illustrates our behavioral model and the right column illustrates the normative theory. The fitted bounded rationality parameter is given by $\hat{\beta}$ and the fitted degree of fairness concerns is given by $\hat{\lambda}$. Note that $\hat{\lambda}$ is based on both the manufacturer’s and retailers’ decisions, thus it also captures the manufacturer’s beliefs of the retailers’ fairness concerns.

As one can see in Table 5, in all three treatments the behavioral model provides a significantly better fit over the normative theory, evidenced by the higher log-likelihood (LL) values and significant likelihood-ratio (LR) tests (all three $p < 0.01$). It is worth noting that we also estimate nested models with only fairness or only BR and find that the full model provides the best fit (please see

---

6 See Figure EC.C.4 for wholesale prices and Figure EC.C.5 for stocking quantities in the online Appendix.

7 In our estimations we set $\beta = 1$ for DR-M and DR-R within calculation capacity. This does admittedly affect the LL values, but note that even if the LL values differed, the behavioral model likely yields a significantly better fit across all three treatments.
Table 5 Structural Estimation Results for the Behavioral Model and Normative Theory

<table>
<thead>
<tr>
<th></th>
<th>CR</th>
<th></th>
<th>DR-M</th>
<th></th>
<th>DR-R</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Behavioral</td>
<td>Normative</td>
<td>Behavioral</td>
<td>Normative</td>
<td>Behavioral</td>
<td>Normative</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>-</td>
<td>-</td>
<td>157.573</td>
<td>1</td>
<td>23.012</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(19.664)</td>
<td></td>
<td>(2.691)</td>
<td></td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\lambda}$</td>
<td>0.150</td>
<td>0</td>
<td>0.072</td>
<td>0</td>
<td>0.028</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.015)</td>
<td>(0.009)</td>
<td></td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_w$</td>
<td>2.266</td>
<td>3.180</td>
<td>3.481</td>
<td>3.959</td>
<td>3.036</td>
<td>3.245</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.205)</td>
<td>(0.334)</td>
<td>(0.370)</td>
<td>(0.273)</td>
<td>(0.266)</td>
</tr>
<tr>
<td>$\hat{\sigma}_q$</td>
<td>7.745</td>
<td>8.494</td>
<td>15.526</td>
<td>15.433</td>
<td>10.045</td>
<td>9.869</td>
</tr>
<tr>
<td></td>
<td>(1.190)</td>
<td>(1.391)</td>
<td>(1.060)</td>
<td>(0.939)</td>
<td>(0.593)</td>
<td>(0.564)</td>
</tr>
<tr>
<td>LL</td>
<td>-683.917</td>
<td>-733.518</td>
<td>-2337.021</td>
<td>-22022.646</td>
<td>-2117.383</td>
<td>-3377.572</td>
</tr>
<tr>
<td>LR $p$ $-$ val</td>
<td>$p &lt; 0.01$</td>
<td>-</td>
<td>$p &lt; 0.01$</td>
<td>-</td>
<td>$p &lt; 0.01$</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: Non-agreement data are excluded from the estimation. Standard errors in parentheses are derived from bootstrapping 2000 times. “LL” represents log-likelihoods and “LR” represents likelihood ratio tests. In all treatments the behavioral model provides a more favorable fit over the normative model.

Comparing the magnitude of the specific parameter estimates, it appears that fairness concerns are highest in CR, followed by DR-M, and then DR-R. The low estimates of fairness in DR-R are especially interesting considering that it is the only treatment where the retailer and manufacturer have decision rights over a contract parameter: retailers set the transfer price and the manufacturer sets the wholesale price. Hence, procedural fairness may diminish any retailer fairness concerns with respect to profits.

Table 6 illustrates the predictions for our behavioral model given our estimates $\hat{\beta}$ and $\hat{\lambda}$, along with the observed decisions and normative predictions in our experiment. Wholesale prices are conditioned on transfer prices, if applicable, and stocking quantity predictions are conditioned on wholesale prices and transfer prices, if applicable. As one can see, the predictions in the behavioral model for wholesale prices and transfer prices are quite close to the actual data. For instance, in
CR, the observed versus behavioral prediction is 19.54 versus 19.74 for wholesale prices. In DR-M, observed versus behavioral predictions are 21.85 versus 21.81 for transfer prices and 18.69 versus 19.63 for wholesale prices. For DR-R, observed versus behavioral predictions are 7.27 versus 7.10 for transfer prices, and 17.06 versus 18.59 for wholesale prices. Stocking quantities, however, are not always better predicted with our behavioral model, as seen in DR-M and DR-R. To investigate this further, we conducted another estimation where we separated out the manufacturers’ beliefs about retailers’ fairness concerns (i.e., fit contract decisions) from the retailers’ actual fairness concerns (i.e., fit stocking quantities). The results indicate that manufacturers overestimate retailers’ fairness concerns in all three treatments, which accounts for the behavioral model slightly underestimating quantities. Unsurprisingly, the predicted decisions under this alternative estimation with three parameters (not presented) are quite accurate. Nevertheless, going forward we opt for the behavioral model with only two parameters as it is relatively parsimonious and performs quite well at organizing our data, especially contract parameters, which are the focal point of our study.

7. Extension: Revenue-Sharing with Decentralized Retailers

One of our first experimental results was that the decentralized retailer demand aggregation strategies generate higher efficiencies than that of a centralized retailer setting. Yet, with an observed average efficiency of 78.89% in DR-M and 77.15% in DR-R, there is clearly room for improvement. One possible reason for this is because we have only considered simple wholesale price contracts. Therefore, in an effort to answer our final research question and identify how to improve supply chain performance further, we now investigate a revenue-sharing contract with multiple retailers and demand aggregation.

Because of the favorable performance in the DR-M and DR-R settings, we will focus on the two decentralized retailer scenarios. We assume that the revenue share split is exogenous. This not only coincides with past experiments on coordinating contracts (Katok and Wu 2009) but it also ensures that any new treatments will have the same number of decisions as our original experiments. Overall, this section will allow us to accomplish two goals. First, we can determine if revenue-sharing contracts can experimentally achieve higher supply chain efficiency over wholesale price contracts. And second, we can test our behavioral model in an out-of-sample experiment to determine its effectiveness at predicting decisions in a different contract.

7.1. Normative Theory

Here we briefly summarize the normative theory for a revenue-sharing contract with decentralized retailers, which to our knowledge has not been explored in the literature. Consider the normative
theory outlined previously in Section 3.1. When the manufacturer and two retailers are fully-integrated, the total supply chain expected profit is

$$\Pi = \mathbb{E} [p \min(d_i + d_j, q_i + q_j)] - c(q_i + q_j).$$  \hspace{1cm} (12)$$

By the first order condition, the optimal stocking quantity ($q_i^*, q_j^*$) must satisfy

$$\alpha_i(q_i) - \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) = \frac{p - c}{p}. \hspace{1cm} (13)$$

Let $q_i^*$ serve as a benchmark since it is the first-best stocking quantity for a fully-integrated supply chain (i.e., 100% efficiency). Thus, any contract that can incentivize a retailer to order $q_i^*$ can coordinate the supply chain.

While our attention is on the decentralized retailer case, theoretically, it is useful to review the centralized retailer case with revenue-sharing. Let $\phi$ be the revenue-sharing portion for the retailers. With revenue-sharing the centralized retailer expected profit function is given by Equation (14)

$$\hat{\pi}_c = \phi \mathbb{E} [p \min(d_i + d_j, q_i + q_j)] - w(q_i + q_j), \hspace{1cm} (14)$$

and the optimal order quantities ($\hat{q}_c^i, \hat{q}_c^j$) satisfy Equation (15)

$$\alpha_i(q_i) - \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) = \frac{\phi p - w}{\phi p}. \hspace{1cm} (15)$$

By comparing Equation (13) and Equation (15), one can notice that the supply chain can be fully coordinated, i.e., $(\hat{q}_c^i, \hat{q}_c^j) = (q_i^*, q_j^*)$, by setting $w = \phi c$ with an arbitrary allocation of profits between the two parties. Then the retailers’ joint expected profit becomes $\hat{\pi}_r = \phi \Pi(q_i^*, q_j^*)$ and the manufacturer’s expected profit becomes $\hat{\pi}_m = (1 - \phi) \Pi(q_i^*, q_j^*)$. It is important to recognize that for an exogenous $\phi$, a self-interested expected-profit maximizing manufacturer may not set the wholesale price in a way that leads to full coordination. For instance, in the extreme case of $\phi = 1$, the manufacturer will earn zero profit with $w = c$, which is obviously not in their best interest. In general, this result stems directly from the existing literature on revenue-sharing contracts with a single retailer (Cachon 2003).

For the decentralized retailer case with revenue-sharing, retailer $i$’s profit function is expressed by Equation (16)

$$\hat{\pi}_{d,i} = \phi \mathbb{E} [p \min(d_i, q_i) + tT_i + (p - t)T_j] - wq_i. \hspace{1cm} (16)$$

The optimal stocking quantity under Nash equilibrium $(\hat{q}_{d,i}^i, \hat{q}_{d,j}^j)$ is given by Equation (17)

$$\alpha_i(q_i) - \beta_i(q_i, q_j) \left( \frac{t}{p} \right) + \gamma_i(q_i, q_j) \left( \frac{p - t}{p} \right) = \frac{\phi p - w}{\phi p}, \hspace{1cm} (17)$$

and we have the following proposition (please see Appendix A for all proofs).
**Proposition 1.** There exists a unique Nash equilibrium, \((\hat{q}_{i}^{d}, \hat{q}_{j}^{d})\), in a revenue-sharing contract with decentralized retailer demand aggregation.

To show that the revenue-sharing contract can coordinate the supply chain with decentralized retailers, we need to show that there is a possible set of \((w, t)\) such that retailers will order up to \((q^{*}_i, q^{*}_j)\). First the following lemma is needed.

**Lemma 1.** Under a revenue-sharing contract a decentralized retailer \(i\)'s optimal inventory level, \(\hat{q}_i\), is increasing in transfer price \(t\).

Recall our assumption that transfer price \(t\) lies in the range \([0, p]\), then consider two extreme cases. Equation \((17)\) simplifies to Equation \((18)\) when \(t = 0\), and simplifies to Equation \((19)\) when \(t = p\).

\[
t = 0 \Rightarrow \alpha_i(q_i) + \gamma_i(q_i, q_j) = \frac{\phi p - w}{\phi p}
\]

\[
t = p \Rightarrow \alpha_i(q_i) = \frac{\phi p - w}{\phi p}
\]

By comparing Equation \((15)\) and Equation \((18)\), we know that a decentralized retailer \(i\)'s optimal inventory level \(\hat{q}_{i}^{d}\) is less than \(\hat{q}_{i}^{c}\) when \(t = 0\). Similarly, comparison of Equation \((15)\) and Equation \((19)\) indicates retailer \(i\) orders more than \(\hat{q}_{i}^{c}\) when \(t = p\). Since Lemma 1 shows a positive relationship between \(\hat{q}_{i}^{d}\) and \(t\), there must be a coordinating set of a wholesale price and a transfer price lying between 0 and \(p\). Combining Equation \((13)\) and \((17)\), the coordinating wholesale price \(w^{*}\) and transfer price \(t\) must satisfy:

\[
w^{*} = \phi[c - \beta_i(q_i, q_j)(p - t) + \gamma_i(q_i, q_j)t].
\]

However, for a given \(\phi\), as with revenue-sharing contracts with centralized retailers, the fully coordinating wholesale price may not necessarily be optimal for an expected-profit maximizing manufacturer. The example of \(\phi = 1\) is useful again in illustrating this, along with \(t = 0\). In this case the coordinating wholesale price may actually be less than \(c\). This is because decentralized retailers understock, relative to the centralized case, due to the option of transferring units at a low price. As a consequence, to achieve full coordination, the manufacturer must set a wholesale price lower than \(c\) to induce first-best quantities. Clearly a self-interested manufacturer will not choose such a wholesale price. We provide experimental point predictions in the next subsection which demonstrate this.
7.2. Experimental Design & Predictions

For this new experiment we consider DR-M and DR-R under a revenue-sharing contract, referred as DRRS-M and DRRS-R hereafter. All experimental protocols are identical to the original design in Section 4. As for the revenue-sharing ratio $\phi$, we exogenously set it to 0.7. We chose this value for a number of reasons. For example, when $\phi = 0.7$, retailer expected profits are close to the normative predictions without revenue-sharing: 183.32 in DRRS-M versus 199.23 in DR-M and 317.13 in DRRS-R versus 314.48 in DR-R. In addition, it leads to a set of useful point predictions that allow us to test the robustness of our prior results (which we highlight momentarily). All normative predictions for the revenue-sharing treatments are shown in Table 7.

Table 7 Normative Predictions under Revenue-Sharing with Decentralized Retailers

<table>
<thead>
<tr>
<th>Treatment</th>
<th>DRRS-M</th>
<th>DRRS-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply Chain Efficiency</td>
<td>88.39%</td>
<td>80.26%</td>
</tr>
<tr>
<td>Manufacturer’s Profit $\pi_m$</td>
<td>1571.35</td>
<td>1125.35</td>
</tr>
<tr>
<td>Retailer’s Profit $\pi_r$</td>
<td>183.32</td>
<td>317.13</td>
</tr>
<tr>
<td>Transfer Price $t$</td>
<td>30.00</td>
<td>0.62</td>
</tr>
<tr>
<td>Wholesale Price $w$</td>
<td>14.00</td>
<td>10.87</td>
</tr>
<tr>
<td>Stocking Quantity $q$</td>
<td>47</td>
<td>41</td>
</tr>
</tbody>
</table>

Note: Stocking quantity $q$ is rounded to integer. Also, (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.

Rather than develop a set of formal hypotheses for this new experiment, we highlight some details about the predictions in Table 7 row-by-row. First, DRRS-M and DRRS-R are predicted to generate higher efficiency than the wholesale price contract DR-M and DR-R treatments. But, as noted when outlining the normative theory, these efficiencies are less than 100%: 88.39% in DRRS-M and 80.26% in DRRS-R (formerly 83.60% and 69.57%). Second, manufacturers in both revenue-sharing treatments should continue to earn significantly higher profits than retailers. Although, our main experimental results suggest that this disparity might not be as large as predicted. Third, almost identical to our original experiment the optimal transfer price in DRRS-M is $p = 30$ and in DRRS-R is 0.62 but these extreme values are unlikely to be observed. Fourth, wholesale prices are now predicted to be less than the midpoint of the contracting space, 14.00 in DRRS-M and 10.87 in DRRS-R (formerly 21.67 in DR-M and 18.37 in DR-R). These wholesale price predictions will allow us to tease apart whether wholesale prices are indeed being offered in line with fairness, or, whether they are simply anchored on potential midpoints $(p + c)/2 = 17.5$ or $p/2 = 15$. Fifth, similar to predictions in Table 2, the optimal transfer price in DRRS-R is 0.00 in theory, but the retailer expected-profit maximizing transfer price is actually 0.62 due to the integer constraint of quantities in our experiment.
optimal stocking quantities are still predicted to be below 50 such that any potential ordering bias, if present, is similar across treatments.

As with our main experiment, each of these new treatments consisted of 42 participants including undergraduate and graduate students from the same northeast university. One minor difference compared to our original experiment is that we added one comprehension question to the pre-play quiz which related to revenue-sharing. Otherwise all of the same protocols were followed. Average payment per participant was $26.94 and each session lasted for 75 minutes on average.

7.3. Results

The left-hand side of Table 8 presents efficiency and profits in our original decentralized retailer treatments, whereas the right-hand side depicts efficiency and profits in the new revenue-sharing treatments. In addressing our first objective of this section, DRRS-M does indeed achieve a higher efficiency than without revenue-sharing, 82.37% versus 78.89% originally ($p < 0.05$), but it still falls short of the normative prediction of 88.39%. Similarly, in DRRS-R the observed efficiency increases under revenue-sharing: 79.82% versus 77.15% originally ($p < 0.05$). These results are relatively consistent with earlier experimental supply chain studies with a single retailer, where coordinating contracts outperform wholesale price contracts but not as much as predicted (Katok and Wu 2009, Kalkanci et al. 2014).

Turning to each party’s profits in Table 8, it is curious to see that while manufacturers benefit from revenue-sharing, retailers are actually worse off. In sum, it appears that revenue-sharing with decentralized retailer demand aggregation can further improve supply chain efficiency. However, it comes at the expense of retailers. This leads to our final experimental result:

RESULT 5. Under decentralized retailer demand aggregation, revenue-sharing contracts improve supply chain efficiency over wholesale price contracts. Although, this improvement comes at the expense of retailers, who earn lower profits.

To determine how contract terms and quantities are set with respect to the normative theory, in Table 9 we provide transfer price, wholesale price, and stocking quantity details. Comparing the observed decisions and normative predictions to one another (significance given in middle two columns), we see a number of similar patterns with respect to our original experiment. In particular, transfer prices neglect to be set at extreme values (both $p < 0.01$) and wholesale prices are set lower than predicted ($p < 0.01$ in DRRS-M and $p < 0.05$ in DRRS-R). This latter finding suggests that wholesale prices being offered more generously appears to be a robust result, as opposed to an anchoring bias. Also, stocking quantities continue to be set rather well, albeit slightly low if anything, consistent with our original experiment.
### Table 8  
Average Efficiency and Profits between the Original Experiment and Revenue-Sharing Experiment

<table>
<thead>
<tr>
<th></th>
<th>Original Experiment</th>
<th>Revenue Sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DR-M</td>
<td>DR-R</td>
</tr>
<tr>
<td>Supply Chain</td>
<td>78.89%</td>
<td>77.15%</td>
</tr>
<tr>
<td>Efficiency (%)</td>
<td>(0.011)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Manufacturer</td>
<td>1112.52</td>
<td>937.89</td>
</tr>
<tr>
<td>Profit</td>
<td>(30.73)</td>
<td>(21.15)</td>
</tr>
<tr>
<td>Retailer</td>
<td>302.53</td>
<td>375.06</td>
</tr>
<tr>
<td>Profit</td>
<td>(9.73)</td>
<td>(6.95)</td>
</tr>
</tbody>
</table>

Note: Standard errors of mean reported in parentheses are across subjects. Significance of t-tests versus original experiment results given by † $p < 0.01$ and ‡ $p < 0.05$. Revenue-sharing leads to higher efficiency in both treatments, although it comes at the expense of retailer profits.

The final columns of Table 9 allow us to address our second goal in this section and see how our behavioral model fares in an out-of-sample test. Rather than refit our model to the new revenue-sharing data, we opt for a conservative approach and use our previously estimated behavioral parameters $\hat{\beta}$ and $\hat{\lambda}$ in Table 5 to generate predictions for the revenue-sharing scenario. This allows us to determine whether our original estimates can be used to predict decisions in alternative contract and demand aggregation environments.

### Table 9  
Average Contract Prices, Quantities, and Predictions under Revenue-Sharing

<table>
<thead>
<tr>
<th></th>
<th>Observed Results</th>
<th>Normative Predictions</th>
<th>Behavioral Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DRRS-M</td>
<td>DRRS-R</td>
<td>DRRS-M</td>
</tr>
<tr>
<td>Transfer Price</td>
<td>21.20</td>
<td>9.30</td>
<td>30.00†</td>
</tr>
<tr>
<td>Wholesale Price</td>
<td>12.16</td>
<td>11.02</td>
<td>13.00†</td>
</tr>
<tr>
<td>Stocking Quantity</td>
<td>45.15</td>
<td>43.39</td>
<td>47.74</td>
</tr>
</tbody>
</table>

Note: Standard errors reported in parentheses are across subjects. Results for DRRS-R are conditioning on agreement. Normative predictions are conditioning on previous decisions. Behavioral predictions are derived from our behavioral model with fitted $\hat{\beta}$ and $\hat{\lambda}$ of DR-M and DR-R, respectively. Significance of t-tests versus observed results given by † $p < 0.01$ and ‡ $p < 0.05$. The behavioral predictions fare quite well at predicting contract and stocking quantity decisions in this out-of-sample test.

Comparing observed decisions to the behavioral model’s predictions (significance given in last two columns), the observed average transfer price in DRRS-M is very close to the predicted value, 21.20 versus 21.22, and is significantly higher in DRRS-R, 9.30 versus 5.34 predicted ($p < 0.01$).

---

9 We also ran an estimation using the new revenue-sharing data and generated predictions based on these estimates. As one might expect the predictions are quite accurate. Details are available in online Appendix [ECA].
Wholesale prices are predicted well using the behavioral model with no significant differences: 12.16 versus 12.03 in DRRS-M and 11.02 versus 10.94 in DRRS-R. Regarding stocking quantities, there are no significant differences in the observed values and behavioral predictions as well. These results suggest that our behavioral model is relatively useful in predicting decisions in a related out-of-sample environment.

8. Conclusion

In this paper we study multilocation demand aggregation in a supply chain using a behavioral approach. We consider two demand aggregation strategies. In one, retailers act as if they are effectively centralized and share a joint inventory to satisfy combined demand. In the other, decentralized retailers act independently but may transship inventory between one another, at a transfer price per unit, after demand is realized. Unlike past behavioral papers on demand aggregation, we consider a two-stage supply chain where transfer prices, wholesale prices, and stocking quantities are endogenously determined. Through a controlled laboratory experiment with three treatments, we evaluate these two common demand aggregation strategies with human decision makers. One treatment corresponds to the centralized retailer demand aggregation strategy. The other two consider the decentralized retailer strategy which differ as to whether the manufacturer or retailers have the power to set the transfer price.

In terms of relevance, companies are constantly searching for ways to manage the supply-demand mismatch problem. All three of the demand aggregation strategies we investigate, which are designed to alleviate this problem, are feasible in practice. The centralized retailer strategy is well-known in the operations management community and is often treated as the classic method for pooling demand risk [Eppen1979]. Regarding the decentralized retailer strategy, the situation where the manufacturer has the ability to set the transfer price represents a scenario where a manufacturer opts to facilitate transfers among retailers, as they may have better access and visibility into the network of retailers. The other decentralized retailer strategy, where the retailers set the transfer price, represents a situation where the retailers choose to take responsibility of all aspects of any inventory transfers themselves.

One of our main experimental results is that both decentralized retailer strategies generate a higher supply chain efficiency over the standard centralized retailer strategy. Another key insight indicates that decentralized retailers, where the manufacturer sets the transfer price, leads to a win-win outcome over the centralized retailer strategy: both the manufacturer and retailers earn higher expected profits. A third important result is that the decentralized retailer strategy, when the retailers set the transfer price, leads to the most equitable outcomes where both parties earn a higher expected profit than the normative theory predicts. A combination of these observations
indicates that a centralized retailer strategy neglects to generate the highest supply chain efficiency, manufacturer profit, and retailer profit, nor provides the most equitable distribution of profits.

Our analysis of specific contract decisions demonstrates that transfer price decisions do not conform to the extreme normative predictions and that wholesale prices are set too low in all demand aggregation environments. In an effort to account for these decisions, we proceed to develop a simple behavioral model based on fairness-minded retailers and random errors, and find that it can well explain the observed transfer and wholesale prices in our experiment.

Given the favorable performance of the two decentralized retailer settings in our experiment, we then proceed to investigate how revenue-sharing can further improve supply chain efficiency. Among other results, we find that revenue-sharing contracts lead to higher efficiency, although, this improvement comes at the expense of retailers. In terms of contract decisions, we observe a number of deviations which are consistent with our original experiment. We also find that our behavioral model can predict contract decisions accurately in this out-of-sample revenue-sharing environment.

Our primary managerial insight that the decentralized retailers strategy, where the manufacturer sets the transfer price, outperforms the centralized retailers strategy may appear surprising at first. This is especially true considering that in classical operations management theory, aggregating multilocation demands by using centralized inventory is considered to be a preferred strategy. However, our paper provides initial evidence that a decentralized retailer system can outperform a centralized retailer system. The main reason for this is that, compared to a centralized setting, manufacturers can take advantage of setting high transfer prices to incentivize retailers to stock more, leading to higher manufacturer profits and also supply chain efficiency. On the other hand, simultaneously, they can offer lower wholesale prices which benefit retailers. In the end, both parties are better off compared to centralizing inventory.

There are limitations to our study, some of which we consider as opportunities for future research. First, under the decentralized case where the retailers set the transfer price, we assume that the transfer price decision is made prior to the manufacturer’s wholesale price decision. We opted for this not only because it is common in practice, but also, if the pricing sequence is reversed such that manufacturers set wholesale prices first, the profit predictions match those of the centralized retailer setting. Second, another limitation is that we assume retailer demands are independent. While challenging, exploring alternative demand aggregation strategies when there is correlation among retailer demands could be valuable. Finally, 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 multilocation demand would be another exciting opportunity for future research.
Acknowledgments

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

References


Appendix A: Proofs

Proof of Proposition 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. Let \( f(d_i, d_j) \) be the joint probability density function of demands. First define the following marginal probabilities:

\[
\begin{align*}
    b_{ij}^1 &= \int_0^{q_i} f(d_i, q_i + q_j - d_i) \, dd_i, \\
    b_{ij}^2 &= \int_{q_j}^{\infty} f(q_i, d_j) \, dd_j, \\
    g_{ij}^1 &= \int_0^{q_i+q_j} f(d_i, q_i + q_j - d_i) \, dd_i, \\
    g_{ij}^2 &= \int_0^{q_j} f(q_i, d_j) \, dd_j.
\end{align*}
\]

Implicit differentiation of Equation (17) leads to

\[
\frac{\partial q^d_i}{\partial q^d_j} = -\frac{tb_{ij}^1 + (p-t)g_{ij}^1}{p(b_{ij}^2 + g_{ij}^2) + t(b_{ij}^1 - b_{ij}^2) + (p-t)(g_{ij}^1 - g_{ij}^2)}. \tag{A.1}
\]

Equation (A.1) 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 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

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

where \(|H|\) is the positive determinant of the Hessian matrix. For the numerator of Equation (A.2), recall the notations of marginal probabilities in proof of Proposition 1 and we have

\[
\begin{align*}
    \frac{\partial^2 \hat{\pi}_{r,i}^d}{\partial q_i \partial t} &= \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) > 0, \tag{A.3} \\
    \frac{\partial^2 \hat{\pi}_{r,i}^d}{\partial q_i \partial q_j} - \frac{\partial^2 \hat{\pi}_{r,i}^d}{\partial^2 q_i^2} &= pf(q_i) - tb_{ij}^1 - (p-t)g_{ij}^2 > 0, \tag{A.4}
\end{align*}
\]

where \( f(q_i) \) is the probability distribution function of retailer \( i \)'s demand. Inequality (A.4) is derived by \( p \geq t, f(q_i) > b_{ij}^2 \) and \( f(q_i) > g_{ij}^2 \). Therefore, the numerator of Equation (A.2) is greater than 0, which means \( \partial q^d_i / \partial t > 0 \) at equilibrium. □
Appendix B: Behavioral Model Details

Fairness Model under a Wholesale Price Contract

With fairness concerns, a decentralized retailer $i$ is optimizing
\begin{equation}
\hat{u}_{r,i}^{d} = \hat{\pi}_{r,i}^{d} - \lambda (\hat{\pi}_{m} - \hat{\pi}_{r,i}^{d})^{+},
\end{equation}
where $\lambda$ is the degree of retailer $i$’s fairness concerns.

The optimal quantity $(\hat{q}_{i}^{d}, \hat{q}_{j}^{d})$ at equilibrium is given by
\begin{align}
\alpha_i(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} - \lambda \left( \frac{w-c}{1+\lambda} \right) p \quad \text{if } \pi_m - \pi_{r,i}^{d} > 0, \\
\alpha_i(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} \quad \text{if } \pi_m - \pi_{r,i}^{d} \leq 0.
\end{align}

Similarly, centralized retailers $i$ and $j$ in CR are optimizing the joint utility function
\begin{equation}
\hat{u}_{c}^{*} = \hat{\pi}_{c}^{*} - 2\lambda (\hat{\pi}_{m} - \hat{\pi}_{c}^{*}/2)^{+}.
\end{equation}

Note that the fairness term in utility function (B.3) considers only half of the joint retailer profit because when making decisions, retailers observe their individual expected profit instead of joint profit. Therefore, the disutility term is at individual level and is multiplied by 2 in the joint utility function.

The optimal stocking quantity $(\hat{q}_{i}^{d}, \hat{q}_{j}^{d})$ must satisfy
\begin{align}
\alpha_i(q_i) - \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) \left( \frac{p-t}{p} \right) &= \frac{p-w}{p} - \lambda \left( \frac{w-c}{1+\lambda} \right) p \quad \text{if } \pi_m - \pi_{r,i}^{c}/2 > 0, \\
\alpha_i(q_i) - \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) \left( \frac{p-t}{p} \right) &= \frac{p-w}{p} \quad \text{if } \pi_m - \pi_{r,i}^{c}/2 \leq 0.
\end{align}

Taking retailers’ fairness concern into account, the manufacturer maximizes its expected profit by setting $w$ (and $t$), which is $\pi_m = (w-c)(q_i + q_j)$ in CR, DR-M and DR-R.

Fairness Model under a Revenue-Sharing Contract

Similar to the wholesale price contract with fairness concerns, decentralized retailer $i$ is optimizing
\begin{equation}
\hat{\hat{u}}_{r,i}^{d} = \hat{\hat{\pi}}_{r,i}^{d} - \lambda (\hat{\hat{\pi}}_{m} - \hat{\hat{\pi}}_{r,i}^{d})^{+},
\end{equation}
where the manufacturer expected profit $\hat{\hat{\pi}}_{m}$ is given by
\begin{align}
\hat{\hat{\pi}}_{m} &= (w-c)(q_i + q_j) + (1 - \phi)E[p \min(d_i, q_i) + tT_i + (p-t)T_j] \\
&\quad + (1 - \phi)E[p \min(d_j, q_j) + tT_j + (p-t)T_i].
\end{align}

The optimal quantity $(\hat{q}_{i}^{d}, \hat{q}_{j}^{d})$ at equilibrium is given by
\begin{align}
\alpha_i(q_i) - \beta_i(q_i, q_j) \left( \frac{\hat{\lambda} - \hat{\phi}p}{\hat{\lambda} - \hat{\phi}p} \right) + \gamma_i(q_i, q_j) \left( \frac{\hat{\lambda}(p-t) - \hat{\phi}p}{\hat{\lambda} - \hat{\phi}p} \right) &= \frac{\hat{\lambda}(w-c) - (1+\lambda)w}{\hat{\lambda} - \hat{\phi}p} \quad \text{if } \hat{\hat{\pi}}_{m} - \hat{\hat{\pi}}_{r,i}^{d} > 0, \\
\alpha_i(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{\hat{\phi}p - w}{\hat{\phi}p} \quad \text{if } \hat{\hat{\pi}}_{m} - \hat{\hat{\pi}}_{r,i}^{d} \leq 0,
\end{align}

where $\hat{\lambda} = (1+\lambda)\phi$, and $\hat{\phi} = (1-\phi)\lambda$.

The centralized retailer case follows similar logic, which is not presented here since it is not explored in our revenue-sharing experiment.
Electronic Companion

EC.A Additional Estimation Results

Our behavioral model for DR-M and DR-R includes two components: bounded rationality for transfer prices and fairness concerns for wholesale prices and quantities. In Table EC.A.1 we present intermediate estimation results with bounded rationality only and fairness only. Since the full behavioral model for CR only includes fairness, there is no intermediate model in this case. Comparing Table 5 and Table EC.A.1, the full model achieves the highest log-likelihood, indicating that both bounded rationality and fairness are necessary in our model.

| Table EC.A.1 Structural Estimation for Nested Models in DR-M and DR-R |
|-----------------------------|-----------------------------|
|                             | DR-M |                   | DR-R |                   |
|                             | BR   | Fair              | BR   | Fair              |
| $\hat{\beta}$               | 180.276 | 0                   | 22.360 | 0                   |
|                             | (21.665) | (2.523)            |       |                    |
| $\lambda$                   | 0     | 1.000              | 0     | < 0.001            |
|                             | (< 0.001) |                   |       | (< 0.001)          |
| $\hat{\sigma}_w$            | 3.957 | 4.972              | 3.245 | 3.245              |
|                             | (0.363) | (0.425)            | (0.264) | (0.219)            |
| $\hat{\sigma}_q$            | 15.435 | 45.692             | 9.869 | 9.869              |
|                             | (0.957) | (1.956)            | (0.582) | (1.474)            |
| LL                          | -2353.113 | -7198.022          | -2121.385 | -3377.572         |
| LR $p - val$                | $p < 0.01$ |                   |       | $p < 0.01$          |

Note: “BR” is the model with bounded rationality only, and “Fair” is the model with fairness only. Non-agreement data are excluded from the estimation. Standard errors in parentheses are derived through bootstrapping 2000 times.

Tables EC.A.2 and EC.A.3 give in-sample estimation for the two revenue-sharing treatments. Table EC.A.2 shows the estimated behavioral parameters and log-likelihoods of the full behavioral model, restricted behavioral models, and normative model. Similar to the results in the wholesale price contract treatments, the full behavioral model generates the largest log-likelihood. Table EC.A.3 compares observed decisions and predictions. Predictions include in-sample predictions by the behavioral model, out-of-sample predictions from the wholesale price contract estimation, and the normative predictions.
Table EC.A.2  Structural Estimation for DRRS-M and DRRS-R

<table>
<thead>
<tr>
<th></th>
<th>DRRS-M</th>
<th></th>
<th>DRRS-R</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>BR</td>
<td>Fair</td>
<td>Norm.</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>160.212</td>
<td>165.936</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(24.242)</td>
<td>(24.571)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\lambda}$</td>
<td>0.048</td>
<td>0</td>
<td>1.000</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td></td>
<td>($&lt; 0.001$)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\sigma}_w$</td>
<td>2.433</td>
<td>2.574</td>
<td>6.036</td>
<td>2.574</td>
</tr>
<tr>
<td></td>
<td>(5.528)</td>
<td>(0.177)</td>
<td>(0.535)</td>
<td>(0.176)</td>
</tr>
<tr>
<td></td>
<td>(1.182)</td>
<td>(1.175)</td>
<td>(1.694)</td>
<td>(1.202)</td>
</tr>
<tr>
<td>LL</td>
<td>-2294.571</td>
<td>-2302.951</td>
<td>-14684.198</td>
<td>-19106.775</td>
</tr>
<tr>
<td>LR $p$-val</td>
<td>$p &lt; 0.01$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: “Full” is model with bounded rationality and fairness, “BR” is the model with bounded rationality only, “Fair” is the model with fairness only, and “Norm.” is the normative model. Non-agreement data are excluded from the estimation. Standard errors in parentheses are derived through bootstrapping 2000 times.

Table EC.A.3  Transfer Price, Wholesale Price, and Stocking Quantities: Observed Values and Predictions in Revenue-Sharing

<table>
<thead>
<tr>
<th></th>
<th>DRRS-M</th>
<th></th>
<th>DRRS-R</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer Price</td>
<td>21.20</td>
<td>21.22</td>
<td>21.22</td>
<td>30.00</td>
</tr>
<tr>
<td>Wholesale Price</td>
<td>12.16</td>
<td>12.03</td>
<td>12.33</td>
<td>13.00</td>
</tr>
<tr>
<td>Stocking Quantity</td>
<td>45.15</td>
<td>45.86</td>
<td>45.70</td>
<td>47.74</td>
</tr>
</tbody>
</table>

Note: “Obs.” is observed values (conditioning on agreement), “Pred.” is out-of-sample predictions from DR-M and DR-R, “Beh.” is in-sample behavioral model predictions, and “Norm.” is normative predictions.
EC.B 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:

\[
\text{Retailer profit} = 30 \times \text{Units Sold} - \text{Wholesale Price} \times \text{Stocking Quantity} \\
+ (30 - \text{Inventory Transfer Price}) \times \text{Units Received} \\
+ \text{Inventory Transfer Price} \times \text{Units Sent}
\]

\[
\text{Manufacturer profit} = (\text{Wholesale Price} - 5) \times \text{Units Purchased by Both Retailers}
\]

“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.

Retailer 1’s profit: \(30 \times 52 + 20 \times 7 - 15 \times 67 = 695\)

Retailer 2’s profit: \(30 \times 42 + (30 - 20) \times 7 - 15 \times 42 = 700\)

Manufacturer’s profit: \((15 - 5) \times (67 + 42) = 1090\)

Figure EC.B.1  Screenshots of Experiment Interface
(a) Stage 1: Inventory transfer price negotiation
(b) Stage 2: Wholesale price decision
(c) Stage 3: Stocking quantity decision

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.
EC.C  Heterogeneity Analysis

**Figure EC.C.2  Bargaining Time Distribution in CR and DR-R**

(a) Joint Stoking Quantity in CR  
(b) Transfer Price in DR-R

**Figure EC.C.3  Distribution of Transfer Price**

(a) DR-M  
(b) DR-R  
(c) DRRS-M  
(d) DRRS-R