# EXPERIMENTAL STUDIES ON SUPPLY CHAIN CONTRACTING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Rihuan Huang

August 2022

EXPERIMENTAL STUDIES ON SUPPLY CHAIN CONTRACTING

Rihuan Huang, Ph.D.

Cornell University 2022

Supply chain contracting is a classic topic in operations management. While the traditional literature typically assumes a rational decision-maker, the behavioral operations management literature has shown that human decision-makers may exhibit behavioral biases and deviate from optimality. Given that many operational decisions are made by managers in practice, it is essential to understand the role of behavioral factors. This dissertation experimentally studies three problems in supply chain contracting, each of which is summarized as follows.

The first chapter focuses on inventory sharing, comparing alternative inventory-sharing strategies in a two-tier supply chain with an upstream manufacturer and two downstream retailers. In one setting, retailers act as if they are centralized and use a single quantity to fulfill joint demand. In the other, retailers are decentralized and face separate demands, but they can transfer inventory after demands are realized. In the latter, decentralized scenario, whether the manufacturer or retailers have decision authority over the inventory transfer price is also considered. By conducting lab experiments, this work shows that when the retailers are decentralized and the manufacturer sets the transfer price, both retailers and the manufacturer earn higher profits than in the centralized retailer strategy, which runs counter to theory. This chapter is a joint work with Andrew Davis and Douglas Thomas (University of Virginia).

The second chapter turns to supply chain finance and investigates a trade credit contract between a supplier and a financially distressed retailer. The supplier sets a wholesale price for the retailer, who begins with some initial capital and orders from the supplier. The retailer will purchase through trade credit if the initial capital is insufficient and repay the supplier after demand realization. The retailer will go bankrupt if the realized demand is too low. Experimental results suggest that when retailers' risk is medium or high, they significantly understock to maintain lower risk and suppliers offer lower prices (but not necessarily less risk) to retailers. When retailers' risk is very low, they slightly overstock, which brings them additional bankruptcy risk. This chapter is a joint work with Andrew Davis and Kyle Hyndman (University of Texas at Dallas).

The third chapter studies management of responsible sourcing and reputation risk in a two-tier supply chain consisting of an upstream supplier and a downstream buyer. The supplier faces random cost shock which can be mitigated by exerting effort. The supplier's effort has two types which differ in cost and whether it complies with the buyer's responsible standards. A non-compliant effort will incur extra reputation loss for the buyer. The buyer offers a risk-sharing contract: a wholesale price and sharing some of the cost shock with the supplier. While the theory suggests that the supplier always prefers non-compliant effort, experimental results show that 25.82% of suppliers choose compliant effort. More surprisingly, the effort choice is driven only by the wholesale price but not by the cost-shock sharing, which is inconsistent with the theory. This chapter is a joint work with Andrew Davis and Li Chen.

## BIOGRAPHICAL SKETCH

Rihuan Huang received a bachelor's degree in Logistics Management and a master's degree in Management Science and Engineering from Huazhong University of Science and Technology, and a bachelor's degree in Law (second major) from Wuhan University. He received his second master's degree and his doctoral degree in Operations, Technology, and Information Management from the SC Johnson College of Business at Cornell University.

To my wife, Yulu Tan

## ACKNOWLEDGEMENTS

I am very fortunate to have met many wonderful people during my doctoral studies. It would not have been possible for me to finish this dissertation without their love and support, and I extend my gratitude to them.

Having Professor Andrew Davis as my advisor has been one of the most fortunate things in my life. Each time I seek advice, he teaches me everything he knows without reservation. I can clearly see how he aims to train me to be an independent scholar step by step, by letting me have more control of my research projects along the way. I never feel any unnecessary pressure from him, and he always gives me the luxury of having options.

I would also like to thank my other committee members, Professor Li Chen, Professor Vishal Gaur, and Professor Chris Forman, for their guidance during my doctoral studies. Beside them, I am deeply grateful to my co-author, Professor Douglas Thomas, from the University of Virginia. I would not survive the job market without his help. I am thankful to Professor Lawrence Robinson for his help with my first chapter, and to Professor Yao Cui for those meaningful chats during my first year and before I entered the job market.

During the past five years in Ithaca, I also met many amazing students who have made my life more enjoyable. I will always remember those interesting discussions with Ashish Ochani, Jun Oh, and Anugna Gondi. I am also grateful to my friend Leting Zhang for your support and encouragement. I feel deeply blessed to meet my dear OTIM cohort, Xiaoyue Yan, Wee Kiat Lee, Yu Nu, Lorraine Yuan, Yuan Cheng, Xiaoyan Liu, Zhen Lian, Zhijun Hua and Ziwei Zhu.

You are a group of people who always deliver happiness and inspiration. I am sure we will all keep in touch no matter where we end up.

Finally, I would not have been able to undertake this journey without the love, support, and understanding of my family. I thank my wife, Yulu Tan. I love you. A long-distance relationship with a 12-hour time difference is never easy, and it has been even more difficult during the pandemic. As I write this paragraph, we have been apart for almost two years. I know you have been choosing to show your positive side to me and to hide those negative sides to keep me from worrying. From now on, you no longer need to do that, and please let me do everything I can to support you. I am forever indebted to my mother, Zhixing Chen, for everything you have done. You are always caring and supportive without questioning my decisions. I always believe that we are friends rather than mother and son.

I again thank all of those who guided me, helped me, and inspired me in the past five years. I will always remember and appreciate it.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

CHAPTER 1

# INVENTORY SHARING IN TWO-TIER SUPPLY CHAINS

## 1.1 Introduction

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

Retailer inventory-sharing strategies vary in two key dimensions: whether the retailers operate in a centralized or decentralized manner, and, when retailers are decentralized, whether the retailers or manufacturer have decision authority over the inventory transfer price. First, retailers may opt to act as if they are effectively centralized and use a single inventory to satisfy joint demand [e.g., 47], or they may act independently in a decentralized manner, initially serving their own demand and then transferring inventory at a per unit transfer price [e.g., 107]. This centralized/decentralized attribute may be fixed given the supply chain setting, but frequently it may be an organizational choice. Large firms with multiple retail outlets often act in a centralized way, but many choose to grant store managers significant autonomy such that retail outlets effectively operate as if they are decentralized [e.g., 42, 123]. Decentralized inventory sharing has been observed in several industries including automobile [140], steel [107], commodities [103], and fashion [44]. Conversely, franchise networks, such as several of the largest convenience store chains, are technically independent retailers but could coordinate in a centralized manner through the franchisor. Second, once in possession of inventory, retailers may have the right to transfer or resell to other retailers at terms of their choosing, whereas a manufacturer with a strong brand and control over pricing and distribution may be able to dictate terms of inventory transfer between independent, authorized retailers [e.g., 102].

While inventory sharing is a common way to increase supply chain efficiency, it may not be the case that all members of the supply chain benefit from a particular strategy. In a two-tier supply chain, theory predicts that certain retailer inventory-sharing strategies may actually lead to lower retailer profits compared to a setting without any inventory sharing at all. This is largely due to an upstream manufacturer having the ability to set the wholesale price for downstream retailers, and in the decentralized retailer inventory-sharing strategy, different parties having decision authority over the inventory transfer price [e.g., 114]. In this paper, we investigate how the potential benefits of inventory sharing are distributed depending on the centralized/decentralized retailer structure and the decision authority over the transfer price (manufacturer or retailers). We consider a two-tier supply chain where an upstream manufacturer endogenously proposes a wholesale price to two downstream retailers. Because supply chain contract decisions are made by managers [e.g., 42, 123, 141], we employ a behavioral approach, complementing existing theoretical research to understand how, if at all, firms should adopt an inventory-sharing strategy. Specifically, we address the following research questions. First, how do inventory-sharing strategies under centralized and decentralized retailer settings compare in terms of distribution of profits? Second, in the decentralized retailer setting, how are profits distributed when the manufacturer, versus the retailers, has decision authority over the transfer price?

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

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

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

Our study provides insights for both practitioners and researchers. Regarding the former, we first demonstrate that retailers earn a profit that is higher (or at least as high) under all retailer inventory-sharing strategies compared to no inventory sharing. Second, when choosing among the different retailer inventory-sharing strategies, our results indicate that retailers prefer decentralized inventory-sharing strategies to a centralized one. Turning to research, a majority of behavioral supply chain studies explore either a two-tier setting without retailer inventory sharing or a one-tier setting with retailer inventory sharing (focusing on quantity decisions). We extend this literature by investigating different retailer inventory-sharing strategies, including no sharing, in a two-tier supply chain with endogenous wholesale prices. Our work also complements the theoretical work which examines coordination mechanisms for decentralized firms. Notably, past research has shown that certain incentive structures can lead to favorable results for decentralized settings as compared to a centralized one [27]. Our study finds that behavioral tendencies, notably fairness, can lead to similar results, where firm outcomes are better in decentralized settings.

## 1.2 Literature Review

Retailer inventory sharing has been investigated both theoretically and behaviorally in the operations management literature. Here we review this literature beginning with the theoretical research and then proceed to detail related behavioral studies. Because our paper focuses on contracting in a two-tier supply chain, we also highlight the behavioral literature on supply chain contracting.

In terms of theoretical research, it is well-known that a single joint retailer inventory can increase retailer profits compared to a setting where each retailer acts independently, under exogenous prices [47]. Moreover, this "centralized" inventory-sharing strategy has also been shown to benefit the manufacturer in cases [6, 100]. There is also a rich theoretical literature on the decentralized retailer inventory-sharing strategy where retailers share units after demand occurs at a transfer price per unit (i.e., transshipment). In particular, [107] investigates the optimal ordering policy in a multi-period setting and shows that this "decentralized" inventory-sharing strategy is beneficial to retailers. [110] compare ordering decisions by retailers under both the centralized and decentralized retailer inventory-sharing strategies, in an asymmetric two-location system. They consider how the transfer price affects both retailers and find that a coordinating transfer price exists. A recent study by [131] extends the comparison in [110] to a multi-location setting, but assumes zero transfer price and risk-averse retailers measured by conditional value-at-risk (CVaR).

Additional theoretical research has extended this literature by evaluating a two-tier supply chain with endogenous wholesale prices (which we study), which can alter the effects of inventory sharing. For instance, [44] consider the decentralized inventory-sharing strategy with endogenous wholesale prices and find that it often *decreases* retailer profits compared to no sharing. This is because with inventory sharing, retailers are less sensitive to wholesale price increases, which the manufacturer takes advantage of. A more general version of their model is proposed by [137]. Another relevant theoretical paper to our study is [114]. They extend the work of [110] by investigating a two-tier supply chain with endogenous wholesale prices and directly compare the centralized and decentralized inventory-sharing strategies. Because supply chain decisions often involve human managers, recent papers have begun to investigate retailer inventory sharing from a behavioral standpoint. Regarding the centralized retailer inventory-sharing case, [61] examine stocking quantity decisions in a one-tier supply chain with fixed prices. They observe a pull-to-center effect, which can negate the theoretical inventory-sharing benefit when demands are highly correlated. However, we are unaware of any behavioral studies that evaluate centralized retailer inventory sharing in a two-tier supply chain.

There are also a select number of behavioral papers which investigate the decentralized retailer inventory-sharing strategy. These papers also focus exclusively on retailer decisions, most typically order quantities with exogenous prices. For instance, [21] investigate stocking quantity decisions by decentralized retailers where both wholesale prices and transfer prices are fixed. [124] focus on quantity decisions, also under exogenous wholesale and transfer prices, but compare different retailer interactions (e.g., face-to-face). [141] examine stocking quantity decisions under the decentralized retailer inventory-sharing strategy, but allow retailers to decide whether to request/fulfill inventory after demand is realized.

There are two behavioral papers which examine the decentralized retailer inventory-sharing strategy and allow for endogenous transfer prices by retailers. In particular, continuing within a one-tier supply chain, [86] experimentally test the two-retailer model of [110] where both stocking quantities and transfer prices are set by retailers. They develop a $2 \times 2$ between-subjects experimental design which varies whether transfer prices are set before or after demand is realized and whether inventory sharing is voluntary or automatic. Similarly, [74] consider a setting where each retailer decides its own stocking quantity, but retailers are allowed to negotiate the transfer price. They find that transfer prices do not differ across different critical fractiles.

As mentioned previously, our paper differs from the existing behavioral literature by directly comparing alternative retailer inventory-sharing strategies to one another. In addition, we do so with human-decision makers in a two-tier supply chain where a manufacturer endogenously proposes a wholesale price to downstream retailers. Shifting to a two-tier supply chain where each party is interested in maximizing its own profits is not only a significant difference theoretically, but it also means that behavioral studies on supply chain contracting are relevant. To highlight a few examples, [69, 70] investigate how contract complexity, through increasing the number of wholesale prices with quantity breakpoints, affects decisions in a supply chain experiment. [40] study wholesale price contracts, while varying the inventory risk location in the supply chain. [14] examine how to design buyback contracts for irrational newsvendors, and [138] 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 [30].

## 1.3 Normative Theory and Predictions

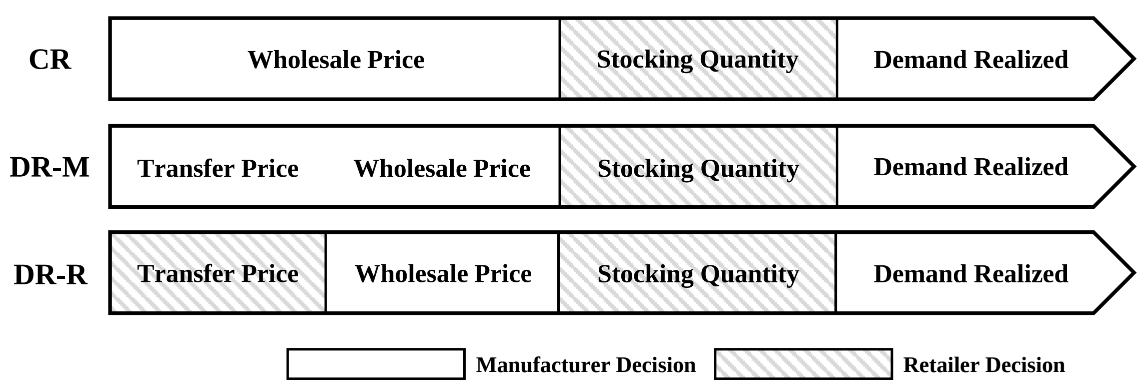
We study a system of one upstream manufacturer and two symmetric downstream retailers. The manufacturer produces a single product at unit cost $c$ and sells it to the two symmetric retailers, indexed by $i$ and $j$, at wholesale price $w$. Each retailer decides a quantity, $q_i$ and $q_j$, purchased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

are independent and follow an identical distribution. Salvage cost is normalized to zero.

For this system, we consider three settings which differ in the inventory-sharing strategy and who has decision authority over the transfer price (if relevant): (1) A no-inventory-sharing "Baseline" setting, (2) a *centralized retailer* inventory-sharing strategy (referred to as CR) where retailers make decisions as if they are a single centralized location and set a joint inventory quantity (there is no transfer price), and (3) a *decentralized retailer* inventory-sharing strategy, where retailers share inventory at a transfer price $t$ per unit, after demand occurs. In this latter decentralized setting, we also consider who has the decision authority over the transfer price. In one case, the manufacturer sets the transfer price $t$ (referred to as DR-M for *decentralized retailers - manufacturer sets the transfer price*). This setting mimics an environment where a retailer, rather than having to reach out to other individual retailers to see if any have excess units or unmet demand, relies on the upstream manufacturer to help facilitate any sharing. Examples of such "dealer-inventory-sharing systems" include Caterpillar, John Deere, and General Motors [140]. In the second case, the retailers negotiate and set the transfer price $t$ (referred to as DR-R for *decentralized retailers - retailers set the transfer price*), where retailers take responsibility over their own inventory sharing.

For the Baseline and CR strategies there are two stages. In stage 1 the manufacturer sets the wholesale price and in stage 2 the two retailers decide an individual or a joint quantity based on the setting. After quantities are determined, demand is realized and any inventory sharing automatically occurs. Under the decentralized inventory-sharing strategies, DR-M and DR-R, we follow [114] and assume that the transfer price is set before the wholesale price, which is common in practice. For example, [97] note that remuneration is decided in advance when firms agree on sharing resources. Also, while there is no difference in predictions in DR-M if we reverse the order of price decisions, if we allow for retailers to set the transfer price after the manufacturer's wholesale price in DR-R, then the normative profit predictions are identical to those in CR. Therefore, the decision sequence in each round of our one-shot environment for DR-M and DR-R consists of three stages. In stage 1, a transfer price is set. In stage 2, the manufacturer sets the wholesale price. In stage 3, the two retailers set stocking quantities. Demands are then realized and inventory sharing takes place. Figure 1.1 illustrates these decisions in all four settings.

Figure 1.1: Decision and Event Sequence for Each Inventory-Sharing Strategy




We note that these four strategies constitute a wide variety of options for retailers and manufacturers. For instance, chain retailers may choose among (a) no inventory sharing (Baseline), (b) a centralized inventory-sharing strategy (CR) or, (c) a decentralized inventory-sharing strategy where they set their own transfer price (DR-R). As another example, competing retailers may choose among (a) no inventory sharing (Baseline), (b) a decentralized inventory-sharing strategy where they act independently but agree to share units at a transfer price per unit (DR-R) or, (c) a decentralized inventory-sharing strategy where they yield control of any inventory sharing to an upstream manufacturer, who may have improved visibility into retailers' inventory levels (DR-M). Last, a powerful manufacturer may choose between (a) coordinating any inventory sharing itself as the upstream party, and thus setting the transfer price (DR-M) or (b) allowing the retailers flexibility to set their own terms and manage it themselves (Baseline, CR, or DR-R).

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

$$\pi_m = (w - c)(q_i + q_j). \tag{1.1}$$

In the following discussion, we refer to Equation (1.1) when showing the manufacturer's profit-maximizing decisions. Unless otherwise noted, we use retailer $i$ when the discussion involves only one retailer and all results apply to retailer $j$ by symmetry. In the following subsections, we use superscripts $b$, $c$, and $d$ for Baseline, centralized retailers (CR), and decentralized retailers (DR-M and DR-R), respectively. Next we show expected profit functions and optimal decisions by backward induction for each setting.

### 1.3.1 No Inventory Sharing - Baseline

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

$$\pi^b_{r,i} = \mathbb{E}[p \min(d_i, q_i)] - wq_i. \tag{1.2}$$

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

$$\alpha(q) = \frac{p - w}{p}, \qquad\qquad w = pq^b_i f(q^b_i) + c. \tag{1.3}$$

### 1.3.2 Inventory Sharing - Centralized Retailers

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

Although the centralized retailers will set a single joint stocking quantity, in the following equations we assume $q_i$ and $q_j$ are each one half of this joint quantity.

The retailer's joint expected profit function is

$$\pi_r^c = \mathbb{E}[p \min(d_i + d_j, q_i + q_j)] - w(q_i + q_j). \tag{1.4}$$

Let $\alpha^c(q_i, q_j)$ be the probability $\Pr(d_i + d_j < q_i + q_j)$, i.e., CDF of the joint demand distribution, and $f^c(d_i, d_j)$ be the corresponding PDF. Solving Equation (1.4) gives the optimal stocking quantity $\left(q_i^c, q_j^c\right)$ derived from Equation (1.5)

$$\alpha^c(q_i, q_j) = \frac{p - w}{p}. \tag{1.5}$$

Given $\left(q_i^c, q_j^c\right)$, the manufacturer's optimal wholesale price is derived from Equation (1.6),

$$w = p(q_i + q_j) f^c(q_i, q_j) + c. \tag{1.6}$$

### 1.3.3 Inventory Sharing - Decentralized Retailers

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

as the transferred quantities from $j$ to $i$. As with past studies, we assume that the transportation cost of shared units is zero.

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

$$\pi_{r,i}^d = \mathbb{E}\left[p \min(d_i, q_i) + tT_i + (p-t)T_j\right] - wq_i. \tag{1.7}$$

In this setting, [110] show that a unique Nash equilibrium exists.[1] The equilibrium stocking quantity $\left(q_i^d, q_j^d\right)$ satisfies Equation (1.8)

$$\alpha(q_i) - \beta_i(q_i, q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i, q_j)\left(\frac{p-t}{p}\right) = \frac{p-w}{p}, \tag{1.8}$$

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

The manufacturer's profit-maximizing wholesale price is derived from Equation (1.9).

$$w = q_i^d[p\alpha(q_i^d) - t\beta_i(q_i^d, q_j^d) + (p-t)\gamma_i(q_i^d, q_j^d)] + c. \tag{1.9}$$

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

[1] See [110] for a general solution with asymmetric retailers.

#### 1.3.3.1 Transfer Price set by the Manufacturer

When the manufacturer sets the transfer price, DR-M, [114] show that the stocking quantity and wholesale price are monotonically increasing in the transfer price. Therefore, the manufacturer prefers a higher transfer price and sets $t = p$.

#### 1.3.3.2 Transfer Price set by the Retailers

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

Last, 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.

### 1.3.4 Experimental Predictions and Hypotheses

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

In all four treatments we use a retail selling price $p$=30 and a manufacturer unit production cost $c$=5. Each retailer faces an integer demand drawn from a uniform distribution between 0 and 100. While our selling price and unit production cost parameters appear to be for a relatively high-margin product, recall that this is across the entire supply chain (the retailer's normative critical fractile is actually less than 50% in all treatments, see quantity predictions in Table 1.1). Importantly, unlike existing behavioral papers on inventory sharing, in all treatments the retailer's critical fractile may vary through $w$ being endogenously set by the manufacturer.

Table 1.1: Normative Theoretical Predictions in the Experiment

| | Baseline | CR | DR-M | DR-R |
|---|---|---|---|---|
| Transfer Price $t$ | - | - | 30.00 | 0.00 |
| Wholesale Price $w$ | 17.50 | 21.67 | 21.67 | 18.33 |
| Stocking Quantity $q$ | 41.67 | 37.27 | 43.03 | 33.33 |
| Manufacturer Profit $\pi_m$ | 1041.67 | 1242.26 | 1434.44 | 888.89 |
| Retailer Profit $\pi_r$ | 260.42 | 207.04 | 199.23 | 314.81 |
| Supply Chain Efficiency (%) | 71.27% | 75.55% | 83.60% | 69.26% |

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

Table 1.1 illustrates the experimental predictions for contract terms, profits, and supply chain efficiency (we provide detailed plots in Appendix A.1). First, beginning with the contract terms in Table 1.1, the predicted transfer price is always extreme, 30 in DR-M and 0 in DR-R. Intuitively, in DR-M, a manufacturer will set $t = p$ to incentivize retailers to order a higher quantity, which earns them a higher profit (i.e., a higher transfer price allows a retailer to earn a higher price on units sent but also requires them to pay more for units received). And, in DR-R, retailers will set $t = 0$, leading to lower stocking quantities (i.e., units sent are less valuable and units received are more affordable). Second, regarding wholesale prices, the predicted values are always equal to or higher than two potential anchor points, $(p + c)/2 = 17.5$ and $p/2 = 15$, which is useful when comparing across treatments. And third, stocking quantities are all predicted to be below 50. Our first experimental hypothesis revolves around these contract-term point predictions:

**Hypothesis 1.1.** *(Contract Terms) Transfer prices, wholesale prices, and stocking quantities will be set such that they coincide with the normative theoretical predictions.*

Turning to the manufacturer and retailer profits in Table 1.1, for manufacturers, it is unsurprising to see that they prefer to have decision authority over the transfer price, leading to the highest profits in DR-M. Overall, the manufacturer's preferred order among the four treatments is predicted as DR-M>CR>Baseline>DR-R. For retailers, they too prefer to have decision authority over the transfer price, yielding the highest profits in DR-R. Interestingly, retailers earn the second-highest profit in Baseline. At first this may seem counterintuitive, as inventory sharing should be beneficial for retailers. While this is true with exogenous wholesale prices it does not necessarily hold in a two-tier supply chain with endogenous wholesale prices. In short, because retailers benefit from inventory sharing, manufacturers are able to charge a higher wholesale price in equilibrium (e.g., 21.67 in CR and DR-M versus 17.50 in Baseline). Across the four treatments, the retailers' preferred order is predicted as DR-R>Baseline>CR>DR-M, and we have:

**Hypothesis 1.2.** *(Profits) Manufacturer profit in the four treatments, from highest to lowest, will be DR-M>CR>Baseline>DR-R. Retailer profit in the four treatments, from highest to lowest, will be DR-R>Baseline>CR>DR-M.*

Continuing with profits in Table 1.1, it is noteworthy that the manufacturer always earns significantly more than the retailer. Among those four treatments, DR-R is the most equitable, thus:

**Hypothesis 1.3.** *(Equity) The manufacturer will always earn a higher profit than the retailer, but DR-R will yield the most equitable distribution of profits between the manufacturer and retailer.*

Last, we develop a hypothesis for supply chain efficiency, which is calculated as the sum of the manufacturer and retailers' expected profits divided by the first-best fully-integrated supply chain benchmark. In the last row of Table 1.1, one can observe that DR-M is predicted to achieve the highest efficiency, followed by CR, Baseline, and DR-R:

**Hypothesis 1.4.** *(Efficiency) Supply chain efficiency in the four treatments, from highest to lowest, will be DR-M>CR>Baseline>DR-R.*

### 1.3.5 Behavioral Discussion

While our experimental hypotheses rely on the normative theory, we would be remiss if we did not highlight that certain behavioral biases may impact decisions (for a summary of biases in operations management in individual decisions, other-regarding behavior, and strategic interactions, please see [18, 39] and [85]). For instance, participants may be susceptible to bounded rationality [e.g., 116] and set transfer prices in a way that do not perfectly coincide with the extreme normative predictions of 30 in DR-M and 0 in DR-R. As another example, in settings where a proposer can make a one-shot offer to a responder, and the proposer is predicted to earn a disproportionately high split of overall profits, experimental studies have found evidence of fairness [108]. In our experiment, this suggests that the manufacturer may, at least qualitatively, set wholesale prices below the normative predictions. Last, existing newsvendor experiments indicate that a pull-to-center bias may push stocking quantities higher than the normative predictions in our setting [111], but at the same time, certain supply chain experiments also find evidence of an understocking bias [e.g., 40], so a directional deviation for quantities is unclear.

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

## 1.4 Experimental Methodology

All four experimental treatments, Baseline, CR, DR-M, and DR-R follow the normative theory and decision sequence outlined in Figure 1.1 in Section 1.3. In particular, in the Baseline treatment each round begins with the manufacturer setting a wholesale price (there is no transfer price). After the wholesale price decision each retailer then independently sets its own stocking quantity. Demand for each retailer is then realized and profits are earned.

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

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

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

occurs, if applicable.

We provide a decision support tool for both roles to minimize complexity and to create a more realistic environment. For instance, there is empirical evidence that managers often use computers as an input for final decisions in supply chain settings. [141] conduct a survey of 54 firms and find that 35 (65%) rely on automated algorithms which are adjusted by human managers and 10 (19%) rely on averages between human decisions and system orders. With this in mind, our decision support tool consists of slide bars and a dynamic figure of expected profits. Specifically, participants can test their decisions by sliding the bar(s), and the expected profits of all three parties will be depicted on the figure. For the transfer price and wholesale price decisions, expected profits are calculated assuming that subsequent decisions are made optimally (which participants are aware of). For stocking decisions in Baseline and CR, there is one bar for the quantity. In DR-M and DR-R, each retailer has two slide bars, one for their own and one for the other retailer's stocking quantity (by checking a box, they can use the tool so that the other retailer's quantity is the best response or they can manually set the other retailer's quantity). In all treatments, the test quantity scroll bar is initially set at the optimal quantity. Please see Appendix A.4.3 for sample instructions and screenshots.

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

Turning to sample sizes, our Baseline, CR, DR-M, and DR-R treatments consist of 30, 57, 60, and 60 participants.[3] We included larger samples sizes for the three inventory-sharing treatments because they have not been investigated before in the laboratory, whereas the Baseline treatment is closely related to existing research between a single manufacturer and single retailer. Also, following existing experimental supply chain and economics research, our study includes university student participants [45, 66], who were recruited from a large university where cash was the only incentive offered. Several studies have shown that

[3]An earlier version of this paper, which was prior to the COVID-19 pandemic, did not include the Baseline treatment and included 42 participants in each of the three treatments (CR, DR-M, and DR-R). During the review process we added the Baseline condition. Due to the pandemic, we ran this treatment synchronously online through Zoom. We provide further details in Appendix A.4.1 (e.g., we put each individual participant in their own private breakout room with an experimenter, required cameras to be on, etc), and here note that these online sessions followed the same protocols and used the same participant pool as our original in-person experiments. Further, to determine whether switching to online methods did not have a meaningful effect on our results, we re-ran a session of each of the original three treatments (CR, DR-M, DR-R), yielding 57, 60, and 60 participants in each. Overall, we found similar results with the online implementation and thus include all data in our analysis.

students make similar decisions as managers in operational settings, such as inventory management and forecasting [e.g., 73, 20, 80]. Nevertheless, we have not seen a supply chain contracting experiment which compares contracting decisions between students and managers, so we recognize this as a limitation.

Our experiment was implemented through oTree [28]. Each session consisted of 12 rounds. In each round, participants were randomly assigned a role and matched with two other participants. This means that both roles (and trios) were randomly determined each round, similar to [101] and [41]. Before a session started, a researcher read through the instructions out loud and answered any questions. Participants were then required to answer several multiple-choice comprehension questions about the game. Participants received cash based on profits from all rounds in the game plus a $7 show-up fee. Average earnings were roughly $25 across all treatments. Each session lasted for 70 minutes on average.

## 1.5 Results

In this section we present our experimental results. Following our hypotheses, we take a bottom-up approach and begin with contract terms in Section 1.5.1. We then proceed to investigate profits and efficiency in Section 1.5.2. In this subsection we also discuss how any deviations in contract terms, relative to theory, account for observed profits and efficiency. For our analysis, the rate of agreements/acceptances by retailers was high, and similar, across treatments. Specifically, the fraction of time that retailers came to an agreement over the stocking quantity in CR and the transfer price in DR-R was 95.18% and 97.08%. These near-100% agreement rates are not particularly surprising as they represent a joint decision rather than a zero-sum negotiation.[4] In addition, the fraction of time retailers accepted the manufacturer's wholesale price and set a positive stocking quantity was also high, and similar, across all four treatments: 98.33% in Baseline, 97.70% in CR, 97.50% in DR-M, and 98.75% in DR-R. Thus, unless otherwise stated, we include all data in our analysis.

Given the panel-structure of our data, unless otherwise noted, we use regression analysis with random effects for all hypothesis tests [64].[5] To provide a specific example, consider our fourth hypothesis on efficiency. Because one manufacturer and two retailers are randomly matched together as a trio and earn the same efficiency (in a round), we only include those observations for the role of the manufacturer for this test. We follow a similar approach any time there is a risk of double or triple counting observations (e.g., efficiency, retailer profit and the joint quantity in CR, the transfer price in DR-R, etc). Also, because each experimental hypothesis consists of a family of multiple comparisons [87], we adjust the critical p-values using Bonferroni corrections, assuming an unadjusted critical p-value of 0.05. [6]

[4]We also did not observe any sort of deadline effect (i.e., a mass of agreements at the end of the two minute time period), which is frequently observed in more adversarial negotiation experiments [109].

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

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

### 1.5.1 Contract Terms

Average observed transfer prices, wholesale prices, and stocking quantities for all four treatments are summarized in the left-hand side of Table 1.2. Beginning with transfer prices for the decentralized retailer inventory-sharing strategies, DR-M and DR-R, one can see that observed transfer prices deviate from their extreme predictions, which is inconsistent with Hypothesis 1.1. In DR-M, manufacturers set the transfer price lower than the normative prediction, 20.26 versus 30 ($p < 0.005$, the corrected critical p-value), whereas in DR-R, retailers set the transfer price higher than the normative prediction, 6.27 versus 0 ($p < 0.005$). These two observations suggest that any bias influencing transfer prices may be present for both manufacturers in DR-M and retailers in DR-R.

To investigate wholesale price and stocking quantity decisions, the right-hand side of Table 1.2 presents the normative predictions conditioned on any previous decisions. For instance, predicted wholesale prices in DR-M and DR-R are conditioned on transfer prices, and all stocking quantities are conditioned on wholesale prices (and transfer prices, if applicable). While all tests are between the observed data and these conditional predictions, we also report the unconditional normative predictions in square brackets.

Wholesale prices, in all three inventory-sharing treatments (CR, DR-M, DR-R), are set significantly lower than the conditional predictions and contradict Hypothesis 1.1: 19.15 versus 21.67 in CR, 18.40 versus 20.41 in DR-M, and 16.78 versus

Table 1.2: Average Contract Prices, Quantities, and Normative Theoretical Predictions

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

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

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

We also conducted a heterogeneity analysis by classifying those participants who did or did not make a particular decision optimally. This is instructive in determining whether any of our aggregate results are driven by a small group of individuals. For each individual participant we conducted a test between their decisions and the conditional predictions. Due to the limited number of observations, we opt for Wilcoxon signed-rank tests and count a significant deviation if there is a difference at the 10% level. Beginning in Figure 1.2a, 81.67% in DR-M and 100% in DR-R of participants set transfer prices that significantly deviated from the normative predictions. In Figure 1.2b, between 40% and 59.65% of participants set wholesale prices that were significantly too low. These first two figures indicate that the average transfer price and wholesale price deviations do not appear to be driven by a small subset of participants. Last, in Figure 1.2c, a majority of participants set quantities in a way that is not significantly different from the conditional prediction, but a reasonable percentage understock in Baseline, CR, and DR-M. This coincides with the aggregate results. Thus we have our first result, which largely rejects Hypothesis 1.1:

**Result 1.1** (Contract Terms)**.** *Transfer prices in the two decentralized inventory-sharing strategies, DR-M and DR-R, deviate symmetrically from the normative predictions. Wholesale prices are set too low in all three of the inventory-sharing environments, CR, DR-M, and DR-R. Stocking quantities are set reasonably well or slightly low in all settings.*

As a final comment regarding contract terms, we also investigated dynamics.

Figure 1.2: Percentages of Participants Deviating in Pricing and Stocking Decisions


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

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

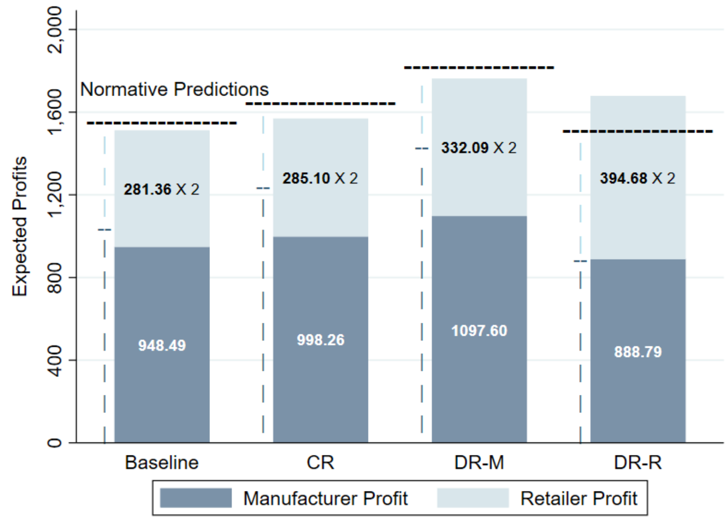
### 1.5.2 Profits and Efficiency

Figure 1.3 depicts the manufacturer expected profit (bottom darker portion), each of the retailer's expected profits (top light portion), and supply chain expected profit (overall height), in all four treatments. While we will compare the observed results to the normative theoretical predictions momentarily, we also include the normative predictions by dashed horizontal and vertical lines.

Beginning with the manufacturer and retailer profits in Figure 1.3, a key result is that the DR-M treatment provides a win-win outcome compared to both the Baseline and CR conditions: retailers earn significantly higher profits, 332.09 versus 281.36 and 285.10 (both $p < 0.004$, the corrected critical p-value), and manufacturers earn significantly higher profits as well, 1097.60 versus 948.49 and 998.26 (both $p < 0.004$). This leads to our second result:

**Result 1.2** (Profits)**.** *A decentralized retailer inventory-sharing strategy where the manufacturer sets the transfer price, DR-M, achieves a win-win outcome compared to the Baseline setting and centralized retailer inventory-sharing strategy, CR, in that both the manufacturer and retailers earn significantly higher profits.*

Figure 1.3: Average Observed Expected Profits




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

To provide further context around this result and Hypothesis 1.2 (Profits), theory predicts manufacturers to earn the highest profit in DR-M (when they have authority over the transfer price). In this sense, the second hypothesis is validated for manufacturers. However, a different picture emerges for retailers. Whereas Hypothesis 1.2 predicts that retailers should prefer an order of DR-R>Baseline>CR>DR-M, the data reject this order and we instead observe DR-R>DR-M>CR≈Baseline. The primary difference being the favorable performance of DR-M over both Baseline and CR, for retailers. A managerial implication of this is that, if retailers are contracting with a powerful manufacturer, enabling inventory sharing while ceding decision authority of the transfer price to the manufacturer will still lead to a relatively high profit. We will discuss this more in Section

1.9.

Proceeding with Hypothesis 1.3 (Equity), another result we can glean from Figure 1.3 is that manufacturers do indeed earn more than retailers, and, DR-R generates the most equitable distribution of profits between the manufacturer and retailers, among the four treatments. This is true in terms of both the percentage profit split and the absolute difference in profits. For instance, the percentage of total supply chain profits that are earned by each retailer in DR-R is 23.5% in DR-R ($^{394.68}/_{(2\times394.68+888.79)}$). Whereas in the other treatments this percentage is between 18.2% and 18.8%. A proportions test indicates that the retailer's share in DR-R is indeed significantly higher than the other treatments (all three $p < 0.008$, the corrected critical p-value). Similarly, the absolute difference in manufacturer profit and average retailer profit is only 494.11 in DR-R, yet the differences are between 667.13 and 765.51 in the other three treatments (all three $p < 0.008$). Therefore Hypothesis 1.3 is supported and we have:

**Result 1.3** (Equity)**.** *Manufacturers always earn more than retailers, but a decentralized retailer inventory-sharing strategy where the retailers set the transfer price, DR-R, achieves the most equitable outcome in terms of distribution of expected profits between the manufacturer and retailers.*

Turning to how profits compare to the normative predictions, the left-hand side of Table 1.3 depicts the observed average profits along with hypothesis tests versus the normative predictions (the latter of which are illustrated in the right-hand side of the table). Regarding the distribution of manufacturer and retailer profits compared to theory, we see a consistent pattern across all treatments: manufacturers earn significantly less (or the same) than the normative predictions and retailers earn significantly more than the normative predictions. In short, outcomes are more equitable than theory predicts (even in DR-R, which should already lead to the most equitable payoffs). This observation will be relevant when we explore behavioral biases in Section 1.7.

Table 1.3: Average Observed Profits, Efficiency, and Normative Theoretical Predictions

| | Observed Results | | | | Normative Predictions | | | |
|---|---|---|---|---|---|---|---|---|
| | Baseline | CR | DR-M | DR-R | Baseline | CR | DR-M | DR-R |
| Manu. | 948.49$^\dagger$ | 998.26$^\dagger$ | 1097.60$^\dagger$ | 888.79 | 1041.67 | 1242.26 | 1434.44 | 888.89 |
| Profit | (21.96) | (25.29) | (22.35) | (20.68) | | | | |
| Retail. | 281.36$^\dagger$ | 285.10$^\dagger$ | 332.09$^\dagger$ | 394.68$^\dagger$ | 260.42 | 207.04 | 199.23 | 314.81 |
| Profit | (7.72) | (7.17) | (7.40) | (7.74) | | | | |
| SC | 68.93% | 71.54% | 80.36%$^\dagger$ | 76.54%$^\dagger$ | 71.27% | 75.55% | 83.60% | 69.26% |
| Effi. | (0.019) | (0.018) | (0.011) | (0.013) | | | | |

Note: "Manu." stands for manufacturer. "Retail." stands for retailer. "SC Effi." stands for supply chain efficiency. Standard errors, across participants, reported in parentheses. Significance of regressions comparing observed versus normative predictions given by $^\dagger$ $p < 0.017$ (the corrected critical p-value). There are significant differences between all profits and efficiencies compared to the normative benchmarks except manufacturer profit in DR-R, and efficiency in Baseline and CR.

Comparing efficiency across treatments (Hypothesis 1.4), in Table 1.3, DR-M and DR-R achieve the highest efficiency (DR-R should, in theory, have the lowest efficiency). While the difference between DR-M and DR-R is not significant, DR-M is significantly higher than the Baseline and CR conditions (both $p < 0.008$, the corrected critical p-value) and DR-R is significantly higher than Baseline ($p < 0.008$). Interestingly, there is no statistical difference between Baseline and CR, where moving from the Baseline to CR should, in theory, increase efficiency. Another notable observation from this table is that the supply chain profit in DR-R is higher than the normative prediction, which we will explore in detail later. Until then, Hypothesis 1.4 is rejected and we have:

**Result 1.4** (Efficiency)**.** *Both decentralized retailer inventory-sharing strategies, DR-M and DR-R, achieve a higher efficiency compared to the Baseline setting (and DR-M is higher than the centralized retailer-sharing strategy, CR). Further, there is no significant increase in efficiency when moving from the Baseline setting to a CR strategy.*

Overall, a combination of the experimental results around profits, equity, and efficiency, provides evidence that both the Baseline setting, where two retailers neglect to share inventory, and the centralized retailer inventory-sharing strategy, CR, fail to perform best across (a) retailer profit, (b) manufacturer profit, (c) equity of profits, and (d) supply chain efficiency. Thus, if a firm is using any of these metrics as their primary criteria for how to share inventory, they should consider a decentralized retailer inventory-sharing strategy, where the retailers or manufacturer have decision authority over the transfer price, depending on the situation.

### 1.5.3 Connecting Contract Terms with Profits and Efficiency

Here we connect our results by summarizing how the the observed contract-term deviations highlighted in Result 1.1 can largely account for Results 1.2, 1.3, and 1.4 (profit, equity, and efficiency). Beginning with Result 1.2, which finds that DR-M generates a Pareto improvement over the Baseline and CR conditions, due to observed wholesale prices being lower than predicted in all treatments, and transfer prices being set too low in DR-M (which benefits the retailer), the manufacturer earns less and the retailers earn more than the normative theory predicts. This, combined with the fact that DR-M is predicted to generate the highest manufacturer profit *and* the highest total supply chain profit, allows manufacturers to effectively redistribute some of their profits to retailers and make both parties better off compared to Baseline and CR (but not DR-R, as it is predicted to generate the highest retailer profit).

For Result 1.3, DR-R generates the most equitable distribution of profits (and more equitable than theory predicts) for two reasons. First, DR-R is predicted to provide the most equitable outcome between parties. Second, manufacturers still offer more generous wholesale prices than theory predicts. A combination of the theoretical prediction and this experimental deviation results in the most equitable split of profits, even more so than theory predicts. Regarding Result 1.4 (that DR-M and DR-R achieved the highest supply chain efficiency), theory predicts DR-M to earn the highest efficiency. Because quantities are set similarly in all treatments (i.e., close to or slightly below the conditional predictions), and both the unconditional and conditional quantity predictions are the highest in DR-M, its observed efficiency is highest. Regarding DR-R, it even outperforms its theoretical prediction (observed efficiency of 76.54% versus prediction of 69.26%) because manufacturers offer lower wholesale prices than optimal *and* retailers set transfer prices higher than optimal. Both of these effects drive quantities higher (observed quantities of 39.80 versus unconditional prediction of 33.33), and hence a higher efficiency than theory predicts.

We provide a summary of these effects in Table 1.4, which illustrates the percentage impact of a particular price and quantity deviation on manufacturer and retailer profits, relative to theory. Each effect is calculated comparing the observed profit to the conditional optimal profit for that decision, divided by the normative prediction. For instance, the wholesale price impact in DR-M, for each respective party, is ${}^{(\pi(w,q(w,t)|t)-\pi(w^*,q(w^*,t)|t)}/_{\pi(t^*,w^*,q(w^*,t^*))}$. Beginning at the bottom of this table, we observe that, unsurprisingly, slightly lower than optimal quantities only have a relatively small impact on both parties' profits (third row). Turning to wholesale prices (second row), we observe that manufacturers decrease their own earnings but greatly increase retailer profits. For instance, manufacturers give up between 3.97% and 10.93% in profits through lower wholesale prices, but this increases retailer profits by 13.23% to 52.38%. Therefore, manufacturers set wholesale prices in a way that is more generous than theory predicts.

Table 1.4: Manufacturer and Retailer Profit Implications from Price and Quantity Deviations

| | Baseline | | CR | | DR-M | | DR-R | |
|---|---|---|---|---|---|---|---|---|
| | Retail. | Manuf. | Retail. | Manuf. | Retail. | Manuf. | Retail. | Manuf. |
| $t$ | - | - | - | - | 26.90% | -13.46% | -3.49% | 11.82% |
| $w$ | 13.23% | -3.97% | 52.38% | -6.28% | 47.75% | -6.44% | 34.14% | -10.93% |
| $q$ | -5.19% | -4.97% | -7.87% | -9.10% | -7.96% | -3.59% | -5.25% | -0.93% |
| Total | 8.04% | -8.95% | 44.50% | -15.37% | 66.69% | -23.48% | 25.39% | -0.04% |

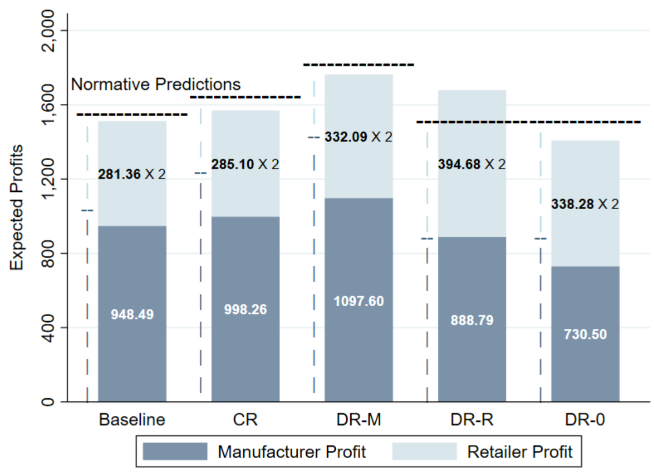
Note: The percentage impact of a particular price and quantity deviation on manufacturer and retailer profits, relative to theory. Each percentage effect is calculated comparing the observed profit to the conditional optimal profit for that decision, divided by the normative prediction. For instance,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

In reviewing the transfer price deviation effects (first row) in Table 1.4, the party with decision authority sets the transfer price in a way that hurts themselves and helps the other party in a significant way. For example, in DR-M, manufacturers give up 13.46% of their profits by setting sub-optimal transfer prices, but this translates into an increase of 26.90% in retailer profits. Ultimately, this helps contribute to the redistribution of wealth from the manufacturer to the retailer and the "win-win" outcome over Baseline and CR. Last, in DR-R, the transfer price deviation by retailers only slightly decreases their own profits, by 3.49%, but the impact of the wholesale price deviation by the manufacturer is much larger, increasing retailer profits by 34.14%, such that the net result is retailers earn more than theory predicts and DR-R becomes quite equitable.

## 1.6 Alternative DR Treatment

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

Figure 1.4: Average Observed Expected Profits in the Main Experiment Plus DR-0




We ran the DR-0 treatment with 57 additional participants. In Figure 1.4 we illustrate the profit results for our original four treatments plus the DR-0 treatment (far-right column of the figure). Comparing DR-0 to Baseline and CR, in both cases we observe that retailers do earn a higher profit in DR-0 (and manufacturers earn less), which supports theory ($p < 0.005$, the corrected critical p-value). Also, comparing DR-0 to DR-R, we see that manufacturers do indeed earn a lower profit in DR-0, which supports theory (730.50 versus 888.89, $p < 0.005$). However, despite the theoretical advantage of DR-0 over DR-R for retailers, DR-0 actually achieves a retailer profit that is significantly less than DR-R experimentally (338.28 versus 394.68, $p < 0.005$). As a consequence, DR-R Pareto dominates DR-0, counter to theory.

Table 1.5: Average Price and Quantity Decisions and Normative Theoretical Predictions in DR-0

| | DR-0 | |
|---|---|---|
| | Observed | Predictions |
| Wholesale Price | $17.35^{\dagger}$ | 18.33 |
| | (0.33) | |
| Stocking Quantity | $31.59^{\dagger}$ | 35.63 |
| | (0.96) | (0.56) |
| | | [33.33] |

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

In Table 1.5 we report average observed wholesale prices and quantities for DR-0, along with the normative predictions. Consistent with the original three inventory-sharing treatments, we find that wholesale prices are again set too low relative to the normative benchmark ($p < 0.005$). However, unlike DR-R, where quantities were set close to optimal, in DR-0 retailers significantly understock relative to the conditional prediction: 31.59 versus 35.63 ($p < 0.005$). This directly accounts for the poor profit performance of DR-0 for both parties: manufacturers offer a more generous wholesale price, thus reducing their own profits, but retailers neglect to capitalize on this and instead significantly understock, hurting both parties' (and the supply chain's) profits. We therefore supplement our previous findings with the following fifth result:

**Result 1.5** (DR-0). *A decentralized retailer inventory-sharing strategy with an exogenous transfer price of zero contributes to significant understocking in quantities. As a result, DR-R achieves a win-win outcome over DR-0: both parties earn a higher profit when the transfer price is endogenously set by retailers compared to when it is exogenously set to zero.*

There are at least two plausible (non-exclusive) factors that may drive the observed understocking bias in DR-0. First, retailers may understock when the transfer price is zero, regardless of how the transfer price is set. Second, retailers may understock when transfer prices are set exogenously. To investigate this, we compare observations in DR-R where the retailers set a transfer price equal to zero, so that the only difference between DR-R and DR-0 is whether the transfer price is set endogenously or exogenously. However, we observe that retailers set a transfer price equal to zero only 2.6% of the time in DR-R, making it difficult to draw conclusions. If we expand this analysis to include any transfer price less than 0.5 (resp. 1), then the number of observations increases to 12.0% (resp. 25.8%). In this case, retailers in DR-R understock by an average of 0.68 units (resp. 1.96 for $t < 1$). This degree of understocking is far less than what we observe in DR-0, which is 4.04. Overall, this analysis suggests that while retailers appear to dislike transfer prices of zero (evidenced by the low frequency of observations with $t = 0$ in DR-R), the excessive understocking in DR-0 may be primarily driven by the fact that such low transfer prices are exogenously set in DR-0. We will explore this further when we fit our behavioral model, in the next section.

## 1.7 Behavioral Model and Estimations

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

In determining which bias to investigate, our previous experimental analysis is instructive. A key finding was that the observed distribution of profits is more equitable than predicted. In particular, we found that manufacturers, who are predicted to earn a significantly larger share of the total profits, offered wholesale prices that are more generous than theory predicts, leading to a more equitable distribution of profits than theory predicts. This indicates that fairness may be influencing decisions [49, 19]. For brevity, below we provide an overview of this behavioral bias and relegate detailed theoretical analyses, estimation procedures, and more results in Appendix A.3.

Fairness has been examined in a number of supply chain studies [e.g., 36, 70, 16]. We follow a similar approach and assume that a retailer (the responder) suffers disutility when their expected profit is less than the manufacturer's profit (we do not consider disutility when the retailer earns more than the manufacturer, which rarely occurred in our data). In addition, while a majority of studies on fairness typically focus on the responding party, there is empirical evidence that proposers in favorable positions often share more of their earnings than theory predicts. This is seen in experimental dictator games, where a proposer who has authority over how much of a surplus to share with a responder routinely offers around 25% of the surplus, as opposed to the equilibrium of zero [e.g., 50, 4]. Therefore, we also consider that the manufacturer suffers disutility when their expected profit is greater than the retailer's. Recall that $\pi_m$ and $\pi^d_{r,i}$ are the expected profit of manufacturer and retailer $i$ under the normative theory in the decentralized setting (the centralized case follows similar logic). Under fairness concerns, the expected utility functions for manufacturer's utility and the decentralized retailer $i$'s utility, are given by

$$u^F_m = \pi_m - \lambda_m(\pi_m - \pi^d_{r,i})^+, \quad u^{d,F}_{r,i} = \pi^d_{r,i} - \lambda_r(\pi_m - \pi^d_{r,i})^+, \tag{1.10}$$

where $\lambda_m$ represents the manufacturer's degree of fairness over advantageous inequality and $\lambda_r$ represents retailer $i$'s degree of fairness concerns over disadvantageous inequality. This formulation is directly related to the well-established fairness model of [49].[7]

Qualitatively, fairness can account for lower wholesale prices by manufacturers, relative to the normative theory. The intuition is straightforward in that fairness-minded manufacturers prefer to offer lower wholesale prices, which equalizes profits. We also find that fairness, according to the model, can account for retailers stocking less than the normative predictions (see Appendix A.3), as

[7] We also considered the ERC model of [19], where manufacturers suffer disutility when earning more than 1/3 of the supply chain profit. It provides a good, albeit slightly worse fit, as the formulation presented. Details are provided in Appendix A.3.

it leads to more equitable profits, but we expect this effect to be small in our estimations given that retailers set quantities well or only slightly low in certain treatments.

Regarding transfer prices, we find that fairness cannot account for the deviation of transfer prices in DR-M or DR-R. Because transfer prices were set too low by manufacturers in DR-M and too high by retailers in DR-R, a symmetric bias may be influencing decisions. Further, the impact of sub-optimal transfer prices on profits is relatively small compared to the deviations in wholesale prices (see Table 1.4). Combining these two observations, we model transfer price deviations with random errors so that they follow a multinomial logit distribution. Transfer prices yielding a higher expected utility are chosen with a higher probability and vice versa, where $\theta$ is the degree of rationality: $\theta \to 0$ is fully rational and $\theta \to \infty$ is fully random decisions [116].

Table 1.6: Aggregate Behavioral Model Estimation Results

| Description | Parameter | Transfer Errors | $\text{Fair}_m$ | $\text{Fair}_r$ | $\text{Fair}_{m,r}$ |
|---|---|---|---|---|---|
| Transfer price errors | $\hat{\theta}$ | 193.383 | 92.563 | 165.504 | 85.814 |
| Manufacturer fairness | $\hat{\lambda}_m$ | - | 0.334 | | 0.333 |
| Retailer fairness | $\hat{\lambda}_r$ | - | - | 0.086 | 0.089 |
| Constrained estimation | $LL_c$ | -13620.24 | -13449.36 | -13566.87 | -13390.01 |
| Sum of separate estimations | $LL_s$ | -13469.88 | -13313.81 | -13404.42 | -13248.71 |

Note: "LL" represents log-likelihoods. Number of observations in each $LL$ is 3370. $\hat{\theta}$ is estimated parameter for transfer price errors. $\hat{\lambda}_m$ and $\hat{\lambda}_r$ are estimated fairness parameters for the manufacturer and retailers, respectively. Full details provided in Appendix A.3.

We fit the model to the data using maximum-likelihood estimation. Table 1.6 provides the results for a model with only transfer price errors (Transfer Errors), each type of fairness separately ($\text{Fair}_m$ and $\text{Fair}_r$), and the "full" model of fairness with both types of fairness ($\text{Fair}_{m,r}$). For each model, we fit the data two ways. The first constrains the parameters to be the same across all five treatments ($LL_c$). The second fits a set of parameters separately for each treatment then sums up the five log-likelihoods ($LL_s$). Both approaches are depicted in Table 1.6. A series of likelihood-ratio tests reveals significant differences across all applicable estimations, such that the full-fairness model with both retailer and manufacturer concerns, $\text{Fair}_{m,r}$, provides the most favorable fit, for both procedures. Aside from this main takeaway, in evaluating the two nested-fairness models, $\text{Fair}_m$ and $\text{Fair}_r$, we see that the log-likelihood is considerably better in $\text{Fair}_m$ (e.g., $LL_c$ of -13449.36 versus -13566.87). This suggests that the inclusion of manufacturer-fairness concerns is especially important to the overall fit.

In the top of Table 1.7 we provide the transfer price, wholesale price, and quantity predictions using the parameter estimates from the full-fairness model, $\text{Fair}_{m,r}$, for each treatment. In an effort to evaluate these predictions we also include the average observed transfer price, wholesale price, and quantities from the data. Beginning with transfer prices in DR-M, the fairness model predicts a price of 20.22 versus 20.26 observed, and in DR-R the prediction is 6.32 versus 6.27 observed. Wholesale prices are also quite accurate, going from left to right in Table 1.7 (Baseline, CR, DR-M, DR-R, DR-0), the fairness model predictions versus observed are: 16.94 versus 16.94, 19.13 versus 19.15, 18.35 versus 18.40, 16.82 versus 16.78, and 17.35 versus 17.35. Last, the predicted quantities are within one unit of the observed quantities, in all five treatments.

Table 1.7: Fairness Predictions and Observed Values and Parameter Estimates, by Treatment

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

Note: Number of observations for Baseline, CR, DR-M, DR-R, and DR-0 are 360, 434, 960, 932, and 684 (3370 total). DR-R and CR are conditioning on agreement. "Fair." is the full fairness model Fair$_{m,r}$. "Obs." is observed average values. $\hat{\theta}$ is estimated parameter for transfer price errors. $\hat{\lambda}_m$ and $\hat{\lambda}_r$ are estimated fairness parameters for the manufacturer and retailers, respectively. Full details provided in Appendix A.3.

The bottom of Table 1.7 shows the parameter estimates by treatment. To begin, the random errors in transfer prices ($\hat{\theta}$) are lower in DR-R than DR-M, which agrees with the analysis in Table 1.4 showing that retailers give up little profit by setting marginally sub-optimal transfer prices. Manufacturer-fairness concerns ($\hat{\lambda}_m$) are generally larger than retailer-fairness concerns ($\hat{\lambda}_r$) in all treatments, but one one must recognize that manufacturers always earn a disproportionately larger split of the overall supply chain profit (i.e., average manufacturer utility is still higher than average retailer utility).

A closer look at the fairness estimates for a given role, across treatments, yields additional insights. Starting with the manufacturer, the estimates are highest in CR and DR-M. This is consistent with the notion that the predicted distribution of profits is largest in these two treatments. Continuing with DR-R and DR-0, the estimates are lower than DR-M and CR, which is natural considering that they should lead to the most equitable distribution of profits. However, whereas theory predicts that DR-R and DR-0 should be identical, we observe that manufacturers have higher fairness concerns in DR-R. This is intuitive if one recalls that, in our DR-R data, retailers first set a noisy transfer price higher than the normative prediction of zero. As shown in our analysis in Table 1.4, this translates into a higher profit for the manufacturer and hence a larger (conditional) predicted difference in profits. Therefore, manufacturer-fairness concerns should indeed be higher in DR-R relative to DR-0. Last, the manufacturer's fairness estimate is lowest in Baseline. One possible explanation for this is that the normative theory, under inventory sharing, predicts that a manufacturer should set a higher wholesale price than in the Baseline setting (i.e., a manufacturer should recognize that retailers have a risk-pooling benefit under inventory sharing and set a higher price). But, our results suggest that a manufacturer does not follow this prescription under retailer inventory sharing, and instead offers a wholesale price that is significantly below what theory predicts, leading to fairness estimates that are higher in the inventory-sharing conditions than the Baseline setting.

A similar story emerges with regards to retailer-fairness concerns across treatments. For instance, retailer-fairness estimates are relatively higher in CR and DR-M, which is expected given that theory predicts a large difference in profits. Comparing the two directly, the higher estimate in CR is likely due to two retailers jointly setting a quantity: if one of them has fairness concerns and the other does not, this could lead to lower quantities and thus a higher fairness estimate. It is also noteworthy that the retailer's estimate in Baseline, which is roughly average across the treatments, is nearly identical to that found in past supply chain experiments with a single retailer (e.g., [40] estimate the value to be 0.050 versus 0.054 in our study).[8] Last, another interesting observation is that DR-0 has the highest retailer-fairness estimate (and far higher than DR-R). Recall at the end of Section 1.6, we conducted an analysis indicating that retailers did not excessively understock in DR-R for transfer prices that were equal to or close to zero. This suggests that retailers react negatively to being required to share inventory at an exogenous transfer price of zero, resulting in them understocking. As a consequence, procedural fairness may be manifesting itself in the higher retailer-fairness estimate for DR-0.[9]

## 1.8 Robustness Check: Revenue Sharing with Decentralized Retailers

As a robustness check of our results, we also investigate an alternative type of contract with inventory sharing, rather than a simple wholesale price contract. In particular, we run two additional treatments which consider revenue sharing between the retailers and the manufacturer. Our objective is to determine whether the behavioral deviations observed in our main experiment hold for a different

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

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

contract under inventory sharing, which has not been explored before.

Given their favorable performance, we consider the two decentralized inventory-sharing strategies with endogenous transfer prices, DR-M and DR-R, but add an exogenous revenue share for the manufacturer of 30%.[10]

### 1.8.1 Normative Theory

We first briefly summarize the normative theory for a revenue-sharing contract with decentralized retailers. Consider the normative theory outlined previously in Section 1.3. By backward deduction, with revenue sharing, retailer $i$'s expected profit function is expressed by Equation (1.11)

$$\tilde{\pi}_{r,i}^{d} = \psi\mathbb{E}\left[p \min(d_i, q_i) + tT_i + (p-t)T_j\right] - wq_i. \tag{1.11}$$

where $\psi$ is the ratio of revenue-sharing. The optimal stocking quantity under Nash equilibrium $(\tilde{q}_i^d, \tilde{q}_j^d)$ is given by Equation (1.12)

$$\alpha(q_i) - \beta_i(q_i, q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i, q_j)\left(\frac{p-t}{p}\right) = \frac{\psi p - w}{\psi p}, \tag{1.12}$$

and we have the following proposition (please see Section 1.8.2 for all proofs).

**Proposition 1.1.** *There exists a unique Nash equilibrium, $(\tilde{q}_i^d, \tilde{q}_j^d)$, in a revenue-sharing contract with the decentralized retailer inventory-sharing strategy.*

[10] An exogenous revenue share coincides with past experiments on revenue sharing with a single retailer (e.g., [73]) and also ensures that the number of decisions are the same as our main experiment.

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

**Lemma 1.1.** *Under a revenue-sharing contract, a decentralized retailer i's optimal inventory level, $\tilde{q}_i^d$, is increasing in the transfer price t.*

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

$$\tilde{\pi}_m = (w - c)(q_i + q_j) + (1 - \psi)\mathbb{E}[p \min(d_i + d_j, q_i + q_j)]. \tag{1.13}$$

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

**Lemma 1.2.** *When the manufacturer optimally sets its wholesale price, its expected profit increases as the transfer price t increases.*

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

In this new experiment we refer to DR-M and DR-R under a revenue-sharing contract as DRRS-M and DRRS-R. All experimental protocols are identical to the original design in Section 1.4. We set the revenue-sharing ratio $\psi$ to 0.7, which generates retailer profit predictions that are close to the normative predictions without revenue sharing. All normative predictions for the revenue-sharing treatments are shown in Table 1.8. Each treatment consisted of 42 participants from the same participant pool as our main experiments. Average payment per participant was $26.94.

Table 1.8: Normative Predictions under Revenue Sharing with Decentralized Retailers

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

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

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

(formerly 83.60% and 69.26%).

We provide contract terms in Table 1.9. Comparing the observed decisions and normative predictions to one another (significance given in the middle two columns), we see the same patterns as in our original experiment. In particular, transfer prices neglect to be set at extreme values (both $p < 0.004$, the corrected critical p-value) and wholesale prices are set lower than predicted (both $p < 0.004$). This latter finding suggests that wholesale prices being offered more generously appears to be a robust result, as opposed to an anchoring bias. Also, quantities are set rather well, albeit slightly low.

Table 1.9: Average Contract Prices, Quantities, and Predictions under Revenue Sharing

| | Observed Results | | Normative Predictions | | Behavioral Predictions | |
|---|---|---|---|---|---|---|
| | DRRS-M | DRRS-R | DRRS-M | DRRS-R | DRRS-M | DRRS-R |
| $t$ | 19.85 | 9.50 | 30.00$^{\dagger}$ | 0.00$^{\dagger}$ | 19.85 | 9.25 |
| | (0.99) | (0.71) | | | | |
| $w$ | 11.91 | 10.87 | 12.85$^{\dagger}$ | 11.74$^{\dagger}$ | 11.85 | 10.86 |
| | (0.37) | (0.29) | (0.11) | (0.07) | | |
| $q$ | 45.64 | 44.68 | 47.19 | 45.30 | 45.75 | 44.06 |
| | (1.27) | (1.05) | (0.71) | (0.42) | | |

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

The left-hand side of Table 1.10 presents profits and efficiency in our original treatments, whereas the right-hand side depicts the new revenue-sharing treatment data. Beginning with each party's profits in Table 1.10, while manufacturers benefit from revenue sharing, retailers are actually worse off. Regarding efficiency, only DRRS-R achieves a higher efficiency than without revenue sharing, 81.94% versus 76.54% originally ($p < 0.008$, the corrected critical p-value). Combining these observations we have:

**Result 1.6.** *Under a decentralized retailer inventory-sharing strategy where the retailers set the transfer price, revenue sharing improves supply chain efficiency over wholesale price contracts, but at the expense of retailers.*

Table 1.10: Average Profits and Efficiency between the Original Experiment and Revenue-Sharing Experiment

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

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

Returning to Table 1.9, the final columns depict treatment-specific predicted decisions from the full-fairness model presented in Section 1.7. Comparing observed decisions to the fairness predictions, there are no significant differences. The observed transfer price in DRRS-M is the same as the prediction, both 19.85, and is very close in DRRS-R, 9.50 versus 9.25. Wholesale prices are 11.91 versus 11.85 in DRRS-M and 10.87 versus 10.86 in DRRS-R. Regarding stocking quantities, they are 45.64 versus 45.75 in DRRS-M and 44.68 versus 44.06 in DRRS-R.

Overall, in this alternative contract setting, we observe similar deviations in decisions as in our main experiment and similar-quality predictions from our fairness model, providing further support for fairness as an explanation for the observed deviations.

### 1.8.2 Proofs

*Proof of Proposition 1.1* To prove there is a unique Nash equilibrium, we have to show the slope of the reaction function is monotonic with an absolute value less than 1. Recall that $f^c(d_i, d_j)$ is the joint probability density function of demands. First define the following marginal probabilities:

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

Implicit differentiation of Equation (1.12) leads to

$$\frac{\partial \tilde{q}_i^d}{\partial \tilde{q}_j^d} = -\frac{tb_{ij}^1 + (p-t)g_{ij}^1}{p(b_{ij}^2 + g_{ij}^2) + t(b_{ij}^1 - b_{ij}^2) + (p-t)(g_{ij}^1 - g_{ij}^2)}. \tag{1.14}$$

Equation (1.14) is a special case of Equation (11) in [110], which has been shown that the slope of the reaction function is non-positive and less than 1 in absolute value. □

*Proof of Lemma 1.1* The proof follows the proof of Lemma 1 in [114]. At equilibrium, the Implicit Function Theorem and the symmetry of retailer $i$, $j$ lead to

$$\frac{\partial \tilde{q}_i^d}{\partial t} = \frac{(\partial^2 \tilde{\pi}_{r,i}^d / \partial q_i \partial t)[(\partial^2 \tilde{\pi}_{r,i}^d / \partial q_i \partial q_j) - (\partial^2 \tilde{\pi}_{r,i}^d / \partial^2 q_i^2)]}{|H|}, \tag{1.15}$$

where $|H|$ is the positive determinant of the Hessian matrix. For the numerator of Equation (1.15), recall the notations of marginal probabilities in proof of Proposition 1.1, and we have

$$\frac{\partial^2 \tilde{\pi}_{r,i}^d}{\partial q_i \partial t} = \beta_i(q_i, q_j) + \gamma_i(q_i, q_j) > 0, \tag{1.16}$$

$$\frac{\partial^2 \tilde{\pi}_{r,i}^d}{\partial q_i \partial q_j} - \frac{\partial^2 \tilde{\pi}_{r,i}^d}{\partial^2 q_i^2} = pf(q_i) - tb_{ij}^2 - (p-t)g_{ij}^2 > 0, \tag{1.17}$$

where $f(q_i)$ is the probability density function of retailer $i$'s demand. Inequality (1.17) is derived by $p \geq t$, $f(q_i) > b_{ij}^2$ and $f(q_i) > g_{ij}^2$. Therefore, the numerator of Equation (1.15) is greater than 0, which means $\partial \tilde{q}_i^d / \partial t > 0$ at equilibrium. □

*Proof of Lemma 1.2* The derivative of manufacturer expected profit function (1.13) with respect to transfer price $t$ is:

$$\frac{\mathrm{d}\tilde{\pi}_m}{\mathrm{d}t} = \frac{\partial \tilde{\pi}_m}{\partial t} + \frac{\partial \tilde{\pi}_m}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial \tilde{\pi}_m}{\partial \tilde{q}^d}\frac{\partial \tilde{q}^d}{\partial t}. \tag{1.18}$$

Note that in Equation (1.18) the subscript of retailer is dropped as quantities are equal at equilibrium. The first term of the right hand side is zero as $t$ is cancelled out at equilibrium. The second term is also zero when $w$ is optimal. Equation (1.18) can be rewritten as

$$\frac{\mathrm{d}\tilde{\pi}_m}{\mathrm{d}t} = 2(w-c)\frac{\partial \tilde{q}^d}{\partial t}. \tag{1.19}$$

From Lemma 1.1 we know that $\partial \tilde{q}^d / \partial t > 0$. Because $w > c$, we have $\mathrm{d}\tilde{\pi}_m / \mathrm{d}t > 0$. □

## 1.9 Discussion

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

### 1.9.1 Managerial Implications

From a managerial perspective, it is unlikely that firms will have the ability to choose among all four of the different inventory-sharing strategies we explore. However, at a minimum, we posit that retailers are able to consider at least one (or more) of these strategies. Thus the first key question for retailers is *if* they should adopt an inventory-sharing strategy. Our results suggest that the answer is yes: observed retailer profits are higher (or at least as high) under all inventory-sharing strategies compared to the Baseline environment (see Table 1.3). Given this observation, the next question becomes *how* retailers should adopt inventory-sharing strategies. First, consider the choice between centralized and decentralized inventory sharing. Our results indicate that retailers prefer decentralized inventory-sharing strategies to a centralized one. Second, regarding the decision authority over the transfer price, our results are consistent with theory and indicate that retailers prefer to negotiate the transfer price rather than have it determined by another party such as the manufacturer.

Turning to the manufacturer, our results are consistent with theory and indicate that manufacturers prefer serving decentralized retailers when they can set the transfer price, although it is worth noting that there is a large gap between the normative manufacturer profit prediction and the observed manufacturer profit (Table 1.3). So, while manufacturers still prefer this setting to all others, they are less effective than they could be at using wholesale and transfer pricing power to extract higher profits. Our results are also consistent with normative theory, in that manufacturers are worst off when retailers are decentralized and have authority over the transfer price. These results suggest that manufacturers should exert effort to try to control the terms of inventory transfer when possible.

### 1.9.2 Contribution to Literature

From a research perspective, our study contributes to the behavioral literature on supply chain contracting and inventory sharing. While a majority of initial supply chain contracting experiments considered a single manufacturer and single retailer (e.g., [69, 15, 138], see [30] for a summary), more recently, a number of important studies have extended such a setting and allowed for multiple retailers and inventory sharing. Given that they are the first to evaluate such a setting experimentally, these works consider a one-tier supply chain and focus on order quantities [e.g., 61, 141] and transfer prices [e.g., 86, 74].

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

In many ways our work complements the existing theoretical research that studies coordination mechanisms for decentralized units. For instance, [27] and [11] investigate coordination issues between a "revenue-maximizing" marketing department and a "cost-minimizing" manufacturing department, within an organization. They find that allowing the two departments to operate in a decentralized manner, with correctly set penalty terms, can achieve outcomes that are equal to or even better than when the two departments are centralized. Our study complements such research in finding that similar outcomes can be achieved by behavioral tendencies. Of course, we would be remiss to say that we are the first supply chain experiment to find evidence of fairness. However, by conducting a novel experiment on a two-tier supply chain with alternative inventory-sharing strategies, we are able to dig deeper and identify the specific impacts that such behavioral biases have on contracts and profits.

## 1.10 Conclusion

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

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

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

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

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

CHAPTER 2

# TRADE CREDIT AND BANKRUPTCY RISK IN SUPPLY CHAINS

## 2.1 Introduction

Insufficient cash flow has always been a critical challenge for firms, especially for small-to-medium enterprises (SMEs). Therefore, firms often seek financing. According to the World Trade Organization [130], up to 80% of trade is financed by credit or credit insurance. There are two prominent financing sources: financing provided by a third-party financial institution outside of the supply chain, such as a bank, and financing offered by supply chain members. Bank financing has been popular but can also be costly and unavailable to SMEs with a short credit history. As a result, providing financing within supply chains has become increasingly popular in recent years, not only because technology has made such financing much easier [96] but also because it is usually more accessible and affordable. Moreover, such financing can improve performance because the creditor, as a supply chain member, may have better information about the debtors and can make financial and operational decisions in combination [75, 118, 121].

One popular supply chain financing tool is trade credit, i.e., a buyer purchases from a supplier through credit and repays after a period of time, usually after increasing its cash flow from selling products. A common example of trade credit is 2/10 net 30, i.e., a buyer can either pay within ten days and receive a 2% discount, or pay within 30 days without receiving any discount. The full price can be treated as the discounted price plus the financing cost for an additional 20 days. Trade credit provides liquidity to buyers, allowing them to purchase a desired quantity with relatively low financing cost compared with bank financing. It is also an important financing source for firms that are short of bank credit [52]. Trade credit is not just beneficial for retailers with limited cash flow. Even retailers without capital constraints also purchase via trade credit from smaller and weaker suppliers [95]. Suppliers as the financing provider can also take advantage of trade credit. By setting appropriate contract terms, suppliers may be able to indirectly affect the retailer's financing decisions or induce higher ordering quantities from retailers. Sometimes, suppliers with weak market power can also use trade credit as a tool to compete with other suppliers [84].

Although trade credit brings benefits to the supply chain, it also comes with downsides. If retailers default due to demand uncertainty, which is not uncommon, suppliers will incur losses. In this case, retailers who bear limited liability file for bankruptcy, but only partially repay suppliers (as much as they can), not in full. By providing financing to retailers, such suppliers also partially share bankruptcy risk with them. However, compared with a financing provider outside of the supply chain, suppliers may be better informed about a retailer's financial status, such as inferring from the received ordering quantity, to better manage the risk of providing financing. In addition, suppliers may also be able to take advantage of retailers' bankruptcy risk and extract more profits because their contract terms could affect retailers' financing choices.

Given its potential benefits and risks, trade credit has been well studied analytically. In a simple two-tier supply chain consisting of a supplier and a capital-constrained retailer, the normative theory predicts that a retailer with higher bankruptcy risk benefits more from limited liability and should set a higher quantity (than if without bankruptcy risk), making the supplier charge a higher wholesale price to extract more profits. While the normative theory assumes rational decision-makers, it is unclear whether behavioral biases can drive decisions from the optimal predictions. For example, retailers may be influenced by the endowment effect [67] and value their initial capital more. Moreover, if behavioral biases do play a role in decision-making, the magnitude of the biases and how they will change with respect to the degree of bankruptcy risk is unknown. Because these decisions are usually made by managers in practice, it is essential to understand the problem from a behavioral perspective.

In this paper, we investigate how bankruptcy risk affects supply chain decisions and performance. We consider a two-tier supply chain consisting of a supplier (she) and a capital-constrained retailer (he). The supplier offers a per-unit wholesale price to the retailer, who purchases units from the supplier to satisfy random demand. We consider a trade credit contract, i.e., the retailer can choose between purchasing with his limited capital and purchasing through trade credit provided by the supplier. In the latter scenario, the retailer repays the supplier after the demand realization. If demand is too low, the retailer, who bears limited liability, will transfer any remaining cash to the supplier, including any revenue earned from demand, and go bankrupt. To understand the impact of behavioral biases, we adopt an experimental approach, combined with existing analytical results. We conduct controlled lab experiments with human participants to test the normative theory. Specifically, the experiments include three treatments with different levels of the retailer's initial capital, which directly affect the retailer's bankruptcy risk. With this design we aim to answer the following research questions. First, how does a human retailer with bankruptcy risk make quantity decisions compared to theory? Second, how does a human supplier set wholesale prices under a trade credit contract and retailer bankruptcy risk? If we find that decisions indeed deviate from the normative theory, we are also interested in what behavioral bias might be driving the decisions.

Our experimental results indicate that human retailers with medium and low levels of initial capital significantly understock, to ensure a lower bankruptcy risk. Anticipating such understocking behavior by retailers, we find that human suppliers set much lower wholesale prices to induce higher quantities. Conversely, retailers with a high initial level of capital choose to slightly overstock, bringing them unnecessary bankruptcy risk, and suppliers set wholesale prices close to the optimal prediction. As a result, the order of net retailer profit (excluding initial capital) is *reversed* relative compared with the normative prediction. In an effort to account for this deviation, we find that assuming a reference-dependent retailer, only, can well predict the observed decisions, including wholesale prices and stocking quantities. Specifically, the retailer's initial capital, as a simple fixed reference point, is sufficient to fit the decisions well. Test this behavioral model experimentally, which assumes that the supplier is rational but that the retailer is reference dependent, we conduct additional experimental treatments which automate the retailer role (i.e., make decisions in line with the normative theory). As opposed to our original experiments, in these follow-up treatments with automated retailers, we observe that supplier decisions largely conform to the normative predictions. This empirically validates our behavioral hypothesis that reference-dependent retailers are driving the deviations from the normative predictions, for both parties, in our main experiments.

We believe our work contributes to both practitioners and researchers. Regarding the former, we show that risky retailers prefer understocking to reduce bankruptcy risk, rather than taking advantage of limited liability, which contradicts the normative theory. As a result, suppliers should focus on inducing higher quantities when setting contract terms of trade credit. In terms of contribution to research, supply-chain finance has rarely been studied experimentally, despite humans playing an integral role, in practice. We show that behavioral factors, notably reference dependence, play an important role in this setting and can significantly drive decisions away from rational predictions. Therefore, we believe that future behavioral research in the field of supply chain finance is critical, especially given that many financial components remain relatively unexplored, such as interest rates, hedging, and factoring. In addition, our work can complement the existing behavioral operations literature by showing that supply chain decisions can be very different under bankruptcy risk, which is often present in practice.

The rest of the paper is organized as follows. In Section 2.2 we provide a literature review of relevant research and identify our contribution to literature. In Section 2.3, we outline the normative model and theoretical benchmark. In Section 2.4, we design a controlled lab experiment for human participants, and discuss behavioral hypotheses. Experimental results are summarized and analyzed in Section 2.5. In Section 2.6, we build a plausible behavioral model of reference dependence to account for the observed deviation in decisions. In Section 2.7, we conduct additional treatments as a robustness check for our behavioral model. Finally, we summarize our work and discuss managerial implications and research limitations in Section 2.8.

## 2.2 Literature Review

Inventory decisions with trade credit contracts have been studied for decades in the operations management literature, with analytical modeling and empirical research being the primary methodologies. Overall, [113] provide a comprehensive review of the trade credit literature in operations management. Beginning with analytical studies, [59] study the basic lot-size model with trade credit and find that order quantity and payment time have to be jointly decided to reach optimality. [38] find that a capital-constrained newsvendor will borrow funds from a bank and order less than would be ideal if the borrowing cost were not too high. They also show that the channel can be coordinated by a non-linear loan schedule. [81] study the impact of financial constraints on sharing inventory risk in a supply chain. In their results, the existence of financial constraints (for both a supplier and retailer) will make the supplier choose to share inventory risk with the retailer, whereas the supplier always prefers taking full inventory risk when without financial constraints. [82] look at trade credit from a supply chain perspective and show that a supplier can fully coordinate the supply chain (with a retailer seeking financing) using both trade credit and markdown allowance.

Trade credit is also studied in more complex settings. For instance, [56] study a multi-period stochastic inventory model and prove that credit terms only affect the value of parameters but not the structure of the optimal policy. [33] considers a multiple-item setting and finds that financing can distort a retailer's inventory decision. Further, [33] shows that such distortion can be mitigated through supplier financing, as the supplier can observe the actual order quantities prior to setting financing terms. [104] investigate trade credit when competition between suppliers is considered. They prove that trade credit softens price competition and leads to higher equilibrium profits, i.e., a horizontal benefit, whereas most research focuses on vertical benefits of trade credit. On the other hand, [34] point out that supplier competition may also lead to a free-rider problem: trade credit extended by a supplier increases the buyer's cash purchase from another supplier. They show that buyers with diverse suppliers receive less trade credit than those who have more concentrated suppliers. In a multi-period model, [91] consider a two-level trade credit structure, i.e., from supplier to buyer and from buyer to customer, and show that maximizing working capital at the end of the horizon is equivalent to minimizing the total cost within the horizon. [98] shows that when there is competition between buyers who have access to external capital, trade credit can incentivize buyers to order more and compete more aggressively, while buyers also benefit from lower wholesale prices.

In addition to supplier financing, other topics receive interest in the literature. For example, in some cases suppliers may be in a financially distressed position and seek financing from retailers. [121] show that such "buyer financing" can lead to lower financing costs and higher service levels and thus significantly benefit both parties. [118] compare buyer financing with bank financing and find that the former brings more flexibility to the buyer under symmetric information or even better performance when asymmetric information exists. [37] shows that the high implicit interest rate of trade credit is due to the combination of an insurance premium and a default premium. [132] identify two main roles of trade credit insurance: a) smoothing the supplier's cash flows, and b) monitoring the buyer's continued creditworthiness after contracting.

Along with analytical results, there are also an abundance of empirical findings for trade credit in recent years. Although it is common for powerful suppliers to provide trade credit to small buyers, [95] find that even retailers without capital constraints purchase via trade credit from smaller and weaker suppliers. [84] study how, when competition exists, trade credit affects firm performance and find that when trade credit offered by suppliers exceeds industry-average levels, buyers' performance is negatively associated with the amount of trade credit. Although it is well documented that diversification can be more profitable in many circumstances, it can also increase the recovery cost for a financially distressed buyer upon default. Taking advantage of a quasi-natural experimental setting provided by a regulatory shock, [7] study how a distressed buyer's sourcing strategy is affected by bankruptcy risk. They show that capital-constrained buyers are forced to under-diversify compared with those without capital constraints. [8] find that provided and received trade credit have different impacts on firm value: the former has a negative direct and a positive indirect effect, and the latter has the opposite effect. They argue that the main reason is the disparate nature of dependence in the supply chain.

Among the existing literature, two papers are closely related to the normative model in our work: [75] and [133]. Both papers study a supply chain of a supplier and a capital-constrained retailer, with endogenous contract terms. [75] compare bank-financing only and supplier-financing only scenarios. In the former, bank interest is assumed to be competitively priced. In the latter, the supplier provides an early-payment-discount contract, i.e., the retailer can either pay upon delivery of the order at a discounted price or repay the supplier after the demand realization at a regular price. The retailer's financing decision is fully determined by the inventory decision. They show that a capital-constrained retailer always prefers supplier financing, and that the supplier always sets its interest rate at a risk-free rate. In a recent paper, [76] further study the impact of credit ratings on inventory, pricing, and financial terms decisions in the same framework. [133] then extend the model by allowing the retailer to take advantage of both financing sources. In particular, given an early-payment-discount contract, the retailer can determine the quantity purchased at the discounted price and at the regular price, respectively. He can also use a bank loan if his cash is insufficient to cover the discounted-price purchase. In this way, the retailer has to decide how much to borrow from the bank and how much of trade credit should be used, i.e., a portfolio decision. They also validate the model using empirical data.

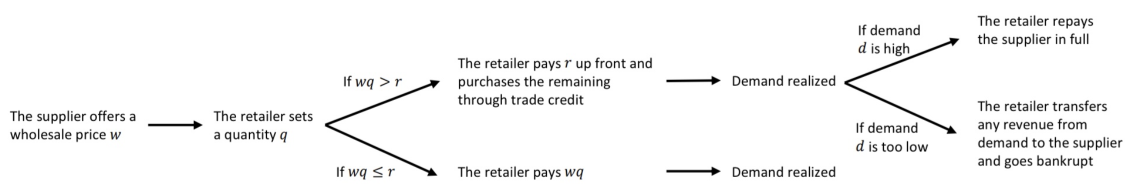
To our knowledge, none of the existing studies in the operations management literature investigate trade credit and bankruptcy risk from a behavioral perspective. Taking this into account, and to better understand the role of behavioral factors in a clean environment, in this paper we simplify the model studied in [75] and [133] and focus on supplier financing. The reasons are twofold: a) existing studies show that supplier financing has an advantage over bank financing [75, 33], and b) it is important to understand how a supplier reacts to a retailer's bankruptcy risk, which affects supply chain performance.

## 2.3 Normative Model

Consider a retailer (he) ordering a product from a supplier (she) at wholesale price $w$ per unit, and selling to the market at unit price $p$. Both the retailer and the supplier are risk-neutral. The supplier produces the product at a per-unit cost $c$. The retailer is capital-constrained and starts with initial capital $r$. Before the selling season begins, the retailer determines a quantity $q$ based on the wholesale price $w$ to satisfy a random demand $d$.

Because the retailer is capital-constrained, he may lack sufficient cash and require financing. Depending on $w, q$ and $r$, there are two scenarios for the retailer's financing outcome. If $wq \leq r$, the retailer's capital is sufficient to cover his purchasing cost. The retailer pays the supplier $wq$ up front. If $wq > r$, the retailer will need financing from the supplier. Specifically, the retailer pays $r$ to the supplier up front, and the rest after the demand realization. If demand is high enough, the retailer can repay the supplier in full, $wq - r$. If demand is low, the retailer transfers all of his revenue from realized demand to the supplier, and earns zero profit, i.e., he goes bankrupt. By bankruptcy, we assume that the retailer bears limited liability. Figure 2.1 summarizes the sequence of events.

Figure 2.1: Sequence of Events




For the normative model introduced in this section, we make some underlying assumptions related to the financial aspects. As such, the model is parsimonious, but captures the key trade-off. First, we assume a perfect market without tax or bankruptcy cost. Second, both parties are creditworthy and the retailer will repay any loan obligations to the extent possible. Finally, there is no information asymmetry, i.e., price, cost and the demand distribution are all common knowledge.

### 2.3.1 Retailer's Decision

Given a wholesale price $w$, the retailer determines $q$ (and the choice of financing) to maximize his expected profit:

$$\mathbb{E}[\pi_r] = \mathbb{E}[p\min(q,d) - wq + r]^+. \tag{2.1}$$

Define $k = (wq - r)^+/p$ as the retailer's bankruptcy threshold, i.e., the minimum demand such that the retailer does not go bankrupt ($k$ will be zero if no financing is needed). Equation (2.1) can be rewritten as

$$\mathbb{E}[\pi_r] = \begin{cases} p\mathbb{E}[\min(d,q)] - p\mathbb{E}[\min(d,k)] & \text{if } wq - r > 0, \quad (2.2a) \\ p\mathbb{E}[\min(d,q)] - wq + r & \text{Otherwise.} \quad (2.2b) \end{cases}$$

Let $f(\cdot)$ and $F(\cdot)$ be the probability density function (PDF) and cumulative distribution function (CDF) of demand $d$, respectively. Define $\bar{F}(q) = 1 - F(q)$; the retailer's optimal quantity derived from Equation (2.2a) and (2.2b) is

$$q^*(w) = \begin{cases} \bar{F}^{-1}((w/p)\bar{F}(k)) & \text{if } wq - r > 0, \quad (2.3a) \\ \bar{F}^{-1}(w/p) & \text{Otherwise.} \quad (2.3b) \end{cases}$$

[75] show that there is a one-to-one match between $w$ and $q^*(w)$. Let $q_l, q_u$ be the two solutions for

$$q\bar{F}(q) = \frac{r}{p}, \tag{2.4}$$

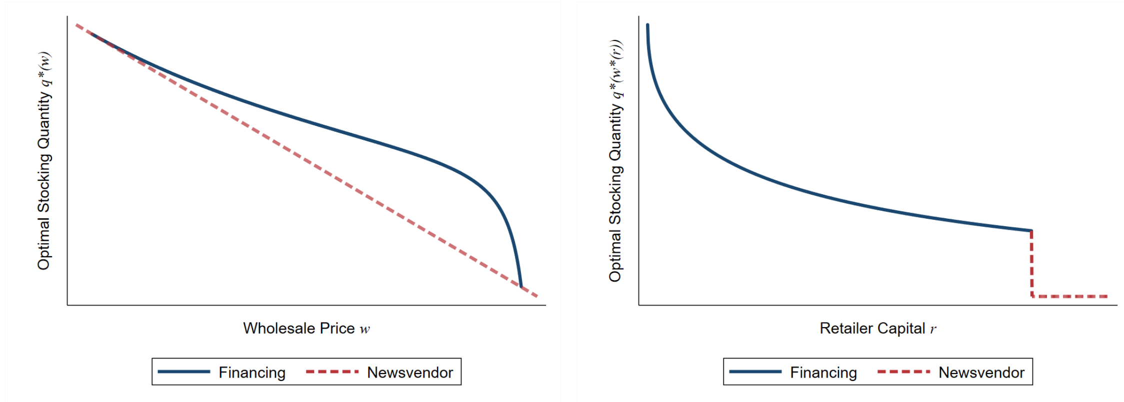
with $q_l \leq q_u$, and define $w_l = p\bar{F}(q_l)$, $w_u = p\bar{F}(q_u)$. The "financing region" corresponding to Equation (2.3a) is equivalent to $w \in (w_u, w_l)$, and the "no-financing region" of Equation (2.3b) is equivalent to $w \in [0, w_u] \cup [w_l, p]$, respectively.

Equation (2.3a) shows that the retailer requires financing when the wholesale price is intermediate. If the wholesale price is too low, the unit purchasing cost is so low that the retailer has enough cash to cover. If the wholesale price is too high, the retailer will order too low a quantity, so that no financing is needed. Note that $(w_u, w_l)$, i.e., the financing region, depends on $r$. In other words, the region is larger when the retailer has less capital and therefore is exposed to higher bankruptcy risk.

Figure 2.2 shows how $q^*$ varies with respect to $w$ and $r$. In Figure 2.2a, supplier financing with limited liability leads to a higher quantity compared with the standard newsvendor quantity. The intuition is that when demand is low, the retailer is actually protected by limited liability (otherwise he would have earned negative profits). Therefore, the retailer will choose a higher quantity in general, which leads to a higher profit (driven by those instances when demand is high). Figure 2.2b shows how $q^*$ varies with respect to retailer capital $r$ given the corresponding optimal wholesale price. Again, the retailer will order more when he has less capital (i.e., more risky), as he is more protected by limited liability. The optimal quantity decreases in $r$ and will become the newsvendor quantity when the retailer has high initial capital (i.e., is sufficiently rich).

Figure 2.2: Optimal Quantity under Supplier Financing

(a) Wholesale Price (b) Retailer Capital




### 2.3.2 Supplier's Decision

We also assume that the supplier is capital-constrained starting with initial capital $r_s$ and bears limited liability as well. However, [75] show that the supplier's optimal decision is independent of $r_s$. Given the retailer's optimal quantity, the supplier's expected profit function can be written as

$$\mathbb{E}[\pi_s(w)] = \begin{cases} p\mathbb{E}[\min(d,k)] - cq + r + r_s & \text{if } q_l < q^*(w) < q_u, \quad (2.5a) \\ (w-c)q + r_s & \text{Otherwise.} \quad (2.5b) \end{cases}$$

To illustrate the supplier's optimal decision, define the following equations:

$$\bar{F}(q) - qf(q) = \frac{c}{p}. \tag{2.6}$$

$$\frac{\bar{F}(q) - qf(q)}{1 - (wqf(k))/(p\bar{F}(k))} = \frac{c}{p}, \tag{2.7}$$

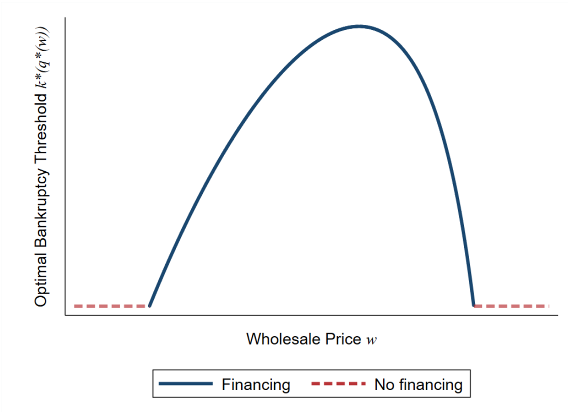
Let $\bar{q}$ and $\hat{q}$ be the solutions to Equations (2.6) and (2.7), respectively. [75] prove that the supplier's optimal wholesale price $w^*$ and corresponding retailer's decision are as follows:

- (No financing) If $\bar{q} \leq q_l$, $w^* = p\bar{F}(\bar{q})$ and the retailer does not need financing.
- (Financing) If $\bar{q} > q_l$, $w^*$ is derived from $\hat{q} = \bar{F}^{-1}(w/p)\bar{F}(k)$ and the retailer needs financing.

To intuitively understand these results, recall that there is a one-to-one match between $w$ and $q^*(w)$. Therefore, the supplier can indirectly determine the retailer's financing decision by setting a wholesale price. The "no financing" case is when it is optimal for the supplier to not finance the retailer. The retailer's optimal decision is $\bar{q}$, which is the standard newsvendor quantity and independent of the retailer's capital $r$. In this case, $\bar{q}$ is smaller than $q_l$ and is not in the financing region $(q_u, q_l)$. The supplier's optimal wholesale price is $w^* = p\bar{F}(\bar{q})$. Such "no financing" case is when the retailer has high initial capital (i.e., is relatively rich). When the retailer has low initial capital (i.e, is relatively poor), the financing region becomes larger and $\bar{q} > q_l$. In this case, $w = p\bar{F}(\bar{q})$ cannot be optimal, and the supplier will set the price $w^*$ following the "financing" case. We refer interested readers to [75] for a general solution.

Assuming the retailer optimally responds to the supplier's wholesale price, the retailer's degree of bankruptcy risk does not vary linearly with respect to the price. Figure 2.3 shows that the retailer's bankruptcy threshold $k(q^*(w))$ exhibits an inverted U-shape in the financing region. Note that this $k(q^*(w))$ can be regarded as

Figure 2.3: Optimal Bankruptcy Threshold with respect to Wholesale Price

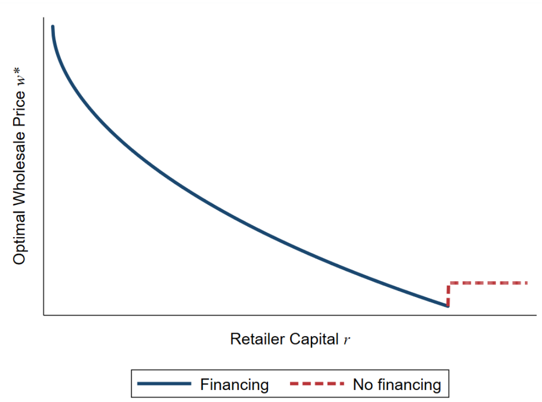



the bankruptcy risk that the supplier offers to the retailer. In Figure 2.3, a certain degree of bankruptcy risk can be achieved through a low or a high wholesale price. This is because the retailer will order more given a lower wholesale price, which results in a similar degree of bankruptcy risk. In other words, the supplier may not be able to reduce the retailer's risk by offering a lower price. We discuss this further when we propose behavioral hypotheses in Section 2.4.2.

Turning to how $w^*$ varies with respect to $r$, Figure 2.4 provides a numerical example. Qualitatively, when in the financing case, $w^*$ decreases in the retailer's capital $r$. In other words, because a retailer with less capital prefers to set a higher quantity, the supplier will set a higher wholesale price to extract more profit. However, the financing $w^*$ can be lower than the no-financing $w^* = p\bar{F}(\bar{q})$, when the retailer's initial capital is high, but not so high that the supplier prefers not to finance him.

While theory predicts that a retailer with higher bankruptcy risk should take

Figure 2.4: Optimal Wholesale Price with respect to Retailer Capital




advantage of limited liability and order a higher quantity (relative to lower bankruptcy risk), it is unclear whether human decision-makers would behave in this way. For instance, a human retailer with high bankruptcy risk may fail to realize the benefits of limited liability and understock to reduce the risk. Moreover, if behavioral biases are driving decisions, how the impact of biases varies with respect to the degree of bankruptcy risk is unknown. Finally, we are also interested in how these biases (if they exist) affect supply chain performance. To investigate the impact of behavioral factors, in the next section we design controlled laboratory experiments with human participants and discuss behavioral hypotheses.

## 2.4 Behavioral Experiment

### 2.4.1 Experimental Design

We conducted controlled laboratory experiments to test the theory introduced in Section 2.3. As we are mainly interested in the retailer's financial risk, we consider two treatments, high risk (HR) and low risk (LR), by varying the retailer's initial capital $r$. In addition, we include a third treatment where the retailer has zero bankruptcy risk predicted by theory, denoted as NR (no risk). Note that NR is essentially a newsvendor case under optimal decisions bankruptcy risk only exists for retailers if they over-order by extreme amounts).

Turning to our parameters, we set the unit retailer selling price $p = 30$ and the supplier production cost $c = 10$. The retailer faces an integer demand drawn from a uniform distribution between 0 and 100. Regarding the supplier's initial capital, we set $r_s = 400$ across all treatments. The reasons are twofold: a) the normative theory shows that the supplier's optimal decision is unaffected by $r_s$, and b) $r_s = 400$ makes supplier bankruptcy rarely happen,[1] which means that we can focus on the retailer's bankruptcy risk. Note that if a supplier goes bankrupt in the game, she is also protected by limited liability, i.e., earns at least a profit of zero.

[1] The average supplier bankruptcy rate in our data is 0.39%.

We set $r$ as 700 (NR), 400 (LR), and 100 (HR). The theoretical predictions of all treatments are shown in Table 2.1. Understanding that the retailer is more protected by limited liability when he faces higher risk, the supplier chooses a higher wholesale price to extract more profit, resulting in a predicted wholesale price of 20 in NR, 21.47 in LR, and 25.87 in HR. The predicted quantity exhibits a similar pattern, as 33.33 in NR, 38.73 in LR, and 42.49 in HR. The retailer's risk is reflected by the bankruptcy threshold $k$. In NR, the retailer has no bankruptcy risk under the optimal decisions ($k = 0$), while it is predicted to be $k = 14.38$ in LR and $k = 33.31$ in HR.

Regarding predicted profits, supplier expected profit increases from 733.33 in NR, to 813.16 in LR, to 907.87 in HR, as the supplier is better able to extract profit from the retailer, (as $r$ decreases). On the other hand, the retailer's predicted expected profit decreases from NR to HR, for both total profit and net profit (excluding initial capital, reported in square brackets). For example, the predicted net retailer profit is 166.67 in NR, 136.44 in LR, and 71.13 in HR. This is due to the increasing wholesale price set by the supplier. Finally, the last row of Table 2.1 shows predicted supply chain profit and net profit. Note that we do not show predicted supply chain efficiency because the benchmark varies from treatment to treatment as total initial capital $r + r_s$ varies, which may lead to inconsistent comparisons.

Each treatment included 60 participants recruited from a large university where cash was the only incentive offered. Each participant was randomly as-

Table 2.1: Experimental Parameter Settings and Predictions

| Treatment | NR | LR | HR |
|---|---|---|---|
| Retailer Initial Capital $r$ | 700 | 400 | 100 |
| Wholesale Price $w$ | 20.00 | 21.47 | 25.87 |
| Quantity $q$ | 33.33 | 38.73 | 42.49 |
| Bankruptcy Threshold $k$ | 0.00 | 14.38 | 33.31 |
| Supplier Expected Profit | 733.33 | 813.16 | 907.87 |
| | [333.33] | [413.16] | [507.87] |
| Retailer Expected Profit | 866.67 | 536.44 | 171.13 |
| | [166.67] | [136.44] | [71.13] |
| Supply Chain Expected Profit | 1600.00 | 1349.60 | 1079.00 |
| | [500.00] | [549.60] | [579.00] |

Note: the supplier's initial capital is $r_s = 400$ across all treatments. Net expected profits (excluding any initial capital) are reported in square brackets.

signed a role and stayed in that role throughout the game. In the experiment, suppliers and retailers were placed in cohorts of six (three of each type), which participants were unaware of. Within each cohort, in each round, pairs of one retailer and one supplier were randomly formed. Both parties were provided with decision support. Specifically, participants could use a sliding bar to test their decisions and observe a plot of the realized profit of both parties for all possible demand realizations. In addition, both parties were shown the probability of retailer bankruptcy. For suppliers, the retailer's realized profits and bankruptcy probability were calculated by assuming an optimal quantity response (which participants were aware of). Note that the word "bankruptcy" was not used in the experiments. Instead, we used the phrases the"possibility that the retailer cannot pay you in full" for suppliers and the "possibility that you cannot repay the supplier in full" for retailers. Finally, retailers were provided a rejection button to use if they disliked a received offer. If a contract was rejected, both parties kept their initial capital.

Our experiment was implemented through oTree [28].[2] Before each session started, a researcher read the instructions aloud and answered any questions. Participants were then required to answer several multiple-choice comprehension questions about the game. They received cash based on profits from a randomly chosen round in the game plus a $7 show-up fee. Average earnings were $22.34 across all treatments. Each session lasted for 70 minutes on average.

### 2.4.2 Behavioral Hypotheses

Although the normative theory provides predictions for a rational decision-maker, existing experimental evidence suggests that human participants can exhibit behavioral biases and deviate from rationality. In this subsection, we discuss three behavioral hypotheses related to the quantity, wholesale price, and bankruptcy risk, followed by how such potential deviations can affect profit outcomes.

Beginning with the quantity, because retailers are endowed with initial capital $r$ in each round, they may exhibit the endowment effect [67], i.e., over-valuing the initial capital, and become loss averse. In other words, retailers will prefer a lower bankruptcy risk and order a quantity that is less than predicted. Moreover, recall that a more risky retailer (i.e., HR) is actually more protected by limited liability

[2]Because of the COVID-19 pandemic, we ran all sessions synchronously online through Zoom.

and should order a quantity that is higher (compared to LR). A human retailer may fail to recognize the benefit of limited liability in this case. As a result, retailers have less incentive to order more, and loss aversion should be more salient when the bankruptcy risk is high. Therefore, we offer the following hypothesis

**Hypothesis 2.1.** *Retailers in LR and HR will set lower quantities to achieve lower bankruptcy risk. The magnitude of deviation is larger in HR than in LR.*

The supplier provides financing to the retailer and bears far less risk. A human supplier may also anticipate understocking behavior by the retailer and thus set a more generous wholesale price to induce a higher quantity. We also expect the magnitude of deviation to be larger in HR than in LR, for three reasons. First, as stated in Hypothesis 2.1, we expect to observe more understocking in HR and suppliers will need to lower wholesale prices more to achieve higher profits. Second, the retailer begins at a disadvantageous position in HR, with less initial capital, and the supplier may prefer a fair outcome [e.g., 36, 70]. Third, the predicted wholesale price in HR is relatively extreme (25.87), and a human supplier may be reluctant to set extreme values based on past lab evidence [e.g., 57]. Thus, we have Hypothesis 2.2:

**Hypothesis 2.2.** *Suppliers in LR and HR will set lower wholesale prices. The magnitude of deviation is larger in HR than in LR.*

Although we expect to see lower wholesale prices in LR and HR, this does not necessarily lead to less bankruptcy risk for the retailer. Recall that Figure 2.3

shows that $k(q^*(w))$ exhibits an inverted-U shape with respect to $w$. In LR and HR, the predicted wholesale prices are on the right side of the peak, and a lower wholesale price can lead to even greater bankruptcy risk. Therefore, suppliers may be facing a trade-off between offering between a lower wholesale price with more bankruptcy risk and a higher price but with less bankruptcy risk. While it is unclear which will be the case, it is unlikely that suppliers offer a higher degree of risk than the normative predictions. Here, we provide a relatively conservative hypothesis:

**Hypothesis 2.3.** *Assuming that retailers set quantities optimally, the bankruptcy threshold in LR and HR will not be higher than the normative predictions.*

Even if we can anticipate these deviations, a lab experiment is still necessary. First, how the resultant profit will change is not fully clear. If both hypotheses 2.1 and 2.2 are supported, we can anticipate that supplier profit will be lower than predicted, as both a lower quantity and a lower wholesale price hurt the supplier (even though it is still optimal for a supplier to decrease the wholesale price, if she expects the retailer to understock). However, how the retailer's profit will change is unknown. Although the retailer can be hurt by his suboptimal ordering, he may also benefit from the lower wholesale price set by the supplier. Second, even in NR, the direction of deviation is unclear. On one hand, NR is mostly a standard newsvendor setting, where human retailers may over-order because of the "pull-to-center" effect [112]. On the other hand, overstocking in our setting can lead to higher bankruptcy risk, which retailers may be averse to. Therefore, we do not develop a hypothesis around NR. We aim to find the answers by conducting a lab experiment.

## 2.5 Results

We first present observed decisions and any deviations, and then we show how these decisions translate into profits. Recall that retailers were allowed to reject a supplier's offer if they found it unfavorable. From the experiment, the rejection rates for the three treatments are 8.67% (HR), 9.00% (LR), and 6.00% (NR). Given the similar rates, rejected data are excluded in the following analysis. Unless otherwise noted, we use regressions with random effects for hypothesis testing.

Before showing the results, we first check whether significant learning is occurring. Overall, there is some learning for the wholesale price but not for the stocking quantity. By comparing observed wholesale prices of the first half and the second half of the experiment (rejected data excluded), suppliers set higher prices in the second half in NR (on average 20.36 vs. 21.07, $p < 0.01$) and LR (18.84 vs. 19.30, $p < 0.01$), but not in HR (19.80 vs. 20.00). Although the difference in LR and HR is significant, the absolute difference is less than 1. Turning to stocking quantities, we do not observe significant learning in terms of quantity deviation ($q - q^*(w)$, observed quantity minus conditionally optimal quantity) in any treatment (4.46 vs. 5.21 in NR, -11.51 vs. -11.34 in LR, -15.82 vs. -14.71 in HR). Given that the learning is slight and not consistent across treatments, we keep the full data for all analyses (rejected data are still excluded).

### 2.5.1 Decisions

Table 2.2 presents the average observed decisions (left side) and the theoretical predictions (right side). For the latter, we provide predictions conditioning on any observed previous decisions and unconditional predictions in square brackets. Beginning with testing Hypothesis 2.1, we first look at the observed quantity and the corresponding bankruptcy risk and observe opposite directions of deviations, relative to the normative predictions. Specifically, retailers significantly understock in HR, 40.55 observed versus 55.76 predicted, and in LR, 35.14 versus 46.46 (both $p < 0.01$), but overstock in NR, 36.38 versus 31.46 ($p < 0.1$). As a result, the actual bankruptcy threshold is significantly lower than predicted in HR and LR (both $p < 0.01$) but higher in NR ($p < 0.01$). Although the absolute deviation is larger in HR than in LR, which we posited in hypotheses 1, we do not find a significant difference in terms of relative deviation (i.e., $q/q^*(w)$). Therefore, Hypothesis 2.1 is partially supported. To summarize, while high-risk retailers understock to reduce their bankruptcy risk, retailers with almost no risk choose to put themselves under bankruptcy risk. Therefore, we have the first result.

**Result 2.1.** *Compared with the normative theory, when retailers have low or moderate levels of initial capital (HR or LR), they order quantities that are too low, which translates into lower bankruptcy risk. When retailers have high levels of initial capital (NR), they order quantities that are too high, which results in higher bankruptcy risk.*

Table 2.2: Observed Decisions and Theoretical Predictions

| | Observed | | | Predicted | | |
|---|---|---|---|---|---|---|
| | NR | LR | HR | NR | LR | HR |
| Wholesale Price | 20.70$^{\dagger}$ | 19.05$^{\ddagger}$ | 19.91$^{\ddagger}$ | 20.00 | 21.47 | 25.87 |
| | (0.34) | (0.28) | (0.47) | | | |
| Quantity | 36.38$^{*}$ | 35.14$^{\ddagger}$ | 40.55$^{\ddagger}$ | 31.46 | 46.46 | 55.76 |
| | (2.68) | (2.17) | (1.77) | (0.68) | (0.60) | (0.68) |
| | | | | [33.33] | [38.73] | [42.49] |
| Bankruptcy Threshold | 5.53$^{\ddagger}$ | 8.98$^{\ddagger}$ | 22.66$^{\ddagger}$ | 0.30 | 15.48 | 32.72 |
| | (1.22) | (1.25) | (1.11) | (0.05) | (0.13) | (0.32) |
| | | | | [0.00$^{\ddagger}$] | [14.38$^{\ddagger}$] | [33.31] |

Note: Standard errors, across subjects, are reported in parentheses. Rejected data are excluded. Predicted quantities and bankruptcy threshold are conditioning on observed wholesale prices. Unconditional normative predictions, when applicable, are reported in square brackets. Significance of regressions with random effects compared with observed versus conditional normative predictions are given by $^{\ddagger}$ $p < 0.01$, $^{\dagger}$ $p < 0.05$ and $^{*}$ $p < 0.1$.

Turning to the wholesale price, the observed value is significantly lower than predicted in HR, 19.91 versus 25.87, and in LR, 19.05 versus 21.47 (both $p < 0.01$), but slightly higher in NR, 20.70 versus 20.00 ($p < 0.05$). In other words, similar to retailers, suppliers exhibit a different direction of deviation in response to retailers' degree of bankruptcy risk. Hypothesis 2.2 is fully supported. Although suppliers in NR set prices slightly higher than normative predictions, note that the absolute difference is relatively small. Combining all these observations yields the following result.

**Result 2.2.** *Compared with the normative predictions, suppliers set wholesale prices too low when retailers have low or moderate levels of initial capital (HR or LR), and they set wholesale prices slightly higher when retailers have high levels of initial capital.*

Table 2.2 also indicates that suppliers do not offer less risky contract terms to retailers, relative to the normative predictions. The conditional predictions of the bankruptcy threshold at the bottom right of Table 2.2 are calculated based on observed wholesale prices and conditionally optimal quantities. Therefore, they can be regarded as the degree of bankruptcy risk offered by suppliers. By comparing conditional bankruptcy thresholds with unconditional ones, we find that suppliers do not offer significantly different degrees of risk in HR alone. In the other two treatments, suppliers' offers lead to more risk. In LR, the average bankruptcy threshold offered is 15.48, which is higher than the theoretical prediction, 14.38 ($p < 0.01$). Similarly, in NR, the average bankruptcy threshold is 0.30, higher than the predicted value, 0.00 ($p < 0.01$). Such findings indicate that Hypothesis 2.3 is not fully supported. Although the difference is statically significant, we again note that the absolute difference is relatively small. Therefore we have Result 2.3.

**Result 2.3.** *Suppliers offer higher bankruptcy risk to retailers with moderate or high initial levels of capital (LR or NR), but the difference is relatively small.*

Next we explore the observed decisions in more depth. Specifically, one might be interested in knowing whether retailers react to wholesale prices differently, across the different treatments. To address this, we conduct further regression analysis of retailer decisions, presented in Table 2.3. The three leftmost columns show treatment-specific results with the observed quantity deviation conditioning on the observed wholesale price $q - q^*(w)$ as the dependent variable. Estimated coefficients indicate that retailers in NR tend to overstock when facing high wholesale prices whereas retailers in HR understock. There are two possible rea-

Table 2.3: Regressions of Deviation of Retailer Decisions

| Dependent Variable | $q - q^*(w)$ | | | $k - k^*(w)$ | | |
|---|---|---|---|---|---|---|
| Treatment | NR | LR | HR | NR | LR | HR |
| Wholesale Price | $0.813^{\ddagger}$ | -0.038 | $-0.430^{\dagger}$ | 0.157 | -0.241 | $-0.734^{\ddagger}$ |
| | (0.194) | (0.273) | (0.214) | (0.116) | (0.157) | (0.146) |
| Lagged Demand | 0.009 | 0.033 | $0.065^{\dagger}$ | 0.012 | $0.021^{*}$ | $0.043^{\dagger}$ |
| | (0.019) | (0.022) | (0.026) | (0.012) | (0.013) | (0.018) |
| Lagged Bankruptcy | 3.412 | -0.578 | $3.187^{*}$ | $7.545^{\ddagger}$ | -0.323 | 1.985 |
| | (2.526) | (2.154) | (1.768) | (1.542) | (1.234) | (1.208) |
| Constant | $-12.483^{\ddagger}$ | $-12.438^{\dagger}$ | $-10.613^{\dagger}$ | 1.147 | -3.171 | 1.841 |
| | (4.319) | (5.433) | (4.749) | (2.463) | (3.124) | (3.223) |
| $N$ | 538 | 520 | 524 | | | |

Note: regressions are run with random effects. Standard deviations of estimated coefficients are reported in parentheses. "Lagged bankruptcy" is a binary outcome of retailer bankruptcy in the previous round.

sons for this deviation pattern: to reduce the difference in realized profits between the two parties or to reduce the bankruptcy risk. Results in the right side of the table show that only the first reason applies to NR and that both apply to HR. The dependent variable of the right three columns is the observed deviation in the bankruptcy threshold $k - k^*(w)$. The coefficient of the wholesale price is not significant in NR, indicating that retailers change little of their bankruptcy risk by responding differently to wholesale prices. However, the coefficient is significant and negative in HR, meaning that retailers do choose lower bankruptcy risk when suppliers set higher wholesale prices. Regarding the coefficients of lagged demand, retailers with a high bankruptcy risk seem to be more sensitive to realized demand, which is understandable. Finally, while not depicted, we do not find that retailers deviated more in quantity or tended to reject offers more after experiencing bankruptcy in any treatment. In addition, there is no significant evidence that suppliers set wholesale prices differently after their paired retailers went bankrupt in the previous round.

### 2.5.2 Profits

Next we show how these decisions translate into profits. Table 2.4 shows the observed average profits on the left side and the theoretical predictions on the right side. Beginning with suppliers, they achieve lower profits in HR and LR, 682.35 versus 907.87 and 672.41 versus 813.16, respectively (both $p < 0.01$). This is understandable given our earlier results, as we observe under-pricing and under-stocking in these two treatments. As for NR, although suppliers achieve slightly higher profits than predicted on average, 757.49 versus 733.33, the difference is not significant.

Table 2.4: Observed Profits and Theoretical Predictions

| | Observed | | | Predicted | | |
|---|---|---|---|---|---|---|
| | NR | LR | HR | NR | LR | HR |
| Supplier Profit | 757.49 | 672.41$^{\ddagger}$ | 682.35$^{\ddagger}$ | 733.33 | 813.16 | 907.87 |
| | (19.75) | (11.24) | (9.29) | | | |
| Retailer Profit | 818.50$^{\ddagger}$ | 596.37$^{\ddagger}$ | 337.55$^{\ddagger}$ | 866.67 | 536.44 | 171.13 |
| | (11.07) | (7.68) | (9.54) | | | |
| Supply Chain Profit | 1574.43 | 1265.71$^{\ddagger}$ | 1015.38$^{\ddagger}$ | 1600.00 | 1349.60 | 1079.00 |
| | (16.72) | (13.75) | (14.07) | | | |

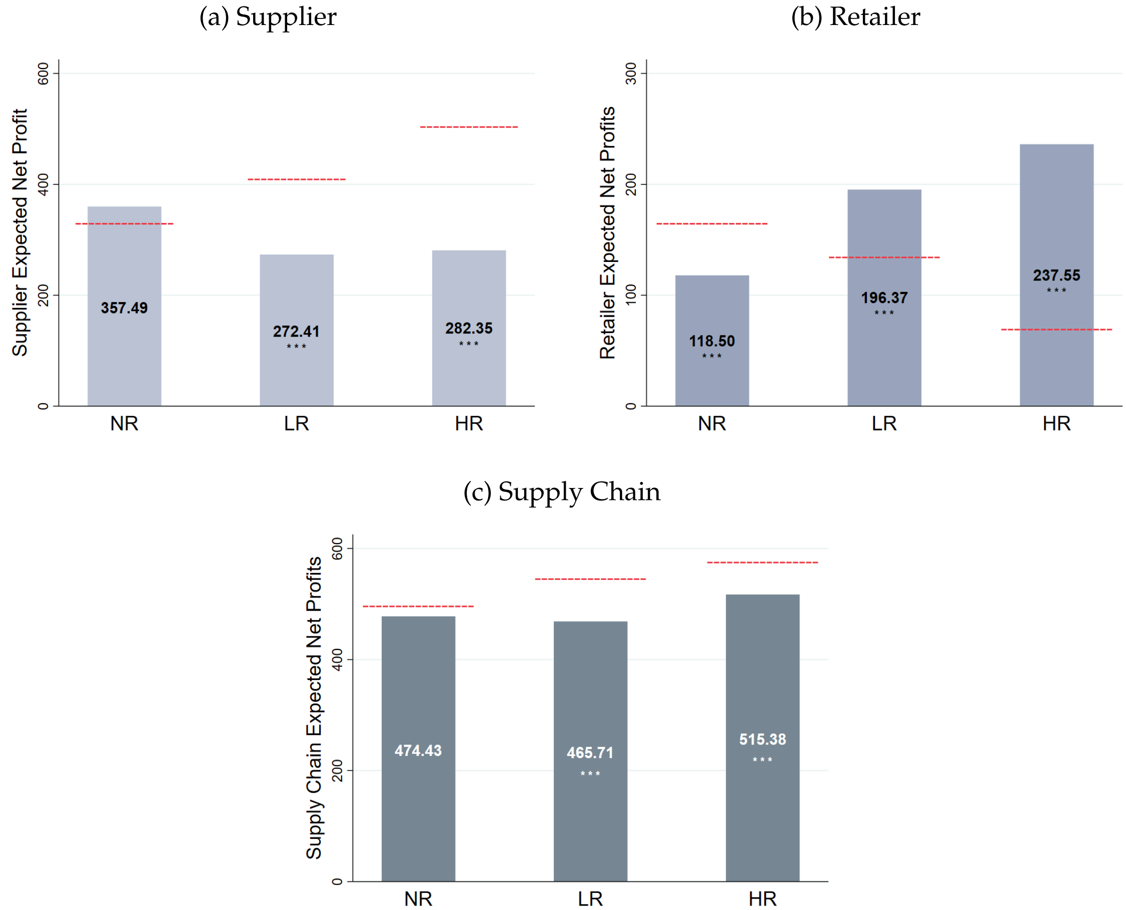
Note: Standard errors, across subjects, are reported in parentheses. Rejected data are excluded. Significance of regressions with random effects comparing observed versus conditional normative predictions are given by $^{\ddagger}$ $p < 0.01$.

In contrast to supplier profit, we observe higher retailer profit in HR and LR, but lower in NR. Specifically, the observed average retailer profit in HR is 337.55, much higher than its theoretical prediction, 171.13 ($p < 0.01$). Similarly, the average retailer profit in LR is 596.37 versus the 536.44 prediction ($p < 0.01$). In NR, retailers achieve an average profit of 818.50, lower than predicted, 866.67 ($p < 0.01$). The lower retailer profits are largely due to the suboptimal stocking shown in Table 2.2, regardless of the direction of deviation.

The last row of Table 2.4 shows the observed average supply chain profit. Although retailers achieve higher profits in HR and LR, the total profits are still lower than the normative predictions (1015.38 vs. 1079.00 in HR and 1265.71 vs. 1349.60 in LR, both $p < 0.01$), due to insufficient stocking. We do not observe significantly different supply chain profits in NR, despite their being slightly lower.

Figure 2.5 presents average net profits and predictions, which allows us to further understand the impact of deviations and to compare treatments. Beginning with supplier net profit, NR achieves higher supplier profit than HR and LR (both $p < 0.05$), which runs counter to the predictions. There is no significant difference between supplier profit in HR and LR. Turning to retailer net profit, we observe an opposite trend compared with the normative predictions. Retailer net profit in HR is higher than that in LR ($p < 0.05$), which is then higher than NR ($p < 0.01$), i.e., HR>LR>NR whereas theory predicts HR<LR<NR. The opposite trend of predicted and observed retailer profits shows how human decision-makers behave differently from the normative theory. To understand the behavioral mechanism behind these deviations, we propose a plausible behavioral model that can capture these deviated decisions in Section 2.6.

Figure 2.5: Average Observed Net Profit




Note: Normative predictions are represented by horizontal red dashed lines. Significance of regressions with random effects comparing observed with normative predictions are given by *** $p < 0.01$.

### 2.5.3 Heterogeneity Analysis

In this subsection we conduct a heterogeneity analysis to examine whether our results are driven by a subset of participants. Figure 2.6 classifies participants into three categories based on whether they consistently set lower (the bottom, lightblue portions in the figure) or higher (the top, light-orange portions) values than optimal, or did not significantly deviate in one direction (the middle, dark-blue portions). Thus, we compare participants' actual decisions with (conditionally) optimal decisions in all rounds,[3] with a Wilcoxon signed-rank test, because of the limited sample size,[4] and a significance level of 5%. Beginning with the wholesale price decision in Figure 2.6a, the majority of suppliers consistently set the wholesale price lower in LR (91.58%) and HR (97.26%), consistent with the average observed results in Table 2.2. However, suppliers in NR exhibit mixed behavior, i.e., 41.67% of them choose to over-price, 20.39% prefer lower prices, and 37.94% stand in between. Turning to the quantity decision of retailers in Figure 2.6b, results are mostly consistent with the average. Specifically, 62.23% of retailers in NR overstock, while 74.73% in LR and 86.13% in HR understock (and thus a lower bankruptcy risk).

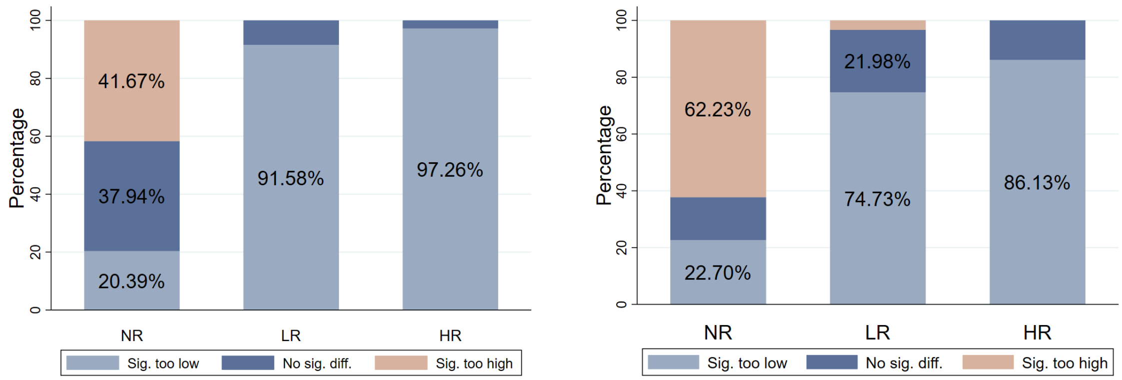
Taking a closer look at the three subgroups of suppliers in NR, the average wholesale prices of the "under-pricing" group, the "in between" group, and the "over-pricing" group are 18.29, 20.15, and 22.28, respectively. The absolute deviation from the optimal, 20, is similar for both the "under-" and the "over-pricing" groups, which can account for the overall average price being just slightly higher than the normative prediction. Although deviations are less consistent in this case, we note that the "over-pricing" group is still the most common type (41.67%) driv-

[3]For the wholesale price, we compare a supplier's actual wholesale price and the theoretically optimal price (e.g., 25.87 in HR). For the quantity, we compare a retailer's actual quantity and the optimal quantity conditioning on the observed wholesale price in each round.

[4]The sample size for each test is 20 before excluding rejected data.

Figure 2.6: Percentages of Participants Deviating in Decisions

(a) Wholesale Price by Supplier (b) Quantity by Retailer




Note: Classification of a participant is based on Wilcoxon signed-rank tests between observed decisions and (conditional) optimal decisions (at 5% level). Percentages below 20% are omitted because of limited space.

ing the average decision.

## 2.6 Behavioral Model: Reference Dependence

So far we have shown that observed decisions significantly deviate from the normative theory. To understand what can account for these deviations, in this section, we investigate a reference-dependence model. Reference dependence has been observed in various operational contexts [e.g., 61, 12, 120]. We first introduce a general framework of reference dependence by [122] and subsequently discuss potential reference points. We then fit these alternative reference points to the data, along with the normative theory, to determine the best-performing model. We show that considering only reference-dependent retailers, but not suppliers, and a simple reference point, are sufficient to capture both parties' observed decisions in our experiments. Thus, our behavioral model is parsimonious and useful in predicting outcomes.

Following the notation of [122], a reference-dependent decision-maker compares its profit to a reference point $\mathscr{P}$ and gains positive (negative) utility if its profit is higher (lower) than $\mathscr{P}$. Based on whether $\mathscr{P}$ is dependent on demand $d$ or any decision, reference points can be categorized by a) a fixed reference point (FRP) that is independent of any value, b) a prospect reference point (PRP) that depends on decisions, and c) a stochastic reference point (SRP) that depends on both realized demand and decisions. Define $\mathscr{D}_{>} = \{d \in [0, a] : \pi(d) > \mathscr{P}\}$ as the domain of demand that leads to a higher realized profit than the reference point, i.e., the gains domain. Similarly, let $\mathscr{D}_{<} = \{d \in [0, a] : \pi(d) < \mathscr{P}\}$ be the losses domain. A general reference-dependent utility function is

$$\mathbb{E}[\pi] + \left[\eta \int_{x \in \mathscr{X}_{>}} (\pi(x) - \mathscr{P}(x)) dF(x) - \lambda \int_{x \in \mathscr{X}_{<}} (\mathscr{P}(x) - \pi(x)) dF(x)\right]. \tag{2.8}$$

The utility in Equation (2.8) consists of expected profit (also called the consumption utility) and the gain-loss utility. The parameters $\eta$ and $\lambda$ are the psychological weights of gains and losses, respectively. We assume $\eta$ and $\lambda$ to be non-negative unless a negative parameter is meaningful.

We work backwards, first investigating the quantity decision of the retailer and then examining the wholesale price decision of the supplier. Based on the observed quantity deviation, it is possible that the supplier assumes the retailer is reference-dependent rather than rational. Therefore, we need to first understand potential reference points for the retailer. For an introduced reference point, we only discuss the main idea in this section; we present the detailed utility function in Appendix B.1.2.

To evaluate a reference point, we fit the observed decisions using maximum-likelihood estimation (MLE). In the estimation, we use truncated normal distributions for both the quantity and wholesale price decisions, with (conditionally) optimal values as the means and estimate the standard deviations. For the quantity, the distribution is truncated at lower bound 0 and upper bound 100 of demand. For the wholesale price, the normal distribution is truncated at unit production cost $c = 10$ and unit selling price $p = 30$. The PDF of a truncated normal distribution $\varphi(x; \mu, \sigma, a, b)$ is defined by

$$\varphi(x; \mu, \sigma, a, b) = \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma\left(\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right)}. \tag{2.9}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and CDF of the standard normal distribution, respectively.

We split the data into a training set (67%) and a testing set (33%).[5] In each treatment, recall that there are 30 participants playing one role (either supplier or retailer). We assign 20 of 30 to the training set and the remaining 10 to the testing set. We first estimate each model using the training set. We then calculate the log-likelihood of the estimated parameters on the testing set and compare across models. Let $I$ be retailers in the testing test (indexed by $i$) and $J$ be suppliers in the

[5]We believe such an out-of-sample validation and comparison is more robust and rigorous than an in-sample approach.

testing test (indexed by $j$), respectively. For retailer $i$ or supplier $j$, some rounds are excluded because of rejection. Therefore, we use $T_i$ and $T_j$ to denote the set of rounds that are included in the estimation, with $t$ being the index. The total log-likelihood functions for retailers and suppliers are shown below.

$$LL_r(\eta_r, \lambda_r, \sigma_r) = \prod_{i \in I} \prod_{t \in T_i} \varphi(q_{it}; \tilde{q}(w_{it}, \eta_r, \lambda_r), \sigma_r, 0, 100). \tag{2.10}$$

$$LL_s(\eta_s, \lambda_s, \sigma_w) = \prod_{j \in J} \prod_{t \in T_j} \varphi(w_{jt}; \tilde{w}(\eta_s, \lambda_s), \sigma_w, 10, 30), \tag{2.11}$$

We use subscript $r$ for all retailer-related estimated variables in Equation (2.10) and $s$ for those related to suppliers in Equation (2.11). In addition, $\tilde{q}(w_{it}, \eta_r, \lambda_r)$ in Equation (2.10) is the optimal quantity of a reference-dependence model conditioning on $w_{it}$, the wholesale price observed by retailer $i$ in round $t$. Similarly, $\tilde{w}(\eta_s, \lambda_s)$ in Equation (2.11) is the optimal wholesale price of a reference-dependence model.

### 2.6.1 Quantity

For the retailer, his initial capital $r$ is a reasonable candidate for the reference point. During the experiments, the retailer was allowed to reject a supplier's offer and keep his initial capital. Therefore, it is likely that the retailer treats $r$ as a target income. Formally, $\mathscr{P}^r = r$ is a FRP which is independent of quantity and demand. As the retailer is essentially making a newsvendor decision under a financial constraint, we also consider expected profit $\mathscr{P}^r = \mathbb{E}[\pi_r(q)]$ as a PRP suggested by [122]. Although in our decision support tool, only realized profit with respect to demand was given instead of expected profit, it was still possible to infer the expected profit from realized profit.

Table 2.5 shows the treatment-specific and total log-likelihoods from the MLE for each reference point and the normative model. Overall, the reference-dependence model, regardless of the reference point, outperforms the normative model in terms of goodness of fit in each treatment. The total log-likelihoods of the two reference-dependence models are -2359.88 (initial capital) and -2360.73 (expected profit), compared with -2436.80 for the normative model. Performance of each reference point slightly varies from treatment to treatment. Beginning with NR, initial capital fits better than the other two models. Turning to LR, expected profit performs the best, but initial capital is just slightly worse. Both reference points are much better than the normative model. Finally, both reference points achieve similar log-likelihoods in HR, both of which are more favorable than that of the normative model. Given that the performance of the two reference points is relatively close in terms of both total likelihood and treatment-specific likelihoods, we also consider both in the following analysis.

Table 2.5: Out-of-sample Model Comparison of Log-likelihood for Retailer

| | Normative Model | FRP: Initial Capital $r$ | PRP: Expected Profit $\mathbb{E}[\pi_r]$ |
|---|---|---|---|
| NR | -876.39 | -860.27 | -866.44 |
| LR | -762.38 | -721.55 | -717.66 |
| HR | -798.03 | -778.07 | -776.60 |
| Total | -2436.80 | -2359.88 | -2360.70 |

Note: Log-likelihoods are calculated on the testing set (33% of the data) with parameters estimated on the training set (67% of the data).

Table 2.6 shows estimated parameters and predicted decisions for the two reference points. As we have shown, they both outperform the normative model using an out-of-sample validation; here we use the full dataset to estimate parameters to obtain better insights from the population. Beginning with predicted quantity, both reference points generate predictions closer to observed values compared with the normative model. Between them, initial capital leads to better predictions despite its slightly worse fit in some treatments. The quantity predicted by initial capital as a reference point versus the average observed value is remarkably close across all treatments (observed values not depicted): 38.76 versus 36.38 in NR, 35.27 versus 35.14 in LR, and 40.44 versus 40.55 in HR. This indicates that a simple FRP is sufficient to capture the observed decisions.

Table 2.6: Estimation Results for Quantity Decisions of Reference-dependent Retailers

| | NR | | LR | | HR | |
|---|---|---|---|---|---|---|
| | FRP: $r$ | PRP: $\mathbb{E}[\pi_r]$ | FRP: $r$ | PRP: $\mathbb{E}[\pi_r]$ | FRP: $r$ | PRP: $\mathbb{E}[\pi_r]$ |
| $\tilde{q}$ | 38.76 | 31.22 | 35.27 | 34.68 | 40.44 | 35.11 |
| Gain: $\hat{\eta}_r$ | 3.362 | 0.009 | 2.815 | 1.294 | 0.789 | 1.050 |
| Loss: $\hat{\lambda}_r$ | 1.651 | 0.019 | 5.412 | 1.722 | 7.137 | 2.105 |

Note: $\tilde{q}$ is the predicted quantity. The average observed quantity is 36.38 in NR, 35.14 in LR, and 40.55 in HR. Two reference points are fitted: i) initial capital $r$ as a FRP, and ii) expected profit $\mathbb{E}[\pi_r]$ as a PRP.

Turning to the estimated parameters in Table 2.6, note that the magnitude of estimated parameters partially depends on the reference point and varies across treatments. For example, different reference points can affect the size of the gains and losses domains in utility function (2.8) and therefore the weights $\eta_r$ and $\lambda_r$. As a result, it is less applicable to compare parameters between treatments. However, we can compare $\hat{\eta}_r$ and $\hat{\lambda}_r$ for each treatment. Beginning with NR, although the estimated $\hat{\eta}_r$ and $\hat{\lambda}_r$ by expected profit are close to zero, initial capital (FRP) estimates a higher $\hat{\eta}_r$ than $\hat{\lambda}_r$ for NR, indicating that retailers in NR care more about gains than losses. As a comparison, both reference points predict the opposite in LR and HR: $\hat{\lambda}_r$ being higher than $\hat{\eta}_r$, which suggests that retailers with higher bankruptcy risk care more about losses than gains. This can also explain the directions of the observed quantity deviations. In NR, retailers assigned greater weight to the gains domain where the demand is high, so they overstocked to earn a higher profit in this domain. On the other hand, retailers in LR and HR weighted the losses domain more and chose to understock. This suggests that the presence of bankruptcy risk (LR and HR versus NR) reverses retailers' preferences around gain and loss domains, which is an important insight.

### 2.6.2 Wholesale Price

We have shown that the reference-dependence model can capture the retailer's decision well. For the supplier's decision, we begin by determining whether only embedding a reference-dependent quantity response in the supplier's problem is sufficient to predict the observed wholesale price deviations. Given that the observed quantities significantly deviate from the normative predictions, it is reasonable to assume that suppliers did not presume that retailers made quantity decisions in a rational way, especially in LR and HR. On top of the retailer response, we assume the supplier is still rational and maximizing her expected profit function, i.e., Equations (2.5a) and (2.5b).

We consider the same two reference points for quantity response discussed in Section 2.6.1: initial capital $r$ as a FRP and expected profit $\mathbb{E}[\pi_r]$ as a PRP. We fit wholesale price decisions using the same approach as quantity decisions: we split the dataset into a training set (67%) and a testing set (33%). Parameters are estimated using the training set, and models are compared using the testing set. For each model, we estimate $\eta_{s-r}$ and $\lambda_{s-r}$ in the retailer's quantity response. Note that subscript $s-r$ is used because these are parameters "believed" by the supplier and may be different from the true parameters of the retailer.

Table 2.7: Out-of-sample Model Comparison of Log-Likelihood for Supplier

| Quantity Response | Normative model | FRP: initial capital $r$ | PRP Expected Profit $\mathbb{E}[\pi_r]$ |
|---|---|---|---|
| NR | -476.29 | -476.26 | -479.42 |
| LR | -464.54 | -401.12 | -401.72 |
| HR | -531.26 | -449.60 | -527.32 |
| Total | -1472.10 | -1326.99 | -1408.46 |

Note: Log-likelihoods are calculated on the testing set (33% of the data) with parameters estimated on the training set (67% of the data). In the models fitted, suppliers are rational but assume that retailers determine quantity following the model in each column.

Table 2.7 compares out-of-sample log-likelihoods of different quantity responses for each treatment. Overall, initial capital (FRP) outperforms expected profit (PRP), as well as the normative model, in all treatments. Regarding treatment-specific log-likelihoods, the normative model fits similarly well as initial capital in NR because the supplier did not deviate from the optimal wholesale price to a large extent (20.70 observed vs. 20.00 predicted). As for the other two treatments, a reference-dependent quantity response is necessary to capture the supplier's decision. In LR, the performance of both reference points is similar and much better than the normative model, with initial capital being marginally better. In HR, initial capital achieves a more favorable log-likelihood than the other two models. Overall, initial capital (FRP) once again fits best in terms of cumulative log-likelihood, across treatments.

Table 2.8: Estimation Results for Wholesale Price Decisions of Suppliers

| | NR | | LR | | HR | |
|---|---|---|---|---|---|---|
| | FRP: $r$ | PRP: $\mathbb{E}[\pi_r]$ | FRP: $r$ | PRP: $\mathbb{E}[\pi_r]$ | FRP: $r$ | PRP: $\mathbb{E}[\pi_r]$ |
| $\tilde{w}$ | 20.64 | 21.01 | 19.01 | 19.40 | 19.79 | 19.75 |
| $\hat{\eta}_{s-r}$ | 0.349 | 0.440 | 3.854 | 0.004 | 0.408 | 2.985 |
| $\hat{\lambda}_{s-r}$ | 0.033 | 0.088 | 7.859 | 0.780 | 8.488 | 7.386 |

Note: $\tilde{w}$ is the predicted wholesale price. The average observed wholesale price is 20.70 in NR, 19.05 in LR, and 19.91 in HR. In the models fitted, suppliers are rational and retailers determine quantity following the model in each column.

Table 2.8 displays predicted wholesale prices and fitted parameters for both reference-dependence models. To fully utilize the data, we conduct the estimations in this table using the full dataset. Compared with expected profit, initial capital as a reference point again generates the best, and fairly accurate, predictions. Specifically, the wholesale price predicted by the initial capital versus average observed value is 20.64 versus 20.70 in NR, 19.01 versus 19.05 in LR, and 19.79 versus 19.91 in HR. Although we may not directly compare the estimated parameters between treatments or models, one can still observe that the estimated $\hat{\eta}_{s-r}$ and $\hat{\lambda}_{s-r}$ share a pattern similar to those in Table 2.6. In particular, $\hat{\eta}_{s-r}$ is larger than $\hat{\lambda}_{s-r}$ in NR, but the opposite is the case in LR and HR. Thus, although the supplier may not accurately anticipate the degree of reference dependence of the retailer, she can largely infer the retailer's preference for gains and losses (which switches in the presence of bankruptcy risk, LR and HR versus NR).

Overall, the estimation results of both quantity and wholesale price decisions show that assuming a reference-dependent retailer can well organize the data. Although there are various candidates for reference points, our results suggest that the retailer's initial capital $r$ is both simple and sufficient to predict the outcomes.

### 2.6.3 Discussion

While we show that reference dependence can well explain the observed deviations, we also acknowledge that there are other possible behavioral factors. Here we provide a brief discussion of these factors, mainly focusing on the retailer side, as we only assume the retailer is reference dependent in our behavioral model.

First, one may argue that the under-stocking behavior in LR and HR is due to probability weighting [54], i.e., the retailer over-estimates the bankruptcy probability and thus reacts by ordering less. However, recall that the retailer's optimal quantity increases in the bankruptcy risk because of limited liability. As a result, the retailer should prefer a higher quantity if he believes the bankruptcy probability is higher, which is inconsistent with our observed deviation. Therefore, probability weighting cannot be the driver of the under-stocking we observe in

HR (and to an extent, LR).

Another possible behavioral bias for the retailer is disappointment aversion [24], i.e., the retailer will feel disappointed upon bankruptcy. In other words, when calculating expected utility, the retailer has different weights on outcomes depending on whether he is bankrupt or not. Disappointment aversion can account for the under-stocking in LR and HR, but cannot explain the over-stocking in NR. Because we aim to find a simple behavioral model to explain all of the observed deviations, we opt for reference dependence which can predict deviations in both directions.

Turning to the supplier, it is possible that the under-pricing in LR and HR is driven by fairness concerns, as the retailer is predicted to earn much less than the supplier. While fairness can indeed predict the lower wholesale price, we have shown that embedding reference-dependent quantity response in the supplier's problem is sufficient to account for the deviation. Therefore, considering fairness adds unnecessary complexity to the behavioral model without improving the predictive ability significantly. Furthermore, recall that in NR, we observed overstocking by retailers and wholesale prices that were slightly higher than the normative prediction, which would be directionally inconsistent with fairness.

## 2.7 Robustness Check: Automated-Retailer Treatment

In Section 2.6.2, we showed that one does not need to model the behavioral bias of supplier in order to capture their decision. Assuming the retailer is reference-dependent can well predict the observed wholesale prices in all treatments. In other words, we implicitly assume that suppliers are rational and are merely anticipating the quantity decisions of reference-dependent retailers. In this section, we empirically test this assumption by conducting three additional treatments where we automate the retailer decision. If supplier decisions become significantly closer to optimal, particularly in LR and HR (where observed wholesale prices deviated significantly from the normative theory), then this behavioral assumption is reasonable. We name the three new treatments NR-A, LR-A, and HR-A, where A represents an automated retailer.

### 2.7.1 Experiment Design and Implementation

Different from the main experiments, participants can only play as a supplier with a computerized retailer who determines the order quantity following the normative theory. Therefore, there is no group interaction and it takes less time to finish a round. We use a within-subject design here so that we can observe a participant's decisions in all three treatments. Specifically, participants went through three treatments in a random order. At the beginning of each treatment, they were notified of the change in the retailer's initial capital. To minimize the ordering effect, we randomly assigned an equal number of participants to each possible sequence of three treatments (six combinations of sequence in total). The treatment sequence was unknown to them before the game started.

Participants were told that the retailer they were playing with was automated and would determine its quantity in an expected-profit maximizing way. We provided the same decision support to participants as in the main experiment, i.e., a realized profit figure with respect to all possible demand levels. Because there is experimental evidence that fairness concerns can exist even when human participants interact with automated players [71], realized retailer profit from the decision support and results were not displayed. Other design elements remained the same as the main experiment, including the participant pool and the protocol.

Regarding sample size, we assign five participants to each sequence of treatments (30 new participants in total). Each treatment has 15 rounds (45 rounds in total). At the beginning of each treatment, they are notified of the retailer's initial capital for the next 15 rounds. We reduced the number of rounds for each treatment because of the limited session length, and we did not observe significant learning in the data. Experiments were implemented virtually through Zoom, and the average earnings were $24.20.

### 2.7.2 Results

Average wholesale prices of the automated-retailer treatments, along with those of the original treatments and normative predictions, are shown in Table 2.9. To better compare decisions from the supplier's perspective, in the second row we show average optimal quantities conditioning on observed wholesale prices. In the third row, we show corresponding expected supplier profits.

Beginning with NR, wholesale prices did not significantly change moving from a human retailer to a computerized retailer or compared with the normative prediction (20.47 in NR-A vs. 20.70 in NR and 20 predicted). This is not surprising, as suppliers had made relatively rational decisions when facing human retailers. Moving to treatments with a higher retailer bankruptcy risk (LR-A and HR-A), suppliers set wholesale prices much closer to optimal when retailers are rational. Specifically, in LR-A there is no significant difference between the observed average wholesale price and the normative prediction (21.21 observed vs. 21.47 predicted). However, the average price in LR-A is much higher than that in LR (21.21 vs. 19.05, $p < 0.01$). Finally, although the average wholesale price in HR-A is lower than predicted (22.46 vs. 25.87 predicted, $p < 0.01$), it is significantly higher (closer to optimal) than that in HR (22.46 vs. 19.91, $p < 0.01$). One possible explanation is that the predicted optimal price of 25.87 is relatively extreme in HR, which makes suppliers less willing to choose. Overall, compared with the original treatments, wholesale prices were indeed set higher in LR-A and

HR-A, but not in NR-A, which is consistent with the estimation results in Section 2.6.2. Summarizing the findings, we have the following result.

**Result 2.4.** *When facing automated retailers, suppliers set wholesale prices significantly higher than when interacting with human retailers who face bankruptcy risk (LR and HR). No significant difference is observed between human and automated retailers in NR.*

Table 2.9 also shows expected supplier profit under the conditionally optimal quantity. From the table, although suppliers set higher prices in LR-A and HR-A, only the latter leads to a higher supplier profit than the original treatments (808.20 in HR-A vs. 764.03 in HR, $p < 0.1$). Nevertheless, the average profit in LR-A is significantly lower than the normative prediction (770.66 in LR-A vs. 813.16 predicted, $p < 0.01$), even though prices were set very close to the optimal price. The reason is twofold: a) while the normative predicted supplier profit is the maximum that a supplier can achieve, the average observed profit is depressed by a few participants who set much lower prices; and b) assuming an optimal quantity response, the sensitivity of supplier profit to the wholesale price is lower in LR than in HR (see Figure B.1a in Appendix for an illustration). As a result, better-set wholesale prices can lead to higher supplier profits in HR-A but not as much in LR-A.

## 2.8 Conclusion

In practice, a firm's financial position is usually intertwined with its operational decisions [53], which suggests the importance of studying them jointly. In this paper, we experimentally investigate how bankruptcy risk affects decision-making and performance under a trade credit contract. Specifically, we consider a two-tier supply chain consisting of a supplier and a capital-constrained retailer. The retailer can purchase from the supplier and pay with his capital up front, or place a higher quantity through trade credit and repay the supplier after the demand realization. If the actual demand is too low for the retailer to repay the supplier in full, the retailer will transfer all of his cash, including initial capital and any revenue from demand, to the supplier and file for bankruptcy. We consider different levels of bankruptcy risk by varying the retailer's initial capital and conduct controlled lab experiments with human participants to observe decisions and performance. In the main experiment, we consider three levels of bankruptcy risk: no risk (NR), low risk (LR), and high risk (HR).

Table 2.9: Average Decisions of Automated-Retailer Treatments

| | Automated Retailer | | | Original Treatment | | | Normative Prediction | | |
|---|---|---|---|---|---|---|---|---|---|
| | NR-A | LR-A | HR-A | NR | LR | HR | NR | LR | HR |
| Wholesale Price | 20.47 | 21.21 | 22.46 | 20.70 | 19.05‡ | 19.91‡ | 20.00 | 21.47 | 25.87‡ |
| | (0.32) | (0.30) | (0.71) | (0.34) | (0.29) | (0.47) | | | |
| Optimal Quantity | 31.88 | 38.14 | 49.16 | 31.15 | 46.23‡ | 55.37‡ | 33.33 | 38.73 | 42.49‡ |
| | (1.09) | (1.22) | (1.61) | (1.18) | (0.85) | (1.13) | | | |
| Supplier Profit | 706.41 | 770.66 | 808.20 | 705.68 | 764.66 | 764.03* | 733.33‡ | 813.16‡ | 907.87‡ |
| (Optimal Quantity) | (5.65) | (7.79) | (22.11) | (4.54) | (7.00) | (13.13) | | | |

Note: Standard errors, across subjects, are reported in parentheses. Rejected data are excluded. Average optimal quantity of original treatments is collapsed within suppliers, which are slightly different from those in Table 2.2. Significance of regressions with random effects compared with decisions of automated-retailer treatments is given by † $p < 0.05$ and ‡ $p < 0.01$.

The normative theory suggests that a retailer with higher bankruptcy risk is more protected by limited liability and should set a higher quantity compared with the case without capital constraints. Knowing this, the supplier will set a higher wholesale price to extract more profit from the retailer. However, our experimental results show that retailers with moderate (LR) and high (HR) bankruptcy risk significantly understock to minimize bankruptcy risk, and that suppliers set lower wholesale prices to induce higher quantities. In contrast, retailers with virtually no bankruptcy risk (NR) choose to overstock, bringing themselves higher bankruptcy probability, and suppliers slightly overprice. As a consequence, a key insight is that observed retailer net profits (excluding initial capital) actually decrease in initial capital, whereas the normative theory predicts the opposite.

To account for these deviations, we develop a simple reference-dependence model and show that the retailer's initial capital as a simple fixed reference point is sufficient to fit both retailers' and suppliers' decisions in the presence of bankruptcy risk. In particular, the reference-dependence model can predict the observed understocking from retailers. Embedding a reference-dependent retailer in the supplier's problem can also fit suppliers' underpricing in LR and HR. Regarding NR, we show that the normative theory is good enough to fit the results. To empirically validate how suppliers anticipate retailers' quantity response, we conduct automated-retailer treatments for each of the original treatments. These new experimental results suggest that wholesale prices remain similar in NR but are significantly higher in LR and HR, supporting our behavioral model.

We believe this paper not only has managerial implications but also contributes to the literature. Beginning with the former, suppliers should be aware that retailers who face bankruptcy risk tend to understock, in order to minimize the likelihood of going bankrupt. Although offering a more generous wholesale price can induce a higher quantity, our results suggest that suppliers may set prices too low and give up too much of their profits. Regarding retailers, although it is reasonable that retailers understock to avoid bankruptcy, they may also fail to realize the potential benefits of limited liability. Turning to our contribution to the literature, to our knowledge, this is the first behavioral investigation in supply chain finance. Our experimental results show that behavioral factors play a significant role in decision-making when bankruptcy risk is involved. Further, note that the deviations are surprisingly large even if we consider only a simplified setting. In other words, we do not consider many of the common components in supply chain finance such as interest rates, bank financing, factoring, and insurance, which could exacerbate these effects. Now that we have established initial results in supply chains with bankruptcy risk, we believe more research considering such complexities is critical to providing a comprehensive understanding of behavioral factors in this field.

There are also limitations in our study that can be interesting opportunities for future research. First, we consider a price-only trade credit contract without an interest rate. In practice, suppliers may also provide a wholesale price and an interest rate, allowing retailers to pay early with a discount or pay later with interest [133]. Such a contract, also known as an early-payment-discount contract, requires suppliers to optimize an interest rate and a wholesale price simultaneously. Experimental research dedicated to studying both price and interest rate would be interesting, especially given that [75] show that a rational supplier should always set her interest rate as a risk-free rate. Second, while supplier financing has become increasingly important, bank financing as a traditional financing tool still prevails in practice. Comparing both financing tools in terms of supply chain performance can generate useful managerial implications. Finally, sometimes suppliers can also be capital-constrained and in favor of offering trade credit for profitability or competition [82]. The supplier alone, or both supplier and retailer, being capital-constrained will be an interesting direction for future research.

CHAPTER 3

# RESPONSIBLE SOURCING AND RISK SHARING

## 3.1 Introduction

As consumers have become increasingly concerned about sustainability and social responsibility, it has been challenging for retailers to manage sustainable and responsible practices behind products. The challenges come not only from retailers themselves but also from their suppliers. In 2019, it was reported that cocoa suppliers to Hershey, Mars, and Nestlé in Western Africa were still using child labor [128]. When it comes to such violations of social responsibility, consumers do not differentiate between where their sources, and retailers are usually held responsible because they are expected to manage their supply chain [65]. Sometimes even violations from deep-tier suppliers can have an impact on retailers. In 2014, an explosion caused by excess dust in a Chinese motor wheels factory killed 146 people [134]. The factory belonged to a supplier of General Motors (GM). Although GM later claimed that the supplier was a tier-2 supplier and had to direct business relationship with GM [115], the claim did not prevent GM from being mentioned in news articles. Most of the time, retailers end up suffering not only from direct losses such as profit loss due to boycotts but, more importantly, from indirect loss such as a damaged public image. In this paper, we refer to such losses from violations as reputation losses. For firms, it is essential to manage reputation risk throughout the supply chain.

Common strategies for managing reputation risk from suppliers, such as dual sourcing [55], supplier auditing [31], supplier certification [35], and supplier penalty [43], aim to avoid or regulate risky suppliers. The availability of a particular strategy depends on the retailer's industry and bargaining power. For example, dual sourcing is less feasible if the number of available suppliers is limited. Moreover, supplier auditing and certification require that the retailer wield sufficient power in the supply chain. In the matter of deep-tier suppliers, it is even more difficult to apply these strategies, as the retailer lacks direct business relationships with them and can only rely on tier-1 suppliers to implement its standards [129].

An alternative idea for managing reputation risk from suppliers is to reduce the cost of taking risk-free actions by offering suppliers a higher price or sharing some of their costs. Consider a chain restaurant brand such as KFC and Burger King who is sourcing chicken from a livestock farm. The farm is facing the risk of livestock diseases (e.g., parasites and bird flu), which will affect its production quantity. The farm can act to prevent these diseases. Some actions, such as using antibiotics and other feed additives that have human health concerns, can cause reputation risk for the restaurant brand [1]. Alternatively, the farm can also choose actions that are risk-free for the brand despite being more costly. For instance, the farm can adopt disease-resistant breeds, improve feed quality, or upgrade farm facilities to avoid contact with wild animals. Given that the farm will incur a sig-

[1] In 2008, Tyson Foods, a meat supplier of KFC, was found to use Gentamicin in poultry whereas the brand labels its product as "raised without antibiotics" [94].

nificant loss if livestock diseases occur, the main trade-off for the farm is reducing the possibility of diseases and minimizing cost input. Therefore, the brand can offer to share the loss to diseases (if they occur) in order to make those risk-free actions more affordable for the farm. We refer to such a strategy as risk sharing hereafter.

In this paper we experimentally study how a retailer (she) can manage reputation risk originating in a supplier (he). The supplier faces a probabilistic cost shock and can exert effort to reduce its probability. There are two types of efforts which differ in their effort cost and whether they comply with the retailer's standards. A non-compliant effort (with low cost) will cause reputation risk to the retailer, whereas a compliant one (with high cost) will not. The retailer offers a risk-sharing contract consisting of a wholesale price and a percentage by which the retailer will share the cost shock if it occurs. Because these decisions are made by managers, we use controlled lab experiments with human participants to study the setting. We aim to answer the following research questions: a) how effective a risk-sharing contract is in inducing compliant effort from the supplier, and b) whether both wholesale price and sharing ratio are necessary for risk sharing to be effective. If our results suggest that risk sharing is not ideal, we are also interested in how the outcome can be improved in terms of achieving more compliant effort.

Experimental results show that human retailers offer suppliers significantly higher wholesale prices but lower sharing ratios compared with theoretical predictions. Regarding supplier decisions, compliant effort accounts for 25.82% of all decisions, while the normative theory suggests that suppliers should never choose compliant effort. Surprisingly, we find that the choice of compliant effort is driven by the wholesale price but not the sharing ratio, which contradicts the theory. As a result, retailers achieve lower profits and supplier profits are higher, leading to a fairer outcome than predicted. We then propose two possible ways to further induce compliant effort. The first is pre-game communication, where the retailer and the supplier can freely communicate at the beginning without making a commitment. The second is unstructured bargaining, where both parties negotiate contract terms and choice of effort. We discuss behavioral hypotheses for each alternative.

We believe our work makes both academic and practical contributions. Starting with the academic aspect, we show that decisions made by human participants significantly deviate from the normative theory, implying that behavioral factors cannot be ignored when studying the problem. Most importantly, given that the theory predicts that responsible sourcing is impossible under this setting, our results suggest that it is possible. We also discuss how to further improve the outcome, and more research in this direction is necessary. Turning to our practical contribution, our findings can help retailers better determine contract terms. Specifically, our results suggest that suppliers care more about wholesale prices than about sharing ratios. This is important for retailers who want to achieve responsible sourcing by implementing a risk-sharing contract with both tier-1 and deep-tier suppliers.

The reset of the paper is organized as follows. In Section 3.2, we provide an extended review of analytical, empirical, and experimental literature in sustainability and social responsibility. In Section 3.3, we outline the analytical model. We then design a controlled lab experiment and discuss behavioral hypotheses in Section 3.4, and we show experimental results and analyses in Section 3.5. We propose two possible ways to further improve responsible outcomes in Section 3.6. In Section 3.7, we summarize our work and discuss opportunities for future research.

## 3.2 Literature Review

Social responsibility and sustainability have received increasing attention in recent decades. In this section, we review relevant studies on these two topics. The review mainly focuses on sourcing, our main interest, but also covers other topics.

Beginning with analytical research, the existing literature considers several popular strategies in managing non-compliance risk such as auditing, certification, dual sourcing, and risk sharing. [31] study the effectiveness of supplier certification, auditing, and contingency payment in mitigating supplier responsibility risk. In practice, the availability of these instruments varies from industry to industry because of the nature of the product and bargaining power of the buyer. They find that instruments are complementary and that when acting alone, certification is the most effective. Continuing with auditing, [25] study how multiple buyers should collaborate in auditing and reduce cost, either through joint auditing or sharing auditing information. [29] also show that buyers sharing auditing costs can improve overall efficiency. However, [48] suggest that cooperation may not improve social responsibility when a firm can benefit from violating other firms. [32] further consider a scenario where auditor and supplier may collude, which weakens the effectiveness of auditing. [99] show that sometimes firms choose to integrate suppliers in order to reduce responsibility for violation risks instead of auditing.

When a less powerful buyer is not able to audit or integrate suppliers, dual sourcing and risk sharing become more attractive. [55] study a buyer sourcing from a responsible but costly supplier and a less costly supplier who may have responsibility for violation risk. The buyer is selling to two segments: socially concerned customers who are willing to pay more for responsibly sourced products and may not purchase if a product is sourced from the risky supplier upon violation. They compare four sourcing and selling strategies and determine when each is optimal. They also show that increasing customers' willingness to pay for responsibility can lead to riskier sourcing but that increasing the cost of violation will always lead to more responsible sourcing. [43] consider a two-tier supply chain of buyer and supplier. Both parties are subjected to operational risk with the same probability, while the supplier can exert "good" or "bad" effort to mitigate the risk. Good and bad effort are defined in the dimension of sustainability compliance. While both types of effort can mitigate operational risk, bad effort will have an impact on the buyer's reputation risk. The authors show that by sharing some of the supplier's loss, the buyer may reduce its reputation risk.

However, these strategies may be less available in a supply chain of more than two tiers because of increased "supply chain distance" and information asymmetry. [63] consider effort decisions to avoid social responsibility violation in a three-tier supply chain. In their model, both tier-0 and tier-1 can make efforts to reduce the violation probability of a tier-2 supplier, and efforts from the tier-0 buyer and the tier-1 supplier are interchangeable. They find that in equilibrium, at most one tier will exert effort. Similarly, [119] consider managing responsibility violation risk in a three-tier supply chain. In their model, only the tier-1 manufacturer can observe whether a tier-2 supplier is risky because of information proximity, and sourcing decisions are made by the manufacturer and are not fully contractible between the buyer and the manufacturer.

Although many studies focus on how firms can detect or avoid responsibility violation, sometimes firms have less incentives to do so. For example, disclosing violations can result in a firm's lower estimated future value. [72] study how mandatory disclosure affects a firm's willingness to learn and reduce its supplier's impact. They find that such disclosure can backfire, as a firm may be reluctant to exert effort to learn about its supplier. Another example is [136]. They consider auditing from a supply chain network perspective. In particular, if a buyer identifies non-compliance in the network, it can either rectify the situation or drop the supplier. Dropping a supplier can change network topology and increase the buyer's cost. The authors show that the buyer's optimal strategy has two stages:

auditing and dropping some suppliers in the first stage, and either auditing the rest or halting the entire auditing in the second stage. In halting auditing, the buyer chooses to tolerate potential non-compliance in the network.

Turning to the empirical literature, [105] find that both the penalty of contract termination after warning and the incentive of supplier training and public recognition can reduce violations. Examining stock prices of Chinese firms, [88] show that the stock market reacts negatively to announcements of environmental incidents. [93] find that firms sometimes learn from inspection experience and improve their environmental performance.

Empirical research also reports the difficulty of managing sustainable practices in multi-tier supply chains. [126] find that higher-tier suppliers tend to address environmental and labor issues passively, which makes them the riskiest part of the supply chain. To build a sustainable network, tier-1 suppliers often play a double agency role: fulfilling tier-0 buyers' sustainability requirements and implementing these requirements with tier-2 suppliers, as discussed by [129]. However, this "cascading" effect often fails in practice, and [125] investigate the reasons behind it. They reveal that tier-1 suppliers' procurement units, who directly engage with higher-tier suppliers, are often left out of tier-0 buyers' sustainability practices. [127] distinguish between environmental and social responsibility and identify drivers for each using multi-method research. They conclude that the former responsibility can be fostered through stakeholder pressures and relational drivers whereas the latter is harder to address. In a survey experiment, [60] find that firms are held accountable upon responsibility violation from suppliers regardless of supply chain distance, firm size, and the strategic importance of the supplied product.

Responsible and sustainable sourcing has been widely studied in experimental literature as well. Building on [55], [92] experimentally investigate how two types of consumer reactions affect responsible sourcing decisions. They find that both encouraging reactions that emphasize the value of responsible sourcing and discouraging reactions that emphasize demand loss upon violations lead to more responsible sourcing decisions, whereas theory predicts that only discouraging reaction is effective. [135] behaviorally compare sourcing decisions under supply disruption risk and responsibility violation risk. Their results show that although over-diversification bias exists among human buyers under two types of risk, diversification is less likely when facing responsibility violation risk.

In addition to sourcing, other topics in sustainability and social responsibility have been studied in the literature, and we provide a quick summary of these topics in this paragraph as it is not the main focus of this paper. The first topic is closed-loop supply chains or reverse logistics where customers can return failed products for remanufacturing [e.g., 106, 117, 22]. The second is consumer communication, i.e., how firms should disclose their sustainable and responsible practices. Firms may have less visibility into their deep-tier suppliers; therefore, disclosure can be costly. However, studies also show that disclosure can lead to more responsible practices [78] or consumer valuations [77, 79]. The last one is waste management in various industries, such as fast fashion [90] and energy [9]. See [83, 2, 10] for comprehensive literature reviews.

Building on the theoretical framework of [43], we experimentally investigate how a human buyer can manage responsible practices from a human supplier. We differ from the existing sourcing literature in the behavioral operations management literature in the following respects. First, existing works focus on the decisions of a single firm, whereas we consider the interaction between buyer and supplier. Second, in the setting of [92] and [135], the buyer can self-manage the reputation risk. In our setting, the buyer's reputation risk is determined by the supplier's decision. Third, existing studies consider sourcing strategies to achieve responsible sourcing, whereas we investigate risk sharing as an alternative strategy.

## 3.3 Normative Theory

Consider a buyer (she) who sells a product to her market and earns revenue $V$. The buyer sources the product from a supplier (he) at wholesale price $W$. The supplier's production cost is normalized to zero, and he may incur a probabilistic cost shock $C$. The supplier can exert effort $e$ to reduce the probability of the shock. There are two types of effort that the supplier can choose from: a) effort complying with the buyer's social responsibility standard, denoted by $c$; and b) effort not complying with the buyer's standard, denoted by $n$. We assume that the supplier can choose only one type of effort. Namely, the supplier first determines the effort type $i = \{c, n\}$ and then the effort level $e$.

Given $i$ and $e$, the probability of the supplier's cost shock is $p_c(i, e) \in [0, 1]$. The probability function $p_c(i, e)$ is decreasing and convex in $e$ for $i = \{c, n\}$, which means more effort leads to a lower probability of cost shock. We assume $\frac{\partial p_c(n,e)}{\partial e} \leq \frac{\partial p_c(c,e)}{\partial e} < 0$ for any $e$, i.e., that non-compliant effort is at least as efficient as compliant effort in reducing the risk. This assumption reflects the practice that the supplier has incentive to take non-compliant actions because of their cost advantage and high accessibility.

Non-compliant effort will have additional impact on the buyer. Consumers are often concerned about a buyer's social responsibility practice throughout the supply chain. If consumers find that a buyer sources from socially irresponsible suppliers, the buyer will be largely held accountable [63]. As a consequence, the buyer may suffer from boycott and have to repair her public relations. Therefore, she will incur a reputation loss $L$ with probability $p_l(i, e)$. We assume that the probability is increasing in non-compliant effort $e$, i.e., $\frac{\partial p_l(n,e)}{\partial e} > 0$. In other words, if more non-compliant actions are taken by the supplier, they are more likely to be discovered by the public. Note that reputation risk will not exist if compliant effort is chosen, i.e., $p_l(c, e) = 0$ for any $e$.

We consider a risk-sharing contract where the supplier's cost shock can be proportionally shared by the buyer. Specifically, the supplier receives a contract $(W, \delta)$ from the buyer where $\delta$ is the buyer's share of the supplier's cost shock. The contract is equivalent to a wholesale price contract if $\delta = 0$. Following [43], we assume ex post limited liability, i.e., a contract is accepted only if ex post profit is always non-negative. Namely, a supplier will only accept a contract if his ex post profit is always greater than zero regardless of the cost shock.

### 3.3.1 Centralized System

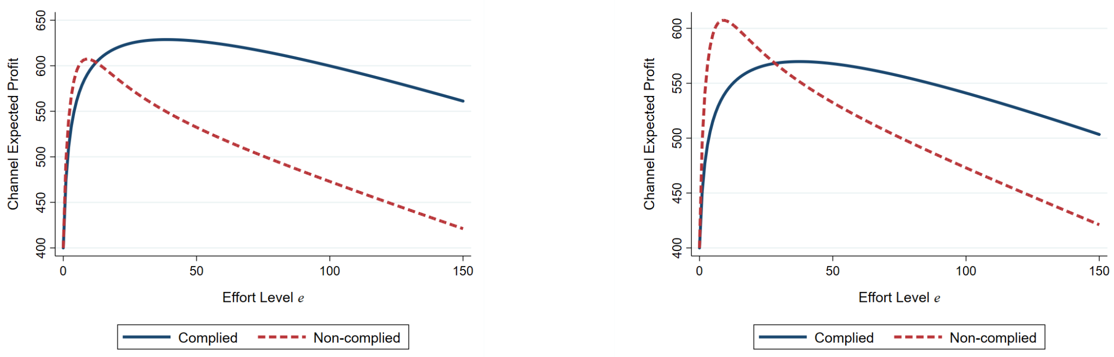
We begin by introducing the centralized case where the buyer and the supplier are integrated. In a centralized system, there is no transfer payment or cost-shock sharing, and the optimization problem for the channel is

$$\max_{i=\{c,n\},e\geq 0} \mathbb{E}\Pi^c = V - p_c(i,e)C - p_l(i,e)L - e. \tag{3.1}$$

One can see from Equation (3.1) that whether compliant or non-compliant effort is preferred by the channel will depend on $p_c(i,e)$ and $p_l(i,e)$, i.e., the effectiveness of both effort types and the impact of non-compliant effort on reputation risk. Figure 3.1 provides numerical examples of expected channel profits with respect to effort level. In Figure 3.1a, compliant effort can maximize the expected channel profit and is better than non-compliant when the effort level is not too low. As a comparison, compliant effort is less effective in Figure 3.1b and the channel prefers non-compliant effort. In addition, when the effort level is high, compliant effort outperforms non-compliant effort because high compliant effort can also reduce the expected cost shock sufficiently, and high non-compliant effort leads to a high expected reputation loss, which makes the channel worse off.

Figure 3.1: Numerical Examples of Expected Channel Profits with respect to Effort Level

(a) Compliant Effort is Preferred (b) Non-compliant Effort is Preferred




Note: $V = 800$, $C = 400$, $L = 200$. $p_c(n, e) = \frac{1}{(1+e)^{0.4}}$ and $p_l(e) = \frac{e^{20}}{(1+e)^{20}}$ for both figures. $p_c(c, e) = \frac{1}{(1+e)^{0.3}}$ for the left figure and $p_c(c, e) = \frac{1}{(1+e)^{0.2}}$ for the right.

### 3.3.2 Decentralized System

For a decentralized system, we introduce the model by backward deduction, beginning with the supplier. Given the buyer's contract $(W, \delta)$, the supplier chooses effort type $i$ and level $e$ to maximize his expected profit function shown below.

$$\max_{i=\{c,n\}, e \geq 0} W - p_c(i, e)(1 - \delta)C - e. \tag{3.2}$$

In equation (3.2), as $\frac{\partial p_c(c,e)}{\partial e} > \frac{\partial p_c(n,e)}{\partial e}$, a rational supplier will always choose $i = n$ instead of $i = c$. The supplier's optimal effort level $e^*(\delta)$ satisfies

$$\frac{\partial p_c(n, e)}{\partial e} = \frac{1}{(\delta - 1)c}. \tag{3.3}$$

Condition (3.3) suggests that the supplier's optimal effort is dependent on the buyer's sharing ratio $\delta$ but not on the wholesale price $W$. Moreover, $e^*(\delta)$ decreases in $\delta$ as $p_c(n, e)$ is convex in $e$, i.e., the supplier has less incentive to reduce the cost-shock risk if the buyer shares a larger portion of it.

Anticipating the supplier's decision, the buyer proposes a contract $(W, \delta)$ to the supplier to maximize her expected profit with constraints:

$$\max_{W\in[0,V],\ \delta\in[0,1]} \mathbb{E}\Pi^b = V - W - p_c(i^*, e^*)\delta C - p_l(i^*, e^*)L \tag{3.4}$$

$$\text{s.t. } W - (1-\delta)C - e^* \geq 0, \tag{3.5}$$

$$W - p_c(i^*, e^*)(1-\delta)C - e^* \geq 0, \tag{3.6}$$

$$(i^*, e^*) \in \arg\max_{i=\{c,n\}, e\geq 0} W - p_c(i, e)(1-\delta)C - e. \tag{3.7}$$

Constraints (3.5-3.6) are the supplier's ex post limited liability and individual rationality constraints, respectively. At optimality, Constraint (3.5) will be binding and Constraint (3.6) can be removed. Note that the supplier's decision depends on $\delta$ alone, and that $W$ can be determined when Constraint (3.5) is binding. Let $\delta^*$ denote the buyer's optimal decision, and then the corresponding wholesale price $W^* = (1-\delta^*)C + e^*$ where $e^*$ is the supplier's optimal decision given $\delta^*$.

While the normative theory suggests the supplier should never choose compliant effort, and his effort level depends only on the buyer's sharing ratio $\delta$, it is unclear whether this is true for human decision-makers. For instance, the supplier exerting non-compliant effort means that the buyer will experience a reputation loss. Past experimental results indicate that social preferences play a role in human decision-making [89]. Further, a human supplier may also care about the wholesale price as form of direct income. In the next section, we design a controlled lab experiment to test the theory and discuss potential behavioral hypotheses.

## 3.4 Behavioral Experiments and Hypotheses

### 3.4.1 Experiment Design

Given the complexity of the normative theory, we simplify the experimental task while we can still capture the main trade-off. First, we reduce the decision space of supplier effort to a binary choice between compliant and non-compliant effort. In other words, the supplier only decides the type of effort, and the corresponding effort cost is fixed with the compliant effort being more costly.[2] Second, the probability of cost shock is the same regardless of effort type, i.e., both types of effort will lead to the same cost-shock probability. Finally, the probability of reputation loss is set to 1. By simplifying the task, we aim to avoid suboptimal decisions due to complexity. After simplification, the supplier is essentially choosing between a costly effort (compliant) and a less costly effort (non-compliant) with negative externality for the buyer (reputation loss). Note that we remove any context of social responsibility throughout the game to avoid the impact of the context. In particular, the reputation loss is referred to as buyer loss, and efforts are described as compliant or not compliant to the buyer's standard. If the context is provided,

[2]Preliminary results from pilot sessions suggest that even limiting each effort type to two levels (four options in total) creates significant complexity for participants.

human participants may be more willing to choose compliant effort simply for in-game positive externalities [3], which is not what we expect to observe.

Regarding parameters, we set the product revenue $V = 800$ and the cost shock $C = 400$. The reputation loss $L$ is set as 200 such that it is sufficiently significant for the buyer. The probability of cost shock is 0.5 regardless of effort type. Regarding effort cost, we set the cost as 110 for compliant effort and 100 for non-compliant effort. The reason for this small cost difference is that preliminary results from pilot sessions indicate that human suppliers are highly sensitive to the cost difference between the two effort types and never choose compliant effort. Thus, we aim to provide a setting in which compliant effort is attractive to at least some participants.

Table 3.1: Experimental Predictions

| | Normative Prediction |
|---|---|
| Wholesale Price $W$ | 100.00 |
| Sharing Ratio $\delta$ (%) | 100.00% |
| Preferred Effort Type | Non-compliant |
| Expected Buyer Profit | 500.00 |
| Expected Supplier Profit | 0.00 |
| Expected Channel Profit | 500.00 |

Note: Product revenue $V$ is 800. Cost shock $C$ is 400. Reputation loss $L$ is 200. The probability of cost shock is 0.5 regardless of effort type. The probability of reputation loss is 1. A rational supplier will always choose non-compliant effort, which costs 100.

Table 3.1 shows theoretical predictions given the parameters. A rational supplier will always prefer non-compliant to compliant effort. Anticipating this, the buyer chooses to share all of the cost shock so that she can offer a lower wholesale price. Recall that the buyer's optimal wholesale price is obtained when Constraint (3.5) is binding. Therefore, the predicted wholesale price is equal to the cost of non-compliant effort, 100. In this case, the buyer takes all profit for the channel and leaves none for the supplier. The intuition is that sharing all cost shock means only a probabilistic loss to the buyer but that a lower wholesale price means keeping more profit for sure.

We also provide decision support to further reduce complexity. The buyer is given two sliding bars to test wholesale price and sharing ratio, respectively. Resultant buyer profit with and without cost shock are updated in real time. The buyer is also allowed to test the supplier's decision, i.e., compliant or non-compliant effort. Turning to the supplier who only has to decide his effort type, we show all the information for each effort type, including realized supplier profit with and without a cost shock, expected supplier profit, expected channel profit, and the corresponding reputation loss (which shows as zero for compliant effort). The supplier is also allowed to reject the buyer's offer. Upon rejection, both parties earn a profit of zero. A tutorial video demonstrating the decision support tool was played before the game began, to help participants understand their tasks.

Before the game, each participant was assigned a role, which remained fixed throughout. Participants played a trial round to help familiarize themselves with the game. The outcome of the trial round did not count towards actual earnings [3]. After the trial round, participants played 15 rounds. We used random matching

[3] Results from the trial round are also not included in our analysis.

for each round, i.e., a buyer was randomly match with a supplier in a round. Regarding sample size, 30 participants were recruited from a large university, and cash was the only incentive offered.

Our experiment was implemented through oTree [28]. because of the COVID-19 pandemic, we ran all sessions synchronously online through Zoom. Before a session started, a researcher read the instructions aloud, played the tutorial videos of decision support, and answered any questions. Subjects received cash based on total profits from the game plus a show-up fee of $10. Average earnings were roughly $30.34. Each session lasted 60 minutes on average.

### 3.4.2 Behavioral Hypotheses

Given that the normative theory assumes that decision-makers are rational, it is likely that behavioral biases drive decisions from optimality. In this subsection, we discuss several potential behavioral hypotheses.

Beginning with the buyer's decision, the normative theory predicts extreme values for both wholesale price and sharing ratio, which leaves the supplier zero profit. However, such extreme predictions are rarely observed in lab settings.[4] Based on past experimental studies, social preferences such as fairness and justice may play a role here as well. On one hand, a buyer may care about the supplier's

[4] For example, recall that the observed transfer prices deviating from extreme predictions in opposite directions in Chapter 1.

payoff and offer a higher price, especially because the buyer is predicted to earn a disproportionately high split of the profits [108]. On the other hand, the buyer may be unwilling to share all of the cost shock and to leave the supplier free of risk [e.g., 62, 139]. Therefore, we expect to see an average sharing ratio of less than than 100%. In sum, we have the first hypothesis.

**Hypothesis 3.1.** *The average wholesale price will be higher than the normative prediction, 100. The average sharing ratio will be lower than the predicted value, 100%.*

Turning to the supplier, while a rational supplier will always choose non-compliant effort, a human supplier may care about the reputation loss for the buyer caused by non-compliant effort. Lab evidence suggests that actions causing negative externalities are less accepted [13]. As a result, we expect to see a non-zero compliant effort rate, as described in Hypothesis 3.2.

**Hypothesis 3.2.** *The percentage of compliant effort chosen among all participants will greater than 0%.*

Because the buyer offers both a wholesale price and a sharing ratio of cost shock, we are also interested in what affects the supplier's choice of compliant effort. The normative theory suggests it should solely depend on the sharing ratio. However, the sharing ratio is meaningful only when the cost shock is realized. Instead, the supplier may treat the wholesale price as a more substantial offer. Here we provide two hypotheses about the potential drivers of compliant effort: Hypothesis 3.3a, which follows the normative theory, and Hypothesis 3.3b, where wholesale price also affects the supplier's choice.

**Hypothesis 3.3a.** *The supplier's choice of compliant effort will depend on the sharing ratio only.*

**Hypothesis 3.3b.** *The supplier's choice of compliant effort will depend on both the wholesale price and the sharing ratio.*

If the above hypotheses are all supported, it is still unknown how the profits of both parties will change. For instance, a higher wholesale price benefits the supplier (but hurts the buyer), while a lower sharing ratio hurts the supplier (but benefits the buyer). Moreover, choosing compliant effort will increase the buyer's profit and channel profits. Therefore, we do not hypothesize about profits and turn to experimental results, shown in the next section, for answers.

## 3.5 Results

In this section we present the experimental results, starting with decisions and then profits. Recall that suppliers were allowed to reject buyers' offers in the experiment; the average rejection rate is 10.22%. For all the results shown in this section, rejected data are excluded unless noted otherwise. We also check whether there is any learning of buyers in early rounds (because suppliers make binary decision only). Results suggest that buyers offer significantly higher prices in early rounds but that the sharing ratio is varies little throughout the game. For all analyses, we include the full data excluding the trial round and use regression with random effects for hypothesis testing.

Beginning with decisions, Table 3.2 shows average decisions of both parties compared with the normative predictions. For the buyer, the average observed wholesale price is much higher than predicted, 410.30 versus 100.00 ($p < 0.01$). Such a significant deviation is understandable, as the predicted price leaves the supplier zero profit. Giving a product revenue of 1,000, this means that buyers are, on average, giving suppliers 41.03% of the channel revenue, which is consistent with typical results from the ultimatum game [58]. Regarding the sharing ratio, the average $\delta$ is 41.22%, significantly lower than the predicted 100% ($p < 0.01$), suggesting that buyers are unwilling to share the entire cost shock. Recall that in the normative model, the optimal wholesale price is derived when Constraint (3.5) is binding given a sharing ratio. The average optimal wholesale price conditioning on the observed sharing ratio is 335.13, which is significantly lower than the observed value, 410.30 ($p < 0.01$). All of these suggest that buyers prefer high prices combined with less risk sharing instead of low prices combined with more risk sharing. Overall, Hypothesis 3.1 is supported, and we have the first result.

**Result 3.1.** *In a two-tier supply chain with risk sharing, compared with the normative predictions, buyers prefer a contract of higher wholesale prices and less risk sharing.*

We also investigate heterogeneity in the buyer's decisions. Figure 3.2 shows the distribution of observed wholesale prices and sharing ratios. In Figure 3.2a, the distribution of wholesale prices is skewed to the left (D'Agostino-Pearson normality test, $p < 0.01$), indicating that the majority of buyers prefer a lower price closer to the optimal. In contrast, the observed sharing ratio in Figure 3.2b is not

Table 3.2: Average Wholesale Price, Sharing Ratio, and Compliant Effort

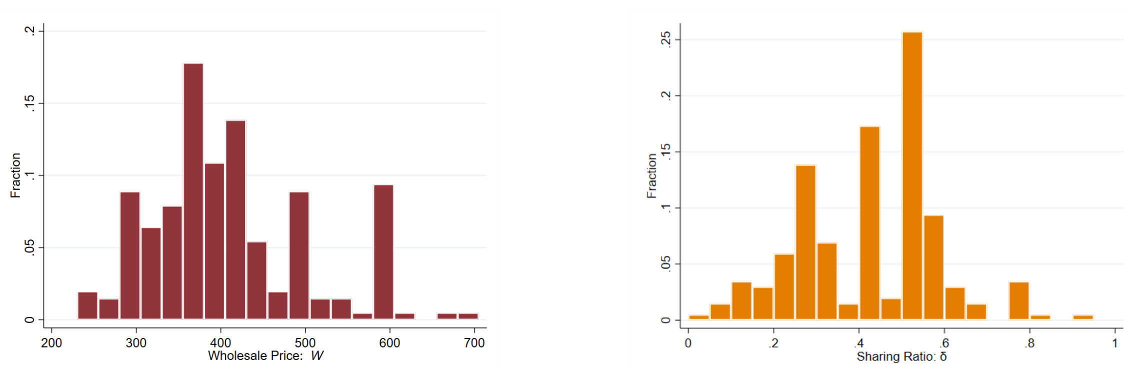
| | Observed | Predicted |
|---|---|---|
| Wholesale Price $W$ | 410.30$^{\ddagger}$ (18.78) | 100.00 |
| Sharing Ratio $\delta$ (%) | 41.22%$^{\ddagger}$ (0.04) | 100.00% |
| Compliant Effort (%) | 25.82%$^{\ddagger}$ (0.07) | 0.00% |

Note: Rejected data are excluded. Standard errors across participants are reported in parentheses. Compliant effort is a binary decision by suppliers. Significance of regressions with random effects comparing observed with normative predictions are given by $^{\ddagger}$ $p < 0.01$.

significantly different from a normal distribution (D'Agostino-Pearson normality test, $p = 0.633$).

Figure 3.2: Distribution of Observed Decisions of Buyers

(a) Wholesale Price (b) Sharing Ratio




Note: Rejected data excluded.

Turning to suppliers, the average compliant rate, i.e., the percentage of compliant effort, is 25.82%, significantly higher than the predicted 0% ($p < 0.01$). While one may not argue that such "compliant rate" is sufficiently high, the result at least suggests that responsible practice is not impossible when behavioral factors are considered and it is contrast to the theory. Overall, Hypothesis 3.2 is supported and we have the following result.

**Result 3.2.** *In a two-tier supply chain with risk sharing, suppliers choose to exert compliant efforts more often than predicted, suggesting that inducing responsible practice from suppliers by risk sharing is possible.*

To look into the heterogeneity of suppliers, in Table 3.3 we calculate average compliant rate at individual level and classify participants into five ranges. For example, a individual compliant rate of 40% means the participant chooses compliant effort in six out of 15 rounds. According to the table, 33% of suppliers strictly followed the normative theory and never chose compliant effort and 67% of suppliers exerted compliant effort at least once. Moreover, the 67% of suppliers are rather evenly distributed in three ranges from 0% to 75%. The classification result shows that the average compliant rate is not driven by a small subset of suppliers.

Table 3.3: Average Compliant Effort at Individual Level

| Range | Fraction of Participant |
|---:|---:|
| 0% | 0.33 |
| (0%,25%] | 0.20 |
| (25%,50%] | 0.20 |
| (50%,75%] | 0.27 |
| (75%,100%] | 0.00 |

Note: Rejected data excluded. Classification is based on average compliant rate at individual level.

To answer what it the driver of compliant effort, we report results of logit regressions with random effects in Table 3.4. The dependent variable is the binary outcome of whether compliant effort is chosen. In addition to wholesale price and sharing ratio, we also include several controlled variables in regressions. The first one is the binary outcome of whether cost shock is realized in the previous round, denoted as "lagged cost shock". This variable is included because experiencing loss from cost shock may affect the supplier's perception of the buyer's sharing ratio. The second one is the supplier's profit in the previous round, denoted as "lagged supplier profit", which is the outcome from the last round.

Table 3.4: Logit Regression of Choice of Compliant Effort with Random Effects

| DV: Compliant Effort | | | |
|---|---|---|---|
| Wholesale Price $W$ | $0.012^{\ddagger}$ | $0.013^{\ddagger}$ | $0.013^{\ddagger}$ |
| | (0.003) | (0.003) | (0.003) |
| Sharing Ratio $\delta$ | 2.241 | 1.699 | 1.745 |
| | (1.644) | (1.798) | (1.816) |
| Lagged Cost Shock | | -0.194 | -0.219 |
| | | (0.511) | (0.528) |
| Lagged Supplier Profit | | | -0.001 |
| | | | (0.003) |
| Constant | $-7.774^{\ddagger}$ | $-8.329^{\ddagger}$ | $-8.256^{\ddagger}$ |
| | (1.689) | (1.980) | (2.014) |

Note: Rejected data are excluded. Standard errors of coefficient are reported in parentheses. The dependent variable is a binary outcome of whether a supplier chooses compliant effort or not in a round. Lagged Cost Shock is a binary outcome of whether cost shock is realized in the previous round.

Estimated coefficients in Table 3.4 suggest that only wholesale price is the main reason for suppliers' choice of compliant effort, and that the impact of sharing ratio is much less significant. The significance of wholesale price is robust when controlled variables are included. This is in contrast to the normative theory, which predicts that the supplier's decision depends on the sharing ratio only and is independent of wholesale price. Therefore, both Hypothesis 3.3a and Hypothesis

3.3b are rejected, and we have the third result, stated below.

**Result 3.3.** *In a two-tier supply chain with risk sharing, the supplier's choice of compliant effect depends on wholesale price only and is not affected by the degree of risk sharing, contradicting the normative theory.*

There are a few possible reasons for Result 3.3. The first is that suppliers treat benefiting from sharing ratios as a probabilistic event because it is only effective when cost shock is realized. On the other hand, wholesale price is a guaranteed revenue for suppliers. Therefore, a wholesale price can represent how "good" the buyer's offer is, which eventually affects the supplier's choice of effort. The second reason is that wholesale price is associated with gains whereas sharing ratio is associated with losses because of cost shock. The classic prospect theory suggests that human decision-makers process gains and losses differently [68]. In this case, gains are more attractive to suppliers.

Table 3.5: Average Expected Profits

| | Observed | Predicted |
|---|---|---|
| Expected Buyer Profit | $355.56^{\ddagger}$<br>(14.06) | 500.00 |
| Expected Supplier Profit | $193.47^{\ddagger}$<br>(5.64) | 0.00 |
| Expected Channel Profit | $545.88^{\ddagger}$<br>(10.67) | 500.00 |

Note: Rejected data are excluded. Standard errors across participants are reported in parentheses. Significance of regressions with random effects comparing observed and normative predictions are given by $^{\ddagger}$ $p < 0.01$.

To show how these observed decisions translate into profits, Table 3.5 presents average observed expected profits of both parties as well as the channel, along with the normative predictions. The average buyer profit is 355.56, significantly lower than the predicted 500 ($p < 0.01$). The lower buyer profit is mainly due to a higher wholesale price, as the benefit of a lower sharing ratio and higher compliant rate are relatively small. Turning to the supplier, the average supplier profit is significantly higher than the normative prediction, 193.47 versus 0.00 ($p < 0.01$), because suppliers benefit from higher wholesale prices. The average channel profit is significantly higher than predicted, 545.88 versus 500.00 ($p < 0.01$). The higher channel profit is due to the higher compliant rate, which leads to less reputation loss. The deviations in profit are not surprising, as the normative predictions are extreme. However, the outcome remains relatively fair, as buyer and supplier achieve 64.56% and 35.44% of the channel profit, respectively. Summarizing all the observations, we have the fourth result.

**Result 3.4.** *In a two-tier supply chain with risk sharing, the outcome is fairer than predicted. In particular, buyer profits and channel profits are higher than and supplier profits are lower than the normative predictions.*

Overall, our experimental results suggest that managing responsible practices from the supplier by a risk-sharing contract is possible when behavioral factors are involved, despite risk sharing making a relatively small contribution. Based on these results, we are interested in further improving the compliant rate, which we discuss in the next section.

## 3.6 Improvement

So far we have shown that the observed decisions and outcomes significantly deviate from the theory. In particular, the observed compliant rate is 25.82%, much higher than the predicted 0%. We are interested in how we can further induce higher compliant rates. In this section, we propose two potential solutions to improve responsible outcomes.

### 3.6.1 Pre-game Communication

In the main experiment, a buyer can only send a take-it-or-leave-it offer to a supplier, who can choose to reject it. No other communication is allowed. As a result, the rejection rate is 10.22%, which suggests that buyers may not always have a good sense of suppliers' willingness to accept. Therefore, allowing pre-game communication may enable the exchange of information and achievement of more responsible practices.

Although costless communication is usually considered cheap talk and should not affect decisions from a rational perspective, in the behavioral economics literature, such communication can affect people's decisions, even if it comes from competitors [26]. Moreover, communication has been shown to be effective in promoting equilibrium, facilitating coordination, and improving efficiency [46, 17, 1] because it can reduce information asymmetry.

Several forms of communication have been investigated, including written communication (with participants remaining anonymous) and face-to-face communication (where interactions may not be no longer anonymous). [51] show that both ways of communication can have different impacts, depending on the context and the decision to be made. In our setting, we use written communication to ensure anonymity consistent with the main experiment.

Regarding experimental design, most protocols will be consistent with the main experiment, such as fixed roles and random matching. At the beginning of each round, a buyer and a supplier have two minutes to communicate via chat box. There is no limitation on communication and no direction is given in instructions. Participants are free to pre-announce their decisions, ask questions, or not communicate at all. After the two minutes, the buyer and the supplier sequentially make decisions, just as in the main experiment. The only difference is that we allow participants to view their chat history during decision-making, such that the effect of communication is not mitigated by participants' memory.

Although there is no formal underlying theory of costless communication, we can discuss and hypothesize the impact of communication based on experimental evidence in the literature. First, we expect to observe higher agreement rates as shown in [108]. During the communication, the buyer can better infer the supplier's favorable range of contract terms and make a better offer. Thus, we offer the following hypothesis:

**Hypothesis 3.4.** *If costless pre-game communication is allowed, the agreement rate will* *be higher than in the absence of communication.*

Second, the outcome may be even fairer than what is observed in the main experiment. [5] show that communication can lead to more egalitarian distributions in the dictator game, which shares some similarity with our setting. Therefore, we offer our second hypothesis:

**Hypothesis 3.5.** *If costless pre-game communication is allowed, the gap between buyers and suppliers will be smaller than in the absence of communication.*

[5] find that the receiver who sends a message to the allocator will receive a higher share. If this is true in our setting, we should see more generous contract terms offered by buyers. However, given that there are two decisions in the contract, it is unknown whether a fairer outcome (if observed) is achieved by a higher price or more risk sharing. Although one may anticipate that a higher price is more likely to be the case based on Result 3.3, communication may also emphasize the impact of risk sharing, which increase the value of the sharing ratio for the supplier. Additional treatments are needed to determine the impact of communication and the underlying mechanism.

### 3.6.2 Bargaining

In the main model, a buyer offers an ultimatum offer to a supplier. Another alternative way that may improve responsible outcomes is unstructured bargaining

[23, Chapter 4.1], i.e., where the bargaining process is free-form and communication is unlimited. Through bargaining, both parties can better understand their partner's demands. In particular, the buyer can better know the supplier's desired contract terms and also ask for compliant effort in return. Unlike pre-game communication, bargaining is more substantial and can largely determine the supply chain outcome. In this subsection, we discuss bargaining predictions and hypotheses using the Nash bargaining framework.

Following the Nash bargaining model, a buyer and supplier jointly maximize their Nash product with the constraints of ex post limited liability and individual rationality.

$$
\begin{aligned}
\max_{W\in[0,V],\ \delta\in[0,1],\ i=\{c,n\}} \mathbb{E}[u^{nash}] = &[V - W - p_c(i,e)\delta C - p_l(i,e)L] \\
&\times [W - p_c(i,e)(1-\delta)C - e] && (3.8)\\
\text{s.t. } & W - (1-\delta)C - e^* \geq 0, && (3.9)\\
& W - p_c(1-\delta)C - e^* \geq 0. && (3.10)
\end{aligned}
$$

Depending on the effort type $i$, the Nash product (3.8) can be rewritten as

$$
\mathbb{E}[u^{nash}] = \begin{cases} [V - W - p_c(c,e)\delta C] \times [W - p_c(c,e)(1-\delta)C - e] & \text{if } i = c, \quad (3.11a) \\ [V - W - p_c(n,e)\delta C - p_l(n,e)L] \times [W - p_c(n,e)(1-\delta)C - e] & \text{if } i = n. \quad (3.11b) \end{cases}
$$

Consider the simplified case in the main experiment: both types of effort are equally effective in reducing the cost-shock probability, and compliant effort (Equation (3.11a)) will always achieve a higher Nash product than non-compliant effort (Equation (3.11b)). Therefore, the theory suggests that moving to unstructured bargaining is more likely to achieve compliant effort.

Turning to experimental design, we again simplify the model to be consistent with the main experiment. Further, we fix the sharing ratio at 40%, which is rounded from the average observed value in the main experiment. Our reason is threefold. First, even after simplification, there are three decision variables (wholesale price, sharing ratio, and effort type) to be negotiated. By reducing one decision variable, we can further simplify the experimental task. Second, there is no unique solution to the simplified model. Precisely, for a given sharing ratio $\delta$, there is always a unique wholesale price $W$ that can globally maximize the Nash product (3.8). Thus, it is difficult to evaluate the bargaining outcome if we allow all three terms to be negotiated. Finally, by fixing the sharing ratio at a level similar to that observed in the main experiment, we can better compare the outcomes between ultimatum setting (the main experiment) and unstructured bargaining. In addition, recall that Result 3.3 suggests that wholesale price is the main driver of the supplier's decision, indicating that wholesale price plays a more important role than the sharing ratio.

Other protocols remain the same as in the main experiment. Regarding bargaining process in each round, the buyer and the supplier engage in a three-minute bargaining period. Either party can propose a wholesale price and effort type to the other party, who can accept or decline the offer. They are also given a chat box to freely communicate. If no agreement is made after three minutes, the round ends with both parties earning a profit of zero. Participants are allowed to test all decisions using a decision support tool similar to the buyer's tool in the main experiment. Note that the decision support tool only provides profit for a participant's own role, but not the other party's. Although knowing the other party's profit is helpful for reaching an agreement, we choose to main consistency with the main experiment because showing the other party's profit per se may cause other-regarding preferences.

Table 3.6: Experimental Predictions of Unstructured Bargaining Compared with Results in the Main Experiment

| | Unstructured Bargaining ($\delta$ = 40%) | Main Experiment Observed [Predicted] |
|---|---|---|
| Wholesale Price $W$ | 575.00 | 410.30 [100.00] |
| Sharing Ratio $\delta$ (%) | 40.00% | 41.22% [100.00%] |
| Compliant Effort (%) | 100.00% | 25.82% [0.00%] |
| Expected Buyer Profit | 345.00 | 355.56 [500.00] |
| Expected Supplier Profit | 345.00 | 193.47 [0.00] |
| Expected Channel Profit | 690.00 | 545.88 [500.00] |

Note: Sharing ratio for unstructured bargaining is fixed at 40%. Normative predictions of the main experiment are displayed in square brackets. Product revenue $V$ is 800. Cost shock $C$ is 400. Reputation loss $L$ is 200. The probability of cost shock is 0.5 regardless of effort type. The probability of reputation loss is 1.

Table 3.6 provides normative predictions for unstructured bargaining along with the average observed results (normative predictions shown in square brackets) in the main experiment. Keeping the sharing ratio at 40%, unstructured bargaining predicts a much higher wholesale price than observed in the main experiment. In addition, the buyer and the supplier should always prefer compliant effort. Therefore, we can hypothesize the change of compliant rate in the bargaining treatment.

**Hypothesis 3.6.** *Moving from ultimatum offer to unstructured bargaining, the supply chain can achieve more compliant effort.*

Table 3.6 also shows that at equilibrium, both parties achieve the same expected profit. In particular, the buyer is predicted to earn a slightly lower profit than observed in the main experiment, 345 versus 355.56, whereas the supplier is predicted to achieve a much higher profit, 345 versus 193.47. The more equal outcome, without hurting the buyer, comes from the higher compliant effort rate, which increases the channel profit. We hypothesize the improved equity in the following hypothesis.

**Hypothesis 3.7.** *Moving from ultimatum offer to unstructured bargaining, the buyer will not earn less profit, but the difference between the buyer's and the supplier's profits will be smaller.*

Although unstructured bargaining predicts that compliant effort is always preferred, it may not hold in the lab, as we have observed multiple deviations from extreme predictions (e.g., transfer price in Chapter 1 and effort type in this chapter). Therefore, an additional lab experiment is needed to test the hypotheses and compare with the main experiment.

## 3.7 Conclusion

In this paper, we study how a retailer manages reputation risk originating from a supplier through risk sharing. The supplier faces a random cost shock and can exert effort to mitigate the risk. In particular, it can choose between effort that complies with the retailer's standards without causing the retailer reputation loss and effort that does not comply with the standards and will cause the retailer reputation loss. The retailer offers a risk-sharing contract consisting of a wholesale price and the portion of the cost shock that will be shared upon realization.

We experimentally investigate the setting, and the results significantly deviate from the normative theory. Beginning with contract terms, the normative theory predicts that the retailer should share all the supplier's entire cost shock and set the wholesale price as low as possible, and that the supplier should always choose non-compliant effort. By contrast, the experimental results show that the retailer places a higher price but offers less sharing of the cost shock. The supplier chooses compliant effort 25.82% of the time, significantly higher than the predicted 0%. However, while the theory predicts that the supplier's effort decision depends only on the sharing of cost but not the wholesale price, our data suggest otherwise: only the wholesale price drives the effort decision. Finally, all these deviations lead to a smaller difference between the retailer's and the supplier's profit, i.e., a fairer outcome.

The data also imply that cost-shock sharing by the retailer may be undervalued by the supplier. In an effort to further improve responsible practice, we discuss the possibility of two alternatives that allow more information exchange: pre-game communication and unstructured bargaining. In the former, the two parties are allowed to communicate and are not required to commit to any decision during communication. In the latter, the retailer and the supplier negotiate contract terms instead of a sequential decision-making in the main experiment. We also provide hypotheses based on the literature and an analytical model.

Overall, our preliminary results suggest that a risk-sharing contract can achieve responsible sourcing despite the risk-sharing component playing a less significant role. However, it is sufficient to argue that it is worth further studying how a retailer can utilize risk sharing to improve responsible outcomes. Regarding future research, it will be interesting and also necessary to investigate risk sharing in multi-tier supply chains. Unlike supplier auditing or certification that cannot be easily implemented with deep-tier suppliers, risk sharing is more feasible, as it offers extra benefits instead of making additional requests. However, the problem will be more complex in multi-tier supply chains. Taking as an example a three-tier supply chain with a tier-2 supplier being the risky one, the retailer will have at her disposal more risk-sharing strategies, such as direct sharing with the tier-2 supplier or delegating the sharing to the tier-1 supplier. Further, the tier-2 supplier may care less about the reputation risk of the retailer because they do not have a direct business relationship. All of these suggest that more research is needed to answer how a retailer can better manage responsible sourcing in multi-tier supply chains.

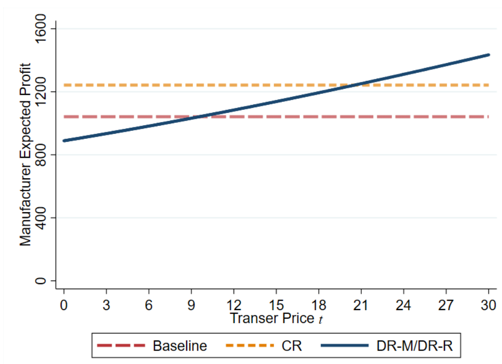
APPENDIX A

# APPENDIX OF CHAPTER 1

## A.1 Supplemental Experimental Prediction Figures

Figure A.1: Manufacturer Expected Profits with respect to the Transfer Price and the Wholesale Price

(a) Transfer Price




(b) Wholesale Price

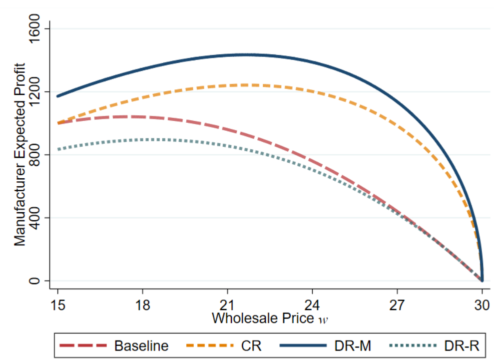
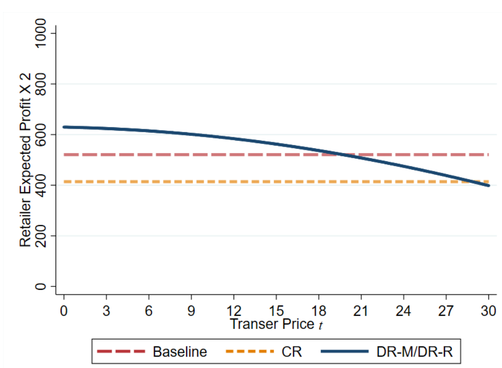



Figure A.2: Retailer Expected Profits with respect to the Transfer Price and the Wholesale Price

(a) Transfer Price




(b) Wholesale Price

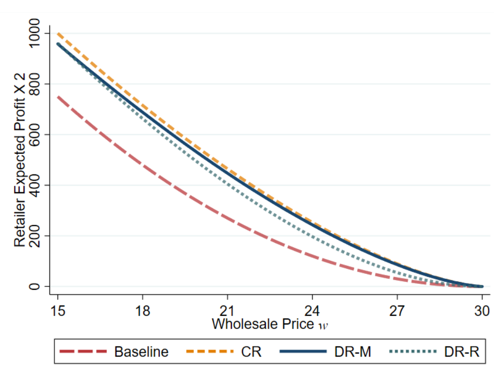
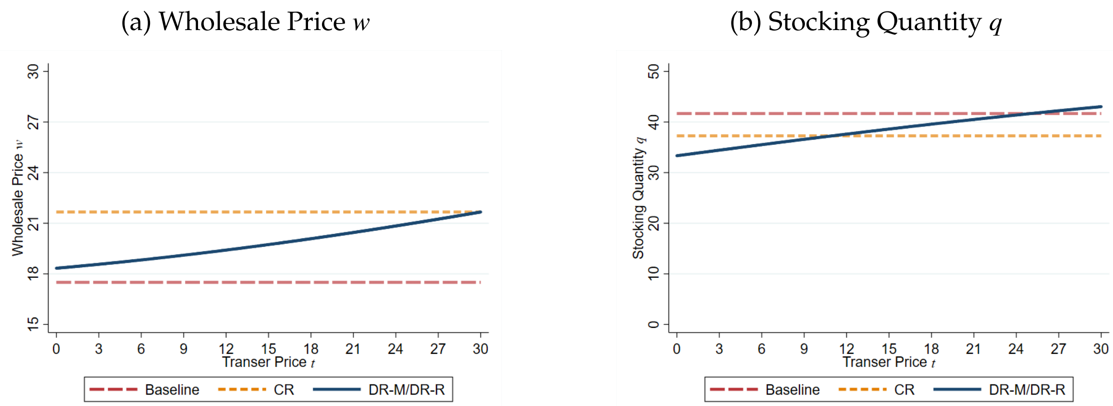


Figure A.3: Wholesale Price and Stocking Quantity with respect to the Transfer Price




## A.2 Power Analysis

Here we provide a power analysis for all hypothesis test results by tables and formal results (1.1-1.5). Due to complexity of conducting power analysis on random effects regression, we show power of t-tests with decisions collapsed within subject (t-tests generally lead to higher p-values in our hypothesis testing). Power levels are reported with critical p-values using Bonferroni corrections, assuming an unadjusted critical p-value of 0.05. The p-values of t-tests are reported in square brackets. Sample sizes are reported in parentheses when applicable.

Table A.1: Power for Table 1.2 and Result 1.1

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

Table A.2: Power for Table 1.3

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

Table A.3: Power for Results 1.2 and 1.4

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

Table A.4: Power for Result 1.3

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

Note: "Retailer's Share" is retailers' share of supply chain profit. "Manufacturer-Retailer Gap" s absolute difference between manufacturer and retailer profits.

Table A.5: Power related to DR-0 and Result 1.5

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

## A.3 Behavioral Model Details

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

### A.3.1 Fairness Model

#### A.3.1.1 Manufacturer Advantageous Fairness

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

$$u_m^{d,F} = \pi_m - \lambda_m(\pi_m - \pi_{r,i}^d)^+, \tag{A.1}$$

where $\lambda_m$ is the manufacturer's degree of fairness concerns. Note that in Equation (A.1) the manufacturer compares its profit with a retailer's normative expected profit without taking retailer's fairness concern into account. This is because our decision support system shows the manufacturer a retailer's optimal quantity and expected profit. Moreover, this allows us to separate the two parties' fairness concerns.

Figure A.4 provides a numerical analysis of how manufacturer advantageous concerns affect its decisions in DR-M and DR-R. While Figure A.4a shows that $\lambda_m$ does not affect the manufacturer's optimal transfer price in DR-M, Figure A.4b shows that a higher $\lambda_m$ makes the manufacturer choose a lower wholesale price in DR-M and DR-R as it achieves a more equitable outcome. In DR-M, one explanation for this result is that a lower wholesale price reduces inequity while keeping the manufacturer expected profit relatively high (since retailers will set a higher quantity in response). On the other hand, reducing the transfer price will lead to lower quantities and have a more significant detrimental impact on the manufacturer's expected profit. Further, the transfer price has less impact on retailer profit compared to the wholesale price. These effects can be seen in Figures A.1 and A.2 in Section A.1.

#### A.3.1.2 Retailer Disadvantageous Fairness

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

$$u_{r,i}^{d,F} = \pi_{r,i}^{d} - \lambda_r(\pi_m - \pi_{r,i}^{d})^{+}, \tag{A.2}$$

where $\lambda_r$ is the degree of retailer $i$'s fairness concerns.

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

$$\begin{aligned}
&\alpha(q_i) - \beta_i(q_i, q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i, q_j)\left(\frac{p-t}{p}\right) = \frac{p-w}{p} - \frac{\lambda_r}{1+\lambda_r}\frac{w-c}{p} && \text{if } \pi_m - \pi_{r,i}^{d} > 0,\\
&\alpha(q_i) - \beta_i(q_i, q_j)\left(\frac{t}{p}\right) + \gamma_i(q_i, q_j)\left(\frac{p-t}{p}\right) = \frac{p-w}{p} && \text{if } \pi_m - \pi_{r,i}^{d} \le 0.
\end{aligned} \tag{A.3}$$

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

$$u_r^{c,F} = \pi_r^{c} - 2\lambda(\pi_m - \pi_r^{c}/2)^{+}. \tag{A.4}$$

Note that the fairness term in utility function (A.4) considers only half of the joint retailer profit because when making decisions, retailers observe individual profit instead of joint profit. Therefore, the disutility term is at the individual level and multiplied by 2 in the joint utility function. Optimal quantities in CR follow a similar logic as DR.

Additional numerical analyses (not presented here) indicate that $\lambda_r$ may potentially lead to a lower transfer price, when the production cost is very high. However this is not applicable in our parameter setting (see Figure A.4a) since the normative theory predicates that retailers should set the transfer price at zero. Turning to the quantity, Figure A.4c shows that a higher degree of disadvantageous fairness concerns results in a lower quantity, which leads to a more fair outcome.

### A.3.2 ERC model

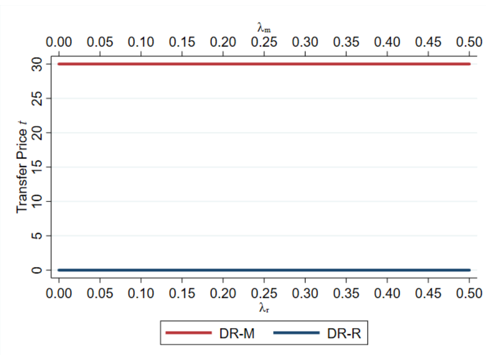
We also consider the equity, reciprocity, and competition (ERC) model in [19] where manufacturers suffer disutility when earning more than 1/3 of the supply chain profit. Here we provide some technical details and later show it fits slightly worse than fairness formulation presented above. In ERC, retailers decide quantities following the normative theory, i.e., maximizing Equation (1.4) in CR and Equation (1.7) in DR. The manufacturer's expected utility function with decentralized retailers is

$$u_m^{d,ERC} = \pi_m - 100\frac{\omega}{3}\left(\frac{\pi_m}{\pi_m + \pi_{r,i}^d + \pi_{r,j}^d} - \frac{1}{3}\right)^2. \tag{A.5}$$

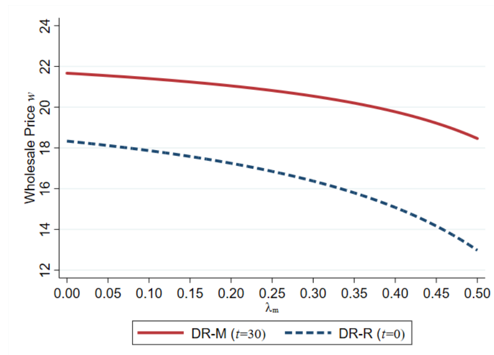
In Equation (A.5), the disutility term reflects concern when the total profit is not evenly distributed among the three parties and $\omega$ is the degree of such concern.

Figure A.4: Impact of Fairness Concerns on Transfer Price, Wholesale Price and Stocking Quantity

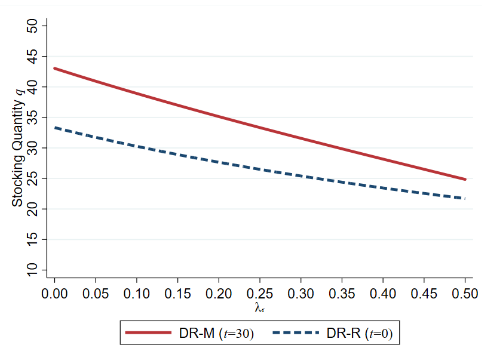
(a) Transfer Price $t$ w.r.t. $\lambda_m$ or $\lambda_r$




(b) Wholesale Price $w$ w.r.t. $\lambda_m$




(c) Stocking Quantity $q$ w.r.t. $\lambda_r$




The number 100 is a scalar given the maximum individual demand in our experiments is 100. The utility function with centralized retailers can be derived by replacing retailer expected profit function in Equation (A.5).

### A.3.3 Random Errors

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

$$\rho_m(t, \theta, \cdot) = \frac{e^{u_m(t,\cdot)/\theta}}{\sum\limits_{t \in \Omega} e^{u_m(t,\cdot)/\theta}}, \quad \rho_r(t, \theta, \cdot) = \frac{e^{u_r(t,\cdot)/\theta}}{\sum\limits_{t \in \Omega} e^{u_r(t,\cdot)/\theta}}, \tag{A.6}$$

where $\theta$ is the degree of rationality: $\theta \to 0$ means the decision maker is fully rational and $\theta \to \infty$ means they make fully random decisions, i.e., each transfer price will be picked with the same probability. Equation (A.6) can be applied to the normative theory or any behavioral model mentioned above by specifying the form of $u_m$ and $u_r$, and corresponding parameters.

### A.3.4 Estimation Methodology and Results

We now structurally estimate the parameters of our behavioral model. A heterogeneity analysis of our data shows that participants make sub-optimal decisions for wholesale prices and stocking quantities around both sides of the optimal points. Therefore, for our maximum-likelihood estimation (MLE) we use truncated normal distributions for these two decisions with conditional optimal values as the means and estimate the standard deviations. For the wholesale price the normal distribution is truncated at unit cost $c = 5$ and unit selling price $p = 30$.

For the stocking quantity the distribution is truncated at lower bound 0 and upper bound 100 of demand for decentralized retailers. For centralized retailers in CR, the upper bound is 200. The probability density function of a truncated normal distribution $\varphi(x;\mu,\sigma,a,b)$ is defined by

$$\varphi(x;\mu,\sigma,a,b) = \frac{\phi(\frac{x-\mu}{\sigma})}{\sigma\left(\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right)}, \tag{A.7}$$

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

Table A.6, A.7, Table A.8 and A.9 show MLE estimation results for Baseline, CR, DR-M, DR-R and DR-0. Random errors model is estimated for DR-M and DR-R only as they include transfer price decisions. Compared to the full-fairness model, ERC fits much worse in CR and slightly worse in three DR treatments.

## A.4 Experimental Details

### A.4.1 Details for Online Experimental Sessions

All online synchronous experimental sessions (one session of each in CR, DR-M, and DR-R, and all sessions in Baseline and DR-0) adhered to the following procedures. A participant logs in to a Zoom session and participates in the study

Table A.6: Maximum-likelihood Estimation Results in Baseline

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

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

Table A.7: Maximum-likelihood Estimation Results in CR and DR-0

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

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

Table A.8: Maximum-likelihood Estimation Results in DR-M

| | Errors | $\text{Fair}_m$ | $\text{Fair}_r$ | $\text{Fair}_{m,r}$ | ERC |
|---|---|---|---|---|---|
| $\hat{\theta}$ | 250.709 | 125.508 | 234.290 | 118.184 | 239.736 |
| | (24.990) | (21.546) | (24.492) | (13.511) | |
| $\hat{\lambda}_m$ | | 0.382 | | 0.382 | |
| | | (0.067) | | (0.032) | |
| $\hat{\lambda}_r$ | | | 0.065 | 0.065 | |
| | | | (0.019) | (0.020) | |
| $\hat{\omega}$ | | | | | 0.354 |
| $\hat{\sigma}_w$ | 3.938 | 3.222 | 3.938 | 3.222 | 3.301 |
| $\hat{\sigma}_q$ | 12.893 | 12.893 | 12.748 | 12.748 | 12.893 |
| $LL$ | -4440.10 | -4399.43 | -4433.17 | -4392.47 | -4405.73 |

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

Table A.9: Maximum-likelihood Estimation Results in DR-R

| | Errors | $\text{Fair}_m$ | $\text{Fair}_r$ | $\text{Fair}_{m,r}$ | ERC |
|---|---|---|---|---|---|
| $\hat{\theta}$ | 19.306 | 33.829 | 20.570 | 35.010 | 16.134 |
| | (1.960) | (3.503) | (2.175) | (6.261) | |
| $\hat{\lambda}_m$ | | 0.332 | | 0.332 | |
| | | (0.022) | | (0.037) | |
| $\hat{\lambda}_r$ | | | 0.017 | 0.012 | |
| | | | (0.013) | (0.013) | |
| $\hat{\omega}$ | | | | | 0.441 |
| $\hat{\sigma}_w$ | 4.156 | 3.544 | 4.156 | 3.544 | 3.507 |
| $\hat{\sigma}_q$ | 9.988 | 9.988 | 9.992 | 9.988 | 9.988 |
| $LL$ | -4123.56 | -4086.90 | -4122.87 | -4086.60 | -4090.46 |

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

in oTree through an internet browser. All participants are required to play on a desktop or laptop computer with a working camera.

1. Each participant is assigned a random 4-digit label prior to a session. This label acts as a participant's Zoom name to ensure anonymity.

2. Participants arriving at the Zoom session first wait in the waiting room. One participant is admitted at a time and verified through a photo ID.

3. Admitted and verified participants wait in a Zoom breakout room until the session begins.

4. Each participant accesses oTree via a unique link sent only to them.

5. After all participants arrive and are checked in, a researcher reads through the instructions and answers any questions in Zoom.

6. The game begins. To ensure that participants cannot see each other during this time, each participant is assigned to an individual Zoom breakout room with a researcher who will answer any questions. To be clear, this requires the experimenter to log into multiple browsers and computers, in order to be "alone" with each individual in each breakout room.

7. During the session, participants are required to be muted and have their cameras turned on.

8. After the session, participants are paid via PayPal.

### A.4.2 Detailed Results for Online Experimental Sessions

In Table A.10 we provide a summary of decisions between in-person versus virtual, along with regressions comparing the two with corrected p-values.

Table A.10: Average Contract Prices and Quantities In-person versus Online

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

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

### A.4.3 Sample Experiment Instructions in treatment DR-R

#### Instructions

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

**Game Overview**

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

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

**Timeline of the Game**

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

**a) If you are manufacturer:**

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

Stage 2: Set a wholesale price.

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

**b) If you are retailer:**

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

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

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

**Inventory Transfer Price Negotiation between Retailers**

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

**Decision Support**

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

**Profit Calculations**

The profit equations are as follows:

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

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

**Results**

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

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

**Example**

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

Decisions:

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

Outcomes:

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

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

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

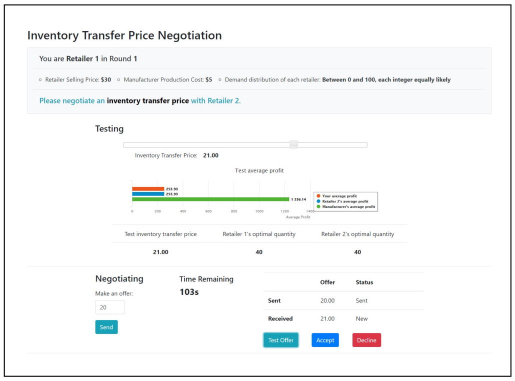
$$\text{Manufacturer's profit: } (\$15 - \$5) \times (67 + 42) = \$1090$$

**Payment**

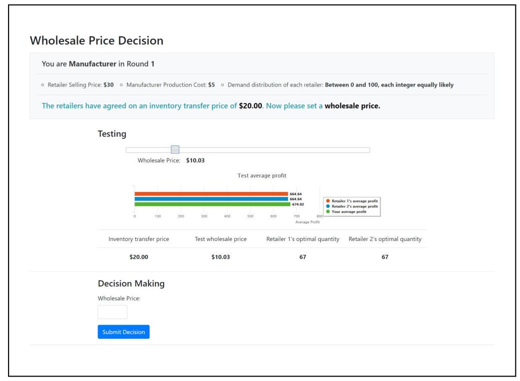
At the end of the session the actual earnings from the game will be converted to US dollars at the rate of 370 laboratory dollars for $1 US dollar. These profits

Figure A.5: Screenshots of Experiment Interface

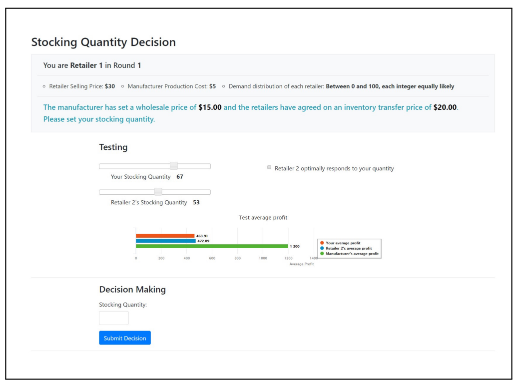
(a) Stage 1: Inventory transfer price negotiation




(b) Stage 2: Wholesale price decision




(c) Stage 3: Stocking quantity decision




will be added to your $7 show-up fee, displayed on your screen, and paid to you in cash at the end of the session.

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

APPENDIX B

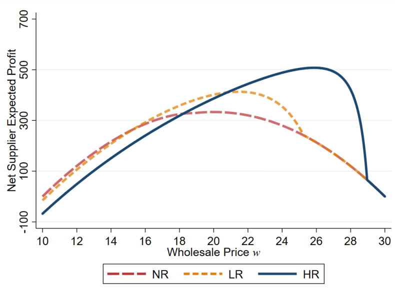
# APPENDIX OF CHAPTER 2

## B.1 Theoretical Details

### B.1.1 Additional Numerical Examples

Figure B.1: Net Expected Profit with respect to Wholesale Price

(a) Supplier




(b) Retailer

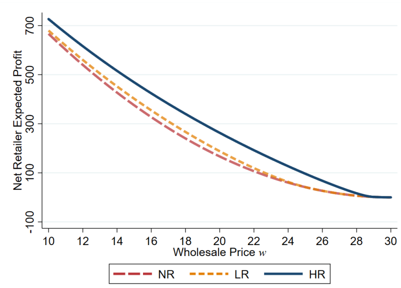
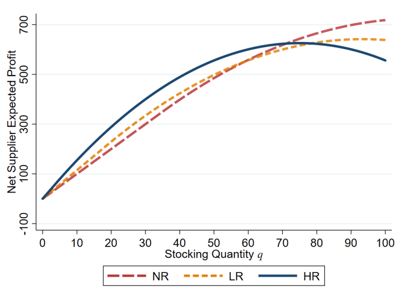
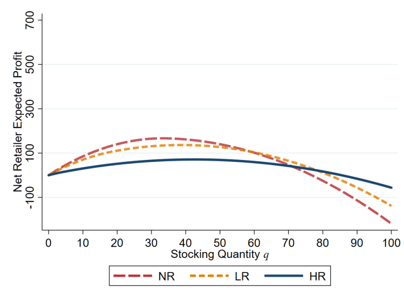



Figure B.2: Net Expected Profit with respect to Stocking Quantity

(a) Supplier




(b) Retailer



## B.1.2 Theoretical Details of the Reference-dependence Model

### B.1.2.1 Fixed Reference Point: Initial Capital

When the retailer's reference point is his initial capital $r$, the retailer is in the gains domain when demand is larger than $\frac{wq}{p}$, and is in the losses domain otherwise. Depending on whether the retailer needs financing or not, his utility function is shown as follows.

- If $wq - r > 0$:

$$\mathbb{E}[u_r^{FRP}(q)] = \mathbb{E}[\pi_r(q)] + \eta \int_{\frac{wq}{p}}^{q} [p\min(x,q) - wq] dF(x) - \lambda \int_0^{\frac{wq}{p}} [r - (px - wq + r)^+] dF(x)$$

- Otherwise:

$$\mathbb{E}[u_r^{FRP}(q)] = \mathbb{E}[\pi_r(q)] + \eta \int_{\frac{wq}{p}}^{q} (p\min(x,q) - wq) dF(x) - \lambda \int_0^{\frac{wq}{p}} (-px + wq) dF(x)$$

### B.1.2.2 Prospect Reference Point: Expected Profit

When the retailer's reference point is his expected profit given a quantity $q$, the reference point $\mathscr{X}(q)$ is dependent on $q$.

- If $wq - r > 0$:

$$\mathbb{E}[u_r^{PRP}(q)] = \mathbb{E}[\pi_r(q)] + \eta \int_{q - \frac{q^2}{2a} + \frac{(r-wq)^2}{2ap^2}}^{q} [p \min(x,q) - wq + r - \mathbb{E}[\pi_r(q)]] dF(x)$$
$$- \lambda \int_{0}^{q - \frac{q^2}{2a} + \frac{(r-wq)^2}{2ap^2}} [\mathbb{E}[\pi_r(q)] - (px - wq + r)^+] dF(x)$$

- Otherwise:

$$\mathbb{E}[u_r^{PRP}(q)] = \mathbb{E}[\pi_r(q)] + \eta \int_{q - \frac{q^2}{2a}}^{q} [p \min(x,q) - wq + r - \mathbb{E}[\pi_r(q)]] dF(x)$$
$$- \lambda \int_{0}^{q - \frac{q^2}{2a}} [\mathbb{E}[\pi_r(q)] - px + wq - r] dF(x)$$

## B.2 Experimental Details

### B.2.1 Sample Experimental Instructions in treatment HR

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 by PayPal at the end of the session. Your earnings will depend on your decisions, the decisions of other participants, and chance. Please keep your camera on and do not unmute yourself throughout the game. If you have any questions, or you accidentally close your browser, please ask the experimenter by chat.

**Game Overview**

You will be randomly assigned to one of two roles for the duration of the session: a retailer or a supplier. The retailer independently purchases units of a product from the supplier at a wholesale price per unit, and sells units to customers for $30 per unit (all $ are laboratory dollars). Customer demand is randomly and independently determined in each round, from 0 to 100, with each integer in that range equally likely. The retailer cannot make extra money from any remaining units.

The supplier starts with $400 cash and produces units at a cost of $10 per unit. The retailer starts with $100 cash to purchase units from the supplier. If the retailer has sufficient cash, it pays the supplier upfront. If the cash is insufficient, the retailer pays the supplier after earning revenue by satisfying the (random) customer demand. If the retailer cannot pay the supplier in full after satisfying its demand, the retailer transfers all its remaining cash to the supplier and earns a profit of $0.

**Timeline in Each Round**

You will play in 20 rounds. Each round has 2 stages. Decisions at each stage are specified as follows.

a) If you are the supplier:

- Stage 1: Set a wholesale price.
- Stage 2: Wait for the retailer to decide its stocking quantity.

b) If you are the retailer:

- Stage 1: Wait for the supplier to decide a wholesale price.

- Stage 2: Set a stocking quantity. If you are not satisfied with the wholesale price set by the supplier, you can reject the wholesale price offer by clicking on the “Reject The Wholesale Price” button without setting a stocking quantity. In this case, both parties keep their initial cash.

**Profit Calculations**

If the retailer can pay the supplier in full:

- Supplier Profit = (Wholesale Price − \$10) × Quantity + \$400
- Retailer Profit = \$30 × Units Sold + \$100 − Wholesale Price × Quantity

If the retailer cannot pay the supplier in full:

- Supplier Profit = \$30 × Units Sold + \$100 + \$400 − \$10 × Quantity
- Retailer Profit = \$0

“Units Sold” is the lower number of demand and quantity. That is, you cannot sell more than the quantity you ordered.

**Decision Support**

In each stage, there will be a testing section and a decision-making section, such that you can test your decision before submission. Screenshots of the two stages are shown below. Slide the scroll bar and you will see actual profits of both parties when demand is 0, 2, 4, ..., 98, 100, respectively. You will also see the likelihood that the retailer will not be able to pay the supplier in full. Note that for the supplier, the profits and the likelihood of the retailer not paying in full in this testing function are based on the assumption that the retailer chooses the stocking quantity to maximize its average profit. The true numbers will be different if the retailer makes a different choice.

**Results**

After both stages, demand will be revealed. Then you will see all of the information of that round in the results page, including your profit and your partner's profit.

This concludes one round. In total there will be 20 rounds. At the beginning of each round, you will be randomly re-matched with another participant. Note that customer demand in one round is completely independent from customer demand in any other round.

Before the game starts, you will be assigned to a breakout room. Note also that you will not be matched with any other participant in the same room.

**Example**

These numbers are simply used to illustrate the sequence of decisions and should not be construed as "good" or "bad" decisions. The supplier has initial cash \$400. The retailer has initial cash \$100.

Decisions & Realized Demand:

- The supplier sets a wholesale price of \$18 (note that decimals are permitted for the wholesale price).

- The retailer sets a stocking quantity of 60. Full payment to the supplier is $18 × 60 = $1,080.

Outcome 1: Realized demand is 68. However, since the retailer sets a stocking quantity of 60, the units sold is 60. The retailer's total cash ($30 × 60 + $100 = $1,900) is sufficient.

- Supplier Profit = ($18 − $10) × 60 + $400 = $880
- Retailer Profit = $30 × 60 + $100 − $18 × 60 = $820

Outcome 2: Realized demand is 12, which is lower than the stocking quantity set by the retailer. The units sold is 12. The retailer's total cash ($30 × 12 + $100 = $460) is insufficient.

- Supplier Profit = $30 × 12 + $100 + $400 − $10 × 60 = $260
- Retailer Profit = $0

**Payment**

At the end of the session, one round will be randomly picked to calculate your actual earnings from the game. Each round is equally likely to be picked. Therefore, it is in your best interest to maximize your profit in each round. The one-round profit will be converted to US dollars at the rate of 50.0 laboratory dollars for 1 US dollar. These profits will be added to your $7.00 show-up fee, displayed on your screen, and paid to you by PayPal after the session. Please fill out the survey at the end of the session to provide your payment information.

### B.2.2 Details for Online Experimental Sessions

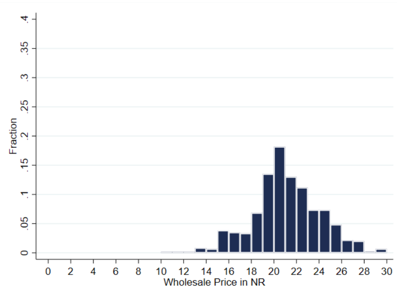
All online synchronous experimental sessions adhered to the following procedures. A participant logs in to a Zoom session and participates in the study in oTree through an internet browser. All participants are required to play on a desktop or laptop computer with a working camera.

1. Each participant is assigned a random 4-digit label prior to a session. This label acts as a participant's Zoom name to ensure anonymity.
2. Participants arriving at the Zoom session first wait in the waiting room. One participant is admitted at a time and verified through a photo ID.
3. Admitted and verified participants wait in a Zoom breakout room until the session begins.
4. Each participant accesses oTree via a unique link sent only to them.
5. After all participants arrive and are checked in, a researcher reads through the instructions and answers any questions in Zoom.
6. The game begins. To ensure that participants cannot communicate with their paired partners, suppliers and retailers are assigned to two separate Zoom breakout rooms.
7. During the session, participants are required to be muted and have their cameras turned on.
8. After the session, participants are paid via PayPal or Venmo.

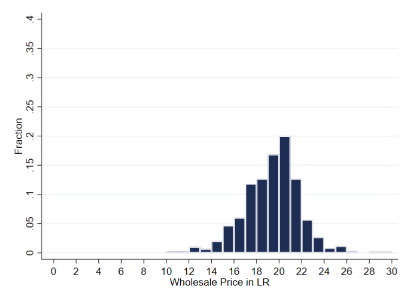
### B.2.3 Distribution of Decisions

Figure B.3: Distribution of Observed Wholesale Prices

(a) NR




(b) LR




(c) HR

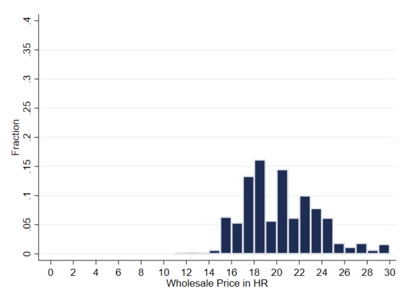
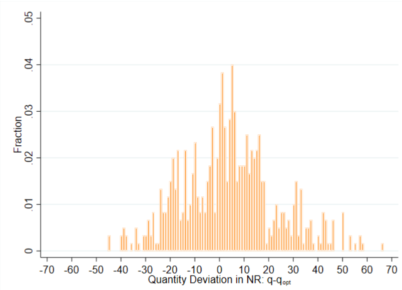
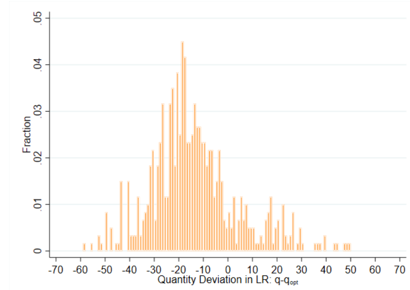


Figure B.4: Distribution of Observed Quantity Deviations

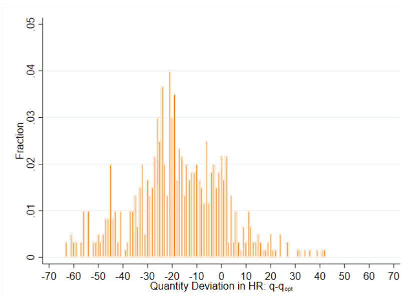
(a) NR




(b) LR




(c) HR



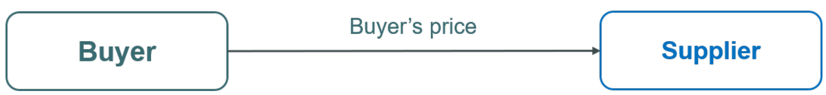
APPENDIX C

# APPENDIX OF CHAPTER 3

## C.1 Sample Experimental 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 after 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 two-player game consisting of a buyer and a supplier. You will be randomly assigned a role, which will remain fixed, for the duration of the game. Specifically, your assigned role for this game will be: Buyer. The buyer purchases one unit of a product from the supplier, and automatically sells it for 1000(*all* are laboratory dollars). The supplier is subject to a cost shock of $400, and can exert effort to reduce the possibility of cost shock to 0.5%. There are two types of effort, depending on whether the effort is considered "compliant" or "non-compliant". Based on what type of effort is chosen, there will be the following outcomes:

a) If compliant effort is chosen:

- The supplier may incur a cost shock of $400.

b) If non-compliant effort is chosen:

- The supplier may incur a cost shock of $400.
- The buyer incurs an additional loss of $200 (referred to as "buyer loss").

The cost of each effort type, cost-shock possibility and buyer losses are shown in the following table:

| | Compliant Effort (C) | Non-compliant Effort (N) |
|---|---|---|
| Effort cost | $110 | $100 |
| Possibility of cost shock | 50% | 50% |
| Buyer loss | $0 | $200 |

The buyer can decide to share a certain percentage of the supplier's cost shock of $400, if it occurs. See below for more explanation and example.

**Timeline in Each Round**

You will play in 15 rounds. Each round has 2 stages. Decisions at each stage are specified as follows.

a) If you are the buyer:

- Stage 1: Set a price to the supplier and a percentage of cost shock (from 0% to 100%) that you will share with the supplier.
- Stage 2: Wait for the supplier to make decisions.

b) If you are the supplier:

- Stage 1: Wait for the buyer to decide a wholesale price.
- Stage 2: Choose an effort type (compliant or non-compliant). If you are not satisfied with the contract terms set by the buyer, you can click on the yellow reject button without making decisions. In this case, both parties earn a profit of $0.

**Profit Calculations**

Buyer profit =

- No cost shock: $1000 – Price to the supplier – Buyer loss
- Cost shock realized: $1000 – Price to the supplier – $400 × Buyer's percentage – Buyer loss

"Buyer loss" will be $0 if compliant effort is chosen.

Supplier Profit =

- No cost shock: Price to the supplier - Effort cost
- Cost shock realized: Price to the supplier – Effort cost – $400 × (100% – Buyer's percentage)

**Example**

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

Decisions:

- The buyer sets a price of \$400 to the supplier and shares 60% (\$240) of the supplier's cost shock.
- The supplier exerts a non-compliant effort of \$100. The possibility of cost shock is 50% and the buyer loss is \$200.

Cost shock \$400 (50% chance)

| Buyer | Supplier |
|---|---|
| $\$400 \times 60\% = \$240$ | $\$400 - \$240 = \$160$ |

Buyer profit =

- No cost shock (50% chance): \$1000 – \$400 – \$200 = \$400.
- Cost shock realized (50% chance): $\$1000 - \$400 - \$400 \times 60\% - \$200 = \$160$.

Supplier Profit =

- No cost shock (50% chance): \$400 – \$100 = \$300.
- Cost shock realized (50% chance): $\$400 - \$100 - \$400 \times (100\% - 60\%) = \$140$.

**Decision Support & Practice Round**

There will be a testing section and a decision-making section, such that you can test your decision before submission. Before the game starts, the experimenter will play a video to further explain the testing section for each role.

Before the game starts, you will play a practice round with a randomly matched participant so that you can familiarize yourself with the game. Your outcome from this round will not count towards your earnings.

**Results**

After all decisions are made, occurrence of the cost shock (and any buyer loss) will be realized. Then you will see all information of that round in the results page. This concludes one round. In total there will be 15 rounds. At the beginning of each round, you will be randomly re-matched with another participant.

**Payment**

At the end of the session the actual earnings from the game will be converted to US dollars at the rate of 190 laboratory dollars for 1 US dollar. These profits will be added to your $10 show-up fee, displayed on your screen, and paid to you after the session by Venmo, PayPal or Amazon gift card. You will see a link to payment survey where you can provide your payment information.

# BIBLIOGRAPHY

[1] Marina Agranov and Chloe Tergiman. Communication in multilateral bargaining. *Journal of Public Economics*, 118:75–85, 2014.

[2] Vishal V Agrawal, Atalay Atasu, and Luk N Van Wassenhove. Om forum—new opportunities for operations management research in sustainability. *Manufacturing & Service Operations Management*, 21(1):1–12, 2019.

[3] James Andreoni. Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, 110(1):1–21, 1995.

[4] James Andreoni, William T. Harbaugh, and Lise Vesterlund. Altruism in experiments. In Steven Durlauf and L.E. Blume, editors, *Behavioural and Experimental Economics*, pages 6–13. Palgrave Macmillan, London, 2010.

[5] James Andreoni and Justin M Rao. The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of public economics*, 95(7-8):513–520, 2011.

[6] Ravi Anupindi and Yehuda Bassok. Centralization of stocks: Retailers vs. manufacturer. *Management Science*, 45(2):178–191, 1999.

[7] Karca D Aral, Erasmo Giambona, and Ye Wang. Buyer's bankruptcy risk, sourcing strategy, and firm value: Evidence from the supplier protection act. *Management Science*, Articles in Advance, https://doi.org/10.1287/mnsc.2021.4270, 2021.

[8] Vivek Astvansh and Niket Jindal. Differential effects of received trade credit and provided trade credit on firm value. *Production and Operations Management*, 31(2):781–798, 2022.

[9] Barış Ata, Deishin Lee, and Mustafa H Tongarlak. Optimizing organic waste to energy operations. *Manufacturing & Service Operations Management*, 14(2):231–244, 2012.

[10] Atalay Atasu, Charles J Corbett, Ximin Huang, and L Beril Toktay. Sustainable operations management through the perspective of manufacturing & service operations management. *Manufacturing & service operations management*, 22(1):146–157, 2020.

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

[12] Opher Baron, Ming Hu, Sami Najafi-Asadolahi, and Qu Qian. Newsvendor selling to loss-averse consumers with stochastic reference points. *Manufacturing & Service Operations Management*, 17(4):456–469, 2015.

[13] Abigail Barr and Danila Serra. The effects of externalities and framing on bribery in a petty corruption experiment. *Experimental Economics*, 12(4):488–503, 2009.

[14] Michael Becker-Peth, Elena Katok, and Ulrich W. Thonemann. Designing

buyback contracts for irrational but predictable newsvendors. *Management Science*, 59(8):1800–1816, 2013.

[15] Michael Becker-Peth, Elena Katok, and Ulrich W Thonemann. Designing buyback contracts for irrational but predictable newsvendors. *Management Science*, 59(8):1800–1816, 2013.

[16] Ruth Beer, Ignacio Rios, and Daniela Saban. Increased transparency in procurement: The role of peer effects. *Management Science*, 67(12):7511–7534, 2021.

[17] Andreas Blume and Andreas Ortmann. The effects of costless pre-play communication: Experimental evidence from games with pareto-ranked equilibria. *Journal of Economic theory*, 132(1):274–290, 2007.

[18] Gary E. Bolton and Yefen Chen. Other-regarding behavior: Fairness, reciprocity, and trust. In Karen Donohue, Elena Katok, and Stephen Leider, editors, *The Handbook of Behavioral Operations*, chapter 6, pages 199–236. Wiley, 2019.

[19] Gary E. Bolton and Axel Ockenfels. ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90(1):166–193, 2000.

[20] Gary E. Bolton, Axel Ockenfels, and Ulrich W. Thonemann. Managers and students as newsvendors. *Management Science*, 58(12):2225–2233, 2012.

[21] A J A Bostian, Charles A Holt, Sanjay Jain, and Kamalini Ramdas. Is trans-

shipment a behaviorally-robust risk-pooling strategy?, 2012. Working paper, London Business School.

[22] Andre P Calmon and Stephen C Graves. Inventory management in a consumer electronics closed-loop supply chain. *Manufacturing & Service Operations Management*, 19(4):568–585, 2017.

[23] Colin F Camerer. *Behavioral game theory: Experiments in strategic interaction*. Princeton university press, 2011.

[24] Colin F Camerer and Teck-Hua Ho. Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, 8(2):167–196, 1994.

[25] Felipe Caro, Prashant Chintapalli, Kumar Rajaram, and Chris S Tang. Improving supplier compliance through joint and shared audits with collective penalty. *Manufacturing & Service Operations Management*, 20(2):363–380, 2018.

[26] Timothy N Cason, Roman M Sheremeta, and Jingjing Zhang. Communication and efficiency in competitive coordination games. *Games and Economic Behavior*, 76(1):26–43, 2012.

[27] Muruvvet Celikbas, J. George Shanthikumar, and Jayashankar M. Swaminathan. Coordinating production quantities and demand forecasts through penalty schemes. *IIE Transactions*, 31(9):851–864, 1999.

[28] Daniel L Chen, Martin Schonger, and Chris Wickens. otree – an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016.

[29] Jiayu Chen, Anyan Qi, and Milind Dawande. Supplier centrality and auditing priority in socially responsible supply chains. *Manufacturing & Service Operations Management*, 22(6):1199–1214, 2020.

[30] Kay-Yut Chen and Diana Wu. Buyer-supplier interactions. In Karen Donohue, Elena Katok, and Stephen Leider, editors, *The Handbook of Behavioral Operations*, chapter 13, pages 459–487. Wiley, 2019.

[31] Li Chen and Hau L Lee. Sourcing under supplier responsibility risk: The effects of certification, audit, and contingency payment. *Management Science*, 63(9):2795–2812, 2017.

[32] Li Chen, Shiqing Yao, and Kaijie Zhu. Responsible sourcing under supplier-auditor collusion. *Manufacturing & Service Operations Management*, 22(6):1234–1250, 2020.

[33] Jiri Chod. Inventory, risk shifting, and trade credit. *Management Science*, 63(10):3207–3225, 2016.

[34] Jiri Chod, Evgeny Lyandres, and S Alex Yang. Trade credit and supplier competition. *Journal of Financial Economics*, 131(2):484–505, 2019.

[35] Charles J Corbett. Global diffusion of iso 9000 certification through supply chains. *Manufacturing & Service Operations Management*, 8(4):330–350, 2006.

[36] Tony Haitao Cui, Jagmohan S Raju, and Z John Zhang. Fairness and channel coordination. *Management Science*, 53(8):1303–1314, 2007.

[37] Vicente Cunat. Trade credit: suppliers as debt collectors and insurance providers. *The Review of Financial Studies*, 20(2):491–527, 2007.

[38] Maqbool Dada and Qiaohai Hu. Financing newsvendor inventory. *Operations Research Letters*, 36(5):569–573, 2008.

[39] Andrew M. Davis. Biases in individual decision-making. In Karen Donohue, Elena Katok, and Stephen Leider, editors, *The Handbook of Behavioral Operations*, chapter 5, pages 149–198. Wiley, 2019.

[40] Andrew M Davis, Elena Katok, and Natalia Santamaría. Push, pull, or both? a behavioral study of how the allocation of inventory risk affects channel efficiency. *Management Science*, 60(11):2666–2683, 2014.

[41] Andrew M Davis and Stephen Leider. Contracts and capacity investment in supply chains. *Manufacturing & Service Operations Management*, 20(3):403–421, 2018.

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

[43] Vibhuti Dhingra and Harish Krishnan. Managing reputation risk in supply chains: The role of risk sharing under limited liability. *Management Science*, 67(8):4845–4862, 2021.

[44] Lingxiu Dong and Nils Rudi. Who benefits from transshipment? exogenous vs. endogenous wholesale prices. *Management Science*, 50(5):645–657, 2004.

[45] Karen Donohue, Elena Katok, and Stephen Leider, editors. *The Handbook of Behavioral Operations*. Wiley, 2019.

[46] John Duffy and Nick Feltovich. Do actions speak louder than words? an experimental comparison of observation and cheap talk. *Games and Economic Behavior*, 39(1):1–27, 2002.

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

[48] Xin Fang and Soo-Haeng Cho. Cooperative approaches to managing social responsibility in a market with externalities. *Manufacturing & Service Operations Management*, 22(6):1215–1233, 2020.

[49] Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114(3):817–868, 1999.

[50] Robert Forsythe, Joel L. Horowitz, N.E. Savin, and Martin Sefton. Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3):347–369, 1994.

[51] Norman Frohlich and Joe Oppenheimer. Some consequences of e-mail vs. face-to-face communication in experiment. *Journal of Economic Behavior & Organization*, 35(3):389–403, 1998.

[52] Mariassunta Giannetti, Mike Burkart, and Tore Ellingsen. What you sell is what you lend? explaining trade credit contracts. *The Review of Financial Studies*, 24(4):1261–1298, 2011.

[53] Jarrod Goentzel and James B. Rice. Guest voices: Managing supply chains is intertwined with financial management. https://www.wsj.com/articles/guest-voices-managing-supply-chains-is-intertwined-with-financial-management-1437662343, 2015. Accessed Auguest 10, 2019.

[54] Richard Gonzalez and George Wu. On the shape of the probability weighting function. *Cognitive psychology*, 38(1):129–166, 1999.

[55] Ruixue Guo, Hau L Lee, and Robert Swinney. Responsible sourcing in supply chains. *Management Science*, 62(9):2722–2744, 2016.

[56] Diwakar Gupta and Lei Wang. A stochastic inventory model with trade credit. *Manufacturing & Service Operations Management*, 11(1):4–18, 2009.

[57] Haresh Gurnani, Karthik Ramachandran, Saibal Ray, and Yusen Xia. Ordering behavior under supply risk: An experimental investigation. *Manufacturing & Service Operations Management*, 16(1):61–75, 2014.

[58] Werner Güth, Rolf Schmittberger, and Bernd Schwarze. An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, 3(4):367–388, 1982.

[59] Charles W Haley and Robert C Higgins. Inventory policy and trade credit financing. *Management science*, 20(4-part-i):464–471, 1973.

[60] Julia Hartmann and Sabine Moeller. Chain liability in multitier supply chains? responsibility attributions for unsustainable supplier behavior. *Journal of operations management*, 32(5):281–294, 2014.

[61] Teck-Hua Ho, Noah Lim, and Tony Haitao Cui. Reference dependence in multilocation newsvendor models: A structural analysis. *Management Science*, 56(11):1891–1910, 2010.

[62] Elizabeth Hoffman and Matthew L Spitzer. Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *The Journal of Legal Studies*, 14(2):259–297, 1985.

[63] Lu Huang, Jing-Sheng Song, and Robert Swinney. Managing social responsibility in multitier supply chains. *Manufacturing & Service Operations Management*, Articles in Advance, https://doi.org/10.1287/msom.2021.1063, 2022.

[64] Kyle Hyndman and Matthew Embrey. Econometrics for experiments. In Karen Donohue, Elena Katok, and Stephen Leider, editors, *The Handbook of Behavioral Operations*, chapter 2, pages 35–88. Wiley, 2019.

[65] Greg Johnsen. Manufacturers should start taking responsibility for suppliers. https://www.vox.com/2015/4/7/11561198/its-time-for-

manufacturers-to-start-taking-responsibility-for, 2015. Accessed April 10, 2022.

[66] John H. Kagel and Alvin E. Roth, editors. *The Handbook of Experimental Economics, Vol 2*. Princeton University Press, Prince, NJ, 2017.

[67] Daniel Kahneman, Jack L Knetsch, and Richard H Thaler. Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, 5(1):193–206, 1991.

[68] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific, 2013.

[69] Basak Kalkanci, Kay-Yut Chen, and Feryal Erhun. Contract complexity and performance under asymmetric demand information: An experimental evaluation. *Management Science*, 57(4):689–704, 2011.

[70] Basak Kalkanci, Kay-Yut Chen, and Feryal Erhun. Complexity as a contract design factor: A human-to-human experimental study. *Production and Operations Management*, 23(2):269–284, 2014.

[71] Basak Kalkancı, Kay-Yut Chen, and Feryal Erhun. Complexity as a contract design factor: A human-to-human experimental study. *Production and Operations Management*, 23(2):269–284, 2014.

[72] Basak Kalkanci and Erica L Plambeck. Managing supplier social and en-

vironmental impacts with voluntary versus mandatory disclosure to investors. *Management Science*, 66(8), 2020.

[73] Elena Katok, Doug Thomas, and Andrew Davis. Service-level agreements as coordination mechanisms: The effect of review periods. *Manufacturing & Service Operations Management*, 10(4):609–624, 2008.

[74] Elena Katok and Sebastián Villa. Centralized or decentralized transfer prices: A behavioral approach for improving supply chain coordination. *Manufacturing & Service Operations Management*, 24(1):143–158, 2022.

[75] Panos Kouvelis and Wenhui Zhao. Financing the newsvendor: supplier vs. bank, and the structure of optimal trade credit contracts. *Operations Research*, 60(3):566–580, 2012.

[76] Panos Kouvelis and Wenhui Zhao. Who should finance the supply chain? impact of credit ratings on supply chain decisions. *Manufacturing & Service Operations Management*, 20(1):19–35, 2018.

[77] Tim Kraft, León Valdés, and Yanchong Zheng. Supply chain visibility and social responsibility: Investigating consumers' behaviors and motives. *Manufacturing & Service Operations Management*, 20(4):617–636, 2018.

[78] Tim Kraft, León Valdés, and Yanchong Zheng. Motivating supplier social responsibility under incomplete visibility. *Manufacturing & Service Operations Management*, 22(6):1268–1286, 2020.

[79] Tim Kraft, León Valdés, and Yanchong Zheng. Consumer trust in social responsibility communications: The role of supply chain visibility. *Production and Operations Management*, Articles in Advance, https://doi.org/10.1111/poms.13808, 2022.

[80] Mirko Kremer, Enno Siemsen, and Douglas J. Thomas. The sum and its parts: Judgmental hierarchical forecasting. *Management Science*, 62(9):2745–2764, 2016.

[81] Guoming Lai, Laurens G Debo, and Katia Sycara. Sharing inventory risk in supply chain: The implication of financial constraint. *Omega*, 37(4):811–825, 2009.

[82] Chang Hwan Lee and Byong-Duk Rhee. Trade credit for supply chain coordination. *European Journal of Operational Research*, 214(1):136–146, 2011.

[83] Hau L Lee and Christopher S Tang. Socially and environmentally responsible value chain innovations: New operations management research opportunities. *Management Science*, 64(3):983–996, 2018.

[84] Hsiao-Hui Lee, Jianer Zhou, and Jingqi Wang. Trade credit financing under competition and its impact on firm performance in supply chains. *Manufacturing & Service Operations Management*, 20(1):36–52, 2018.

[85] Stephen Leider. Behavioral analysis of strategic interactions: Game theory, bargaining, and agency. In Karen Donohue, Elena Katok, and Stephen Lei-

der, editors, *The Handbook of Behavioral Operations*, chapter 7, pages 237–286. Wiley, 2019.

[86] Shan Li and Kay-Yut Chen. The commitment conundrum of inventory sharing. *Production and Operations Management*, 29(2):353–370, 2020.

[87] John A. List, Azeem M. Shaikh, and Yang Xu. Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22:773–793, 2019.

[88] Chris KY Lo, Christopher S Tang, Yi Zhou, Andy CL Yeung, and Di Fan. Environmental incidents and the market value of firms: An empirical investigation in the chinese context. *Manufacturing & Service Operations Management*, 20(3):422–439, 2018.

[89] Christoph H Loch and Yaozhong Wu. Social preferences and supply chain performance: An experimental study. *Management Science*, 54(11):1835–1849, 2008.

[90] Xiaoyang Long and Javad Nasiry. Sustainability in the fast fashion industry. *Manufacturing & Service Operations Management*, 24(3):1276–1293, 2022.

[91] Wei Luo and Kevin H Shang. Managing inventory for firms with trade credit and deficit penalty. *Operations Research*, 67(2):468–478, 2019.

[92] Mahdi Mahmoudzadeh and Enno Siemsen. Reliably responsible: Structuring supply chains for socially responsible behavior. *Georgia Tech Scheller College of Business Research Paper*, (18-38), 2019.

[93] Vidya Mani and Suresh Muthulingam. Does learning from inspections affect environmental performance? evidence from unconventional well development in pennsylvania. *Manufacturing & Service Operations Management*, 21(1):177–197, 2019.

[94] Dan Mitchell. Tyson stops claiming chicken is free of antibiotics. https://www.cbsnews.com/news/tyson-stops-claiming-chicken-is-free-of-antibiotics/, 2008. Accessed April 10, 2022.

[95] Justin Murfin and Ken Njoroge. The implicit costs of trade credit borrowing by large firms. *The Review of Financial Studies*, 28(1):112–145, 2015.

[96] Robert Murphy. How supply chain finance is offering companies a new cash source. https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2018/04/09/how-supply-chain-finance-is-offering-companies-a-new-cash-source/, 2018. Accessed Auguest 6, 2019.

[97] James A Narus and James C Anderson. Rethinking distribution: adaptive channels. *Harvard Business Review*, 74:112–120, 1996.

[98] Jie Ning. Strategic trade credit in a supply chain with buyer competition. *Manufacturing & Service Operations Management*, Articles in Advance, https://doi.org/10.1287/msom.2022.1103, 2022.

[99] Adem Orsdemir, Bin Hu, and Vinayak Deshpande. Ensuring corporate social and environmental responsibility through vertical integration and hori-

zontal sourcing. *Manufacturing & Service Operations Management*, 21(2):417–434, 2019.

[100] Özalp Özer. Replenishment strategies for distribution systems under advance demand information. *Management Science*, 49(3):255–272, 2003.

[101] Ozalp Ozer, Yanchong Zheng, and Kay-Yut Chen. Trust in forecast information sharing. *Management Science*, 57(6):1111–1137, 2011.

[102] V. Paddy Padmanabhan, Nils Rudi, and Ilia Tsetlin. Who prices? manufacturer's versus retailers' control of retail prices, 2010. Working paper, INSEAD.

[103] Seung Jae Park, Guoming Lai, and Sridhar Seshadri. Inventory sharing in the presence of commodity markets. *Production and Operations Management*, 25(7):1245–1260, 2016.

[104] Heikki Peura, S Alex Yang, and Guoming Lai. Trade credit in competition: A horizontal benefit. *Manufacturing & Service Operations Management*, 19(2):263–289, 2017.

[105] Angharad H Porteous, Sonali V Rammohan, and Hau L Lee. Carrots or sticks? improving social and environmental compliance at suppliers through incentives and penalties. *Production and Operations management*, 24(9):1402–1413, 2015.

[106] Saibal Ray, Tamer Boyaci, and Necati Aras. Optimal prices and trade-in

rebates for durable, remanufacturable products. *Manufacturing & Service Operations Management*, 7(3):208–228, 2005.

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

[108] Alvin E. Roth. Bargaining experiments. In John Kagel and Alvin E. Roth, editors, *Handbook of Experimental Economics*, pages 253–348. Princeton University Press, Princeton, NJ, 1995.

[109] Alvin E. Roth, J.Keith Murnighan, and Francoise Schoumaker. The deadline effect in bargaining: Some experimental evidence. *American Economic Review*, 78(4):806–823, 1988.

[110] Nils Rudi, Sandeep Kapur, and David F Pyke. A two-location inventory model with transshipment and local decision making. *Management Science*, 47(12):1668–1680, 2001.

[111] Maurice E. Schweitzer and Gérard P. Cachon. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, 46(3):404–420, 2000.

[112] Maurice E Schweitzer and Gérard P Cachon. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, 46(3):404–420, 2000.

[113] Daniel Seifert, Ralf W Seifert, and Margarita Protopappa-Sieke. A review of trade credit literature: Opportunities for research in operations. *European Journal of Operational Research*, 231(2):245–256, 2013.

[114] Jing Shao, Harish Krishnan, and S Thomas McCormick. Incentives for transshipment in a supply chain with decentralized retailers. *Manufacturing & Service Operations Management*, 13(3):361–372, 2011.

[115] Norihiko Shirouzu and Ran Li. Gm looks for alternative supplies in china after deadly factory blast. https://www.reuters.com/article/us-china-factory-blast/gm-looks-for-alternative-supplies-in-china-after-deadly-factory-blast-idUSKBN0G307420140803, 2014. Accessed April 10, 2022.

[116] Xuanming Su. Bounded rationality in newsvendor models. *Manufacturing & Service Operations Management*, 10(4):566–589, 2008.

[117] Ravi Subramanian and Ramanath Subramanyam. Key factors in the market for remanufactured products. *Manufacturing & Service Operations Management*, 14(2):315–326, 2012.

[118] Christopher S Tang, S Alex Yang, and Jing Wu. Sourcing from suppliers with financial constraints and performance risk. *Manufacturing & Service Operations Management*, 20(1):70–84, 2018.

[119] Sammi Tang and Jing-Sheng Jeannette Song. The effect of partial delegation and information proximity on multi-tier responsible sourcing. *University of Miami Business School Research Paper*, (3396843), 2020.

[120] Necati Tereyağoğlu, Peter S Fader, and Senthil Veeraraghavan. Multiattribute loss aversion and reference dependence: Evidence from the performing arts industry. *Management Science*, 64(1):421–436, 2018.

[121] Tunay I Tunca and Weiming Zhu. Buyer intermediation in supplier finance. *Management Science*, 64(12):5631–5650, 2018.

[122] Bhavani Shanker Uppari and Sameer Hasija. Modeling newsvendor behavior: A prospect theory approach. *Manufacturing & Service Operations Management*, 21(3):481–500, 2019.

[123] Karel H Van Donselaar, Vishal Gaur, Tom Van Woensel, Rob ACM Broekmeulen, and Jan C Fransoo. Ordering behavior in retail stores and implications for automated replenishment. *Management Science*, 56(5):766–784, 2010.

[124] Sebastián Villa and Jaime Andrés Castañeda. Transshipments in supply chains: A behavioral investigation. *European Journal of Operational Research*, 269(2):715–729, 2018.

[125] Verónica H Villena. The missing link? the strategic role of procurement in building sustainable supply networks. *Production and Operations Management*, 28(5):1149–1172, 2019.

[126] Verónica H Villena and Dennis A Gioia. On the riskiness of lower-tier suppliers: Managing sustainability in supply networks. *Journal of Operations Management*, 64:65–87, 2018.

[127] Verónica H Villena, Miriam Wilhelm, and Cheng-Yong Xiao. Untangling drivers for supplier environmental and social responsibility: An investigation in philips lighting's chinese supply chain. *Journal of Operations Management*, 67(4):476–510, 2021.

[128] Peter Whoriskey and Rachel Siegel. Cocoa's child laborers. https://www.washingtonpost.com/graphics/2019/business/hershey-nestle-mars-chocolate-child-labor-west-africa, 2019. Accessed April 10, 2022.

[129] Miriam M Wilhelm, Constantin Blome, Vikram Bhakoo, and Antony Paulraj. Sustainability in multi-tier supply chains: Understanding the double agency role of the first-tier supplier. *Journal of Operations Management*, 41:42–60, 2016.

[130] WTO and IFC. Trade finance and the compliance challenge: A showcase of international cooperation, 2019. Accessed December 25, 2020.

[131] Chaolin Yang, Zhenyu Hu, and Sean X Zhou. Multilocation newsvendor problem: Centralization and inventory pooling. *Management Science*, Articles in Advance, https://doi.org/10.1287/mnsc.2019.3532, 2020.

[132] S Alex Yang, Nitin Bakshi, and Christopher J Chen. Trade credit insurance: Operational value and contract choice. *Management Science*, 67(2):875–891, 2021.

[133] S Alex Yang and John R Birge. Trade credit, risk sharing, and inventory financing portfolios. *Management Science*, 64(8):3667–3689, 2018.

[134] Rose Yu. China dust blast explosion death toll nearly doubles. https://www.wsj.com/articles/china-dust-blast-explosion-death-toll-nearly-doubles-1419957615, 2014. Accessed January 22, 2022.

[135] Vincent Junhao Yu, Karen L. Donohue, and Karthik Natarajan. Sourcing under supply disruption and responsibility violation risks: A behavioral investigation. *Working paper*, 2020.

[136] Han Zhang, Goker Aydin, and Rodney P Parker. Social responsibility auditing in supply chain networks. *Management Science*, 68(2):1058–1077, 2022.

[137] Jun Zhang. Transshipment and its impact on supply chain members' performance. *Management Science*, 51(10):1534–1539, 2005.

[138] Yinghao Zhang, Karen Donohue, and Tony Haitao Cui. Contract preferences and performance for the loss averse supplier: Buyback versus revenue sharing. *Management Science*, 62(6):1734–1754, 2016.

[139] Yinghao Zhang, Karen Donohue, and Tony Haitao Cui. Contract preferences and performance for the loss-averse supplier: Buyback vs. revenue sharing. *Management Science*, 62(6):1734–1754, 2016.

[140] Hui Zhao, Vinayak Deshpande, and Jennifer K Ryan. Inventory sharing and rationing in decentralized dealer networks. *Management Science*, 51(4):531–547, 2005.

[141] Hui Zhao, Liang Xu, and Enno Siemsen. Inventory sharing and demand-side underweighting. *Manufacturing & Service Operations Management*, 23(5):1005–1331, 2021.