INCENTIVE MECHANISMS FOR INVENTORY CONTROL AND PROMOTION PLANNING IN VALUE CHAIN MANAGEMENT

by

Bharadwaj Kadiyala

APPROVED BY SUPERVISORY COMMITTEE:

Özalp Özer, Chair

Alain Bensoussan

Dorothée Honhon

A. Serdar Şimşek

In the loving memory of my grandfather, Viswanadha Sastry Kadiyala (1936–2015).

INCENTIVE MECHANISMS FOR INVENTORY CONTROL AND PROMOTION PLANNING IN VALUE CHAIN MANAGEMENT

by

BHARADWAJ KADIYALA, BTech, MS

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN

MANAGEMENT SCIENCE

THE UNIVERSITY OF TEXAS AT DALLAS

December 2017

ACKNOWLEDGMENTS

I am deeply grateful to my advisor, Özalp Özer, whose support and direction made this dissertation possible. His guidance, feedback, and approach to research have also shaped my research philosophy. Among many others things, I am thankful to him for providing exciting research opportunities, for encouraging me to take up challenging projects, but most of all, for always believing that it can be done. That is a lesson I always hope to carry with me.

I could not have asked for a more supportive dissertation committee. I am grateful to Alain Bensoussan for his guidance on our joint work. His relentless pursuit of knowledge and excellence has always inspired me to work harder. I am very thankful to Dorothée Honhon, for introducing me to exciting research topics through her course, and later, for going out of her way to help me during the PhD program. I would like to thank Serdar Şimşek for making sure I was always comfortable to bounce off any ideas with him. I gratefully acknowledge my previous mentor, Savaş Dayanık at Bilkent University, who continues to inspire me. I owe a great deal to R. B. Lenin, then a faculty at DA-IICT, for teaching the undergraduate class on probability models with such great enthusiasm.

I also want to acknowledge the extensive support and guidance I received from the other faculty in the operations group at UT Dallas. I would like to specially thank Ganesh Janakiraman for helping me navigate PhD program in the initial years. I want to thank Milind Dawande for offering many helpful suggestions at critical points during the program. I also want to thank Anyan Qi, for his open-door policy, from which I benefited on multiple occasions. He always offered to help like a friend.

As I reflect back, I cherish all the wonderful friendships I made during my five years at UT Dallas. I want to thank Manmohan Aseri, Yang Bo, Wei Chen, Blair Flicker, Shivam Gupta, Samuel Jayrajan, Joohyun Kim, Ismail Kirci, Harsha Manchiraju, Cheng Nie, Sina Shokoohyar, and Yulia Vorotyntseva for their support and kindness. A special note of thanks to, Harish Guda and Xiao Zhang, for *always* being there for me. I also acknowledge the excellent administrative support provided by Amanda Besch and Karen Decker.

I cannot possibly thank enough my parents, Somayajulu and Nagamani Kadiyala, and my sister, Lakshmi Chaturvedula, for their unconditional love and support. Last, but not the least, I want to acknowledge my dear wife, Vinila Sista, who has been the pillar of support right through my PhD program. Being with her has made the journey all the more fun.

September, 2017

INCENTIVE MECHANISMS FOR INVENTORY CONTROL AND PROMOTION PLANNING IN VALUE CHAIN MANAGEMENT

Bharadwaj Kadiyala, PhD
The University of Texas at Dallas, 2017

Supervising Professor: Özalp Özer, Chair

Today's value chains consist of multiple economic agents interacting at different levels so that a product or service is delivered to its end customers in an efficient manner. This dissertation explores the role of incentives in better coordinating such complex value chain operations on the supply and demand sides of the market. In particular, we study information sharing and promotion mechanisms, which firms can use to facilitate demand information sharing to better manage inventory upstream and boost customer demand by providing temporary price incentives in a sales promotion downstream, respectively. We adopt a data-driven approach in that we propose statistical frameworks that incorporate real-time data, and which can be used by decision makers to design and measure the impact of offering such incentives.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

In this dissertation we study the role of economic incentives in a value chain and how firms can offer such incentives at various stages in the value chain to better match supply and demand. One can conceptualize a value chain as consisting of multiple interfaces, at which information, products, or money are exchanged between trading parties. Broadly, these interfaces can be classified as upstream and downstream. Upstream interface refers to the operations that go into manufacturing product or offering a service. Downstream interface is where consumption of the end product or service occurs. We consider two novel incentive mechanisms in this study, one which facilitates demand information sharing to better manage inventory upstream and one which offers price-related incentives (in a sales promotion) to boost customer demand downstream.

The rapid economic growth in the past few decades, fueled by globalization and easy access to information through the internet, has brought forward a plethora of innovations in the way firms are managing their operations. In the context of information sharing and inventory management, for example, vendor-managed inventory (VMI) agreement pioneered by Wal-Mart and P&G in the late 1980's have focused on building inter-firm collaborations by centralizing inventory control. A consequence of such collaborative practices has been that the value chains have become longer upstream. While VMI has enabled firms to manage operations at a global scale, it does not directly take into account the objectives of the various players along the value chain.

In Chapter 2, we study an inventory control problem faced by an upstream supplier who is in a VMI agreement with a retailer. VMI partnership provides the supplier a unique opportunity to manage inventory for the supply chain, in exchange for point-of-sales (POS) and inventory level information from the retailer. However, as is increasingly the case in the retail industry, big-box retailers capture and analyze customer purchasing behavior beyond the traditional POS data. Such analysis provides the retailer access to market signals that are otherwise hard to capture using POS information. In the absence of a credible channel for communication between the parties, the value of these signals to the supply chain operations can be limited. We demonstrate and quantify the implication of the incentive issue in VMI that renders communication of such important market signals as non-credible. To help institute a sound VMI collaboration, we propose a dynamic inventory mechanism for the supplier, to manage inventory and information in the supply chain. The proposed mechanism combines the ability of the supplier to learn about market conditions from POS data (over multiple selling periods) and to dynamically determine when to acquire his demand information. We show that the dynamic mechanism significantly improves the supplier's expected profit and increases the efficiency of the overall supply chain operations under a VMI agreement. We also show that inventory decisions serve a strategic purpose in addition to their classic role of satisfying customer demand.

In addition to supply-side inventory management, firms also actively engage in directly impacting the demand-side of the market by running sales promotions. As defined by Blattberg and Briesch (2012), sales promotion is a temporary price incentive offered to customers to stimulate demand. Firms do so for a number of reasons, including clearing end-of-season inventory, competitive pressures, or increasing customer base. At the heart of sales promotion mechanism is the idea that *temporarily* lowering prices provides an instantaneous incentive to induce lower-valuation customers to purchase the product. The key to a successful sales promotion, therefore, lies in understanding how much customers value consuming the product and the underlying factors that drive their consumption. With the advent of internet based economy and the ability to store and process large amounts of data, identifying customer-level factors that drive demand has become easier. However, one unintended consequence of the deluge of promotion activity over the years has been that firms now face a challenge in selling products at their regular price.

In Chapter 3, we study the impact of providing customers a delayed incentive—in the form of a gift card—for spending above an expenditure threshold on *regular* priced products. Such a promotion, popularly known as a gift card promotion, is widely used by department and consumer electronic stores. The general perception among retailers is that gift card promotion is profitable because it boosts customer expenditure during the promotion, and promotion costs are incurred only if the gift cards are redeemed in the future. The fact that increased expenditure during the promotion, if any, is realized from purchase of regular priced products, adds to the appeal of a gift card promotion. On the flip side, however, the promotion could backfire by providing customers with “free money”, if they would have made the purchase regardless of the promotion, or if customers reduce their expenditure towards the lower expenditure threshold to maximize gains from the promotion. Our main objective in this study is to quantify the effectiveness of gift card promotion in boosting customer expenditure during the promotion and later during redemption. To this end, we collaborate with a major U.S.-based department store, which runs several gift card promotions on its online channel annually.

We find that the gift card promotion impacts customer response (purchase and expenditure decisions) through participation in the promotion and through mere exposure to the advertisement of the promotion. On average, customer expenditure increases by 31.45% (or $198.64) during a gift card promotion, of which 96.34% can be attributed to participation in the promotion and remainder to the advertisement effect of the promotion on customers who do not participate in the promotion. Likewise, customer purchase probability increases by 17.54% during the promotion, which is mainly driven by the participation in the promotion. In addition, we find that not only are gift cards effective in bringing customers back to the store, but they also induce customers to spend more. On average, customers spend $525.28, more than the face value of the gift card and this effect positively correlates with the face value of the gift card. Therefore, redemption of gift cards can be profitable for the retailer offsetting any promotional costs incurred. For further discussion and insights, we also refer the reader to Kadiyala et al. (2017).

We conclude our key findings and contributions to the existing literature in Chapter 4.

CHAPTER 2

A STRATEGIC APPROACH TO COLLABORATIVE INVENTORY MANAGEMENT

2.1 Introduction

We study the interaction between a supplier and a retailer who operate within a collaborative partnership agreement, such as vendor-managed inventory (VMI). Under this agreement the supplier (she) takes the sole responsibility, including financial and operational control, of inventory in the supply chain. The retailer (he) takes the responsibility of store level execution to satisfy end customer demand as much as possible. The retailer uses information technology, such as EDI, to share customer sales information through point of sales (POS) data and inventory levels with the supplier at the end of each selling period. The POS data help the supplier to improve her demand forecasts for future periods, thereby also improving her inventory replenishment process over time. Practitioners and scholars have shown that centralized inventory control together with information sharing (e.g., VMI) allows supply chains to be more efficient and responsive to customer needs (see, for example, Aviv 2004; Simchi-Levi et al. 2008 for an extensive review of this literature). However, recent empirical and anecdotal evidence also suggests that VMI type agreements have also proved difficult to maintain over multiple planning horizons (e.g., Kouvelis et al. 2006, Brinkhoff et al. 2015). One often cited reason for such failed relationships has been incentive misalignment and declining of trust among firms implementing VMI, which manifests in the following fashion. The retailer, owing to his proximity and close relationship with his customers, obtains new demand information that is over and beyond POS data. The supplier could improve her forecasts, and hence, her inventory decisions by learning about the retailer's private demand information. However, the retailer faces a conflict of interest in credibly revealing his private information even in a long-term/multi-period relationship. The leftover inventory is the supplier's (and not the retailer's) liability. Hence, the retailer is always better off depicting a positive outlook of the market to ensure sufficient inventory during all selling periods (provided that this information is perceived reliable) and always demands for more supply. The resulting relationship, therefore, often boils down to only exchanging POS data. In addition POS data often do not contain lost sales, which neither party observes. Unaccounted lost sales results in supplier carrying less inventory, leading to further lost sales. Such lack of coordination often diminishes the retailer's patience for relinquishing control of his inventory, leading to a "lose-lose" outcome for both parties.[1] For such dynamic settings, we propose and study an inventory-mechanism that the supplier could use to improve her forecasts over time, by accounting for unobserved lost sales as well as enable credible sharing of the retailer's demand information, while maximizing expected profit over a finite planning horizon.

In an effort to make their operations lean, while maintaining high customer service, big-box retailers have increasingly formed collaborative relationships with their suppliers. For example, Wal-Mart and Procter & Gamble, Tesco and Nestlé, Northern Foods and Sainsbury's, are involved in VMI practices (see Lee et al. 1997; Watson 2005, The Grocer 2009). In a typical VMI agreement, the retailer relinquishes control of ordering decisions and sometimes also the financial responsibility of the inventory at his location to the supplier (as in consignment shipping). For example, Wal-Mart only owns its products briefly as they pass through the check-out scanner (see pg. 156 in Simchi-Levi et al. 2008). Centralizing inventory control by moving it up the supply chain, closer to the source, counters Bullwhip

[1] In the classical case, Hammond (2006) describes the stern opposition to VMI practice from Barilla's distributors. The article points out the difficulty Barilla had in incorporating promotional data, that is separate from the usual EDI information, into their forecasting process. Giorgio Maggiali, then Director of Logistics at Barilla, noted,

> *"We're grappling with how to treat these promotions in our operations planning processes, including forecasting, manufacturing, and logistics."*

Ineffectively managing inventory lead to disappointment of the distributors over VMI implementation, and eventually falling out of the relationship.

effect in at least two ways (Lee et al., 1997). First, variability in the demand seen by the supplier is not compounded by the retailer's ordering decisions. Second, the flow of POS data upstream improves demand forecasts, leading to fewer stock-outs at the retail store. Technology companies such as Dell and Apple have managed to avoid selling through resellers by vertically integrating with the downstream, thus minimizing the Bullwhip effect. They, however, also practice VMI with their upstream suppliers (Lee et al. 1997; Katariya et al. 2014).

In a VMI[2] agreement, the retailer shares POS information with the supplier (Dong et al., 2014). In addition to POS data, the retailer often has access to his customers' information that could be beneficial for the supplier in making replenishment decisions. For example, the retailer often obtains customer purchase data via loyalty/reward membership programs, and runs promotion events long after an initial VMI agreement is established. Analyzing such data provides the retailer a better estimate of potential demand for the product. If the supplier were to ask for such information, the retailer may find it profitable to report high demand for the product in an ongoing selling season (even when the information suggests otherwise). High inventory levels enable the retailer to possibly sell more in the later periods as well as improve customer perception of the store, without incurring stocking costs. Anticipating this incentive, the supplier may discount or even disregard such information (even when the retailer provides accurate information as there is no way to verify its accuracy).

The classic VMI setting does not provide the supplier and the retailer with a means to credibly share demand information beyond the POS data. As a result, Blackhurst et al. (2006) report, signicant gaps exist between potential and realized benefits of VMI. Scholars and managers attribute inaccurate forecasts as one of the key reasons for inappropriate levels of inventory at downstream locations. Although the supplier could learn about the customer demand through the POS data, such learning could potentially take many periods of sales.

[2]Henceforth, we consider a VMI setting in which inventory is managed in a consignment fashion.

In the meantime, the supplier continues to maintain higher/lower than necessary inventory levels. Consequences of improperly stocked shelves could be dire. Recently, *Forbes (2014)* reported that Wal-Mart, a pioneer in VMI, is loosing around $3 billion owing to out-of-stock items. The following excerpt illustrates another example for why the retailer may be reluctant to truthfully share demand information with the supplier:

> *... Sainsbury's discovered that a cereal brand called Grape-Nuts was worth stocking despite weak sales because the shoppers who bought it were extremely loyal to Sainsbury's and often big spenders. (The Guardian 2013)*

Sainsbury's in this case is likely not going to share low demand information with the supplier of Grape-Nuts if both firms are in a VMI type relationship.

Another inherent problem that hinders the supplier from improving her demand forecasts in aforementioned relationships is due to unobserved lost sales. When customers do not find a product on the shelf, they typically leave the store without purchasing and informing the retailer.[3] Thus, neither the retailer nor the supplier (vendor) observes demand that is not satisfied. This censoring of demand information creates the following inventory-information trade-off for the supplier. By maintaining higher inventory levels, the supplier has a greater chance to record an uncensored demand realization. This accuracy of information helps the supplier to improve her demand forecasts for future periods. The improvement in forecasts, however, comes at the risk of carrying excess inventory into future periods.

Researchers also note that the value of POS data when the supplier does not know the retailer's demand model can be limited (Chen and Lee, 2009). This information asymmetry

[3]A high-end retailer, such as Wholefoods, may occasionally ask at the cashier whether the customer found everything they were looking for. Yet, this practice is neither common at majority of the stores, such as Wal-Mart, nor consistent and accurate enough to be useful for inventory replenishment systems. Even if one assumes the retailer can collect all lost sales information, communicating such data still runs into the aforementioned incentive problem. That is, the retailer has incentive to claim high lost sales to induce the supplier to carry more inventory.

coupled with the inability to share local knowledge with the centralized decision maker can lead to overall poor performance (Aviv, 2002). For example, Spartan Stores ended its VMI program a year after its inception citing the supplier's inability to take into account promotional events at the retail store (Mathews, 1995). In fact, after taking off in early 90s, VMI practice faced tough opposition in industry, resulting from frictions between[4] the supply chain members participating in it. However, IT infrastructure developments and widespread use of data analytics to drive business decisions, has led to a resurgence in VMI implementations in the past decade (Aquino, 2009). Hence, for VMI partnerships to have sustained success, it is important to examine and align incentives of firms under this agreement, so that firms can credibly share and use retailer's demand information.

The above observations motivate us to study the following questions: How should the supplier dynamically manage centralized inventory in a lean fashion when lost sales are unobserved and the retailer has private demand information, over multiple-selling periods? Can the supplier use her inventory decisions to gain long-term strategic leverage in her partnership with the retailer? Is there a mechanism that the supplier could use (within the VMI framework) to credibly elicit demand information from the retailer while effectively managing inventory over a finite planning horizon? To address these questions, we propose a solution approach for the supplier that combines dynamic inventory control with mechanism design.

2.2 Literature Review

Many operations management scholars have explored and documented the benefits that accrue from the practice of information sharing in decentralized supply chains (Lee et al.,

[4] Although our emphasis is on incentive issues between supply chain members in VMI agreement, we note that it is possible that similar issues arise within a firm, as is illustrated in the following example. The monthly incentives of the sales department at Barilla were tied to the volume of sales of the product. This lead to the sales representatives at Barilla to exaggerate demand to improve their chances of end-of-month bonuses (Hammond, 2006).

2000; Cachon and Fisher, 2000; Aviv, 2001; Ren et al., 2010; Ha et al., 2011; Shang et al., 2016). The value of information sharing, in particular, of demand forecasts within the supply chain has been shown to play an important role in determining success/failure of collaborative partnerships such as—VMI and collaborative planning, forecasting and replenishment (CPFR), to name a few (Aviv, 2002, 2007; Chen and Lee, 2009). Aviv (2007) shows that supply chain characteristics such as—the retailer's ability to observe superior market signals and the supplier's agility in production, contribute to a win-win situation in a collaborative forecasting partnership. We note that these partnerships improve visibility of POS information upstream (which is verifiable) and/or centralizing replenishment processes. Improved visibility of demand in turn helps the supplier resolve some of demand variability over the planning horizon, albeit rather slowly. The question of whether these data-rich environments induce credible sharing of valuable information that is private and unverifiable, such as the retailer's subjective assessment of demand, is a natural extension to this line of investigation.

Researchers have provided several contractual remedies to alleviate the credibility issue that may arise when self-interested firms report demand forecasts. Cachon and Lariviere (2001) consider capacity decisions under demand information asymmetry. In their model, the retailer has a more accurate demand forecast and the forecast sharing game is modeled via a signaling game. In a separating equilibrium of such a game, demand forecast is shared credibly. Özer and Wei (2006) model the forecast sharing game using a screening and a signaling model. They propose capacity reservation contracts and advance purchase contracts that enable credible information sharing. Li and Zhang (2008) consider an interesting extension in which the supplier elicits demand information from multiple retailers. Babich et al. (2012) design a buyback contract when the retailer possesses private demand information. One of our contributions in this study is to extend this stream of literature by considering a multi-period inventory model in which the supplier improves her demand forecasts over time by incorporating historical POS data.

Few papers in literature consider an incentive problem in a dynamic setting: Zhang and Zenios (2008); Zhang et al. (2010); Oh and Özer (2013); Feng et al. (2015). Dynamic refers to the possibility of, evolution of information asymmetry and the solution being history dependent. Zhang and Zenios (2008) study long-term dynamic contracts that are offered once, at the beginning of planning horizon. Zhang et al. (2010) study dynamic short-term contracts that are offered in every period. The retailer's inventory level is his private information and demand information is common knowledge. In contrast, our focus is on long-term contracts in the presence of demand information asymmetry. The timing of contracts is another feature that distinguishes the dynamic aspect of our solution approach. Oh and Özer (2013) determine the supplier's one-shot capacity decision, for a single period of demand realization, using a (dynamic) mechanism, where the more informed party plays an active role. We consider a multi-period inventory problem, where demand is realized in each period and leftover inventory is carried forward. Feng et al. (2015) model a dynamic bargaining game between a buyer and a seller, in which the buyer is privately informed about his demand. The negotiation in their model continues until an agreement, on quantity and payment, for the trade of a product is reached. We introduce an important dimension to this stream of literature, by studying an incentive problem in a dynamic learning environment. The ability of the supplier to learn through her actions (inventory decisions) provides a basis for comparing the value of learning and value of screening in a supply chain setting.

The dynamic nature of demand information asymmetry described above, arises from the fact the supplier (statistically) updates her demand forecasts using the periodic POS data. The statistical evolution of demand forecasts has been modeled in literature using various approaches such as, time series, martingale method of forecast evolution (see Aviv 2001, 2002, 2007 and references there in), and Bayesian inference. We adopt the Bayesian approach (Scarf, 1959, 1960; Azoury, 1985; Lovejoy, 1990). In particular, parameters of the demand distribtuion are the retailer's private information. The supplier knows the family of demand distributions but only has a distributional knowledge of its parameters. Thus, the supplier updates his demand forecast over time following Bayes rule using the POS data, i.e., censored demand information. Due to information censoring, forecast evolution in our problem resembles that of the unobserved lost-sales Bayesian inventory problem (Lariviere and Porteus, 1999; Chen and Plambeck, 2008; Chen, 2010; Bisi et al., 2011). A noteworthy aspect of our Bayesian forecast evolution model is that, inventory decisions made by the supplier determine the extent of censoring of demand data in each sales period. Thus, the evolution of forecasts is *endogenized* through the supplier's inventory decisions.

2.3 The Model

We consider a supply chain consisting of a supplier and a retailer, participating in a VMI agreement. The supplier is responsible for periodically producing and maintaining on-hand inventory, over the remaining planning horizon consisting of N selling periods. At the beginning of a selling period $n \in \{1, \cdots, N\}$, the supplier decides on production quantity for the period, at unit cost c. The quantity produced is delivered to the retailer before demand is realized. The retailer then satisfies demand, to the extent possible from the inventory on-hand. Unmet demand is *lost* and neither the supplier nor the retailer observes these lost sales. For every unit sold, the retailer earns r from the customer and pays a wholesale price w to the supplier. The supplier is liable for the leftover inventory. Hence, the supplier incurs a unit holding cost, h on the leftover inventory that is carried over to the next period. At the end of the planning horizon, i.e. in period N, the leftover inventory is salvaged by the supplier, at the cost of its production, c.

Demand in each period is i.i.d. and is generated from a non-negative distribution, $G(z)$, $z \geq 0$. Both the supplier and the retailer are uncertain about demand, prior to its realization, in each selling period. To make better inventory decisions over time, the supplier obtains demand forecasts through the planning horizon. However, the retailer acquires additional demand information given his proximity to customers. This information could contain some useful market signal such as, an indicator of average market size for the remaining N periods. Using this information, the retailer is able to accurately estimate some parameter ξ of demand distribution, such that larger ξ represents larger average demand (recall that Sainsbury's private information about size of demand). Thus, we assume demand is stochastically ordered in the following sense, $\xi_1 \leq \xi_2$ implies $G(z|\xi_1) \geq G(z|\xi_2)$ for all $z \geq 0$. Therefore, the retailer's demand information comprises of complete knowledge of the underlying demand distribution. The supplier, however, consolidates her prior demand information (for the remaining horizon) in form of a belief (p.d.f.), π, over $\Theta := [\underline{\xi}, \overline{\xi})$, the set of values ξ takes. Information from previous selling periods and initial market research could be summarized to develop the prior belief.

At the beginning of a selling period n, the supplier raises on-hand inventory level from x_n to y_n. Demand for that period, D_n is then realized, but the supplier only observes the POS information, that is, $z_n := \min\{y_n, D_n\}$ for that period. Using this information, the supplier updates her belief[5] about ξ using the Bayes rule as follows.

$$
\begin{aligned}
\pi_{n+1}(\xi) &= \mathbf{1}_{\{z_n = y_n\}} \cdot \frac{\overline{G}(y_n|\xi)\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \overline{G}(y_n|\eta)\pi_n(\eta)\mathrm{d}\eta} + \mathbf{1}_{\{z_n < y_n\}} \cdot \frac{g(z_n|\xi)\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} g(z_n|\eta)\pi_n(\eta)\mathrm{d}\eta} \\
&= \mathbf{1}_{\{z_n = y_n\}} \cdot \pi^c_{n+1}(\xi|y_n) + \mathbf{1}_{\{z_n < y_n\}} \cdot \pi^e_{n+1}(\xi), \quad 1 \leq n \leq N,
\end{aligned} \tag{2.1}
$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function, $\pi_1 = \pi$, and $\overline{G}(\cdot) = 1 - G(\cdot)$. The first term in Equation (2.1), $\pi^c_{n+1}(\xi|y_n)$, is the posterior when the demand realization in the current period is greater than the on-hand inventory level. The second term, $\pi^e_{n+1}(\xi)$, is the posterior when the supplier observes the exact demand realization. We differentiate the notation of π^e_{n+1} from π^c_{n+1} to emphasize the dependence of the posterior on the on-hand inventory level y_n, when the demand realization is *censored*.

[5] $\pi_n(\xi) := \mathbb{P}\{\xi \in \mathrm{d}\xi|\mathcal{Z}_{n-1}\}$, where $\{\mathcal{Z}_{n-1}\}$ is the natural filtration associated with sales process z_{n-1}, $n \geq 2$.

The supplier's expected profit[6] in period n after bringing the on-hand inventory level to y_n is given by

$$\begin{aligned}\mathbb{E}_{\xi,\,D_n}\big[w\min\{y_n,\,D_n\}-c(y_n-x_n)-h(y_n-D_n)^+\big] &= cx_n+(w-c)y_n-(w+h)\cdot\mathbb{E}_{\xi,\,D_n}[y_n-D_n]^+\\ &= cx_n+(w-c)y_n-(w+h)\int_0^{y_n}Q_n(z)\,\mathrm{d}z,\end{aligned} \tag{2.2}$$

$$\text{where}\quad Q_n(z):=\int_0^z q_n(u)\,\mathrm{d}u\ \text{ and }\ q_n(z):=\int_\Theta g(z|\xi)\pi_n(\xi)\mathrm{d}\xi \tag{2.3}$$

denote the posterior predictive (demand) distribution and density, respectively, of demand in period n. The expectation in the supplier's profit function is with respect to random demand and the unknown market signal, ξ, that is private information to the retailer.

The retailer's expected profit, in period n, given his demand information ξ, is

$$(r-w)\cdot\mathbb{E}_{D_n}\big[\min\{y_n,\,D_n\}\big]=(r-w)\bigg(y_n-\int_0^{y_n}G(z|\xi)\,\mathrm{d}z\bigg).$$

We note that the retailer's profit is increasing in on-hand inventory level y_n. Therefore, the retailer has incentive to report optimistic demand to induce the supplier to allocate high on-hand inventory in each period. At the beginning of period $n+1$, the supplier updates on-hand inventory level x_{n+1} to $[y_n-D_n]^+$ and her belief over ξ to π_{n+1} following Equation (2.1).

[6] We note that a penalty cost of p per unit of lost sales can be included in the supplier's profit function as follows.

$$\begin{aligned}&\mathbb{E}\big[w\min\{y_n,\,D_n\}-c(y_n-x_n)-h(y_n-D_n)^+-p(D_n-y_n)^+\big]\\ &=cx_n+(w-c)y_n-(w+h)\mathbb{E}[y_n-D_n]^+-p\mathbb{E}\big[D_n-y_n+(y_n-D_n)^+\big]\\ &=cx_n+(w+p-c)y_n-(w+h+p)\mathbb{E}[y_n-D_n]^+-p\mathbb{E}\big[D_n\big].\end{aligned}$$

Given π_n, the last term is a constant in the objective function the supplier maximizes in period n. Thus, introducing penalty cost does not affect the insights generated in the study.

2.3.1 The *Learn and Screen* Approach

We propose a 'learn and screen' approach to help the supplier with the joint problem of inventory control, demand estimation, and the incentive problem, that arises in a VMI framework. The sequence of events in the learn and screen approach is as follows. At the beginning of period $n \in \{1, 2, \cdots, N\}$, the supplier decides between— *learning* on her own or *screening* the retailer. (i) If the supplier opts for the former, she raises her inventory level upto y_n while statistically improving her belief over ξ via the Bayes law. The problem proceeds to the next period. (ii) Otherwise, the supplier decides to screen the retailer at the beginning of period n and offers a menu of long-term contracts, $\{S(\xi|\, x_n,\, \pi_n),\, P(\xi|\, x_n,\, \pi_n)\}_{\xi\in\Theta}$. We note that these contracts are a function of on-hand inventory level as well as the updated posterior belief. The retailer decides whether or not to accept a contract from this menu. If he accepts and chooses the contract $S(\hat{\xi}|\, x_n,\, \pi_n)$, $P(\hat{\xi}|\, x_n,\, \pi_n)$, the supplier procures inventory for the remaining planning horizon following the base-stock level $S(\hat{\xi}|\, x_n,\, \pi_n)$. In other words, she produces enough to bring the inventory level y_n to $S(\hat{\xi}|\, x_n,\, \pi_n)$ in each period. The retailer pays $P(\hat{\xi}|\, x_n,\, \pi_n)$ to the supplier at the beginning of every period or equivalently, the retailer pays a one-time discounted lump sum after accepting one of the contracts. The retailer then satisfies realized demand to the extent possible by procuring at w per unit and selling at r; the supplier updates inventory, and this repeats next period.

In an ongoing VMI agreement, the retailer over time is in a better position than the supplier to assess market conditions. However, the inability to credibly communicate his market assessment, when approached by the supplier, creates a tension in the supply chain. The retailer tries to push the supplier into maintaining higher inventory levels, and the supplier cannot ascertain if the retailer is right. Such tensions within VMI agreements have been commonly observed in practice. For example, Spartan Stores shut down VMI agreements after a year, blaming the supplier's inability to forecast accurately (Mathews, 1995). Offering screening contracts within an existing VMI agreement provides a credible way to communicate demand information and hence, has potential to alleviate such tensions. In addition to designing these contracts, in the learn and screen approach, the supplier also determines when to offer them in a long-term, multi-period relationship. If the retailer does not accept the menu of contracts offered, the supply chain relationship ends, and the supplier and the retailer make profit by pursuing their options outside the VMI relationship. The supplier's outside option is normalized to zero and the retailer's to $\Pi^r_{\min}(n)$. The retailer earns his profit by using his shelf space to stock a different product for the remaining horizon. Thus, timing of the contracts also impact value of the outside option for the retailer.

We highlight three important benefits of the proposed contract terms here: (i) The ongoing VMI agreement between the parties is unaffected if the retailer agrees to one of the base-stock levels in the menu. The ownership of inventory continues to remain with the supplier and the contracts merely act an instrument to facilitate credible communication of demand forecasts. (ii) The form of the contract is optimal (i.e., best among all possible forms) because the supplier faces the classical periodic-review inventory control problem with lost sales, once demand information is (and can be) credibly shared. For such an inventory problem Karlin and Scarf (1958) have shown the optimality of base-stock policy, thus justifying the contract terms. (iii) Monitoring the contract terms, once it is accepted, requires minimal effort. The supplier collects a one-time payment from the retailer and the retailer periodically monitors the inventory level maintained by the supplier. Current VMI frameworks such as, PeopleSoft Enterprise Inventory and Fulfillment Management by Oracle, already implements this feature. For example, an automated message is delivered to the retailer as soon as inventory is replenished on his shelf (see pg. 1040, Oracle 2009). In all subsequent periods, the retailer pays a unit wholesale price w on the units procured. Thus the financial transactions between the supplier and the retailer remain unaffected following the period in which the contracts are offered.

The Contract Design Problem.

Suppose that the supplier offers the menu of contracts, $\{S(\cdot),\, P(\cdot)\}$, at the beginning of period n. If the retailer chooses a particular contract $S(\tilde{\xi})$, $P(\tilde{\xi})$ from this menu, the supplier delivers inventory following a base-stock level $S(\tilde{\xi})$. The supplier offers base-stock levels that are at least as much as the on-hand inventory level x_n. Such a commitment from the supplier assures the retailer that enough inventory will be available on the shelf, if he decides to participate in the screening mechanism. Hence, this commitment facilitates easier implementation of the screening contracts. Thus $y_j = S(\tilde{\xi})$, $j \geq n$, and the inventory evolves as follows:

$$x_{j+1} = [S(\tilde{\xi}) - D_j]^+, \quad n \leq j \leq N.$$

Given the menu of contracts and the retailer's choice of a contract from the menu, the expected profit of the supplier, the retailer of type ξ, and the total supply chain, over the remaining horizon are

$$\Pi_n^s\big(S(\tilde{\xi}),\, P(\tilde{\xi})\big|\, x_n,\, \pi_n\big) = \sum_{i=n}^{N} \alpha^{i-n}\mathbb{E}_D\Big[w \min\{S(\tilde{\xi}),\, D_i\} - c(S(\tilde{\xi}) - x_i) - h(S(\tilde{\xi}) - D_i)^+ + P(\tilde{\xi}) + \alpha^{N-n+1}c\big(S(\hat{\xi}) - D_N\big)^+\Big],$$

$$\Pi_n^r(S(\tilde{\xi}),\, P(\tilde{\xi}),\, \xi) = \sum_{i=n}^{N} \alpha^{i-n}\mathbb{E}_D\Big[(r-w)\min\{S(\tilde{\xi}),\, D_i\} - P(\tilde{\xi})\Big], \quad \text{and}$$

$$\Pi_n^{tot}\big(S(\tilde{\xi})\big|\, x_n,\, \pi_n\big) = \sum_{j=n}^{N} \alpha^{j-1}\mathbb{E}_{\xi,D}\Big[r\cdot\min\{S(\tilde{\xi}),\, D_j\} - c(S(\tilde{\xi}) - x_j) - h(S(\tilde{\xi}) - D_j)^+ + \alpha^{N-n+1}c\big(S(\tilde{\xi}) - D_N\big)^+\Big],$$

respectively. Note that $\tilde{\xi}(\xi) : \Theta \to \Theta$ is the retailer's response function and $\alpha \in [0,\, 1)$ is the discount factor. Excess inventory is the supplier's responsibility. Hence, only the supplier benefits from salvaging the leftover inventory at the end of the planning horizon.

Each menu, $\{S(\cdot),\, P(\cdot)\}$ determines a Bayesian game in which the retailer chooses, in equilibrium, a contract that maximizes his total expected profit over the remaining planning horizon:

$$\Pi_n^r\big(S(\tilde{\xi}),\, P(\tilde{\xi}),\, \xi\big) = \max_{\eta}\, \Pi_n^r(S(\eta),\, P(\eta),\, \xi), \quad \forall\, \xi \in \Theta. \tag{IC}$$

To ensure participation of the retailer the supplier guarantees at least his reservation profit

$$\Pi_n^r(S(\xi),\, P(\xi),\, \xi) \geq \Pi_{\min}^r(n), \quad \forall\, \xi \in \Theta. \tag{PC}$$

Otherwise, no contract is preferable to the retailer. The supplier's incentive problem in period n can be summarized as follows:[7]

$$\tilde{\Pi}_n^{sr}\big(x_n,\, \pi_n\big) := \max_{S(\cdot),\, P(\cdot)}\, \mathbb{E}_{\xi}\big[\Pi_n^s\big(S(\xi),\, P(\xi)|\, x_n,\, \pi_n\big)\big];\ \text{subject to } S(\cdot) \geq x_n, \text{ (IC), and (PC).} \tag{2.4}$$

The Bayesian Inventory Control-Optimal Stopping Problem.

Let $\tilde{V}_n(x_n,\, \pi_n)$ denote the maximum profit the supplier could make during periods n through N using the learn and screen approach. Recall that $(x_n,\, \pi_n)$ represents the on-hand inventory in the retail store and the supplier's belief about the demand, respectively, at the beginning of period n. The supplier decides whether to continue learning about demand through POS information or screen the retailer for better demand information. The optimality equations for $n = 1,\, 2,\, \cdots,\, N$ determines this trade-off,

$$\tilde{V}_n(x_n,\, \pi_n) = \max\left\{\tilde{\Pi}_n^{lr}(x_n, \pi_n),\ \tilde{\Pi}_n^{sr}(x_n,\, \pi_n)\right\}, \tag{2.5}$$

where $\tilde{V}_{N+1}(x_{N+1},\, \pi_{N+1}) := cx_{N+1}$ for all x_{N+1}, π_{N+1}. $\tilde{\Pi}_n^{sr}(\cdot,\, \cdot)$ is defined in Equation (2.4) and

$$\tilde{\Pi}_n^{lr}(x_n, \pi_n) := \max_{y \geq x_n}\, \tilde{L}_n(y,\, \pi_n)$$

[7]The *revelation principle* (Myerson, 1981) ensures that, without loss of generality, the supplier can restrict herself to those menus of contracts in which, the retailer truthfully communicates her private assessment of demand, in Bayesian equilibrium. Thus, we narrow our search for the optimal menu to those menus of contracts, for which $\tilde{\xi} = \xi$ is the best response function for the retailer of type ξ.

$$:= \max_{y\ge x_n} \Big\{ cx_n + (w-c)y - (w+h)\int_0^y Q_n(z)\,\mathrm{d}z + \alpha\big(1-Q_n(y)\big)\tilde{V}_{n+1}\big(0,\,\pi^c_{n+1}(\cdot|y)\big)$$
$$+\,\alpha\int_0^y q_n(z)\tilde{V}_{n+1}\big(y-z,\,\pi^e_{n+1}\big)\,\mathrm{d}z\Big\},\quad \text{for all } y\ge x_n,$$

where Q_n, q_n are defined in Equation (2.3). The first term on the rhs of Equation (2.5) is the maximum profit obtained by the supplier if she decides to continue learning via the Bayesian updating in period n. The second term is the profit obtained from screening the retailer in period n. The first three terms of $\tilde{L}_n(y,\,\pi_n)$ represent the myopic profit from raising on-hand inventory level to y_n in period n (as defined in Equation (2.2)). The last two terms correspond to future profit stream, depending on the demand observation in period n being censored or uncensored.

The optimal time to offer the menu of contracts is $\tau := \min\{n:\ 1\le n\le N,\ \tilde{\Pi}^{lr}_n(x_n,\,\pi_n)\le \tilde{\Pi}^{sr}_n(x_n,\,\pi_n)\}$. If $\tilde{\Pi}^{lr}_n(x_n,\,\pi_n) > \tilde{\Pi}^{sr}_n(x_n,\,\pi_n)$ for all $1\le n\le N$, we set $\tau = N+1$ which means it is never optimal for the supplier to offer the menu of contracts in any period.

The DP in Equation (2.5) can be simplified (see Appendix A.2 for details) following the transformation $V_n(x,\,\pi) := \tilde{V}_n(x,\,\pi) - cx_n$ and $\Pi^{sr}_n(x,\,\pi) := \tilde{\Pi}^{sr}_n(x,\,\pi) - cx$. The resulting optimality equations for $n = 1, \cdots, N$ are

$$V_n(x_n,\,\pi_n) = \max\Big\{\Pi^{lr}_n(x_n,\pi_n),\ \Pi^{sr}_n(x_n,\,\pi_n)\Big\}, \tag{2.6}$$

where $V_{N+1}(x_{N+1},\pi_{N+1}) := 0$ for all x_{N+1},π_{N+1} and

$$\Pi^{lr}_n(x_n,\pi_n) := \max_{y\ge x_n} L_n(y,\,\pi_n) \tag{2.7}$$
$$:= \max_{y\ge x_n} \Big\{ (w-c)y - (w+h-\alpha c)\int_0^y Q_n(z)\,\mathrm{d}z + \alpha\big(1-Q_n(y)\big)V_{n+1}\big(0,\,\pi^c_{n+1}(\cdot|y)\big)$$
$$+\,\alpha\int_0^y q_n(z)V_{n+1}\big(y-z,\,\pi^e_{n+1}\big)\,\mathrm{d}z\Big\},\quad \text{for all } y\ge x_n.$$

2.3.2 Relation to the Bayesian Inventory Problem

We remark that if the supplier does not screen the retailer's private information, that is if she sets $\tau = N+1$, then her problem in Equation (2.5) is to solve the Bayesian inventory problem with unobserved sales. Even this simpler dynamic program is difficult to solve for two reasons. First, the objective function, $\tilde{L}_n(y, \pi_n)$ is not concave in y in general. In particular future beliefs are affected by inventory decisions through $\pi^c_{n+1}(\cdot|y)$ in Equation (2.1). Hence, a simple policy such as a state-dependent base-stock policy need not be optimal. For an example we refer the reader to Theorem 2(i) in Bisi et al. (2011). Second, to update the belief over ξ at the end of period n, one needs to keep track of π_1 and the entire history of sales observations $(z_1, \cdots, z_n)$. Hence, the state space of the DP grows with time and one quickly runs into the curse of dimensionality. To overcome the analytical and computational challenges posed by the Bayesian inventory problem with unobserved lost sales, researchers have focused on the newsvendor class of distributions (see Lariviere and Porteus 1999; Bisi et al. 2011). These demand distributions possess several desirable statistical properties when demand observations are censored by the on-hand inventory level—as is the case in the classical newsvendor problem.

Definition 2.3.1. *A cummulative distribution, $F(z|\xi)$, $z \geq 0$ belongs to the newsvendor family (denoted henceforth as $\mathcal{N}$) if it can be expressed as $1 - e^{-\frac{t(z)}{\xi}}$, where ξ is the parameter and $t(z)$ is a non-negative increasing function.*

Given n sales realizations $z_1, \cdots, z_n$, of which m are uncensored demand observations, the two-dimensional sufficient statistic for the newsvendor likelihood and the unknown parameter ξ is $\left(m; \sum_{i=1}^{n} b(z_i)\right)$. That is, all the information contained in the sample $\{z_i\}_{i=1}^{n}$ regarding the unknown parameter ξ, can be summarized by the two numbers. Thus, the state space of the DP can be reduced to three variables. Braden and Freimer (1991) were the first to identify[8] the newsvendor family of distributions. Further they present several distributions, including Weibull, that belong to the newsvendor family. We add a new dimension to this

[8]They comment that (pg. 1390, Braden and Freimer 1991): *"If a modeler feels that no member of the families we characterize is a reasonable approximation, then he will almost surely encounter serious analytic and computational problems if his data include censored observations."*

classical Bayesian inventory problem by incorporating demand information asymmetry in decentralized supply chains.

2.4 Analysis

To solve for the supplier's optimal strategy we solve the problem by backwards induction. In §2.4.1 we first characterize the optimal contracts offered by the supplier in any given period. In §2.4.2 we study the impact of the supplier's inventory decisions on the screening contracts. In §2.4.3 we characterize the optimal time to offer the menu of contracts.

2.4.1 Optimal Menu of Contracts

In a VMI relationship, the supplier's belief about market conditions evolves over time. This aspect of VMI adds a dynamic learning dimension to the adverse selection problems seen in operations management/economics literature. Hence profit from offering screening contracts to the retailer depends on the period in which they are offered. Another interesting feature of the VMI framework is that the supplier has control over inventory decisions, which in turn drive her belief. This gives rise to a dynamic incentive problem in which the supplier faces a trade-off between learning and offering screening contracts. In Lemma 2.4.1, we use the classical approach (Mirrlees, 1971) of solving this incentive problem by expressing (IC) as a differential equation that governs the marginal informational rent offered to the retailer.

Lemma 2.4.1. *A menu of contracts $\{S(\cdot),\, P(\cdot)\}$ is feasible, i.e. the menu satisfies (IC) and (PC), if and only if the menu also satisfies the following:*

(i) The expected profit (informational rent) of the retailer for the remaining periods is given by

$$\Pi_n^r(\xi) := \Pi_n^r\big(S(\xi),\, P(\xi),\, \xi\big) = \Pi_{\min}^r(n) - \gamma(n)\cdot(r-w)\int_{\underline{\xi}}^{\xi}\int_0^{S(\eta)}\frac{\delta}{\delta\eta}G(z|\eta)\mathrm{d}z\mathrm{d}\eta, \quad \forall\, \xi, \tag{2.8}$$

and $\gamma(n) := \frac{1-\alpha^{N-n+1}}{1-\alpha}$.

(ii) The base-stock level $S(\xi)$ *is increasing.*[9]

We use the above characterization of (IC) and (PC) to reformulate and solve for the supplier's dynamic contract design problem and obtain her resulting optimal profit.

Lemma 2.4.2. *The following optimization problem determines the optimal menu of contracts.*

$$\tilde{\Pi}_n^{sr}(x_n,\, \pi_n) = cx_n - \Pi^r_{\min}(n) + \gamma(n) \max_{\left\{\substack{S(\cdot) \\ s.t.\ S' \geq 0,\, S(\underline{\xi}) \geq x_n}\right\}} \int_\Theta \pi_n(\xi) \cdot H(S(\xi),\, \xi|\pi_n)\,\mathrm{d}\xi, \quad (2.9)$$

$$\textit{where}\ \ H(S,\, \xi|\pi_n) := (r-c)S - (r+h-\alpha c)\int_0^S G(z|\xi)\mathrm{d}z + \frac{(r-w)}{\lambda_n(\xi)}\int_0^S \frac{\partial}{\partial \xi} G(z|\xi)\mathrm{d}z \quad (2.10)$$

and $\lambda_n(\xi) := \frac{\pi_n(\xi)}{\int_\xi^{\bar{\xi}} \pi_n(\eta)\,\mathrm{d}\eta}$ *is the failure rate function corresponding to the pdf,* π_n.

Solving the above problem determines the optimal menu of contracts to be offered to the retailer, if the supplier decides to screen the retailer in period n. The similarity with the classical mechanism design solution approach ends here. First note that the supplier's (principal's) belief process, π_n, is dynamically updated using the Bayes rule (see Equation 2.1). Thus, the information structure in our problem deviates from the classical principal-agent problem, in which the principal's belief remains static. Second the supplier's dynamic *virtual surplus*, $H(S, \xi|\pi_n)$ (defined in Equation (2.10)), depends on the initial prior information π, historical POS data $z_1, \cdots, z_{n-1}$ and inventory decisions $y_1, \cdots, y_{n-1}$, through the updated failure rate function $\lambda_n(\cdot)$. The function $\lambda_n(\xi)$ measures the supplier's subjective belief that the parameter of demand distribution is ξ, conditioned on it being at least ξ. Note that, larger ξ translates to having greater average demand. Inventory decisions and POS data prior to screening determine the evolution of $\lambda_n(\cdot)$ and hence, also determine the dynamic learning aspect of the learn and screen mechanism.

[9]We use increasing/decreasing in the weak sense throughout the chapter.

Third, consider the maximization problem in Equation (2.9) without the monotonicity constraint, $S' \geq 0$. The standard solution approach involves showing that the optimal base-stock menu for this relaxed problem is increasing in the retailer's type, thus establishing its optimality for the constrained optimization problem. We define $\hat{S}_n(\xi|\pi_n) := \arg\max_S\{H(S,\xi|\pi_n)\}$. However, unlike in the classical mechanism design problems, the function $H(S,\xi|\,\pi_n)$, which results from a multi-period lost-sales inventory problem (after accounting for the retailer's informational rent), is not concave in S for every type ξ, and a given prior π_n. Thus, we look for weaker structural properties, such as unimodality of $H(\cdot,\,\xi|\pi_n)$, that would ensure first order conditions are not only necessary, but also sufficient for existence and uniqueness of $\hat{S}_n(\xi|\pi_n)$. To this end, in Theorem 2.4.1, we characterize a family of demand distributions for which $H(S,\,\xi|\pi_n)$ is unimodal in S. Lemma A.3.1 in the appendix provides conditions to verify unimodality.

Theorem 2.4.1. *Suppose that the demand distribution, $G(z|\xi)$, is from the exponential family (see §3.4 Berger and Casella 2002) i.e., $g(z|\xi) = h(z)\cdot c(\xi)\cdot e^{-t(z)\cdot w(\xi)}$, $z \geq 0$ and functions $h(\cdot)$, $t(\cdot)$, $c(\cdot)$ and $w(\cdot)$ are differentiable where, $h(\cdot)$, $t(\cdot)$ are defined over $\mathbb{R}^+$ and $c(\cdot)$, $w(\cdot)$ are defined over Θ. If c, w are decreasing and t is increasing, then the following statements hold for any belief π_n.*

1. *The family of demand distributions, $\big\{G(z|\xi)\big\}_{\xi\in\Theta}$, is stochastically increasing.*

2. *$H(\cdot,\,\xi|\,\pi_n)$ is unimodal for all $\xi \in \Theta$.*

3. *$S_n^*(\xi|x_n,\,\pi_n) := \max\{\hat{S}_n(\xi|\,\pi_n),\,x_n\}$, is the maximizer for $H(S,\,\xi|\,\pi_n)$ over $S(\underline{\xi}) \geq x_n$, where $\hat{S}_n(\xi)$ solves the following first order condition.*

$$(r-c) - (r+h-\alpha c)G(S|\xi) + \frac{(r-w)}{\lambda_n(\xi)}\cdot\frac{\partial}{\partial\xi}G(S|\xi) = 0 \tag{2.11}$$

Normal, Gamma, Weibull are some distributions that satisfy the sufficient conditions (all with unknown *scale* parameter) in Theorem 2.4.1. Unimodality of $H(\cdot,\xi|\pi_n)$ ensures a simple characterization of the menu of base-stock levels. Such a characterization emphasizes the practical value of offering the menu of contracts to the retailer in a VMI relationship. Once the retailer accepts a contract, the inventory policy for the remaining horizon is a simple base-stock policy. Otherwise, the optimal inventory levels determined by the VMI manager have a complex history-dependent structure. We note that the characterization in Theorem 2.4.1 also include[10] the newsvendor family of distributions, which have been widely used in unobserved lost-sales Bayesian inventory literature (see discussion in §2.3.2). In Theorem 2.4.2 we provide sufficient conditions that guarantee monotonicity of the base-stock levels $\hat{S}_n$ and characterize the optimal menu of contracts for the incentive problem in Equation (2.4).

Theorem 2.4.2. *Suppose that π_n has IFR property, $\big\{G(\cdot|\xi)\big\}_\xi \subset \mathcal{N}$, and $\frac{r-c}{r+h-\alpha c} \leq 1 - e^{-2}$. Then the menu of base-stock levels $S_n^*(\xi)$ is increasing. Therefore, $\{S_n^*(\cdot),\, P_n^*(\cdot)\}$, is the optimal menu of contracts where*

$$P_n^*(\xi) := (r-w)\bigg(S_n^*(\xi) - \int_0^{S_n^*(\xi)} G(z|\xi)\,\mathrm{d}z\bigg) + (r-w)\int_{\underline{\xi}}^{\xi}\int_0^{S_n^*(\eta)} \frac{\partial}{\partial\eta}G(z|\eta)\,\mathrm{d}z\,\mathrm{d}\eta - \frac{\Pi^r_{\min}(n)}{\gamma(n)}. \tag{2.12}$$

Henceforth, we denote $S_n^*(\xi) = S_n^*(\xi|x_n,\, \pi_n)$, to simplify notation. The evolution of $\lambda_n(\cdot)$, via (censored) Bayesian updating determines the dynamic aspect of the optimal contracts, (2.11)–(2.12). Note that $\lambda_{n+1}(\xi) > \lambda_n(\xi)$ means the supplier is more confident in period $n+1$ that the underlying market signal is ξ, given it is at least ξ, compared to the previous period. Given this updated belief, it can be verified from Equation (2.11) that type-ξ retailer is offered a higher base-stock level in period $n+1$ as opposed to period n. The evolution of failure rate function depends on the supplier's historical inventory decisions and POS information. This dependence of contract structure on demand information acquired through inventory control, ties together the supplier's learning and screening problems. We further

[10]Setting $w(\xi) = c(\xi) = \xi^{-1}$ and $h(z) = t'(z)$, we obtain the newsvendor family.

investigate the impact of inventory decisions in learn and screen approach, on the supply chain operations in §2.4.2.

The first condition in Theorem 2.4.2 (π_n has IFR property), is a standard assumption in static, exogenous information mechanism design problems (see pg. 156, Tirole 2002). The mechanism we consider has a dynamic, *endogenous*[11] information structure with learning (see Bergemann and Välimäki 2006; Zhang et al. 2010; Roesler 2014; Birge and Keskin 2015 for related literature). The subtle issue here is the preservation of IFR property under Bayesian learning. The newsvendor family of distributions is particularly useful here. These distributions not only have a finite-dimensional sufficient statistic, but also the inverse gamma distribution is a conjugate prior for the newsvendor likelihood and this motivates the second condition. The inverse gamma distribution has IFR property for most of its parameter space. Since the updated posterior is also inverse gamma, IFR property is preserved. The last sufficient condition is on the cost parameters is satisfied in low to medium[12] margin industry. Although this assumption might seem restrictive at first, note that lower margins are typically a characteristic of the fast moving consumer goods in the retail industry (see *Ernst & Young 2013*). We also remark that these conditions are sufficient but not *necessary.*

It can be verified from Equation (2.12) that monotonicity of the menu of optimal payments, $P_n^*(\cdot)$ follows from monotonicity of menu of base-stock levels. Thus, offering a payment schedule $P_n^*(S_n^*)$, such that type-ξ retailer chooses order-upto level, $S_n^*(\xi)$ and hence, pays $P_n^*(S_n^*(\xi))$, implements the menu of contracts $\{S_n^*(\cdot), P_n^*(\cdot)\}$. The retailer does not need to explicitly communicate his private information using the payment schedule, rather communicates the inventory level he finds suitable given his market demand. Generally, in VMI relationships, it is considered a good practice for the retailer to guide the supplier in

[11]Note that the supplier's action $(y_n - x_n)$ influences the information signal (periodic sales, z_n) observed, and together they govern her belief process at the beginning of the following period.

[12]The condition is met if, for example, the total margin in the supply chain, $\frac{r-c}{r}$, is less than 87%.

making inventory decisions by communicating about changing market conditions.[13] While this communication is informal, the screening mechanism proposed ensures that the retailer stands monetarily accountable for providing his inputs.

2.4.2 Impact of Dynamic Inventory Decisions on The Optimal Contracts

In this section we show how the supplier's inventory decisions prior to offering contracts affect the optimal contracts she offers to the retailer. By maintaining higher inventory levels the supplier can obtain potentially uncensored POS information to improve her knowledge about market conditions. We show that obtaining superior quality demand information (uncensored POS data), may or may not translate into offering base-stock levels closer to the channel-coordinating level in the following period. That is, whether the retailer has the incentive to choose a base-stock level (and a corresponding payment) that is closer to the supply chain-coordinating level depends also on the magnitude of the (uncensored) POS information and the inventory level on hand. However, censored POS information unilaterally pushes the optimal menu of base-stock levels away from the coordinating level. Thus, in addition to the number of selling periods in which the supplier invests to learn about demand, censoring and magnitude of demand information gathered in these periods, plays a key role in determining the value of screening to the supplier.

Given the significant negative (tangible, and otherwise) impact of lost sales on the customers' perception of the retail store, we anchor our insights in this section around this event. We proceed by first understanding the impact of inventory decisions in a VMI agreement on the supplier's belief process and later, understand its implications on the optimal menu of contracts offered to the retailer. In Theorem 2.4.3, we show that by Bayesian updating, the supplier's posterior belief is stochastically ordered, in the sense of failure

[13] See http://www.vendormanagedinventory.com/pitfalls.php.

(hazard) rate,[14] with respect to the prior belief. The direction of ordering depends on whether the demand realization in a period is censored or not.

Theorem 2.4.3. *We have the following properties.*

1. *If the sales observation in period* n *is censored, i.e.,* $z_n = y_n$*, then* $\lambda_n(\xi) \geq \lambda_{n+1}(\xi|y_n)$ *and* $S_n^*(\xi) \geq S_{n+1}^*(\xi)$.

2. *Suppose the sales observation in period* n *is uncensored, i.e.,* $z_n < y_n$.

   (i) If $x_n \leq \hat{S}_n(\underline{\xi})$*, then for types* $\xi \geq t(z_n)$*,* $\lambda_n(\xi) \leq \lambda_{n+1}(\xi)$ *and* $S_n^*(\xi) \leq S_{n+1}^*(\xi)$.

   (ii) If $x_{n+1} \leq x_n$*, then for types* $\xi < t(z_n)$*, such that* $g(z_n|\xi) < g(z_n|\overline{\xi})$*, we have* $\lambda_n(\xi) > \lambda_{n+1}(\xi)$ *and* $S_n^*(\xi) > S_{n+1}^*(\xi)$.

The intuition behind the failure rate ordering in Theorem 2.4.3 Part 1 is that, a censored observation suggests to the supplier that the average market size must be larger. Thus, given the average market size is at least ξ, the supplier is more confident that it is greater than ξ, following a period of censored demand realization. An interesting consequence of this ordering is that the optimal menu of base-stock levels offered in the following period is smaller. Intuitively, one would expect the supplier to offer higher base-stock levels in the following period to account for her updated belief in higher demand. However, a careful analysis and thought reveals that the result is in fact the opposite. An empty shelf at the end of a period, signals possible lost sales in that period to the retailer. Therefore, in the following period, the retailer has reason to want the supplier to stock more inventory then the previous period. That is, the retailer has greater incentive than the previous period to choose a base-stock level meant for a larger market. Such exaggeration has also been observed in VMI practice and often leads to tensions between the two parties. Realizing this incentive

[14] Let G, H be two continuous, differentiable distribution functions and μ_1, μ_2, denote their failure rate functions, respectively. Then $G \leq_{fr} H$ iff $\mu_1(x) \geq \mu_2(x)$ for all $x \geq 0$.

issue exists, the supplier lowers the optimal menu of base-stock levels from the previous period for all types of the retailer (see Equation 2.11), to preserve incentive compatibility of the contracts (Theorem 2.4.3, Part 1). In addition to lost revenues in the current period, lost sales (using learn and screen approach) also negatively impact the menu of contracts offered in the following period. Thus, from the point of view of the supply chain, offering contracts following a period with lost sales, hurts both the retailer's and the supplier's profit. The supplier, in this case, contracts upon inventory levels lower than what is necessary to meet demand.

The impact of uncensored demand on the retailers incentives in the following period is determined by the size of the demand observation and the on-hand inventory level. Given the average market size is at least ξ, a small demand realization increases the supplier's confidence that the average market size is ξ, in the following period. This intuition explains the failure rate ordering in Part 2(i) Theorem 2.4.3. Low on-hand inventory level in addition to a small, uncensored demand realization implies that the supplier offers higher base-stock levels to larger markets in the following period. Increasing the optimal menu of base-stock levels in the following period, deters the retailer with larger market size ($\xi > t(z_n)$) from choosing base-stock level meant for a smaller market. The situation gets interesting when the on-hand inventory level is high enough that the supplier does not replenish in period n (i.e., $y_n = x_n$).[15] As the size of demand realization increases, the supplier's confidence that the average market size is ξ reduces in the following period. In the event of a large uncensored demand realization, the retailer's incentives mimic his incentives following a censored demand realization. The supplier, therefore, resorts to lowering the menu of base-stock levels to preserve incentive compatibility of the menu. Hence, censoring of demand, size of demand realization, and on-hand inventory level, impacts the suppliers knowledge about market conditions and the optimal contracts offered.

[15]Note that, $y_n = x_n \implies x_{n+1} \leq x_n$.

Symmetric Demand Information.

To understand how inventory decisions impact efficiency of screening contracts in a VMI relationship, we consider symmetric information setting here. In this setting, the supplier and the retailer have identical information about demand. That is, market signal, ξ is common knowledge, while demand realization in a period is still unknown to both players at the beginning of the period. In this case, the supplier faces the classical lost-sales inventory control problem (Karlin and Scarf, 1958). The optimal base-stock level for this problem is determined using the critical fractile, $S^{sb}(\xi) := G^{-1}\left(\frac{w-c}{w+h-\alpha c}\middle|\xi\right)$, which depends on the underlying market conditions, ξ. Thus, the wholesale price contract, commonly used within the VMI agreements, prohibits the supplier from coordinating the supply chain. In Theorem 2.4.4, we characterize linear base-stock level contracts that coordinate this decentralized supply chain and examine its properties.

Theorem 2.4.4. *Under symmetric demand information, the linear base-stock level contract with base-stock level,* $S^{fb}(\xi) := G^{-1}\left(\frac{r-c}{r+h-\alpha c}\middle|\xi\right)$ *and marginal price,* $p^{fb} := \frac{(r-w)(h+c(1-\alpha))}{r+h-\alpha c}$ *coordinates the channel. That is, this contract maximizes the total supply chain profit.*

Since the supplier has all the demand information she needs to make inventory decisions, she no longer learns from POS data, and contracts upon the coordinating base-stock level. The learn and screen approach in the symmetric case boils down to determining the optimal time to offer the linear contracts. Since the supplier can arbitrarily share the total supply chain profit, she does not lose any informational rent to the retailer. Therefore, postponing offering contracts does not benefit the supplier in terms of rent-extraction, and hence, offers the (coordinating) contract at the beginning of the planning horizon.

The Evolution of The Optimal Menu of Contracts.

Here, we show how inventory dynamics in a VMI agreement affects the structure of the optimal contracts offered in the learn and screen approach. As a benchmark we compare the optimal contracts with the linear coordinating contracts in the symmetric case.

Theorem 2.4.5. *We have the following properties.*

1. *The marginal price paid by the retailer is always greater than the first best price, i.e.,* $\frac{\mathrm{d}P_n^*(S_n^*)}{\mathrm{d}S_n^*} \geq p^{fb} = \frac{(r-w)(h+c(1-\alpha))}{r+h-\alpha c} > 0$, *for all* $\xi \in \Theta$.

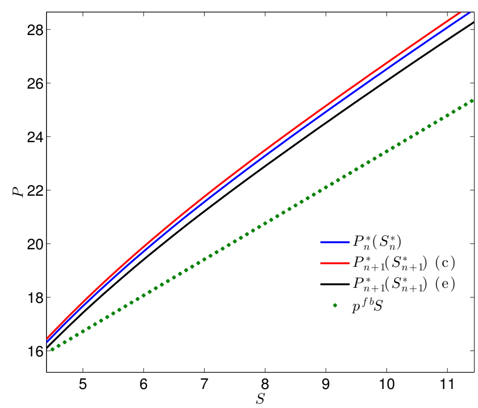
2. *If demand realization in period* n *is censored, the marginal price the retailer pays for the base-stock level increases in the following period, i.e.,*

$$\frac{\mathrm{d}P_{n+1}^*(S_{n+1}^*)}{\mathrm{d}S_{n+1}^*} \geq \frac{\mathrm{d}P_n^*(S_n^*)}{\mathrm{d}S_n^*}.$$

3. *If demand realization in period* n *is uncensored and* $x_n \leq \hat{S}_n(\underline{\xi})$, *then the retailer with larger market demand pays lower marginal price in the following period,*

$$\frac{\mathrm{d}P_n^*(S_n^*(\xi))}{\mathrm{d}S_n^*(\xi)} \geq \frac{\mathrm{d}P_{n+1}^*(S_{n+1}^*(\xi))}{\mathrm{d}S_{n+1}^*(\xi)}, \text{ for all } \xi \geq t(z_n).$$

Part 1 of the above theorem shows that the retailer pays higher marginal price than in the symmetric demand information setting, to convince the supplier about the credibility of his demand information. This monetary commitment on the part of the retailer assures credibility of the information he shares. Parts 2 and 3 show the impact of inventory dynamics on the marginal price. Censored POS information negatively impacts the retailer (and the supply chain). It results in the retailer choosing a lower base-stock level (Part 1 of Theorem 2.4.3) and paying more for the contract, in the following period. However, an uncensored, small demand realization benefits the retailer and the supply chain. In particular, the retailer pays a lower marginal price and reserves more suitable base-stock level, given his market conditions. Lower marginal price provides sufficient incentives to deter the retailer with larger market from choosing a low base-stock level. These observations corroborate the discussion of Theorem 2.4.3. Figure 2.1 illustrates dynamics of the structure of optimal contracts, as driven by inventory decisions and POS data. Theorems 2.4.3 and 2.4.5 describe the strategic



(c), (e) denote optimal contracts, following a period of censored or uncensored, demand realization in period n, respectively. Demand is exponentially distributed with mean ξ, $\Theta = [2, 15]$, and prior is inverse Gamma distribution. See §A.4.2 for details on belief updating.

Figure 2.1: Dynamics of Optimal Contracts

role of inventory decisions in an on-going VMI relationship. The inventory decisions prior to screening the retailer determine the supplier's future belief about demand and also the structure of the optimal contracts offered.

2.4.3 Timing of Contracts in VMI

Suppose that the supplier starts period n with inventory level, x_n and belief, π_n. The supplier has to decide between offering the optimal menu of contracts or continue learning via Bayesian updating. The value of exercising the first option is $\Pi_n^{sr}(x_n, \pi_n)$ and the value of continuing to learn is $\Pi_n^{lr}(x_n, \pi_n)$, defined in Equations (2.4) and (2.7), respectively. By delaying offering contracts, the supplier gives herself a chance to improve her knowledge about market conditions, contingent on her prior inventory decisions. If after many selling periods, the supplier is able to accurately estimate market conditions (i.e., there is no longer information asymmetry), she can extract all the informational rent by screening the retailer (see Theorem 2.4.4 for the coordinating contract). Thus, delaying has potential benefits in terms of greater rent extraction. However, until the contracts are offered, the supplier makes inventory decisions based on her limited knowledge of demand, which could potentially result in lost sales over several periods.

The above trade-off drives the supplier's inventory decisions and the timing of screening contracts. Accurately timing when to offer the optimal menu of contracts can mitigate the tension that may exist normally exist between firms participating in VMI agreements. Theorem 2.4.6 below provides provides a partial characterization of the optimal time to offer the contracts.

Theorem 2.4.6. *For $n \in \{1, \cdots, N\}$, we have the following.*

1. $\Pi_n^{sr}(x_n, \pi_n)$ *is decreasing in* x_n.

2. *If* $x_n > K_n$*, then* $\tau \geq n + 1$.

The on-hand inventory level serves as an indicator for the supplier in determining whether it's better to postpone offering contracts. With high on-hand inventory level, the supplier incurs holding cost for excess inventory regardless of screening the retailer. The supplier could therefore benefit from learning while inventory level is high, and eventually offering the menu of contracts when she is better informed about market conditions.

2.5 Value of Learn and Screen Approach

Here, we examine the value of three approaches the supplier could use to manage inventory and information in a VMI framework. (1) In the *learn* approach ($\tilde{V}_1^{lr}$), which represents the status quo, the supplier invests in statistically improving her demand forecasts using POS data, until the end of planning horizon. Therefore, she adds value to her operations by incorporating learning. Note that, $\tilde{V}_n^{lr}$ can be computed using the recursion in Equation (2.5) by setting $\tilde{\Pi}_n^{sr} = -\infty$ for $n = 1, \ldots, N$. (2) In the *screen* approach ($\tilde{V}_1^{sr}$), the supplier offers the optimal menu of contracts to the retailer at the beginning of the planning horizon. $\tilde{V}_1^{sr}$ can be computed by setting $\tilde{\Pi}_1^{lr} = -\infty$ in Equation (2.5). The screen approach quantifies the value of information sharing in a VMI setting. However, in this approach, the supplier undermines her ability to improve her forecasts over time. Note that in both of these approaches, the supplier's timing of contracts is determined ahead of the planning horizon. (3) In the *learn and screen* approach ($\tilde{V}_n$), the supplier dynamically evaluates on-hand inventory level and her belief about market conditions, to determine whether to offer or postpone the contracts. Thus, timing of contracts is a strategic decision in this approach. The learn and screen approach also quantifies: (i) the value $\left(\frac{\tilde{V}_1 - \tilde{V}_1^{lr}}{\tilde{V}_1^{lr}} \times 100\right)$, strategically screening the retailer adds to the status quo operations; and (ii) the value $\left(\frac{\tilde{V}_1 - \tilde{V}_1^{sr}}{\tilde{V}_1^{sr}} \times 100\right)$, dynamic learning adds to the screen approach. In addition, we also report the centralized supply chain profit ($\tilde{V}_n^{cs}$) as a benchmark. In the centralized setting, there is a single decision maker for the supply chain. This decision maker faces the classical lost-sales inventory problem with complete demand information. The optimal base-stock level for this case is characterized in Theorem 2.4.4 and the value function is computed using $\tilde{V}_n^{cs}(x, \pi) := cx + V_n^{cs}(x, \pi)$ and Equation (A.6).

To evaluate performance of these three inventory management approaches, we consider a two-point prior and exponential demand distribution. In particular, the supplier believes that average demand is either *high* ($\overline{\xi}$) or *low* ($\underline{\xi}$). Her initial prior is denoted by p, representing her subjective belief that demand is high in the ongoing season. The state space of the dynamic program in Equation (2.6) for this case has two variables in it—inventory level (x_n) and probability of high type demand (p_n). We direct the reader to §A.4.1 for more details.

We calibrate double marginalization (DM) and degree of information asymmetry (DIA) in the supply chain using $\frac{w-c}{r-c}$ and $\rho := \frac{\mathrm{Var}_\pi(D)}{\mathbb{E}_\pi[\mathrm{Var}(D|\xi)]}$, respectively. Smaller the ratio $\frac{w-c}{r-c}$, greater is the impact of double marginalization on the supply chain operations. With a small $\frac{w-c}{r-c}$, the supplier realizes a relatively smaller margin on each unit of the product sold, and hence, would maintain inventory level farther away from the supply chain coordinating inventory level. ρ is a measure of the variance in demand seen by the supplier relative to the expected variance in demand seen by the retailer. Since $\mathrm{Var}_\pi(D) = \mathrm{Var}_\pi(\mathbb{E}[D|\xi]) + \mathbb{E}_\pi\big[\mathrm{Var}(D|\xi)\big]$,

$$\rho := \frac{\mathrm{Var}_\pi(D)}{\mathbb{E}_\pi[\mathrm{Var}(D|\xi)]} = 1 + \frac{\mathrm{Var}_\pi(\mathbb{E}[D|\xi])}{\mathbb{E}_\pi[\mathrm{Var}(D|\xi)]} = 1 + \frac{\mathrm{Var}_\pi(\xi)}{\mathbb{E}_\pi[\xi^2]} = 1 + \frac{\mathbb{E}_\pi[\xi^2] - \mathbb{E}_\pi^2[\xi]}{\mathbb{E}_\pi[\xi^2]} = 2 - \frac{\mathbb{E}_\pi^2[\xi]}{\mathbb{E}_\pi[\xi^2]},$$

where the third equality follows from $D|\xi \sim \exp(\xi)$. From the above it follows that, $\rho \in [1, 2]$. When $\rho = 1$, the supplier and the retailer have the same demand information, while $\rho = 2$ corresponds to the highest degree of information asymmetry in the supply chain. We consider two (low and high) instances of ρ in Tables 2.1 and 2.2, representing (low and high, respectively) degree of information asymmetry, given the supplier's prior information about average demand.

The following parameter values remain fixed throughout this section: $r = 12$, $c = 3$, $h = 2$, $\alpha = 0.8$, $\Pi_r^{\min} = 0, \underline{\xi} = 2$, and $x_1 = 8$. The remaining parameters are varied as follows: $w = \{10, 6\}$, denoting low (L) and high (H) double marginalization; $\overline{\xi} = \{6, 12\}$, representing smaller (Table 2.1) and larger (Table 2.2) of demand variability; and $N = \{3, 5, 10\}$. We vary the supplier's initial prior to illustrate the affect of ρ on the supplier's expected profit.

In Tables 2.1 and 2.2, we highlight how double marginalization and degree of information asymmetry, together with the duration of the ongoing VMI agreement (short- vs long-term), plays a crucial role in determining the value of learning and strategically screening the retailer. Table 2.2 (where $\overline{\xi} = 12$) corresponds to a market in which, there is greater variation in the supplier's knowledge about average demand compared to Table 2.1 (where $\overline{\xi} = 6$). We note that degree of information asymmetry evolves endogenously in the VMI agreement, while

Table 2.1: Value of Learn and Screen Approach under Low Demand Variability

N	DM	DIA	$\tilde{V}_1^{lr}$	$\tilde{V}_1^{sr}$	$\tilde{V}_1$	$\tilde{V}_n^{cs}$	$\frac{\tilde{V}_1-\tilde{V}_1^{lr}}{\tilde{V}_1^{lr}}(\%)$	$\frac{\tilde{V}_1-\tilde{V}_1^{sr}}{\tilde{V}_1^{sr}}(\%)$
3	L	L	62.96	72.35	72.35	83.68	14.91	0
		H	38.04	43.33	44.07	52.10	15.85	1.70
	H	L	29.17	54.38	54.38	83.68	86.44	0
		H	15.40	37.34	37.34	52.10	142.39	0
5	L	L	78.08	90.61	93.18	107.27	19.34	2.84
		H	46.29	50.63	55.17	65.97	19.06	8.84
	H	L	33.09	65.85	70.23	107.27	112.24	6.65
		H	16.85	42.38	42.38	65.97	151.47	0
10	L	L	96.14	112.43	119.29	135.64	24.08	6.10
		H	56.68	59.35	70.06	83.05	23.61	18.04
	H	L	38.10	79.56	93.86	135.64	146.33	17.97
		H	19.63	48.40	52.69	83.05	168.43	8.86

$(\underline{\xi}, \overline{\xi}) = (2, 6)$; $\rho = 1.1$ and $\rho = 1.25$ correspond to initial priors $p = 0.75$ and $p = 0.25$, respectively.

Table 2.2: Value of Learn and Screen Approach under High Demand Variability

N	DM	DIA	$\tilde{V}_1^{lr}$	$\tilde{V}_1^{sr}$	$\tilde{V}_1$	$\tilde{V}_n^{cs}$	$\frac{\tilde{V}_1-\tilde{V}_1^{lr}}{\tilde{V}_1^{lr}}(\%)$	$\frac{\tilde{V}_1-\tilde{V}_1^{sr}}{\tilde{V}_1^{sr}}(\%)$
3	L	L	107.26	135.87	135.87	154.35	26.67	0
		H	37.46	46.45	46.45	56.37	24.02	0
	H	L	45.69	103.71	103.71	154.35	127.04	0
		H	14.31	41.16	41.16	56.37	187.54	0
5	L	L	139.44	178.12	178.12	204.20	27.74	0
		H	46.62	54.93	58.70	72.35	25.90	6.85
	H	L	54.58	133.82	145.82	204.20	167.19	8.97
		H	15.85	47.63	47.63	72.36	200.56	0
10	L	L	178.15	228.62	230.90	263.88	29.61	1.00
		H	58.56	65.07	75.86	92.01	29.54	16.59
	H	L	65.54	169.81	200.80	263.88	206.38	18.25
		H	19.03	55.38	60.34	92.01	217.09	8.95

$(\underline{\xi}, \overline{\xi}) = (2, 12)$; $\rho = 1.1$ and $\rho = 1.51$ correspond to initial priors $p = 0.85$ and $p = 0.14$, respectively.

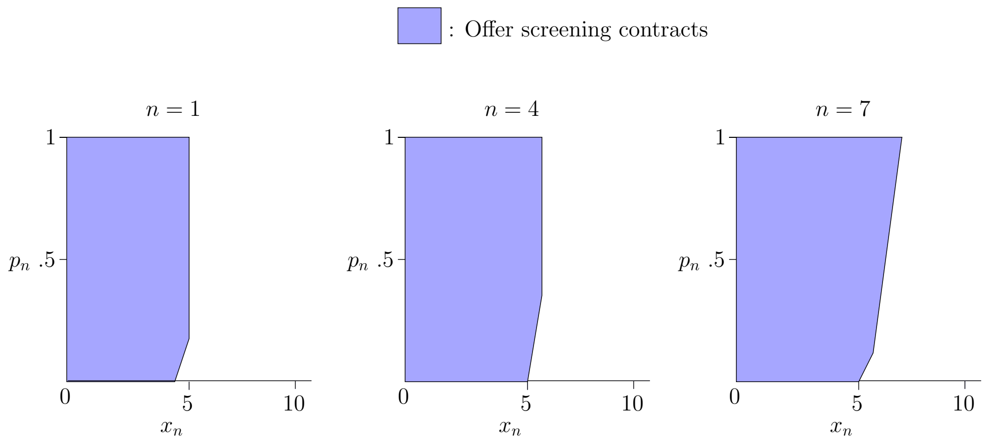
double marginalization, demand variability, and length of planning horizon are exogenously determined, ahead of the season.

In markets with high double marginalization, the optimal menu of contracts provides a way for the supplier to realize part of the retailer's margin. Thus, V_1^{sr} and V_1 are significantly higher than V_1^{lr} in these conditions. However, in the fast-moving consumer good (FMCG) retail industry (Sainsbury's example fits this segment), the retailer is squeezed for profits (see *Ernst & Young 2013*). The low double marginalization condition abstracts this industry. For these market conditions, we observe that quality of the supplier's prior information (ρ) and length of planning horizon for the VMI relationship (N), play an important role in determining value of strategically screening the retailer. When the planning horizon is short ($N = 3$), the supplier gains significantly (24% to 37%) over the learn approach by contracting with the retailer. The value of learning is more pronounced when ρ is higher. The supplier's profit using the screen approach, V_1^{sr}, is close (and in most cases equal) to the learn and screen approach, V_1, in such cases.

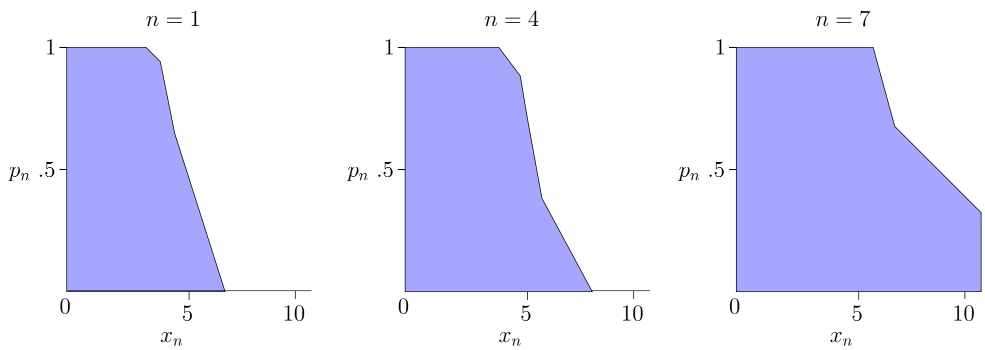
As the planning horizon for the VMI relationship gets longer, value of learning, V_1^{lr}, approaches V_1^{sr}. However, under high ρ conditions, V_1^{lr} approaches V_1^{sr} faster than compared to low ρ conditions. When ρ is high (low), percentage gain from offering the menu of contracts reduces from 37% to 8% (24% to 22.5%) as N increases from 3 to 10 (see Table 2.1). These observations suggest to the supplier that the quality of prior knowledge about demand plays a critical role in determining the value of learning versus screening.

The last two columns of Tables 2.1 and 2.2 quantify, the value strategically screening the retailer adds to the status quo operations and the value dynamic learning adds to the screening mechanism, respectively. The value of the former strongly dominates the latter, thus indicating to the supplier that unless the VMI agreements are of a long-term nature, it is in her best interest to offer the contracts early.

On the supplier's side (vendor's), entering into a VMI relationship entails setting up necessary IT infrastructure that enables sharing of POS information and also, reorganizing workforce to meet added responsibility of monitoring inventory. Thus, the decision to enter into a VMI agreement is typically a (long-term) strategic one. Therefore, long planning horizon ($N = 10$) and low double marginalization, best models the VMI environment. For precisely these market conditions, the learn and screen approach adds significant value to the supply chain operations. Comparing $\tilde{V}_1$ and $\tilde{V}_n^{cs}$ in Tables 2.1 and 2.2, we also note that the supplier makes upto 60% to 76% of centralized supply chain profit using the learn and screen approach.



(a) Contract offering region under low double marginalization and low demand variability.

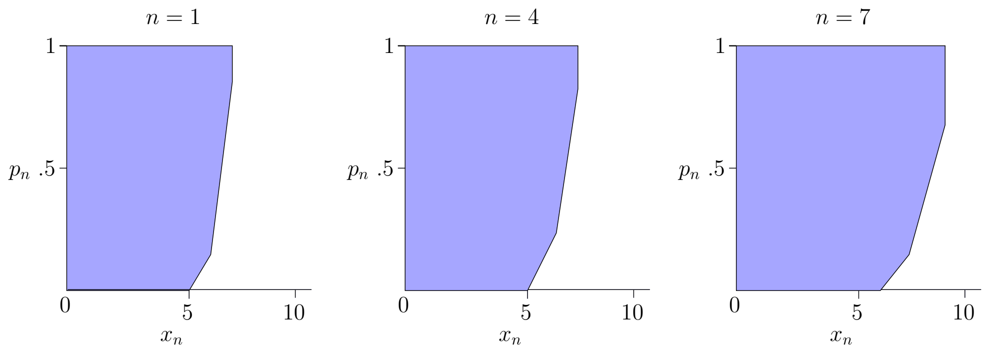


(b) Contract offering region under high double marginalization and low demand variability.

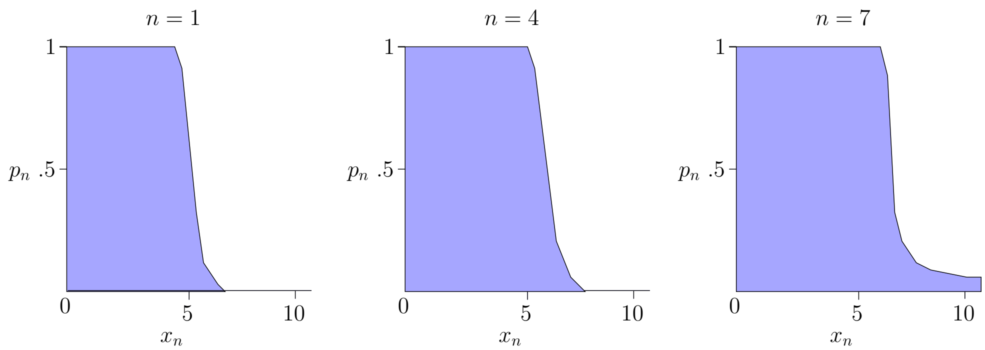
Figure 2.2: Optimal Timing of Screening Contracts in VMI

2.5.1 The Optimal Time to Offer Contracts

We examine the optimal time for the supplier to offer screening contracts across different market characteristics. Figures 2.2a–2.2b (market with low demand variability) and 2.3a–2.3b (market with high demand variability) illustrate the optimal regions to offer contracts, in periods 1, 4, and 7, when $N = 10$. Figures 2.2a and 2.3a (2.2b and 2.3b) represents the case when the extent of double marginalization in the market is low (high).



(a) Contract offering region under low double marginalization and high demand variability.



(b) Contract offering region under high double marginalization and high demand variability.

Figure 2.3: Optimal Timing of Screening Contracts in VMI

The supplier offers contracts if her on-hand inventory level, x_n and her belief about demand, p_n fall within the shaded region. A larger region indicates that there is a greater likelihood that contracts are offered in the current period. In all the illustrations, likelihood of the supplier offering contracts increases over time (un-shaded region shrinks). This increase is because a longer planning horizon lends greater value to investing in learning about market conditions. Hence, the value of postponing contracts reduces over time. In the last period, the value of learning is the least and hence, we observe that the supplier is always better off screening the retailer.

The figures also illustrate that for a given on-hand inventory level (resp., belief), the optimal stopping time has a state- and time-dependent threshold structure in the supplier's belief (resp., on-hand inventory level). At the beginning of each period, the supplier observes her inventory level (x_n), and if it is less than a certain threshold, assesses her belief about market conditions (p_n) to determine whether to offer or postpone contracts. Interestingly, the extent of double marginalization (Figures 2.2a and 2.2b; Figures 2.3a and 2.3b) plays an important role in determining if the threshold structure (in the supplier's belief), is lower or upper cut-off type, for a given inventory level.

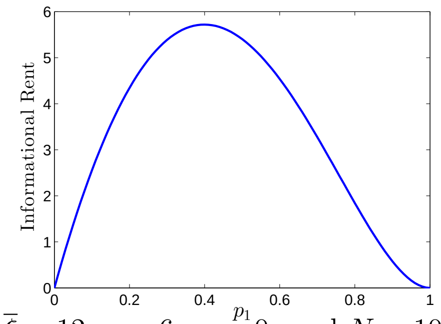
Next, we compare markets where the supplier faces different levels of variability in demand. Comparing Figures 2.2a and 2.3a, reveal that the supplier does not hesitate to offer contracts even when on-hand inventory level is high because the market size is potentially larger in 2.3a. Figures 2.2a–2.3b also illustrate that reducing on-hand inventory in the supply chain before seeking a quantity commitment from the retailer is beneficial for the supplier. This reduction serves two purposes. First, the supplier can use the existing on-hand inventory strategically to improve her knowledge about market conditions. Second, by offering contracts when inventory level is low, the supplier makes it easier for the retailer to commit to base-stock levels that are higher than on-hand inventory level. The retailer is assured inventory levels ex-post will not be smaller than the current on-hand inventory levels.

2.5.2 Role of Time in Rent-Extraction

The two-point prior considered in this section, also enables us to illustrate the role of time in rent-extraction in closed form. Specifically, we illustrate how much the supplier gains by improving her belief before offering contracts. We refer the reader to §A.4.1 for the analytical derivations supporting our discussion in this section. The supplier offers two contracts $(\underline{S}_n^*, \underline{P}_n^*)$ and $(\overline{S}_n^*, \overline{P}_n^*)$, if she decides to screen the retailer in period n. The informational rent (defined in Equation (2.8)) of the retailer in the two-point prior case simplifies to

$$\mathbb{E}_\xi[\Pi_n^r(\xi)] = \gamma(n)(r-w)p_n \int_0^{\underline{S}_n^*} \big(G(z|\underline{\xi}) - G(z|\overline{\xi})\big)\mathrm{d}z,$$

where $\underline{S}_n^* = \max\{x_n, \hat{\underline{S}}_n\}$ and $e^{-\frac{\hat{\underline{S}}_n}{\underline{\xi}}} - e^{-\frac{\hat{\underline{S}}}{\overline{\xi}}} \frac{p_n(r-w)}{p_n(r-w)+(1-p_n)(r+h-\alpha c)} = \frac{(1-p_n)(h+c(1-\alpha))}{p_n(r-w)+(1-p_n)(r+h-\alpha c)}$. By delaying to offer the contracts, the supplier potentially improves her knowledge about market



$\overline{\xi} = 12, w = 6, x_1 = 0,$ and $N = 10.$

Figure 2.4: Information Rent as a Function of the Supplier's Belief

conditions, i.e., p_n tends away from 0.5 (ρ tends to 1). Belief $p_n = 0.5$ essentially means the supplier does not have much information about market size. Figure 2.4 illustrates that lower degree of information asymmetry results in lower informational rent offered to the retailer.

2.6 Conclusion

We investigate the incentive issue that arises in collaborative supply chain agreements such as, vendor-managed inventory. The incentive issue arises from the fact that the supplier makes inventory decisions for the supply chain and also bears the cost of holding excess inventory. Thus, although the retailer has superior demand information, he only benefits from projecting higher demand to the supplier. To resolve the incentive issue, we propose and characterize a dynamic inventory-mechanism that enables the supplier to better manage inventory and information in the supply chain. In this mechanism, the supplier learns about market conditions via inventory decisions (and POS data) and then, offers a menu of screening contracts to the retailer. The retailer communicates his private demand information by choosing the base-stock level from the menu of contracts that is most appropriate, given the market conditions. In exchange for maintaining inventory at the mutually agreed upon level, the retailer pays a one-time fee.

Two elements of the VMI agreement that impact the outcome of the proposed approach are, the censored nature of the periodic POS data and the duration for which the VMI agreement is signed for. We show that both the quality (censored or not) and the magnitude of demand realizations play a key role in the evolution of the supplier's belief process. The supplier's updated belief process in turn determines the structure of the screening contracts offered. For example, censored demand observation exacerbates the retailer's incentive to report larger market size. Consequently, to ensure credible communication, the supplier lowers the menu of base-stock levels offered in the following period. The proposed mechanism, therefore, allows us to explore the dynamic interplay between inventory decisions and evolution of incentives in a VMI agreement. Thus, our findings highlight the strategic aspect of inventory management.

The learn and screen approach also emphasizes the importance of dynamically (i.e., depending on current on-hand inventory level and belief about demand) determining the *right* time to offer the screening contracts. Higher on-hand inventory levels and longer planning horizon, imply there is greater value to postponing the contracts. We illustrate how the threshold-type structure of optimal time to offer contracts depends on market characteristics such as, double marginalization and market-size variability. In addition, we also quantify the value of learning and the value of strategically screening the retailer in VMI agreements. We extend some of the findings in this chapter in Bensoussan et al. (2017), where the retailer's profit function, if he rejects the contract, is modeled endogenous to the system.

From the supply chain point-of-view, there are two compelling reasons to adopt VMI. First, there is a single inventory manager for the supply chain, closer to the upstream. Second, the inventory manager has access to the periodic point-of-sale information. Both these reasons significantly combat the well-documented Bullwhip effect in supply chains. However, there remain unresolved issues, such as credibly sharing demand information beyond POS data, between the supply chain members in VMI. As a result, important demand information that is observed locally at the retail store could be ignored by the VMI manager in making inventory decisions. This lack of credible communication between VMI members has lead to tensions and eventually, falling out of VMI. Learn and screen approach addressees this key issue and provides a channel for credible communication within an existing VMI agreement. From an operations view of implementing learn and screen approach, it can be noted that the VMI manager requires minimal additional monitoring effort to oversee the approach.

CHAPTER 3

VALUE OF DELAYED INCENTIVE: AN EMPIRICAL INVESTIGATION OF GIFT CARD PROMOTIONS

3.1 Introduction

Consumer sales promotions are widely used by retailers and manufacturers alike to stimulate short-term sales. In the last decade, retailers have significantly increased spending on sales promotions, thus underscoring the need to incentivize shoppers to spend more in a competitive environment. The nature and timing of the incentive, however, varies widely across different promotional mechanisms and has implications on the process of promotion planning, on how customers respond, and on total sales. Gift card promotion is a widely used promotional mechanism, which incentivizes customers to spend more than an expenditure threshold on regular priced products, by rewarding them with a (promotional) gift card[1] for future purchases. This is in stark contrast to discount promotions which offer instantaneous reward by lowering sales price. The potential to stimulate sales during the promotion and later, at the time of redemption, has led several major department and consumer electronic stores to regularly run gift card promotions (see Figure 3.1). The main objective of this study is to empirically test and quantify the benefits of gift card promotion. We achieve this objective by collaborating with a major fashion retailer, who regularly runs these promotions on their online channel.

The recent proliferation of gift card promotions can be attributed to its potential for having a sustained positive impact on the retailer's bottom-line. To participate in the promotion, customers spend beyond an expenditure threshold on regular priced products. For example, the gift card promotion on the bottom left in Figure 3.1 has multiple expenditure

[1]Gift cards in this study refer to promotional gift cards, i.e., the ones obtained through a promotion and not the gift cards which customers can buy at a retail store.

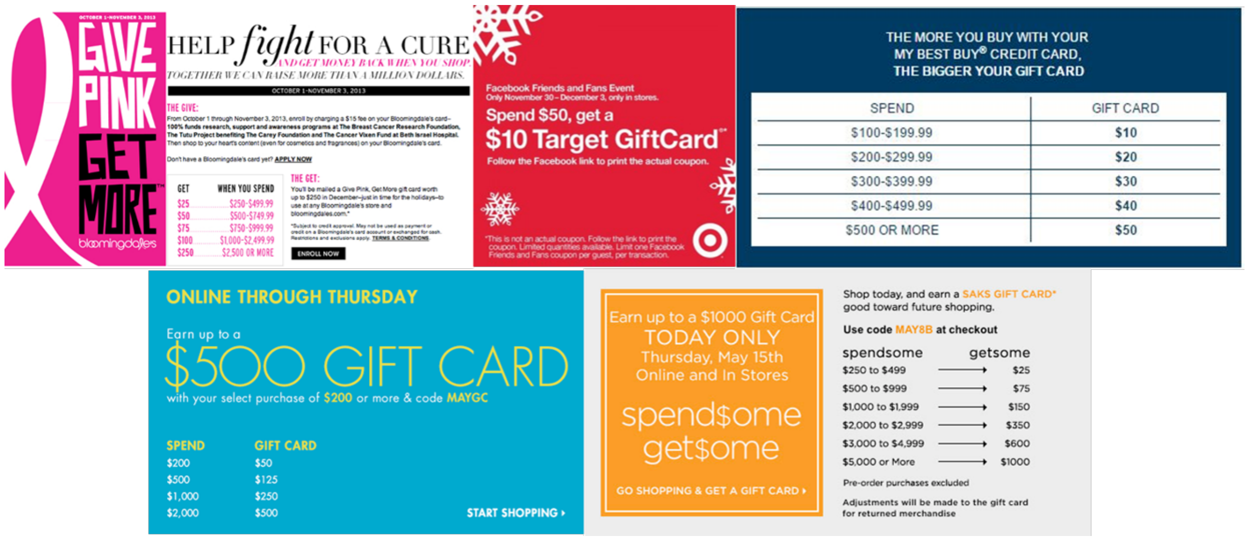


Figure 3.1: Gift card promotions offered at various department and consumer electronics stores.

thresholds set at \$200 through \$2000 to qualify for a gift card of face values ranging from \$50 through \$500. The thresholds are devised in a way to encourage customers to spend more to receive the incentive in the promotion. As a result, the threshold-effect of the promotion could potentially drive sales during the promotion. More importantly, these sales are realized at higher margins. This is pertinent, particularly in the fashion retail industry, where the average discount has gone up from 38% to 60% in the last decade. In 2012, then still J.C. Penney's CEO Ron Johnson, reported that the company was selling fewer than one out of every 500 items at regular price (Kapner, 2013). Apart from potentially boosting profit during the promotion, our conversations with executives at the partner fashion retailer indicate that selling at regular price also improves the customers' long-term perception of quality/value of the product. Hence the retailer/brand is more likely to sustain higher willingness to pay among its customers in the future. For the same reason, gift card promotion is also popular among manufacturers and hence, is easier to implement as a store-wide event by the retailer.

The temporal separation between the promotion and the redemption of the gift card earned generates, what is known as, *slippage*. Slippage refers to the gift cards that go unredeemed. Slippage directly contributes to the retailer's profit equation by lowering their promotion costs. The value of other delayed incentive promotions, such as mail-in rebates,[2] has primarily relied on slippage.[3] Although redemption rates of promotional gift cards have been observed to be higher than that of rebates, yet, close to 50% of gift cards go unredeemed (Long, 2015). Redemption rate at the major department store we collaborated with was around 47%, lending further support to the general observation.[4]

The delayed incentive aspect of a gift card promotion which is tied to the retail store can potentially generate incremental sales even after the promotion ends. This is because customers are required to revisit the store to make a purchase to redeem their gift card. While the retailer incurs cost equivalent to the value of the gift cards that are redeemed, some of this cost is recovered if the customers are induced to spend more than the value of the gift card. Evidence for such gift card-induced spending comes from popular press (Quinn, 2014). GiftCardGranny.com, an online marketplace for discounted gift cards, reports that close to 70% of the customers spend 38% more than the face value of a gift card. In summary, the general consensus among marketing managers is that redemption of promotional gift cards can be profitable regardless of whether the gift cards are redeemed or not.

Gift card promotion positively impacts the consumption utility of customers by providing an incentive to purchase products earlier in the season at regular price. From a rational economic theory standpoint, customer response to this incentive depends on the tradeoff between the net utility associated with potentially increasing their planned expenditure and the expected utility they derive from redemption of the gift card in the future.[5] Increasing

[2]Mail-in rebate is a sales promotion in which customers need to mail the rebate form along with their receipt to receive a rebate check in the future.

[3]Redemption rates for rebates have been observed to be close to zero (Lu and Moorthy, 2007).

[4]Promotional gift cards, unlike non-promotional gift cards (i.e., which can be directly bought at a store), are valid for a short time period (Federal Reserve System, 2010) This limited validity could also exacerbate the number of gift cards that are not redeemed.

[5]Since redemption of the gift card is uncertain, the expected value of a gift card is less than its face value.

expenditure to participate in the promotion reduces the customers' disposable budget in the short-term. However, purchasing a product earlier in the season enables fashion-sensitive customers to use it for a longer duration, thus boosting their consumption utility. Purchasing earlier in the season also alleviates risk associated with inventory unavailability if customers postponed their purchase to the markdown season. This is especially important for fashion or consumer electronic goods, which have limited initial inventory and replenishment lead times are typically long. For example, in the coats product category, Soysal and Krishnamurthi (2013) estimate that about 80% of the sales were accounted by fashion-sensitive buyers while the remainder were accounted by price-sensitive buyers. However, only 43% of total sales came at regular prices. This statistic suggests that, providing incentives (such as, for example, gift cards) could motivate more fashion-sensitive buyers to purchase at regular price.

Our main research objectives are threefold. First, we empirically test and quantify the impact of gift card promotion for the retailer, during the promotion. In particular we test whether customers indeed increase their expenditure to participate in the promotion rather than being given a gift card for a purchase they would have anyway made. The latter amounts to giving away free money. Likewise, when customers return to redeem their gift cards, we test whether in fact redemption is beneficial for the retailer. If customers spend close to the face value of the gift card, running a gift card promotion amounts to giving away a discount (albeit delayed in time) on a fashionable product during the season. Such an outcome can in fact hurt the retailer's sales because they lose a potential sale from a fashion-sensitive customer who might be willing to buy the product at regular price without the promotion. Second, we investigate if redemption of gift cards can induce customers to spend more. Finding evidence in support or to the contrary, has implications on the value of slippage in context of gift card promotion.

Third, while the main focus of our investigation is in measuring the effect of gift card promotion on those who participate and later, redeem, our empirical context of online retail also lends itself to studying the impact of advertising the promotion to customers using the e-mail and website channels. On the one hand, there is recent empirical evidence for strong advertisement effect of promotions resulting from mere exposure to the promotion in offline and online settings (Venkatesan and Farris, 2012; Sahni et al., 2016). On the other hand, researchers have also found an inverted-U shaped relationship between frequency of e-mail campaigns and purchase probability. For example, more frequent marketing e-mails has been shown to create a spamming effect (Kumar et al., 2014).

The data-related challenge in accurately measuring the impact of gift card promotion is to have access to disaggregate customer purchase data during the promotion (qualification and redemption stages) and during remainder of the year. Customer decisions during the qualification stage correspond to their purchase, participation (i.e., meeting requirements of the promotion), and their expenditure. Customer decisions during the redemption stage correspond to their redemption (i.e., using the gift card) and expenditure decisions. The difference between aggregate sales during and outside a gift card promotion provides a rough measure of the aggregate effect of the promotion. However, this approach can be misleading because customers might follow different strategies in the way they respond to the promotion. Customers could be displacing (postpone or accelerate) their purchase instance from outside the promotion to a promotion. Not accounting for such customer purchasing behavior could significantly overestimate the effect of the promotion. The other data-related challenge pertains to be able to match the customers' response during the promotion and their response during redemption.

We obtained a novel dataset from the online channel of a major U.S.-based department store, which regularly features in the Fortune 500 list of companies. Gift card promotions are a major promotional event for this retailer and are used only on their online channel.[6] Roughly a quarter of their total revenues are realized on their online channel and average

[6]Customers can qualify and later, redeem their gift card, only on the retailer's website.

daily sales during a gift card promotion is 43% greater than the average daily sales during the rest of the year. Since qualification and redemption stages can only occur on the retailer's website, the data captures the customers' actions and tradeoff fairly accurately. The dataset we build for analysis is comprehensive in that, it combines various aspects of online customer shopping process. In particular, we obtained four datasets from the retailer which provide details of: customer website browsing activity, customer sales, detailed description of e-mails sent by the retailer, and the gift cards customers qualified for.

We jointly model customer choice—purchase-participate during qualification stage and redeem or not during redemption stage—and resulting expenditure decisions using a limited dependent variable framework (Lee, 1983; Maddala, 1983). This is a flexible framework which accounts for the bias due to self-selection in the observed customer expenditure. This bias arises because customer choice underlying the observed expenditures is not random, but in fact, a result of (latent) utility maximization. For example, a customer who is more likely to return to the store would have greater utility of participation in the promotion, and hence, also more likely to spend more to participate in the promotion. However, we, as researchers, only observe the expenditure corresponding to the chosen alternative (with highest utility) rather than the customer decision making process. This interdependency is parametrized by specifying a joint distribution between the unobserved factors in the utility specification and the expenditure specification.

Our empirical analysis suggests that delayed incentive in the form of a gift card has a significant positive impact on customer expenditure during the qualification and redemption stages of the promotion. Customers who participate in a gift card promotion spend, on average, about 2.9 times that of customers who do not participate. But looking at the customer expenditure during the promotion in isolation can be misleading for two reasons. First, the promotion impacts baseline expenditure, as most customers are exposed to the promotion either through an e-mail, website banner ad, or other third-party websites. Second, only customers whose expenditure exceeds the threshold participate in the promotion, creating an upward bias in the observed expenditures of these customers. To overcome this, we first build an aggregate model combining the promotion and non-promotion periods. This model estimates that customer expenditure increases by 31.45% or $198.64, during a promotion. Combining this result with the promotion period analysis, puts the effect of promotion on customer expenditure in perspective. Of the $198.64, roughly 96.34% can be attributed to the increase in expenditure motivated by participation in the promotion and the remainder to the advertisement of the promotion. Therefore advertising products included in the promotion can attract greater attention from customers resulting in a purchase, even if they do not participate in the promotion. Likewise, gift card promotion boosts net revenues by increasing customer purchase probability by 17.54% during the promotion. However, unlike the effect of advertisement on expenditure, the positive effect on purchase probability is limited to customers who participate in the promotion. To capture the impact of redemption on customer response, we model the customer expenditure net off the gift card face value during the redemption stage of the promotion. We find strong evidence for gift card-induced spending. Customers who redeem their gift cards spend, on average, $525.28 more than the face value of the card.

Our response models also shed light on how gift card promotion impacts customer purchasing behavior in the context of fashion and apparel industry. We find that customer utility from their previous purchase initially increases with time and then continually drops. The point at which customer utility starts to drop, defined as the purchase cycle, is when customers are back on the market. This customer behavior manifests itself as a U-shaped relationship between purchase probability and time since their last purchase. Estimation of the customer choice models indicate that the purchase cycle of customers who choose to participate is lower than other customers. This observed self-selection suggests that gift card promotion can potentially differentiate customer population based on the inherent heterogeneity in customer responsiveness to purchasing products at regular price. This finding has important implications on the types of industries most suitable for gift card promotion.

Based on our extensive interactions with the customer analytics group at the retailer, we believe our analysis can be incorporated by managers at various stages of the promotion planning process. The customer response model we build can be used to better allocate funds for running gift card promotion; suppress marketing e-mails by targeting more responsive customers groups, and in better timing the promotions. Further, our finding that redemption of gift cards can in fact be beneficial for the retailer, suggests that reminding customers about their gift card through an e-mail, can boost the overall impact of the promotion. This finding and its implication, is contrary to the current practice. Our partner retailer does not send reminder e-mails for unredeemed gift cards, based on the general belief that redemption only increases promotion expenses.

The remainder of the chapter is structured as follows. Section 3.2 reviews extant literature; Sections 3.3 and 3.4 discuss the various stages in a gift card promotion and develop hypotheses about customer response at the qualification and the redemption stages. Section 3.5 provides details about the data. Section 3.6 presents the empirical model, estimation strategy, and the discussion of main results. We summarize our key findings for retailers and conclude in Section 3.8.

3.2 Literature

This study contributes in particular to the literature on sales promotion and more generally, to the field of pricing and revenue management. The extant empirical literature on sales promotions can be broadly categorized into two streams—building customer response models to quantify the impact of sales promotion (Chintagunta, 1993; Van Heerde et al., 2003) and using these response models to optimize various aspects including, budget allocation, timing, design, and targeting of the promotion planning process (Gönül and Hofstede, 2006; Khan et al., 2009; Ferreira et al., 2016; Cohen et al., 2017; Baardman et al., 2017). The emphasis in the first stream has been on isolating the impact of temporary price change during a promotion on what (brand-choice probability), when (purchase incidence probability), and how much (quantity) customers buy in the consumer packaged goods industry. We add to this literature in two ways. First, to the best of our knowledge, this is the first field study of quantifying the impact of delayed incentive promotion on customers response using field data. Gift card promotions are unique in that the price during the promotion remains unchanged and are generally implemented as a store-wide event. Second, we consider fashion and apparel industry, in which customers are known to exhibit significantly different consumption patterns compared to packaged goods. For example, packaged goods, unlike fashion products, can be bought in bulk and stored for future use and their consumption is relatively stable, enabling customers to stockpile or accelerate their purchase during the promotion.[7] As pointed out by Blattberg and Briesch (2012) and other recent review studies, there is limited empirical evidence of impact of sales promotions outside of consumer packaged goods industry.

Khouja et al. (2011) is the other study of gift card promotion, in which the authors analytically solve for the optimal design of gift card promotion. They show that gift card promotion mechanism enables customers with higher redemption probability or greater budget, to participate in the promotion. However, customer purchase decision is not modeled in the paper and the emphasis is on customer stockpiling, which, as discussed above, can be product category specific. Mail-in rebate is the other delayed incentive promotion which has received more attention in literature. The value of mail-in rebate has been attributed to the significant slippage observed in the retail industry. Various theories, such as post-consumption utility arbitrage (Chen et al., 2005), uncertain redemption costs (Lu and Moorthy, 2007), or present-biased preferences (Gilpatric, 2009) have been proposed to explain this phenomenon

[7] Senior management at the retailer we collaborated with agree that such customer responses have limited applicability in the context of our study.

and verified through analytical models. However, the value of gift card promotion does not necessarily arise due to slippage. In fact, as tested in the study, redemption can be profitable for the retailer if the gift card induces customers to spend more.

Sales promotions are typically accompanied by an advertisement on various channels, such as television, radio, e-mails, banner ads in newspaper or websites. Apart from communicating the promotion-specific information, research has shown that, promotional advertisements are also effective in reminding customers about the retailer/brand (Neslin, 2002; Venkatesan and Farris, 2012). In the context of online promotions, the e-mail channel has been widely used by retailers due to lower advertisement costs (Kumar et al., 2014). Recent empirical evidence suggests that the advertisement effect of communicating online discount promotion through an e-mail can dominate the effect of the promotion itself. For example, Sahni et al. (2016) find that the increase in average expenditure of customers who do not participate in the promotion accounts for over 90% of gains. In our empirical context, we find limited evidence for advertisement effect. Customers on average spend 4% more during the promotion period compared to non-promotion period, even when they do not participate. However, their purchase probability is not significantly affected compared to the non-promotion period. Therefore, our findings suggest that depending on the promotional mechanism and the retailer's e-mail marketing strategy, the magnitude of the advertisement effect may vary significantly.

The context of our study and our findings also have implications on the pricing and revenue management literature. Several papers in this literature consider the retailer's problem of pricing seasonal (fashion-like) goods to extract maximum customer surplus, with inventory considerations and endogenous demand. The classical view of sales promotion is that it increases customer surplus extraction by charging higher prices to less price-sensitive customers and temporarily lowering prices for the more price-sensitive customers (Varian, 1980; Narasimhan, 1984). This problem has been extended in various directions to incorporate customer purchase timing decision (Su, 2007; Cachon and Swinney, 2009, 2011; Soysal and Krishnamurthi, 2012) and behavioral regularities into the utility specification (Nasiry and Popescu, 2012; Özer and Zheng, 2016). Implicit in these models is the notion that customer's valuation (or reservation price) captures their price and fashion sensitivity. However, the latter stems from the consumption of the product rather than their budgetary constraints as the former does. While Soysal and Krishnamurthi (2012) model the time dependency of consumption utility, they consider it homogeneous in the population. Our findings suggest that gift card promotions can incentivize customer who have a shorter purchase cycle (or are more fashion-sensitive) to participate in the promotion by purchasing products at regular price.

3.3 Research Setting and Stages of Gift Card Promotion

We collaborated with is a U.S.-based department store that sells fashion and apparel products, such as men's and women's clothing, and accessories (see Figure 3.2 for the top selling categories). The store is often featured in the Fortune 500 list of companies and its name is kept confidential due to a non-disclosure agreement. Henceforth, we will refer to this store as the retailer. The retailer operates both brick-and-mortar outlets and an online channel. Our focus is on the online channel, on which the retailer runs roughly 17 gift card promotions annually in the U.S.. These promotions last on average 3 days. Gift card promotion is an important sales event for the retailer as is evidenced by sales spike during the promotion in Figure 3.3 and are used periodically throughout the year.

To implement gift card promotion the retailer needs to make long-term decisions, such as budget allocation and scheduling of the promotion, ahead of time. The promotion budget accounts for the redemption costs. There are roughly one or two gift card promotions each month and this has been the practice over the period of data we collected.

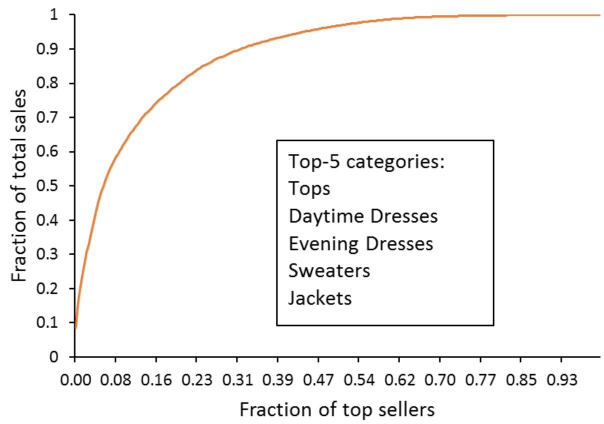


Figure 3.2: Percentage of sales corresponding to top-selling items on the website during the period 2012–2015.

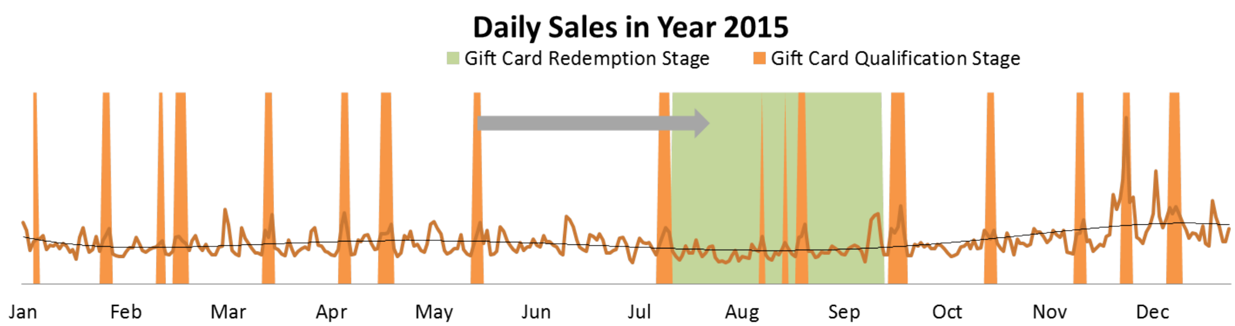


Figure 3.3: Daily sales during the year 2015.

There are three types of gift card promotions implemented by the retailer. Private gift card promotions are advertised only through an e-mail. Public gift card promotions are such that, in addition to an e-mail, there is also a banner-ad on the website advertising the promotion. Hybrid gift card promotion start out as private but are treated as public after a short duration. The gift card promotion e-mail for all types of gift card promotions is sent to customers on the morning of the promotion. Prior to this e-mail, the retailer does not make any announcements of an upcoming promotion to its customers. Therefore, any customer who accesses the website during the promotion can become aware of a public or hybrid gift card promotion as opposed to a private gift card promotion, which only those customers that receive the e-mail can become aware of. In addition, the website highlights all the products that are included in the promotion. This includes most, if not all, products that are not on sale. Note also that, only customers who visit the website through the e-mail would be exposed to the highlighted products included in a private gift card promotion.

The retailer obtains customer e-mails when they register on the website. The registration process also allows customers to choose the frequency with which they receive marketing e-mails from the retailer. This frequency can range anywhere between zero–one e-mail per week to two–three per day. Note however that, timing of the marketing e-mails is determined by the retailer. All customers that click on any marketing e-mail from the retailer in the previous year receive the gift card promotion e-mail for public and hybrid events.[8] Customers who also made a purchase in the past year, receive the promotion e-mail for private events.

Customers who qualify for a gift card during a promotion receive an e-mail, with the gift card code in it, eight weeks[9] after the promotion ends. The gift cards are valid for a period of three months after which the codes expire.

The customer response (decisions) can be classified into two stages: qualification and redemption, which are described in more detail below.

3.3.1 Qualification Stage

The qualification stage begins with the retailer advertising[10] promotion-specific information to its customers using an e-mail or via website banner ads (depending on the type of the promotion). Figure 3.1 in the introduction illustrates promotion-specific information contained in those e-mails or the website banner-ads. The e-mail subject line typically reads, "Up to

[8]Customers who opt-out of receiving marketing e-mails from the retailer are excluded from the list.

[9]Eight weeks is also the return period employed by the retailer. This delay prevents customers from gaming the promotion—by returning the product after qualifying for a gift card.

[10]Sometimes this information can also be obtained from third-party websites. For example, Retailmenot.com, a U.S.-based internet company, aggregates coupons, discounts, gift card promotions, and other types of promotional offers from various retailers.

a \$500 Gift Card", which indicates the highest face value of a gift card that can be earned in the promotion. The advertisement also provides information about different expenditure thresholds and corresponding face values of the gift card that can be earned. There are two actions that customers have to take to qualify (participate in) for the promotion. First, the customer has to spend more than the expenditure threshold on regular priced items. Second, the customer needs to enter a promotion code at the checkout (such as "GETGC"). Because of the second action, not all customers who spend more than the expenditure threshold on regular priced items automatically qualify for a gift card. The promotion advertisement also specifies that the customer would receive the gift card in an e-mail about eight weeks after the promotion ends. Apart from a select few brands, gift card promotion includes all regular-priced products on the website. All products that are included in the promotion are marked as such. We were assured by the retailer that exclusions, if any, were negligible. In other words, gift card promotions are a website-wide event for the retailer.

3.3.2 Redemption Stage

The redemption stage lasts for a three-month period during which the gift cards can be used by the customer. Qualified customers receive a gift card code in an e-mail eight weeks after the promotion ends. This e-mail subject line typically reads "Your Gift Card is Here". All the customers who qualify for a gift card receive this e-mail on the same day. Not all customers who qualify for a gift card redeem them during the redemption stage. Among others reasons, the slippage can be explained by the forgetful nature of customers or because they loose track of the e-mail (with gift card code) among other marketing e-mails received from the retailer.

3.4 Theory and Hypotheses

Rational economic theory posits customers as expected utility maximizers who determine their purchase and expenditure decisions based on their net utility of consumption subject to a budget constraint (Mas-Colell et al., 1995). Gift card promotion impacts the net utility by providing customers a delayed incentive in the form of a gift card. Customer utility from participation in the gift card promotion arises from the consumption of the purchased product and from the perceived value of the gift card to be redeemed in the future. The disutility is tied to the fact that customers have to spend more than the expenditure thresholds on regular priced products during the qualification stage and remain cognizant, of the gift card that is received in the future and of the economic tradeoffs made while qualifying for it. In the following two sub-sections we derive implications of offering customers such a delayed incentive, on their purchasing behavior during qualification and redemption stages of the promotion, in the context of fashion and apparel (F&A) industry.

3.4.1 Impact during Qualification Stage

The vast literature on sales promotions extensively documents that customers alter their purchasing behavior in a promotion by taking into account the net utility they would derive from their current and future consumption (Gupta, 1988; Bell et al., 1999; Van Heerde et al., 2003; Su, 2010). In the analysis of coupons in the consumer packaged goods (CPG) industry (such as, yogurt, coffee, etc.), research has shown three types of customer response accounting for the sales bump during the promotion. In response to temporary discounts, customers tend to stockpile for future consumption, accelerate their future purchase, or switch to the promoted brand. As a result, most studies report increased sales during the promotion. In contrast, the price of a product included in the gift card promotion does not change and most, if not all brands, are included in the promotion. The context of F&A industry also differs from the CPG industry in two important respects. Unlike CPG products, fashion products are seasonal with shorter lifecycle, have unstable demand, and cannot be inventoried for future use (Su, 2010). Therefore, traditional view of sales promotion, as a price discriminatory mechanism, cannot explain customer response to gift card promotion. To understand the impact of gift card promotion, we first consider the pricing strategy adopted by retailers in the F&A industry. Retailers markdown prices towards the end of season for a product, to incentivize price-sensitive customers to make a purchase. Forward-looking customers anticipate these markdowns and hence, typically postpone their purchase to the end of season. Such customer response has been widely documented in the fashion industry and has been central to the pricing and revenue management models. However, customers tradeoff several risks by delaying their purchase to the markdown season. For example, the product might go out of stock before the markdown season or might not be marked down enough. In addition, markdown season also signals fading out of past trends. This implies that, customers have a limited time post-markdown, to consume the product before, eventually, fashion trends change (Soysal and Krishnamurthi, 2013). Gift card promotion positively impacts this tradeoff for customers who derive greater utility from consuming the product during regular season[11] by providing a gift card. Therefore, we expect that the average purchase probability of customers during the promotion to be higher.

Hypothesis 1. *Customers are more likely to make a purchase during gift card promotion period compared to non-promotion period.*

In addition to purchasing products at regular price, customers also need to spend beyond an expenditure threshold to qualify for a gift card. The tiered structure of the expenditure thresholds has important implications on the customer expenditure. We can analyze the two possible outcomes by considering a representative customer who decides to make a purchase during the promotion. First, suppose that the customer's intended expenditure on regular-priced products already exceeds the lowest threshold. This customers already qualifies for a gift card without increasing their expenditure. However, the tiered structure means

[11]Products part of the gift card promotion are considered to be in a regular season because these products are offered at regular price even outside the promotion.

that relative value of a gift card is maximized when the realized expenditure is slightly above an expenditure threshold. This implies that customers can reduce their expenditure to the nearest lower threshold, or spend more to exceed the nearest greater threshold.[12] By reducing their expenditure, customers might allocate their expenditure between the qualification and redemption stages. Since, the redemption stage is two months out, customers have to sacrifice on their current consumption utility and face risks associated with potential stockouts as discussed above. As a result, benefit associated lowering expenditure to participate in a gift card promotion—while it exists—also lowers the utility obtained from participation. Therefore, we expect this effect to have limited validity in our context. The other alternative for this customer is to increase intended expenditure, to the nearest greater threshold, to participate in the promotion. This option allows the customer to not compromise on their current consumption utility while participating in the promotion. Therefore, for a customer who already exceeds the lowest expenditure threshold, we expect the customer to either increase their expenditure or leave it unchanged to participate in the promotion. On average, this translates to a positive effect on the customer expenditure.

Next, consider the case where the customer's intended expenditure is lower than the lowest expenditure threshold. This customer can qualify for a gift card by increasing their expenditure. For customers who value buying products at regular price, this creates a tradeoff. Increasing their expenditure allows them to increase consumption utility by purchasing a product which is still in fashion. Since most brands are included in the promotion run by the retailer, the promotion could also incentivize customers who are brand conscious to participate in it. Therefore, for such customers, we expect them to increase their expenditure to participate in the promotion. Based on a stockpiling argument, Khouja et al. (2011) analytically show that there exists a *gift card pull*. That is, when the intended expenditure of

[12]For example, suppose the promotion specifies the following two expenditure thresholds and corresponding gift card face values: \$100 – \$10 and \$200 – \$20. For a customer who is intending to spend \$190, the relative value of the gift card increases from 5.2% to 10% by spending \$10 more or by spending \$90 less.

customer is lower than an expenditure threshold, but greater than a level, they increase their expenditure to qualify for a gift card.

Creating purchasing thresholds to induce customers to increase expenditure has also been implemented in other promotional contexts, such as free-shipping offers.[13] Stevens and Banjo (2014) point out that free-shipping offer with expenditure threshold have become commonplace in retail industry and has resulted in customers increasing their expenditure to qualify for the offer. For example, shoppers spent an average $124 more to avail the free shipping offer according to the Japan-based online retailer, Rakuten. Therefore, from a rational economic theory standpoint and anecdotal evidence, we expect that the average customer expenditure during a gift card promotion increases due to participation in the promotion.

Hypothesis 2. *Customers spend more during gift card promotion period compared to non-promotion period and the incremental expenditure can be primarily attributed to participation in the promotion.*

The arguments presented in support of Hypotheses 1 and 2 address the scenario in which customers are more likely to make a purchase and increase their expenditure to participate in the promotion. An alternate argument that could explain the increased likelihood of purchase, can be attributed to the advertisement effect of the promotion. In particular, the advertisement of the promotion serves as a reminder to the customer about the store and also highlights products that are part of the promotion. This communication about the promotion happens through a gift card promotion e-mail, banner ad on the website (for public and hybrid promotions), and through third-party blogs or social media website, which may also carry the advertisement. These different channels of advertisements could nudge the

[13]For example, Ann Taylor and Ralph Lauren have a minimum expenditure threshold at $175 and $195, respectively, to become eligible for free shipping offer. The average online shopper has to be $82 to qualify for free shipping.

customer to visit the website, even when they are not inherently interested in participating in the promotion. The website also highlights products which are part of the promotion by clearly marking them. Note that, this includes most, if not all, products, which are not on sale. Highlighting such products can also draw greater customer attention. There is recent evidence pointing to dominant advertisement effect of online discount promotions. On an online ticket resale platform, Sahni et al. (2016) find that around 90% of the total value of the promotion can be attributed to the advertisement effect.

However, magnitude of the advertisement effect of the gift card promotion may also depend on the marketing strategy adopted by the retailer. In our context, customers registered on the website typically choose the volume of marketing e-mails they would like to receive from the retailer. Therefore, this could range anywhere between one per week to all marketing e-mails. Note that, customers are not informed about the maximum number of e-mails at the registration, but based on our conversations with the retailer it is about two–three e-mails per day. Customers who opted in to receiving more e-mails may be subject to potential wearout effect, resulting in customers ignoring e-mails including the promotion e-mail, from the retailer (Bass et al., 2007). Likewise, Kumar et al. (2014) find that the number of marketing e-mails the customer receives has a U-shaped effect on customer opt-out probability. Therefore, while the advertisement effect can be heterogeneous in the population, we expect an overall positive effect on customer purchase and expenditure decisions during the promotion.

Hypothesis 3. *(a) Customers are more likely to make a purchase during gift card promotion period compared to a non-promotion period, even when they do not participate in the promotion.*

(b) Customers spend more during a gift card promotion period compared to a non-promotion period, even when they do not participate in the promotion.

3.4.2 Impact of Delayed Incentive on the Redemption Stage

In addition to the immediate boost in consumption utility, participation in a gift card promotion also rewards customers with a gift card which they could redeem during the redemption stage of the promotion. However, given the temporal separation between the qualification and the redemption stages, customers have only a belief of their future actions. This delayed aspect of the incentive may create an inconsistency between what customers intend to do and what they actually end up doing, during the redemption stage.

The inconsistency between the customer's expectation and the realized outcome, leading to lower redemption rates or slippage, can be particularly advantageous for the retailer. Lower redemption rates lower the cost of implementing the promotion. Retailers have generally used arguments based on low redemption rates to endorse delayed incentive promotions. For example, low redemption rates is a characteristic of mail-in rebate promotion, another popular delayed incentive promotion.

The value of gift card promotion for the retailer, however, does not merely rely on low redemption rates. Customers have to make a purchase at the retail store during the redemption stage to redeem their gift card. If customers spend less than the face value of the gift card, the retailer incurs an equivalent cost. If, however, the gift card induces customers to spend more than the face value, some of the additional sales realized can offset the cost incurred by the retailer (equivalent to the face value of the gift card). Therefore, the value of the gift card promotion during the redemption stage, depends on whether it is able to induce customers to make additional expenditure during the redemption stage. The key to answering this question lies in understanding how customers might perceive the gift card during the redemption stage.

To assess the value of the gift card, customers have to recall the expenditure, and in particular, of any incremental expenditure (if at all), they incurred to qualify for the gift card. Given the temporal separation between the qualification and redemption stages, existing literature suggests otherwise. First, customers tend to recall the product bought rather than the expense incurred (Dickson and Sawyer, 1990). Second, the disutility associated with any incremental expenditure during the qualification stage diminishes over time (see Soman 2001 and references therein). These observations suggest that although customers might be intent on redeeming their (promotional) gift card, they are less likely to recall the tradeoff made during the qualification stage. This slip in memory may influence customers to treat the gift card as a non-promotional gift card, leading to a splurge in their spending during the redemption stage. Online gift card marketplaces, such as, Giftcards.com[14] and Giftcardgranny.com,[15] report that customers tend to overspend the face value of gift cards by at least 20%. Based on the arguments presented above, we hypothesize that customers spend more than the face value of the gift card if they redeem them.

Hypothesis 4. *Qualified customers who redeem their gift card, spend more than the face value of the gift card during the redemption stage of the promotion.*

3.5 Data

We obtained a collection of datasets from a major department store's (retailer) online platform, detailing the various aspects of the customer shopping process and engagement with the retailer (see § 3.3 for a description of the retailer). The collection constitutes several interrelated datasets corresponding to customer's, website activity, sales, e-mail communication (with the retailer) and qualification of gift cards. We were also provided a comprehensive promotion calendar indicating the days on which a gift card promotion was held. All of these datasets are identifiable at customer-level (using their e-mail[16]). The

[14]Retrieved from https://www.giftcards.com/gcgf/discount-giftcard-fraud on September 21, 2017.

[15]Retrieved from https://www.giftcardgranny.com/statistics/ on September 21, 2017.

[16]The customer e-mail ids were encrypted to protect personally identifiable information.

datasets span the period between January 2012 and December 2015. To avoid potential data discrepancy issue in the website dataset (as communicated by the retailer), we used the data since July 2012 for our empirical analysis. We provide the details of these datasets next.

Website activity. This dataset includes, customer's website activity information, such as the date of their website visit, source website which led them to the retailer's website, various webpages (category, product, checkout, purchase) the customer visited, and in what sequence. This dataset also includes information about whether these customers entered a promotion code during gift card promotional events.

Sales. This dataset contains information about purchases made on the retailer's website, including details about the product, quantity bought, selling price, and the payment method (Visa, MasterCard, Amex, store credit card) that was used at the checkout. The payment method field also indicates if a customer redeemed a promotional gift card during the checkout. However, the dataset does not indicate the promotion in which the redeemed gift card was earned.

E-mail communication. This dataset records all e-mail communication by the retailer with its customers. Specifically, this dataset includes, the date on which the retailer sent out the e-mail, the subject line, the set of customer email-ids it was sent to and what actions the customers took (opened, clicked, or opted out). Based on the subject line, these e-mails were categorized into three—gift card promotion related, other promotion related, and advertisement e-mails. Gift card promotion e-mails comprises of e-mails that inform customers about a gift card promotion event. Other promotion e-mails inform customers about discount offers or any other sales event. Advertisement e-mails include e-mails that advertise a specific brand, new arrivals or e-mails confirming customer purchases.

Promotion information. For each gift card promotion, this dataset specifies the expenditure thresholds, corresponding gift card face values, and the dates of the promotion.

Qualified customers. This dataset contains the list of customers who qualified for a promotional gift card, identifier of the gift card promotion in which it was earned, and the face value of the gift card.

3.5.1 Key Independent Variables

To measure the impact of the promotion on customer purchasing behavior, it is important to isolate the affects resulting from customer's past purchases, website activity, and the marketing efforts of the retailer. These variables capture latent customer consumption patterns driven by their budgetary considerations and their search costs associated with browsing activity. For example, customers less constrained by budgets are more likely to spend more on the website, but could have higher search costs, resulting in infrequent website visits. Such heterogeneity at customer-level is captured using independent variables which we classify as customer, promotion, and communication characteristics.

We divide the four years, for which the data was obtained (01/2012–12/2015), into contiguous periods, $t = 1, \ldots, 77$; that corresponds to the days of gift card promotion (indicated $\mathsf{Promo}_t = 1$) and the days between two gift card promotions (indicated $\mathsf{Promo}_t = 0$). The number of days in each period is indicated by $\mathsf{Period\_Length}_t$. The average promotion period lasts 2.8 days and average non-promotion period lasts 18.4 days. For each promotion, we also determine the redemption stage that starts eight weeks after the promotion period ends and lasts for three months. Note that, the redemption stage of a promotion could overlap with multiple promotion and non-promotion periods.

The customer and communication characteristics are computed in a way that they accommodate annual (such as, holiday) shoppers while at the same time are also recent enough to potentially explain current purchasing behavior. Therefore, the customer's activity in the 13 months prior to each promotion and non-promotion period is used to calibrate the independent variables. The analytics team at the retailer also considered 13- instead of a 12-month window more accurate, since some holiday shoppers, for example, shop at the beginning of December one year, followed by shopping at the end of December next year. Since we need 13 months historical data to initialize our independent variables, we considered the period between July 2012 and September 2013 for initialization and the period between September 2013 and December 2015 (38 promotion periods) for estimation. For each customer and time period (promotion or non-promotion), we compute the customer and communication characteristics.

Customer characteristics (CustomerChar). This category includes customer purchasing and website visit characteristics, such as the number of days since their last purchase (Purchase_rec), the number of days on which they made a purchase (Purchase_freq), and their total expenditure (Exp_annual) are computed.[17] Likewise, the number of days since the customer's last website visit (Web_rec) and the number of days on which they made a visit (Web_freq) capture their website activity. We aggregate frequency measures at a daily level. That is, all purchases and website visits on the same day are counted as one. We also note that the recency measures are upper bounded by 13-month (396 days). We infer customer choice of marketing e-mail frequency from the number of e-mails they receive from the retailer in the past two weeks. The categorical variable (Email_cat) with three levels (low, medium, high) identifies their preference. The low category includes customers who received less than one e-mail per week, medium category includes customers who receive between one and seven e-mails per week. The rest are categorized as high level. This categorization is based on the information that the retailer sends two or three marketing e-mails to customers who opted to receive all e-mails from the retailer. The low category also includes customers who opted out of the e-mail program.

Communication characteristics (CommChar). This category includes recency of the marketing e-mails received by customers prior to a period. The marketing e-mails are classified

[17]We also computed magnitude of their last purchases. However, it turned out to be highly correlated with their annual expenditure, and hence was excluded in the analysis.

as related to gift card promotions ($\mathsf{GCPromo\_}$), related to other promotions ($\mathsf{Promo\_}$), and the remaining advertising communication ($\mathsf{Comm\_}$). The gift card promotion related e-mails inform customers about the promotion (including the expenditure thresholds and gift card face values) on the first day of that promotion and sometimes a reminder on the last day. Other promotions that are frequently run (but do not overlap with the gift card promotions) by the retailer are discount events. Recency measures for each category of the e-mails received by the customer are computed as described above. In addition to these disaggregate measures, we also define the following aggregate measure.[18]

$$\mathsf{Email\_sent\_rec} = \min\left\{\mathsf{GCPromo\_email\_rec}, \mathsf{Promo\_email\_rec}, \mathsf{Comm\_email\_rec}\right\}.$$

Promotion characteristics ($\mathsf{PromotionChar}$). For each promotion period, t, this category includes the lowest expenditure threshold ($\mathsf{Lowest\_threshold}_t$), the (perceived) discount,[19] defined as

$$\mathsf{Discount}_t = \frac{\mathsf{Lowest\_face\_value}_t}{\mathsf{Lowest\_threshold}_t} \times 100,$$

number of thresholds in the promotion ($\mathsf{Nbr\_tiers}_t$), and the type of the promotion ($\mathsf{Type}_t \in \{\mathsf{Public}, \mathsf{Private}, \mathsf{Hybrid}\}$). Table 3.1 presents the summary statistics and correlations between these variables.

Gift card promotion e-mail ($\mathsf{GC\_email}$). Registered customers who opted to receive e-mails from the retailer receive the gift card promotion e-mail depending on the type of the promotion (as described in § 3.3). We use this variable to identify customers who received the e-mail for a gift card promotion.

Other controls. The category ($\mathsf{OtherControls}$) include month (Month_t), year (Year_t). In addition to capture the markdown season, we use an indicator variable ($\mathsf{Markdown}$). The

[18]We note that the aggregate e-mail variables are highly correlated with the other promotion e-mail variables due to the frequently held discount events on the website.

[19]For some gift card promotions the discount varies with the expenditure thresholds, in which case we use the highest perceived discount among all the thresholds.

Table 3.1: Summary statistics and correlations between promotion characteristics.

	Mean	Std. Dev.	(1)	(2)	(3)	(4)
Discount	18.05	6.09	1			
Nbr_tiers	4.12	0.82	-0.7231***	1		
ln(Lowest_threshold)	5.71	0.40	-0.1455***	-0.2728***	1	
Period_Length	3.10	0.85	-0.065***	-0.034***	-0.4071***	1
Email_cat (low)	0.45	0.50				
Email_cat (med)	0.05	0.22				
Email_cat (high)	0.50	0.50				
GC_email	0.63	0.48				
Private	0.14	0.35				
Public	0.72	0.45				
Hybrid	0.13	0.34				

*** $p < 0.001$

Note: $N = 471,852$.

markdown season occurs between June–July (Spring) and November–January (Fall). Note that, we omit monthly control while including the markdown variable to cleanly capture the effect of markdowns on customer response.

For the analysis we considered a random sample of 21,767 customers who were monitored over the 77 time periods. Table 3.2 summarizes their characteristics during the 400 days prior to the start of a promotion or a non-promotion period. Notably, there is marginally higher purchase and website activity for these customers prior to the start of a promotion period compared to the start of a non-promotion period. The correlations between these variables measured using the entire data and promotion-period only are presented in Tables 3.3 and 3.4, respectively. The correlations have identical signs and similar magnitudes between the promotion-period only and aggregate data.

3.5.2 Dependent Variables

The dependent variables during promotion and non-promotion periods are: Purchase (0/1), indicating whether customers made a purchase during the (non-) promotion period, and

Table 3.2: Summary statistics of key independent variables.

	Non-Promotion		Promotion		Aggregate	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Purchase_rec	220.39	147.22	214.68	145.65	217.82	146.54
Purchase_freq	1.79	3.41	1.85	3.42	1.82	3.42
ln(Exp_annual)	4.34	2.91	4.50	2.87	4.42	2.89
Web_rec	149.69	147.58	141.63	143.51	146.05	145.81
Web_freq	18.51	36.10	19.60	37.01	19.00	36.52
Email_sent_rec	116.21	167.67	94.61	156.59	106.46	163.12

Table 3.3: Correlation between key independent variables for the aggregate data.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Purchase_rec	1						
Purchase_freq	-0.4197***	1					
ln(Exp_annual)	-0.7532***	0.5088***	1				
Web_rec	0.6419***	-0.3051***	-0.5079***	1			
Web_freq	-0.2617***	0.5764***	0.3256***	-0.3963***	1		
Email_sent_rec	0.0554***	-0.0193***	0.0112***	0.0963***	-0.1218***	-0.6525***	1

*** $p < 0.001$

Note: $N = 1,04,405$.

Table 3.4: Correlation between key independent variables for the promotion period data.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Purchase_rec	1						
Purchase_freq	-0.4190***	1					
ln(Exp_annual)	-0.7476***	0.5089***	1				
Web_rec	0.6337***	-0.3017***	-0.4934***	1			
Web_freq	-0.2592***	0.5753***	0.3226***	-0.3956***	1		
Email_sent_rec	0.0363***	-0.0127***	0.0349***	0.0672***	-0.1140***	-0.6416***	1

*** $p < 0.001$

Note: $N = 471,852$.

Expenditure, indicating how much they spent in total during the (non-) promotion period. In addition, Participate (0/1) indicates whether the customer participated (i.e., qualified) in the gift card promotion. We aggregate these variables during the promotion and non-promotion periods, respectively, since most customers tend to make a single purchase in these time periods (see Table 3.5).

Figures 3.4 and 3.5 illustrate the distribution of customer expenditure during non-promotion and promotion periods. We use logarithmic transformation of expenditure to

Table 3.5: Summary statistics for the number of purchases per customer (aggregated at daily-level) during promotion and non-promotion periods. More than 95% (85%) of the customers made a single purchase during a (non-) promotion period.

Period	# Purchases	Mean	Std. Dev.
Promotion	9,555	1.032	0.184
Non-Promotion	43,658	1.237	0.757

account for the right-skewed nature of the expenditure distribution. There are three interesting features of the distribution of customer expenditure who participate. First, simple inspection reveals that the average expenditure is shifted to the right compared to promotion period (those who do not participate) or non-promotion period. Second, the probability mass appears to accumulate around the expenditure thresholds (corresponding to spikes). This suggests one of two effects at play: customers increase their expenditure to get a higher valued gift card if their expenditure is slightly lower than the threshold, or they reduce their expenditure if it is slightly more than the expenditure threshold. In both cases, customers would increase the value of the gift card relative to their expenditure. Third, comparing the expenditure distribution between non-promotion and promotion periods (when customer do not participate), we notice that there is a steeper drop in the distribution function immediately following the median in the latter case. This is likely because of the customers who opt to participate in the promotion. Therefore, the promotion is more likely to impact customers whose expenditure is in a certain region (beyond the median value).

For each customer who qualifies for a gift card, we record the following dependent variables during redemption stage: Redeem (0/1), indicating whether they redeemed the gift card and $\mathsf{Expenditure}^r$ denotes their expenditure during the redemption stage which involved redemption of a gift card. To measure the impact of redemption, we consider two dependent variables: net expenditure during the redemption stage, $\mathsf{NetExpenditure} := |\mathsf{Expenditure}^r - \mathsf{GC\_Value}|$ of customers who redeem and the total expenditure of customers who qualified in the promotion, $\mathsf{TotalExpenditure} := \mathsf{Expenditure}^q + |\mathsf{Expenditure}^r - \mathsf{GC\_Value}| \cdot \mathbf{1}_{\{\mathsf{Redeem}=1\}}$. Since customers

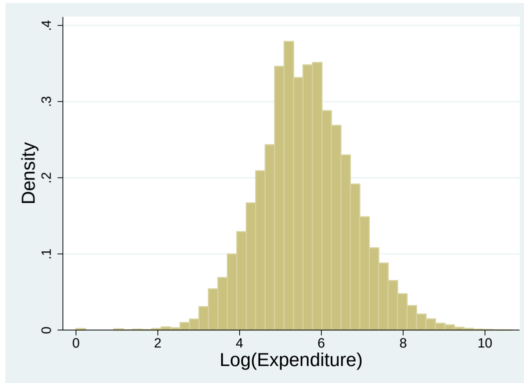
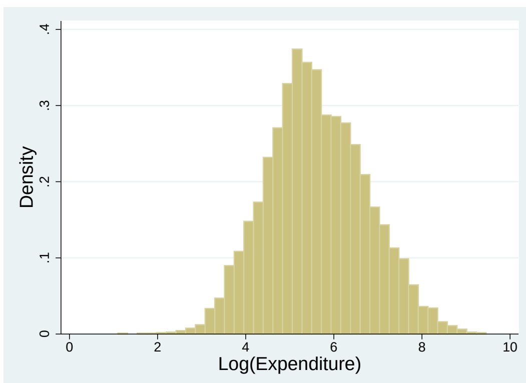
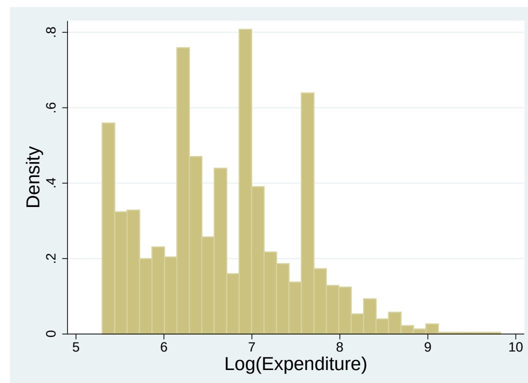


Figure 3.4: Distribution of customer expenditure who made a purchase during non-promotion period.



(a) Distribution of customer expenditure who made a purchase but did not participate in a promotion.

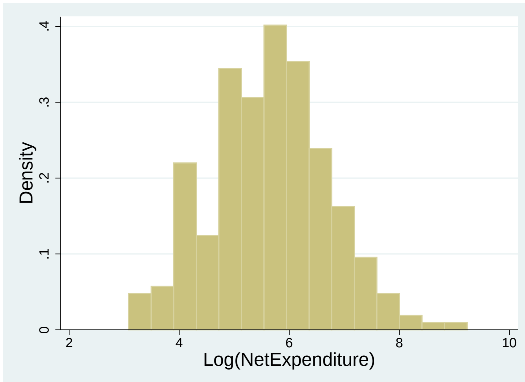


(b) Distribution of customer expenditure who participated in a promotion.

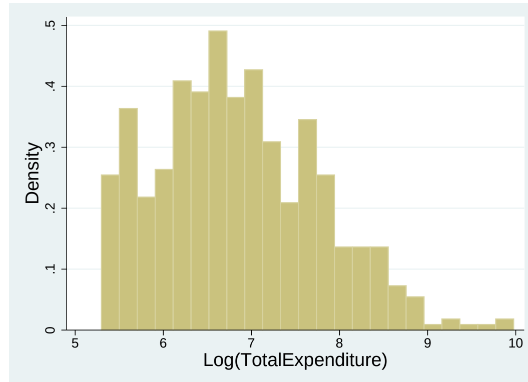
Figure 3.5: Customer expenditure during promotion period.

can redeem their gift cards on multiple occasions, we aggregate all the purchase instances during the three-month period in which the customer redeemed their gift card. Figure 3.6 illustrates the distribution of incremental sales beyond the face value of the gift card during the redemption stage and the total expenditure of qualified customers in the promotion.

To ensure validity and improve explanatory power of our empirical models, we take the following four steps. First, we exclude records from the dataset that do not have any (purchase, e-mail, website) activity during the 13 months preceding the promotion or the non-promotion period (even if those customers made purchases during the period). This is to



(a) Distribution of net customer expenditure of customers who redeem their gift card.



(b) Distribution of total expenditure of qualified customers.

Figure 3.6: Customer expenditure during promotion period.

account for the limitation imposed by our data, which cannot be used to explain shopping behavior of such customers.

Second, given our empirical context of F&A industry, we remark that some customers might be averse to buying products online. One potential reason being, customers tend to value the experience of going to a store, in addition to the product purchased. Shopping in-store also helps customers ensure they buy the right sizes.[20] Therefore, we exclude customers who did not make a purchase during the period (promotion or non-promotion) between January 2013 and December 2015. Since the gift card promotions are only applicable online, these customers are expected to be impacted minimally by them.

Third, only customers potentially exposed to the advertisement of the promotion can be impacted by it. In particular, for private gift card promotions, customers need to receive the e-mail to be potentially aware of the promotion. Therefore for private events, we exclude customer records who did not receive the gift card promotion e-mail.[21] Note that, for public

[20]It is also possible that customers have a higher cost to return an unsatisfactory product (which was bought online) than to visit a brick and mortar store to make the purchase.

[21]It is possible that third-party blogs advertise private events, so customers could become aware through an indirect channel. However, our interactions with the retailer suggests that this effect is minimal and therefore, excluding such records do not alter the main insights.

and hybrid promotions, the promotion is advertised via an e-mail and a banner ad visible to any visitor to the website. Since, the percentage of customers that received the gift card promotion e-mail for public and hybrid events is 57% and 56.9%, respectively, we treat both types of events as public events.

Fourth, for the redemption stage, we consider customers who qualified for a gift card between February, 2015–December, 2015. The corresponding redemption stage run between April, 2015–May, 2016. This is because a key component of the redemption analysis is to detect if a customer used a promotional gift card during their purchase in the redemption stage. A marker for such an event in the website log-files is available only after April, 2015. Furthermore, of the 655 customers who qualified for gift cards during this period: 286 did not redeem and 255 redeemed their gift card. The remaining 114 customers made purchases during the redemption stage, but there is no marker of gift card redemption. From our interactions with the retailer, it seems likely they in fact redeemed their gift card but were not recorded properly. Since we cannot determine which of those payments involved a gift card, we exclude these 114 customers from our analysis.

3.6 Customer Response Models

In this section we present three econometric models to capture customer response (purchase and expenditure) during the qualification and redemption stages of the promotion. Using these models we test our hypotheses and quantify the impact of gift card promotion. In the aggregate model, we first benchmark the impact of gift card promotion on customer response during the promotion period (qualification stage) relative to non-promotion periods. To further isolate the effects of participation in the promotion and of the advertisement of the promotion, we conduct a promotion period analysis. In the redemption model, we investigate the impact of redemption on the net customer expenditure.

One of the primary concerns in estimating models involving choice and resulting continuous outcomes (such as, choice of profession and earnings) is the inherent self-selection due to unobservable factors. In our empirical context, customer choices include whether to purchase or participate in the promotion and later, whether to redeem their gift cards. The customer decision process underlying these choices is unobservable, rather, we observe their resulting expenditure. To address the selection issue, we consider the limited dependent variable framework (LDV) with selectivity bias, which makes distributional assumptions on the correlation structure between the choice and outcome decisions (Maddala, 1983). We adapt the framework proposed by Zhang and Krishnamurthi (2004), which we explain using the aggregate model. We use the generalized version of the framework to model customer response during the promotion period.

3.6.1 Aggregate Model

To benchmark the impact of the promotion, we utilize customer purchasing behavior during the promotion and non-promotion periods. In particular, the specification for customer i's expenditure during period t is given below.

$$\ln(\mathsf{Expenditure}^*_{it}) = \boldsymbol{\alpha}\mathbf{X}_{it}+\epsilon_{i1t} := \alpha_0+\alpha_1\mathsf{CustomerChar}_{it}+\alpha_2\mathsf{Promo}_t+\alpha_3\mathsf{OtherControls}_t+\xi_{it}. \tag{3.1}$$

The logarithmic transformation accounts for right-skewed nature of expenditure (see Figures 3.4 and 3.5). $\mathsf{CustomerChar}_{it}$ controls for the customer's past purchasing patterns and their budgetary constraints. Controlling for website browsing behavior and their preference of marketing e-mail frequency, captures customer search costs and their deal-seeking nature. Coefficient α_2 captures the aggregate effect of the promotion on customer expenditure. The aggregate effect combines the effect of participation in the promotion and the effect of advertisement of the promotion.

Estimating Equation (3.1) using ordinary least squares (OLS) technique can result in inconsistent estimates. This is because customers do not randomly decide to make a purchase, rather they *choose* to purchase if it maximizes their utility. In otherwords, we are interested in estimating expected customer expenditure devoid of the self-selection issue, modeled by the latent variable $\mathsf{Expenditure}^*_{it}$. However, as researchers we observe customer expenditure conditional on a purchase, $\mathsf{Expenditure}_{it}$. To account for this self-selection we consider a flexible framework (limited-dependent variable) that allows us to jointly model customer purchase and expenditure decisions. This model specifies that customers choose to purchase iff the utility from purchase (U_{i1t}) exceeds the utility of no-purchase (U_{i0t}), which are described below.

$$U_{ijt} := V_{ijt} + \nu_{ijt} := \tilde{\boldsymbol{\alpha}}_{\boldsymbol{i}}\tilde{\mathbf{X}}_{it} + \nu_{ijt}, \quad j = 0, 1 \tag{3.2}$$

We set $V_{i0t} = 0$ to ensure identifiability of the parameters and interpret the coefficients relative to the no-purchase alternative. The utility of purchase is composed of two parts: V_{i1t} is the representative utility function resulting from a purchase and ν_{i1t} are the unobservable factors (to the econometrician) contributing to the utility of purchase. The resulting probability of purchase can be computed as $\mathbb{P}\{\nu^*_{it} < V_{i1t}\}$, where $\nu^*_{it} := \nu_{i0t} - \nu_{i1t}$. Because $V_{i0t} = 0$, ν^*_{it} denotes the net representative utility of no-purchase. In otherwords, customers make a purchase if their net representative utility of no-purchase is less than representative utility of purchase.

The utility of purchase, specified by $\tilde{\mathbf{X}}_{it}$ includes non-linear terms $\mathsf{Purchase\_rec}^2$ and $\ln(\mathsf{Period\_length})$; $\mathsf{Email\_sent\_rec}$, in addition to the variables in $\mathbf{X}_{it}$. Including $\mathsf{Purchase\_rec}$ and $\mathsf{Purchase\_rec}^2$ captures the impact of the most recent purchase on the utility obtained from making the current purchase, which can vary depending on the product category (Neslin et al., 2013). In the context of fashion products, customers may obtain positive utility from repeated use of the purchased product. This is because, customers typically purchase fashion products, such as dresses, for use on certain occasions or to use them a certain number of times. Therefore, customers may initially gain more utility from a previous purchase compared to utility from another purchase. Eventually, as fashion trends change, customers may be more inclined to make another purchase. This is in contrast to the consumer packaged goods industry, where consumption utility is generally steady and purchase decisions are driven from inventory-related effects. That is, with time, as inventory depletes, the probability of another purchase monotonically increases, as illustrated by Khan et al. (2009).

Including $\ln(\mathsf{Period\_length})$ captures potential non-linear effect of length of period on purchase incidence. We expect the probability of a purchase to monotonically increase with Period_length since it offers customers more opportunity to visit the website. This increase need not be linear, since customers who have not purchased for a long duration might have purchased at another store. Therefore, the rate of increase in purchase probability potentially reduces with Period_length tapering[22] off in the end.

The Email_sent_rec variable captures the impact of timing of e-mail marketing efforts on the customer's purchase decision. E-mail channel is generally perceived to be a less costly way for the retailer to motivate its customers to visit its website (Neslin et al., 2013). This is particularly beneficial given the role spontaneity plays in customer's purchase decision. Once on the website, various advertisements and attractive offers, could lead the customer to make an unintended purchase. However, too much of e-mail marketing can also dilute customer's attention to future marketing efforts, thereby having a negative impact on customer purchase (Kumar et al., 2014). While the recency of e-mail can impact the customer decision to visit the website or make a purchase, the recency of e-mail may not play a role in determining the customer expenditure. Once on the website or a purchase a decision has been made, the expenditure decision is driven by budgetary considerations. Hence, we exclude it from the expenditure specification.

[22]For example, $P(t) \propto \dfrac{t^\beta}{1+t^\beta}$, $\beta > 0$, where $P(t), t$ denote $\mathbb{P}\{\mathsf{Purchase} = 1\}$ and Period_length, captures such non-linearity. This relationship implies that $\text{logit}(P) \propto \beta \ln(t)$.

The self-selection bias discussed above arises because unobservable factors affecting purchase decision (ν) are potentially correlated with the unobservable factors affecting the expenditure decision (ξ). If this correlation is not significant, it suggests that the (observable) factors in $\mathbf{X}$ and $\tilde{\mathbf{X}}$ are rich enough to address the selection issue. In this case, one can independently estimate Equations (3.1) and (3.2). There are several approaches proposed in the literature to model the correlation between the choice and outcome variables, depending on the distributional assumptions made on the error terms. We adopt one such framework proposed by Zhang and Krishnamurthi (2004), who model ξ_{it}'s as i.i.d. according to logistic distribution with mean 0 and standard deviation σ^2. Likewise, ν_{ijt}'s are assumed i.i.d. according to standard Gumbel distribution. The latter is a standard assumption leading to the following logistic model: $\mathbb{P}\{\nu^*_{it} < V_{i1t}\} = \frac{e^{V_{i1t}}}{1+e^{V_{i1t}}}$.

The correlation between the purchase and expenditure decisions is modeled by specifying a bivariate logistic distribution over ν^*_{it} and ξ_{it}. The joint distribution function is given by:

$$F_{\nu^*_{it},\xi_{it}}(x,y) = F_{\nu^*_{it}}(x)F_{\xi_{it}}(y) \cdot [1 + \theta(1 - F_{\nu^*_{it}}(x))(1 - F_{\xi_{it}}(y))], \quad \theta \in [-1,1], \tag{3.3}$$

where $\rho = 3\theta/\pi^2$ is the correlation between ν^*_{it} and ξ_{it}. The closed-form expression of the log-likelihood function, for a more general polychotomous choice model, is given in the Appendix B.1. The log-likelihood is maximized to estimate parameters of interest. Further, the above correlation structure easily generalizes to polychotomous choices, which we use to model customer response during the promotion period. This provides a consistent framework throughout our analysis, allowing us to compares results across different models. The other widely used approach to estimate LDV models with self-selection is a two-stage estimation technique, such as, two-stage Heckman model (Heckman, 1979) or extensions to polychotomous choice models as in Trost and Lee (1984). However, as noted by Krishnamurthi and Raj (1988), the MLE approach is more efficient than two-stage estimation technique because information contained in the error terms is used in the estimation of both the choice and outcome equations simultaneously.

3.6.2 Promotion Period Model

In this section we further investigate how participation in the promotion impacts customer expenditure during the promotion. The arguments presented in § 3.4 support the view that gift card promotion incentivizes customers to increase expenditure to participate in the promotion. However, including $\mathsf{Participation}_{ik}$ in the expenditure specification in the aggregate model to test this hypothesis results in an endogeneity problem. Factors that influence participation in the promotion potentially also influence customer's expenditure during the promotion period. This endogeneity issue can be resolved by generalizing the choice model in the aggregate model to also include the participation alternative in the choice process during the promotion period.

We specify[23] the expenditure for customer i during promotion period, $k = 1, \ldots, 38$, as follows.

$$\ln(\mathsf{Expenditure}^*_{ijk}) = \begin{cases} \boldsymbol{\beta}_1 \mathbf{Z}_{i1k} + \xi_{i1k} \\ := \beta_{10} + \beta_{11}\mathsf{CustomerChar}_{ik} + \beta_{12}\mathsf{OtherControls}_k + \xi_{i1k}, & \text{if } j = 1, \\ \\ \boldsymbol{\beta}_2 \mathbf{Z}_{i2k} + \xi_{i2k} \\ := \beta_{20} + \beta_{21}\mathsf{CustomerChar}_{ik} + \beta_{22}\mathsf{PromotionChar}_{ik} \\ \qquad\qquad + \beta_{23}\mathsf{OtherControls}_k + \xi_{i2k}, & \text{if } j = 2, \end{cases} \tag{3.4}$$

where ξ_{ijk} are the unobserved errors, i.i.d. according to logistic distribution with mean 0 and variance δ_j; and $j = 1, 2$ represent purchase (without participation) and participation alternatives, respectively. However, $\mathsf{Expenditure}^*_{ijk}$ is unobservable, rather we observe customer expenditure conditional on an alternative being chosen.

[23] The GC_email variable in the purchase and participate expenditure equations turned out to be insignificant. Dropping them did not alter other coefficients, and hence were excluded in the final model specification.

Customer's decision to purchase (or participate) during the promotion is based on the utility they derive from each alternative. The utility of the three alternatives for customer i during promotion period k is given by:

$$U_{ijk} = V_{ijk} + \nu_{ijk} = \tilde{\boldsymbol{\beta}}_j \tilde{\mathbf{Z}}_{ijk} + \nu_{ijk}, \quad j = 0, 1, 2;\ k = 1, \ldots, 38, \tag{3.5}$$

where $j = 0$ denotes the no-purchase alternative. The specification of the utility associated with each alternative, U_{ijk}, includes promotion characteristics in addition to the customer and communication characteristics. The promotion characteristics account for the fact that gift card promotions with lower expenditure threshold or higher perceived discount potentially boost participation. Lower expenditure threshold makes it easier for customers with lower intended expenditure to participate and higher perceived discount attracts customers who are more price-sensitive to participate in the promotion. The specification above implies that the customer choice process during the promotion periods can be modeled using following probability model.

$$\begin{aligned}
\mathbb{P}\{\mathsf{Purchase}_{ik} = 0\} &= \mathbb{P}\{\nu^*_{i0k} < V_{i0k}\}, \\
\mathbb{P}\{\mathsf{Purchase}_{ik} = 1\} &= \mathbb{P}\{\nu^*_{i1k} < V_{i1k}\}, \quad \text{and} \\
\mathbb{P}\{\mathsf{Participate}_{ik} = 1\} &= \mathbb{P}\{\nu^*_{i2k} < V_{i2k}\},
\end{aligned}$$

where $\nu^*_{ijk} := \max_{l \neq j}\{V_{ilk} + \nu_{ilk}\} - \nu_{ijk}$. We replicate the distributional assumptions made on the error terms in the aggregate model. Following this, the choice probabilities can be computed in closed-form as follows (McFadden, 1974).

$$\begin{aligned}
\mathbb{P}\{\mathsf{Purchase}_{ik} = 0\} &= \frac{e^{V_{i0k}}}{e^{V_{i0k}} + e^{V_{i1k}} + e^{V_{i2k}}} \\
\mathbb{P}\{\mathsf{Purchase}_{ik} = 1\} &= \frac{e^{V_{i1k}}}{e^{V_{i0k}} + e^{V_{i1k}} + e^{V_{i2k}}} \\
\mathbb{P}\{\mathsf{Participate}_{ik} = 1\} &= \frac{e^{V_{i2k}}}{e^{V_{i0k}} + e^{V_{i1k}} + e^{V_{i2k}}}
\end{aligned}$$

We specify two bivariate logistic distributions over ν^*_{ijk} and ξ_{ijk}, $j = 1, 2$, respectively. This allows us to estimate the correlations between purchase and resulting expenditure (ρ_1) and, participation and resulting expenditure (ρ_2), separately. We estimate $\theta_j, j = 1, 2$ from the data, which can be scaled to compute $\rho_j = 3\theta_j/\pi^2$.

3.6.3 Redemption Model

In this section we present the empirical model used to test the impact of redemption on the net expenditure customers make during the redemption stage of the promotion. We define the net expenditure as follows.

$$\mathsf{NetExpenditure}^*_{il} := \begin{cases} \mathsf{NetExpenditure}_{il}, & \text{if } \mathsf{Redeem}_{il} = 1, \\ 0, & \text{if } \mathsf{Redeem}_{il} = 0, \end{cases} \tag{3.6}$$

where $l \in \{1, \ldots, 10\}$ denote the redemption stage and $|x| := \max\{x, 0\}$. The above definition of net expenditure denotes the total realized revenues for the retailer during the redemption stage of the promotion. If customers redeem less than the face value of the gift card, the retailer does not realize any revenues. We specify the regression for the logarithmic transformation of the net expenditure below.

$$\ln(\mathsf{NetExpenditure}^*_{il}) = \boldsymbol{\beta}^r \mathbf{Z}_{il} + \xi^r_{il} := \beta^r_0 + \beta^r_1 \mathsf{CustomerChar}_{il} + \beta^r_2 \ln(\mathsf{GC\_value})_{il} + \xi^r_{il}. \tag{3.7}$$

In addition to the customer purchase and website visit patterns, this specification also includes the face value of the gift card earned by the customer. Essentially, this is to test if the magnitude of the gift card induced expenditure varies with the size of the gift card that is redeemed. To estimate the above regression we calibrate the independent variables up to the beginning of the qualification stage of the promotion. This way the independent variable, ln(Exp_annual) is not highly correlated with the customers' qualification stage expenditure. This circumvents potential endogeneity issues which may result from customers simultaneously determining their qualification and redemption stage expenditures.

To estimate the latent variable NetExpenditure*, we use a LDV framework, such as the one discussed in § 3.6.1. To this end, we the model the redemption stage choice as a result of customers maximizing their utility. The utility associated with redemption, U^r_{il}, includes Email_sent_rec in addition to the customer characteristics and face value of the gift card, which are also included in the expenditure specification.

3.7 Results and Managerial Insights

We first present the results from the estimation of the three models and test the hypotheses. We summarize some of our findings briefly for managers.

3.7.1 Aggregate Model

The results from estimation of the aggregate model (Equations 3.1, 3.2) are presented in Table 3.6. The coefficient of Promo in the logistic regression suggests that the purchase probability of customers increases during a promotion. The probability of a purchase during a promotion period is 5.63% which is 17.54% greater than the non-promotion period purchase probability. This increase in purchase probability is the aggregate effect of customers who participate in the promotion and otherwise. Therefore, our findings support Hypothesis 1. The promotion mechanism and its communication to customers through e-mail or website banner ad plays the dual role of informing interested customers about the promotion offer and potentially reminds other customers about the store. Because gift card promotion e-mails are sent less frequently compared to other marketing e-mails, we expect the e-mail to have a positive impact on customer response. Customers who do not get the e-mail, yet visit the website (through alternate channels such as, Google search or direct website visit) are exposed to the banner-ad which occupies a huge portion of the website.

Table 3.6: Estimation results of the aggregate model.

	ln(Expenditure)	Purchase
Constant	5.10646***	-5.29308***
	(0.03987)	(0.03895)
Purchase_rec	0.00234***	-0.01698***
	(0.03585)	(0.07295)
Purchase_rec^2		0.00004***
		(0.07435)
Purchase_freq	0.00453***	0.07426***
	(0.00122)	(0.00136)
ln(Exp$_a$nnual)	0.18655***	0.02354***
	(0.00434)	(0.00339)
Web_rec	-0.00006	0.00145***
	(0.02483)	(0.01783)
Web_freq	-0.00195***	0.00642***
	(0.00016)	(0.00013)
Email_cat (med)	0.05497†	0.09655***
	(0.03084)	(0.02545)
Email_cat (high)	-0.03544*	0.13425***
	(0.01434)	(0.01493)
Email_sent_rec		0.99128***
		(0.01550)
Promo	0.27348***	0.19461***
	(0.01658)	(0.02041)
ln(Period_length)		0.99923***
		(0.00897)
Markdown	-0.09394***	0.18003***
	(0.01295)	(0.01066)
2014	0.00865	-0.29874***
	(0.01946)	(0.01537)
2015	0.05076**	-0.56113***
	(0.01907)	(0.01508)
ρ	-0.3040	
- ln($\mathcal{L}$)	250,999	
N	1,045,405	

Standard errors in parentheses

† $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The signs and significance of Purchase_rec and Purchase_rec^2 indicates a U-shaped relationship between purchase probability and time since last purchase (see Figure 3.7).[24] This observation supports the theory that customers initially gain utility from the repeated use of their previously purchased product. Over time however, that utility starts to wane down potentially due to change in fashion trends. At this point, customers are more likely to return to the store for another purchase. We define the point at which the purchase probability starts to increase as the purchase cycle. The purchase cycle for shoppers at this retailer is estimated to be, on average, 217.7 days.[25] In addition to the above, there is another effect which could potentially accentuate the U-shaped relationship. In our empirical context customers place their orders online. Therefore, the uncertainty about whether customer is satisfied with the product is only resolved after it is received in the mail. In case they are not satisfied with the product they would return it for another purchase. This implies that, on average the probability of a (re-)purchase is higher in the first few days after a purchase compared to a later day.

The coefficient of Web_freq indicates that customers who frequently visit the website are more likely to make a purchase, but end up spending lesser. This result can be explained by the fact that frequent visitors to the website tend to have lower opportunity cost of time and hence, are more likely to be price sensitive. The coefficient of Email_sent_rec, suggests that it is beneficial for the retailer to send marketing e-mails less often. The coefficient of Markdown variable is intuitive in that, customers are more likely to make a purchase during markdown season, but spend less due to the deep discounts offered.

The coefficient of Promo in the ln(Expenditure) equation suggests a strong positive effect of the promotion period on customer expenditure. The average customer expenditure during

---

[24]Gönül and Shi (1998) also report a U-shaped relationship between purchase probability and purchase recency in the context of catalog retailer.

[25]Computed as $\frac{0.01698}{2\times 0.00004}$.

a promotion period is \$830.24, which is 31.45% greater than the average expenditure during a non-promotion period.[26] This increase can be explained in part due to participation in the promotion, which require customers to spend beyond an expenditure threshold on regular priced products or the advertisement effect of highlighting products that are included in the promotion. In § 3.6.2, we dissect this increase in expenditure during promotion period due to participation in the promotion and the advertisement effect. We also note that, the measure of correlation between the purchase and expenditure decision, captured by ρ, turns out to be significant. That is there are unobservable factors that are simultaneously impacting both purchase and expenditure decisions.

Our next goal is to isolate the advertisement effect of the promotion during the qualification stage (see Hypothesis 3). By definition, advertisement effect impacts customers who were potentially exposed to the promotion but did not participate in the promotion. Therefore, we exclude customers who participated in the promotion to re-estimate Equations (3.1)-(3.2). This approach is similar to the one adopted by Sahni et al. (2016). The results of the estimation are presented in Table 3.7.

The coefficient of Promo in the Purchase equation, which captures the advertisement effect of the promotion on customer purchase decision, is not significant. Therefore we do not find evidence in support of Hypothesis 3(a). This finding suggests that the cumulative impact of advertising through website banner-ads and e-mail may not increase purchase probability of customers who do not participate in the promotion. We note that it is difficult to isolate the effect of the two channels because customers who receive the e-mail and visit the website may also be impacted by the banner-ad. The coefficient of Promo in the Expenditure equation, which captures the advertisement effect of the promotion on customer expenditure decision, is positive and significant. The average customer expenditure during a promotion who do

[26]We use the fact that if $\ln(X)$ has a logistic distribution with mean μ, then X has a log-logistic distribution with mean e^{μ}.

not participate is \$680.22. However, this model does not explicitly account for the selection bias that results from customers choosing to purchase without participating in the promotion. Therefore, we defer this analysis to the promotion period analysis, where we explicitly model customer's decision to purchase or participate.

Table 3.7: Estimating advertisement effect.

	ln(Expenditure)	Purchase
Constant	5.09481***	-5.28896***
	(0.04060)	(0.03943)
Purchase_rec	0.00234***	-0.01695***
	0.03648	(0.07424)
Purchase_rec^2		0.00004***
		(0.07563)
Purchase_freq	0.00514***	0.07522***
	(0.00125)	(0.00139)
ln(Exp_annual)	0.18719***	0.01901***
	(0.00441)	(0.00346)
Web_rec	-0.00005	0.00149***
	(0.02522)	(0.01813)
Web_freq	-0.00198***	0.00649***
	(0.00016)	(0.00013)
Email_cat (med)	0.05076	0.09549***
	(0.03140)	(0.02589)
Email_cat (high)	-0.04699**	0.12762***
	(0.01457)	(0.01514)
Email_sent_rec		1.00095***
		(0.01567)
Promo	0.07448***	-0.01387
	(0.01795)	(0.02115)
ln(Period_length)		0.99939***
		(0.00902)
Markdown	-0.08802***	0.18167***
	(0.01316)	(0.01088)
2014	0.01908	-0.28117***
	(0.01985)	(0.01570)
2015	0.06190**	-0.54830***
	(0.01949)	(0.01542)
ρ	-0.3040	
$-\ln(\mathcal{L})$	242,519	
N	1,043,813	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.7.2 Promotion Period Model

The results of the estimation of promotion period model is presented in Table 3.8. We quantify the impact of participation in the promotion as the difference between the expected expenditure of customers if they participated in the promotion and if they made a purchase without participation. This difference $\mathbb{E}[\mathsf{Expenditure}^*_{i2} - \mathsf{Expenditure}^*_{i1}] = e^{\boldsymbol{\beta}_2 \mathbf{Z}_{i2}} - e^{\boldsymbol{\beta}_1 \mathbf{Z}_{i1}}$, is averaged across all customers.[27] We find that, on average, customers who participate in the promotion spend \$2,033.2, which is about 2.97 times the average expenditure of customers who make a purchase without participation. Such a significant increase in expenditure can be explained partly by the fact that customers need to buy regular-priced products which are significantly more expensive than products on sale. In general, the retailer offers upto 75% discount for products on sale. For the retailer, this finding suggests that getting the customer to participate in the promotion, provides a significant boost to their revenues during the promotion period. Average expenditure of customers who do not participate in the promotion is \$684.77, which is 8.42% greater than the average expenditure during non-promotion period. Therefore, we find strong evidence for Hypothesis 3(b). This effect may be attributed to the fact that products that are included in the gift card promotion are highlighted on the website, attracting greater customer attention even if they do not end up participating in the promotion.

Attributing the difference between the average expenditure of customers who participate and otherwise, to the promotion can be misleading for retailers. This is because, participation in the promotion potentially alters expenditure distribution of customers whose expenditure is close enough to one of the expenditure thresholds (compare Figures 3.4 and 3.5). We address this issue by first computing the increase in mean of the expenditure distribution between

[27]Note that, including the selection bias terms $\mathbb{E}[\xi_1 | U_1 > \max\{U_2, U_0\}]$, $\mathbb{E}[\xi_2 | U_2 > \max\{U_1, U_0\}]$ would capture the effect of the different unobservable characteristics in addition to the effect of participating in the promotion (Trost and Lee, 1984).

the promotion and non-promotion periods using the aggregate model. Next, based on the promotion period model, we attribute a fraction of the increase in the mean to participation. In our context, we find the mean expenditure increases by \$198.65 during the promotion. Of this, we attribute 96.34%[28] to participation in the promotion. That is, average increase in the expenditure due to participation is \$191.38 compared to the non-promotion period. This finding supports Hypothesis 2.

Purchase cycle of customers who do not participate in the promotion is around 206.21 days and of those who participate is around 180.19 days. That is, purchase probability of customers (as a function of purchase recency) who participate starts to increase prior to the point at which the same happens for customers who do not participate in the promotion (see Figure 3.7). One possible explanation would be that customers with shorter purchase cycle are more likely to redeem their gift cards and hence, gain greater utility from participation. The other potential explanation for this result could due to the mechanism driving participation in the promotion. Gift card promotion rewards customers who buy products at regular price. Therefore, the selection of products included in the promotion can be considered contemporary fashion compared to others on sale. Gift card promotion potentially incentivizes customers who gain greater utility from consuming fashionable products. By virtue of their inherent fashion-sensitiveness, these customers are prone to making more frequent purchases resulting in shorter purchase cycle.

The coefficient of ln($\mathsf{Exp\_Ann}$) can be interpreted as the marginal change in probabilities of purchase and participation for a percentage increase in the annual expenditure.[29] To place

---

[28]The effect of participation is computed as $\frac{\Delta_p}{\Delta_p+\Delta_a} \times 100$, where $\Delta_p = 2,033.2 - 631.59 = 1401.61$ and $\Delta_a = 684.77 - 631.59 = 53.18$.

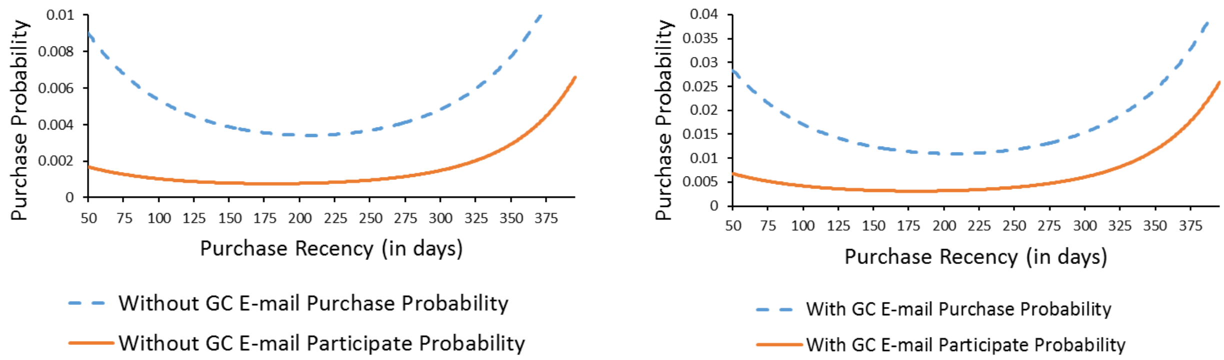
[29]All the coefficients in the choice process are to be interpreted as the impact on the utility of choosing the alternative relative to the utility of choosing the base case (no-purchase option). There is, however, no direct interpretation of these coefficients on the unconditional probability of choosing the alternative. In fact, the sign of of the coefficient does not necessarily indicate the direction of the relationship (Wooldridge, 2002). Therefore, we average (across all customers) the marginal effect (ME) of changing an independent

Table 3.8: Estimates of the promotion period model.

	ln(Expenditure$_1$)	ln(Expenditure$_2$)	Purchase	Participate
Constant	4.98714***	4.81846**	-7.61758***	-4.12780***
	(0.10962)	(1.48133)	(0.54796)	(1.21803)
Purchase_rec	0.00255***	0.37397*	-0.01650***	-0.01694***
	(0.09519)	(0.17956)	(0.18844)	(0.38133)
Purchase_rec^2			0.00004***	0.00005***
			(0.19389)	(0.39321)
Purchase_freq	-0.00494†	0.00852†	0.04611***	0.03574***
	(0.00277)	(0.00499)	(0.00229)	(0.00397)
ln(Exp_annual)	0.20053***	0.09290***	0.10667***	0.23366***
	(0.01142)	(0.02148)	(0.00844)	(0.01631)
Web_rec	-0.00001	0.00022	0.00074***	-0.00069**
	(0.06838)	(0.13585)	(0.04978)	(0.10251)
Web_freq	-0.00142***	-0.00124†	0.00462***	0.00391***
	(0.00038)	(0.00074)	(0.00028)	(0.00053)
Email_cat (med)	0.16544*	0.09966	0.29316**	-0.29384
	(0.07826)	(0.15600)	(0.09664)	(0.23127)
Email_cat (high)	-0.01195	-0.03598	0.05390	-0.01818
	(0.03772)	(0.07785)	(0.29579)	(0.57201)
Email_sent_rec			0.00235***	0.00063†
			(0.06429)	(0.13384)
GC_email			0.39750***	0.65538***
			(0.10477)	(0.18583)
Private		0.01769	0.15936**	-0.56802***
		(0.19110)	(0.06054)	(0.15835)
ln(Lowest_threshold)		0.47182**	0.09706	-0.77425***
		(0.17889)	(0.06028)	(0.14038)
Nbr_tiers		-0.05234	0.07647*	0.12554†
		(0.09783)	(0.03555)	(0.07613)
Discount		-0.03376**	0.02078***	0.03332***
		(0.01135)	(0.00443)	(0.00909)
ln(Period_length)			1.27033***	0.73746***
			(0.05953)	(0.13126)
GC_email × Email_sent_rec			0.00881***	0.00990***
			(0.12377)	(0.22445)
GC_email × Email_cat (med)			-0.10203	0.40629
			(0.14994)	(0.30463)
GC_email × Email_cat (high)			-0.05502	-0.17192
			(0.30913)	(0.59232)
Markdown	-0.06602†	0.09535	0.21968***	-0.07993
	(0.03402)	(0.07692)	(0.02894)	(0.05873)
2014	0.06488	0.23943*	-0.14119**	-0.52725***
	(0.05252)	(0.09978)	(0.04296)	(0.07868)
2015	0.10655*	0.20578*	-0.32602***	-0.66339***
	(0.05119)	(0.09403)	(0.04211)	(0.07475)
ρ_1	-0.3040***			
ρ_2	-0.3040***			
$-\ln(\mathcal{L})$	53,720.11			
N	471,852			

Standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$



(a) When customers do not receive promotion e-mail.

(b) When customers receive promotion e-mail.

Figure 3.7: Purchase probability as a function of purchase recency. Note that all continuous independent variables are set to their average values observed in the sample, Year $=2015$, Private, Email_Cat =high, and Markdown to 0.

the marginal change in perspective, we report it as a percentage change from the baseline probability. We find that a percentage increase in annual expenditure increases probabilities of purchase and participation by 10.43% and 23.13%, respectively. The relatively stronger effect on participation probability suggests that customers with greater spending power are more likely to be flexible to increase their expenditure to participate in the promotion.

The coefficients of other marketing e-mails, in contrast to gift card e-mail, have a negative impact on customer purchase probability during the promotion. In particular, fewer marketing e-mails, sent farther out from the promotion, boosts the probability of making a purchase or participating in the promotion. Taken with the above observation, it suggests that the impact of e-mail channel depends not only on the contents, but also on the timing and frequency.

variable, Z_j on probability of participation of customer i as follows.

$$\frac{\partial}{\partial Z_j}\mathbb{P}\{\mathsf{Participate}=1\} = \pi_{i2}\cdot\{\beta_{2j}-(\beta_{0j}\pi_{i0}+\beta_{1j}\pi_{i1}+\beta_{2j}\pi_{i2})\} = \pi_{i2}\cdot\{\beta_{2j}-(\beta_{1j}\pi_{i1}+\beta_{2j}\pi_{i2})\},$$

where $\pi_{i2}=\widehat{\mathbb{P}}\{\mathsf{Participate}=1\}$, $\pi_{i1}=\widehat{\mathbb{P}}\{\mathsf{Purchase}=1\}$ are the predicted probabilities of participation and purchase, respectively.

Coefficients of promotional characteristics, such as the lowest expenditure threshold and the perceived discount have intuitive explanations. Lower expenditure thresholds reduces the hurdle for participation and hence, increases probability of participation and lowers probability of purchase without participation. For a dollar decrease in the lowest expenditure threshold (from the average \$328), the participation probability increases by .24% and drops the purchase (without participation) probability by 0.03%. Similarly, increasing the perceived discount by one percentage point increases the participation probability by 3.29% and purchase probability by 2.04%. Promotions with better deals attract greater customer attention, even if they do not end up participating in the promotion. The coefficient of Markdown variable indicates that holding gift card promotion during a markdown season can lower participation as more customers prefer instantaneous reward of buying a product on sale, instead of a delayed reward.

3.7.3 Redemption Model

In this section we test the hypothesis regarding the impact of gift card redemption on the customer expenditure during the redemption stage of the promotion. The results from estimation of the redemption model is presented in Table 3.9. We find that on average, net expenditure of customers (after excluding gift card face value) during the redemption stage is \$525.28. On average the face value of the gift card is around \$180. Therefore, we find a significant evidence for the gift-card induced expenditure during the redemption, supporting Hypothesis 4. We also find that the effect of gift card induced expenditure is positively impacted by the face value of the gift card. The log-log specification indicates that for a percentage increase in the face value of the gift card, the net expenditure increases by .62%. Therefore, customers with a gift card face of greater value are likely to splurge more during the redemption stage.

Redeeming a gift card requires customers to keep track of their e-mails and visit the website during the redemption stage. The opportunity cost of time involved in performing

Table 3.9: Estimates of redemption model.

	$\ln$(NetExpenditure$_0$)	Redeem
Constant	1.96152**	-1.01784
	(0.64097)	(0.70119)
Purchase_rec	0.00142	-0.00099
	(0.47388)	(0.47520)
Purchase_freq	0.02175	-0.01136
	(0.02187)	(0.01617)
ln(Exp_Ann)	0.10193†	-0.04311
	(0.05965)	(0.05789)
Web_rec	0.00083	-0.00159†
	(0.37254)	(0.38249)
Web_freq	0.00093	0.00506*
	(0.00232)	(0.00234)
GC_face_value	0.61781***	0.34981***
	(0.10238)	(0.10051)
Email_sent_rec		0.00038
		(0.39415)
Email_cat (med)	-0.04120	-0.16883
	(0.43954)	(0.52498)
Email_cat (high)	0.18001	-0.66873*
	(0.20663)	(0.32991)
ρ	-0.3040***	
N	541	
$-\ln(\mathcal{L})$	747	

Standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The baseline probability of redemption is 47.13% .

these activities is captured in the website visit variables. The positive sign of Web_Freq and negative sign of Web_Rec suggests that customers who are website-savvy have lower opportunity cost of time and hence, more likely to redeem their gift card. The face value of the gift card has a strong positive effect on the probability of redemption. For every 1% increase (roughly \$1.73) in the face value of the gift card, the redemption probability increases by 17.24% from the baseline redemption probability.[30] The negative sign of Email_cat_high indicates that customers who opt to receive more marketing e-mails from the retailer have a

[30]The utility specification implies that $\mathbb{P}\{\mathsf{Redeem}_{ijl} = 1\} = \frac{e^{\tilde{\boldsymbol{\beta}}_1^r \tilde{\mathbf{Z}}_{il}^r}}{1+e^{\tilde{\boldsymbol{\beta}}_1^r \tilde{\mathbf{Z}}_{il}^r}}$. Therefore $\frac{\partial}{\partial Z_j}\mathbb{P}\{\mathsf{Redeem}_{ijl} = 1\} = \tilde{\beta}_j^r \cdot \pi_0^r \cdot \pi_1^r$, where $\pi_0^r := \widehat{\mathbb{P}}\{\mathsf{Redeem} = 1\}$ and $\pi_1^r := 1 - \pi_0^r$ are predicted no-redemption and redemption probabilities, respectively. Therefore, for every unit change in Z_j, probability of redemption changes by $\tilde{\beta}_j^r \cdot \pi_0^r \cdot \pi_1^r$ percentage points for each customer. We report averages across all customers.

lower redemption probability. As discussed above, receiving more marketing e-mails increases the cost associated with having to track the e-mail with gift card code in it.

While incremental expenditure beyond the face value of the gift card is beneficial for the retailer, it may also result from customers redistributing their expenditure between the qualification and redemption stages of the promotion. This means that customers spend more than the face value of the gift card because they spent less during the qualification stage. To test if this effect is under play, we measure the impact of redemption on the total expenditure customers make as a part of the promotion, defined as, $\mathsf{TotalExpenditure}^*$. There are two sources of selection bias here, during the qualification and redemption stages. Therefore, models presented in § 3.6, which account for contemporaneous selection bias are no longer applicable. Instead, we take a matching approach (based on propensity score) to account for selection bias based on observable factors (Rosenbaum and Rubin, 1983). The details of the implementation are presented in Appendix B.4. We estimate the average effect of redemption on total customer expenditure in the promotion (for customers who redeem) as \$427.6 ($p < .05$). This suggests that the customers who later redeem their gift card may spend a little lesser than customers who do not. However the magnitude of it is significantly smaller than the incremental expenditure beyond the face value during the redemption stage. Therefore our analysis suggests that the effect of customers distributing their expenditure between the two stages of the promotion has limited validity in the context of fashion industry.

3.7.4 Managerial Insights

Incremental customer expenditure due to the promotion, is an important metric to measure the return on investment of the promotion. Our analysis of gift card promotion indicates that retailers should factor into this calculation the threshold structure and delayed redemption aspect of gift card promotion. The threshold structure impacts only a portion of the customer expenditure distribution and hence, comparing expenditure of customers who participate in the promotion and otherwise, can significantly exaggerate this value. The empirical approach outlined in the study provides a straightforward method for managers to determine how much of the increased customer expenditure can be attributed to the promotion. While unredeemed gift cards are generally profitable for gift card promotion, we find that customers who redeem spend significantly more than customers who do not, even after accounting for the gift card face value. Therefore, it may be beneficial for the retailer to send regular reminders about an unredeemed gift card.

In context of fashion and apparel industry, we find that customers who participate in the promotion have a shorter purchase cycle, i.e, they are generally back on the market earlier than others. One plausible explanation is that some of these customers obtain greater utility from consuming products in-fashion (which are usually at regular price). Gift card promotion provides an incentive to the price-sensitive customers within this segment to make a purchase during the promotion. Therefore, retailer may find that gift card promotion is an effective mechanism in categories where there is heterogeneity in fashion-sensitivity of customer population. This could potentially also explain the widespread usage of gift card promotion by fashion and consumer electronic retailers.

Results from the promotion response model quantify the elasticity in participation and resulting expenditure, with respect to promotion characteristics, such as lowest expenditure threshold and perceived discount offered in the promotion. On the one hand, lower expenditure threshold boosts participation but may not result in increased revenues since customers may not increase their expenditure to participate. On the other hand, higher expenditure threshold can detract customers from participation although those that participate may significantly increase their expenditure. This tradeoff suggests that there is an optimal value, maximizing revenues resulting from participation in the promotion.

Our results also suggest that gift card promotions are more effective when they are held during the regular season, as opposed to markdown season, to avoid potential cannibalization.

This is because customers tend to prefer instantaneous reward obtained by purchasing a deeply discounted product rather than a delayed incentive obtained by participating in a gift card promotion.

We find evidence for a positive effect of advertising gift card promotion on customers who are inherently not interested in participating in the promotion. However, this effect is limited to the customer expenditure rather than the purchase probability of customers. This finding suggests that the retailer may benefit from highlighting products which are on regular price during the promotion. Such an advertisement may increase customer attention towards the regular priced products during the promotion, potentially resulting in a sale.

3.8 Conclusion

Gift card promotions are a new promotion vehicle, quickly gaining traction in the retail industry. They work by providing customers a delayed incentive to buy products at regular price. The incentive is a gift card that is redeemable at the retail store in the future. The appeal of running a gift card promotion is that it encourages customer to spend more on regular priced products and also locks their future expenditure at the retail store. In fashion industry, where product life cycles are short, margins are high, gift card promotions hold a lot of promise for retailers. Our main objective has been to test whether the touted benefits of gift card promotions—due to which retailers are increasingly offering them—are in fact valid or not. To this end, we analyze a novel dataset from a major U.S.-based department store.

Based on existing theory, we hypothesize and find significant positive impact of participation in gift card promotion and of the advertisement effect of the promotion, on the customer response. Gift card promotion boosts customer purchase probability and their expected expenditure during the promotion. In particular, we find that gift card promotion increases customer purchase probability by 17.54% and customer expenditure by 31.45%. Majority of the realized benefits can be attributed to the participation in the promotion. This is in sharp contrast to other recent studies, which have found a dominant advertisement effect in the context of online sales promotions.

The unique aspect of the gift card promotion is that customers need to make additional purchases in the future to redeem their gift card. We find that this redemption mechanism is profitable for the retailer, both in terms of increasing customer responsiveness and their redemption expenditure. Close to 50% of customers return to the website to redeem their gift cards. These customers also spend significantly more than the face value of the gift card during the redemption stage. This positive effect of redemption is robust even after taking into the customer expenditure during the qualification stage. This provides strong evidence for gift-card induced spending. That is, customers who redeem their gift cards do not necessarily treat it as a promotional reward during the redemption stage. Rather they are likely to treat as a gift leading to greater expenditure.

Our analysis of gift card promotion suggests several opportunities to further optimize the planning and implementation of gift card promotions. First, our observation that gift card promotions can incentivize fashion-sensitive customers to participate in the promotion implies that there is scope for targeting the promotion to a smaller customer segment. Doing so can further increase the effectiveness of the promotion. Second, it would be worth exploring how the benefit of gift card promotion varies with the fashion-sensitivity in the product category. This can potentially explain why gift card promotions are more commonly offered by consumer electronic and fashion retailers compared to grocery stores. Third, our results indicate that retailers may benefit from better design of gift card promotion. This study estimates the aggregate effect of all the expenditure thresholds on customer response. However, it would also be interesting to explore how different thresholds contribute to the overall lift due to the promotion. Understanding the effect of using multiple thresholds may assist retailers in better designing (for example, using a non-linear incentive scheme) gift card promotions.

CHAPTER 4

CONCLUSION

In this dissertation, we investigated how firms can better manage their operations by providing incentives to different economic agents along the value chain. We first considered the inventory control problem faced by retailers and their suppliers in the upstream value chain. In particular we consider how firms can maximize their gains from participating in a collaborative inventory management practices, such as vendor-managed inventory. We propose a learn and screen mechanism in this setting, which allows suppliers to incorporate what they learn from demand realizations into designing contracts for retailers. These contracts facilitate integration of local knowledge available with the retailer into a centralized inventory management process. The learn and screen mechanism not only stipulates the contract terms but also the optimal time at which to offer the contracts in an ongoing supply chain relationship. Our results indicate that incorporating the learn and screen approach can create a win-win outcome by significantly boosting the supplier's profits while also improving the gains for the retailer.

The second incentive scheme we explored in this dissertation is a gift card promotion mechanism which is being widely used by retailers today. Unlike standard price promotions mechanisms, the price of a product remains unchanged in a gift card promotion. Rather the incentive offered to the customer is delayed. In collaboration with a major U.S.-based fashion retailer, we empirically investigate (in an online context) if indeed delayed incentives boost customer demand during the promotion. To accurately benchmark the performance of the promotion, we model customer decision making process using detailed information about their online shopping behavior during promotion and non-promotion periods. Our empirical models strongly suggest that delayed incentive in the form of a gift card can be effective in increasing customer expenditure and their purchase probability during the promotion. Interestingly, we also note that the benefit of delayed incentive for the retailer can extend beyond the promotion period. Customers who return to redeem their gift card tend to spend significantly more than the value of the gift card. While there are significant benefits associated with gift card promotions, we also note that these can be contingent on the type of product category for which they are implemented.

In summary, this dissertation extends the current body of work relating to the role of incentives in value chain management in two important ways. First, we provide frameworks for incorporating real-time data into design and implementation of incentive mechanisms. Second, this work highlights that timing of the incentive provided to economic agents can be a strategic decision for firms.

APPENDIX A

APPENDIX TO CHAPTER 1

A.1 Glossary of Notation

Cost related parameters		**Demand and forecasts**	
r	unit retail price	ξ	demand characteristic
w	unit wholesale price	$\Theta := [\underline{\xi}, \overline{\xi}]$	possible values ξ takes
c	unit cost of production	$q_n(\cdot), Q_n(\cdot)$	predictive demand density and distribution in period n, respectively
h	unit holding cost per period	$g(\cdot\|\xi), G(\cdot\|\xi)$	p.d.f. and c.d.f. of demand, respectively
α	discounting factor	$\pi_n(\cdot)$	supplier's belief (p.d.f.) beginning of period n
N	number of selling periods		

Decision Variables	
y_n	inventory level in period n after ordering decision
$\{S_n^*(\cdot), P_n^*(\cdot)\}$	optimal menu of contracts offered in period n
$S^{fb}(\cdot)$	optimal base-stock levels under symmetric demand information using the linear coordinating contract
p^{fb}	coordinating price under symmetric information
$S^{sb}(\cdot)$	optimal base-stock levels under symmetric demand information
τ	optimal time to offer screening contracts

A.2 Simplification of the DP

Using the definitions of transformations V_n, Π_n^{sr}, we can simplify $\tilde{L}_n(y, \pi_n)$ as follows

$$\tilde{L}_n(y, \pi_n) := cx_n + (w-c)y_n - (w+h)\int_0^y Q_n(z)\,\mathrm{d}z + \alpha \cdot \overline{Q}_n(z) \cdot \tilde{V}_{n+1}\big(0, \pi_{n+1}^c(\cdot|y)\big)$$
$$+ \alpha \int_0^y q_n(z) \cdot \tilde{V}_{n+1}\big(y-z, \pi_{n+1}^e\big)\,\mathrm{d}z$$

$$
\begin{aligned}
&= cx_n + (w-c)y_n - (w+h)\int_0^y Q_n(z)\,\mathrm{d}z + \alpha\cdot\overline{Q}_n(z)\cdot V_{n+1}\big(0,\,\pi^c_{n+1}(\cdot|y)\big) \\
&\qquad + \alpha\int_0^y q_n(z)\cdot\big[V_{n+1}\big(y-z,\,\pi^e_{n+1}\big) + c(y-z)\big]\,\mathrm{d}z \\
&= cx_n + (w-c)y_n - (w+h)\int_0^y Q_n(z)\,\mathrm{d}z + \alpha\cdot\overline{Q}_n(z)\cdot V_{n+1}\big(0,\,\pi^c_{n+1}(\cdot|y)\big) \\
&\quad + \alpha\int_0^y q_n(z)\cdot V_{n+1}\big(y-z,\,\pi^e_{n+1}\big)\,\mathrm{d}z + c\alpha\int_0^y q_n(z)\cdot(y-z)\,\mathrm{d}z \\
&= cx_n + (w-c)y_n - (w+h-\alpha c)\int_0^y Q_n(z)\,\mathrm{d}z + \alpha\cdot\overline{Q}_n(z)\cdot V_{n+1}\big(0,\,\pi^c_{n+1}(\cdot|y)\big) \\
&\quad + \alpha\int_0^y q_n(z)\cdot V_{n+1}\big(y-z,\,\pi^e_{n+1}\big)\,\mathrm{d}z \\
&= cx_n + L_n(y,\,\pi_n)
\end{aligned}
$$

A.3 Proofs

Proof. Proof of Lemma 2.4.1. Let $\{S(\cdot),\,P(\cdot)\}$ satisfy (IC) and (PC). The (PC) implies that $\Pi^r_n(S(\underline{\xi}),\,P(\underline{\xi}),\,\underline{\xi}) \geq \Pi^r_{\min}(n)$. Since the supplier's profit is increasing in $P(\underline{\xi})$ and the retailer's is decreasing in $P(\underline{\xi})$, the supplier does not violate the constraints by setting $P(\underline{\xi})$ in such a way that $\Pi^r_n(S(\underline{\xi}),\,P(\underline{\xi}),\,\underline{\xi}) = \Pi^r_{\min}(n)$. The retailer's total expected profit can be simplified as follows.

$$
\begin{aligned}
\Pi^r_n\big(S,\,P,\,\xi\big) &= \frac{(1-\alpha^{N-n+1})}{1-\alpha}\cdot\Big\{\mathbb{E}\big[(r-w)\cdot\min\{S,\,D_i\}\big] - P\Big\} \\
&= \frac{(1-\alpha^{N-n+1})}{1-\alpha}\cdot\Big\{(r-w)\cdot\Big[S - \int_0^S G(z|\xi)\mathrm{d}\,z\Big] - P\Big\}. \qquad \text{(A.1)}
\end{aligned}
$$

From the above equation we get, $\dfrac{\mathrm{d}\Pi^r_n(S,P,\xi)}{\mathrm{d}\xi} > 0$. Then, for any $\xi_1,\,\xi_2 \in \Theta$ such that $\xi_1 > \xi_2$, $\Pi^r_n(S,P,\xi_1) > \Pi^r_n(S,P,\xi_2)$ for any S, P. In particular taking supremum (over the entire menu offered) on the rhs of the last inequality first and then on the lhs, we get $\Pi^r_n(S(\xi),P(\xi),\xi_1) > \Pi^r_n(S(\xi),P(\xi),\xi_2)$. Thus as long as $\Pi^r_n(S(\underline{\xi}),P(\underline{\xi}),\underline{\xi}) = \Pi^r_{\min}(n)$, the remaining (PC) constraints are redundant.

It follows from Equation (A.1), (IC), and the envelope theorem

$$
\begin{aligned}
\frac{\partial}{\partial\zeta}\Pi^r_n(S(\zeta),P(\zeta),\zeta) &= \frac{\partial}{\partial\zeta}\Pi^r_n(S(\eta),P(\eta),\zeta)\Big|_{\eta=\zeta} \\
&= \frac{(1-\alpha^{N-n+1})}{1-\alpha}\cdot\frac{\partial}{\partial\zeta}\Big[(r-w)\cdot\Big\{S(\eta) - \int_0^{S(\eta)} G(z|\zeta)\,\mathrm{d}z\Big\} - P(\eta)\Big]\Big|_{\eta=\zeta}
\end{aligned}
$$

$$= -\frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)\cdot \int_0^{S(\eta)} \frac{\partial}{\partial \zeta} G(z|\zeta)\,\mathrm{d}z\Big|_{\eta=\zeta}.$$

This gives,

$$\Pi_n^r(S(\xi),P(\xi),\xi) = \Pi_{\min}^r(n) - \frac{(1-\alpha^{N-n+1})}{1-\alpha}\cdot(r-w)\int_{\underline{\xi}}^{\xi}\int_0^{S(\eta)} \frac{\partial}{\partial \eta} G(z|\eta)\,\mathrm{d}z\,\mathrm{d}\eta.$$

To prove (ii), note that $\frac{\mathrm{d}^2\Pi_n^r(S,P,\xi)}{\mathrm{d}S^2} = -\frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)g(S|\xi) < 0$ and $\frac{\mathrm{d}^2\Pi_n^r(S,P,\xi)}{\mathrm{d}S\mathrm{d}\xi} = -\frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)\cdot\frac{\partial G(S|\xi)}{\mathrm{d}\xi} > 0$. Then $S'(\cdot) > 0$. Else for $\xi_1 > \xi_2$,

$$0 = \frac{\partial \Pi_n^r(S,P,\xi_1)}{\partial S}\Big|_{S=S(\xi_1)} > \frac{\partial \Pi_n^r(S,P,\xi_1)}{\partial S}\Big|_{S=S(\xi_2)} > \frac{\partial \Pi_n^r(S,P,\xi_2)}{\partial S}\Big|_{S=S(\xi_2)}.$$

The first equality follows from (IC) and the two inequalities due to the signs of the second derivates. The last inequality contradicts (IC), since $S(\xi_2)$ is the maximizer for type ξ_2.

To prove (i) and (ii) imply (IC) and (PC)

$$\begin{aligned}
&\Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\xi)\\
&= \int_{\underline{\xi}}^{\xi} \frac{\mathrm{d}}{\mathrm{d}x}\Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),x)\,\mathrm{d}x + \Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\underline{\xi})\\
&= \int_{\underline{\xi}}^{\hat{\xi}} \frac{\mathrm{d}}{\mathrm{d}x}\Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),x)\,\mathrm{d}x + \int_{\tilde{\xi}}^{\xi} \frac{\mathrm{d}}{\mathrm{d}x}\Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),x)\,\mathrm{d}x + \Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\underline{\xi})\\
&= \Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\tilde{\xi}) - \Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\underline{\xi}) + \int_{\tilde{\xi}}^{\xi} \frac{\mathrm{d}}{\mathrm{d}x}\Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),x)\,\mathrm{d}x + \Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\underline{\xi})\\
&= \Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\tilde{\xi}) - \frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)\int_{\hat{\xi}}^{\xi}\int_0^{S(\tilde{\xi})} \frac{\delta}{\delta x}G(z|x)\mathrm{d}z\,\mathrm{d}x\\
&= \Pi_n^r(S(\tilde{\xi}),\ P(\tilde{\xi}),\ \tilde{\xi}) - \frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)\int_{\tilde{\xi}}^{\xi}\left[\int_0^{S(\tilde{\xi})} \frac{\delta}{\delta x}G(z|x)\mathrm{d}z - \int_0^{S(x)} \frac{\delta}{\delta x}G(z|x)\mathrm{d}z\right]\mathrm{d}x\\
&\qquad - \frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)\int_{\tilde{\xi}}^{\xi}\int_0^{S(x)} \frac{\delta}{\delta x}G(z|x)\mathrm{d}z\,\mathrm{d}x\\
&= \Pi_n^r(S(\tilde{\xi}),\ P(\tilde{\xi}),\ \tilde{\xi}) - \frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)\int_{\tilde{\xi}}^{\xi}\left[\int_0^{S(\tilde{\xi})} \frac{\delta}{\delta x}G(z|x)\mathrm{d}z - \int_0^{S(x)} \frac{\delta}{\delta x}G(z|x)\mathrm{d}z\right]\mathrm{d}x\\
&\qquad + \Pi_n^r(S(\xi),\ P(\xi),\ \xi) - \Pi_n^r(S(\tilde{\xi}),P(\tilde{\xi}),\tilde{\xi})\\
&= \Pi_n^r(S(\xi),P(\xi),\xi) - \frac{(1-\alpha^{N-n+1})}{1-\alpha}(r-w)\int_{\tilde{\xi}}^{\xi}\int_{S(x)}^{S(\tilde{\xi})} \frac{\delta}{\delta x}G(z|x)\mathrm{d}z\mathrm{d}x.
\end{aligned}$$

When $\xi > \tilde{\xi}$, $S(x) > S(\tilde{\xi})$ for all x in $(\tilde{\xi}, \xi)$. Hence the value of the second integral is positive, thus it is not optimal for the retailer of type ξ to choose $S(\tilde{\xi}), P(\tilde{\xi})$. By a similar argument, one can rule out the case $\xi < \tilde{\xi}$. Together they imply (IC). The (PC) can obtained by setting $\xi = \underline{\xi}$ in (i) of the lemma. $\square$

Proof. Proof of Lemma 2.4.2. The retailer of type ξ accepts contract $S(\xi)$, $P(\xi)$, therefore $y_j = S(\xi)$, for all $j \geq n$. The total surplus generated in the supply chain from periods n through N is (we denote $S(\xi) = S$)

$$\begin{aligned}
&\sum_{j=n}^{N} \alpha^{j-1} \Big[r \cdot \min\{S,\, D_j\} - c(S - x_j) - h(S - D_j)^+ \Big] + \alpha^{N-n+1} c[S - D_N]^+ \\
&= cx_n + \sum_{j=n}^{N} \alpha^{j-1} \cdot \Big[(r-c) \cdot S - (r+h) \cdot [S - D_j]^+ \Big] + c \sum_{j=n}^{N-1} \alpha^{j} [S - D_j]^+ + \alpha^{N-n+1} c[S - D_N]^+ \\
&= cx_n + \sum_{j=n}^{N} \alpha^{j-1} \cdot \Big[(r-c) \cdot S - (r+h-\alpha c) \cdot [S - D_j]^+ \Big].
\end{aligned}$$

The total expected profit of the supply chain is

$$\begin{aligned}
\Pi_n^{tot}(S(\xi)|\, x_n, \pi_n) &= cx_n + \sum_{j=n}^{N} \alpha^{j-1} \mathbb{E}_D \Big[(r-c) \cdot S(\xi) - (r+h-\alpha c) \cdot [S(\xi) - D_j]^+ \Big] \\
&= cx_n + \gamma(n) \cdot \Big[(r-c) \cdot S(\xi) - (r+h-\alpha c) \cdot \int_0^{S(\xi)} G(z|\xi)\, \mathrm{d}z \Big]. \qquad \text{(A.2)}
\end{aligned}$$

Any incentive compatible contract can be equivalently characterized in terms of the profits they generate for the retailer Equation (2.8). When the retailer's type is ξ, the supplier's profit for any incentive-compatible menu $\{S(\cdot),\, P(\cdot)\}$ is the difference in profits, $\Pi_n^{tot}(S(\xi)|x_n, \pi_n) - \Pi_n^r(S(\xi),\, P(\xi),\, \xi)$. The supplier has only a belief π_n on ξ, hence her expected profits are (using Equations (A.2) and (2.8)) $\mathbb{E}_\xi\big[\Pi_n^{tot}(S(\xi)|x_n, \pi_n) - \Pi_n^r(S(\xi),\, P(\xi),\, \xi)\big] = cx_n - \Pi_{\min}^r(n) + \gamma(n) \cdot \mathbb{E}_\xi \Big[(r-c)S(\xi) - (r+h-\alpha c) \cdot \int_0^{S(\xi)} G(x|\xi)\, \mathrm{d}x + (r-w) \cdot \int_{\underline{\xi}}^{\xi} \int_0^{S(\eta)} \frac{\partial}{\partial \eta} G(z|\eta) \mathrm{d}z\, \mathrm{d}\eta \Big]$. The following simplification is needed for the last term:

$$\int_{\underline{\xi}}^{\bar{\xi}} \Bigg(\overbrace{\int_{\underline{\xi}}^{\xi} \mathrm{d}\eta \int_0^{S(\eta)} \frac{\partial}{\partial \eta} G(z|\eta) \mathrm{d}z}^{u} \Bigg) \cdot \overbrace{\pi(\xi) \mathrm{d}\xi}^{\mathrm{d}v}$$

$$
\begin{aligned}
&= \int_{\underline{\xi}}^{\xi} \pi(\eta)\,\mathrm{d}\eta \cdot \int_{\underline{\xi}}^{\xi} \mathrm{d}\eta \int_0^{S(\eta)} \frac{\partial}{\partial\eta} G(z|\eta)\mathrm{d}z \bigg|_{\underline{\xi}}^{\overline{\xi}} - \int_{\underline{\xi}}^{\overline{\xi}} \int_{\underline{\xi}}^{\xi} \pi(\eta)\,\mathrm{d}\eta \cdot \int_0^{S(\xi)} \frac{\partial}{\partial\xi} G(z|\xi)\mathrm{d}z\mathrm{d}\xi \\
&= \int_{\underline{\xi}}^{\overline{\xi}} \mathrm{d}\eta \int_0^{S(\eta)} \frac{\partial}{\partial\eta} G(z|\eta)\mathrm{d}z\mathrm{d}\eta - \int_{\underline{\xi}}^{\overline{\xi}} \int_{\underline{\xi}}^{\xi} \pi(\eta)\,\mathrm{d}\eta \cdot \int_0^{S(\xi)} \frac{\partial}{\partial\xi} G(z|\xi)\mathrm{d}z\mathrm{d}\xi \\
&= \int_{\underline{\xi}}^{\overline{\xi}} \int_{\xi}^{\overline{\xi}} \pi(\eta)\,\mathrm{d}\eta \cdot \int_0^{S(\xi)} \frac{\partial}{\partial\xi} G(z|\xi)\mathrm{d}z\mathrm{d}\xi.
\end{aligned}
$$

□

Lemma A.3.1. *Let $f : [a,\, b] \to \mathbb{R}$ be a twice differentiable function over $(a,\, b)$. If the following two conditions hold: (i) $\exists\, \hat{x} \in (a,\, b)$ such that $f_{xx} < 0$ for all $x < \hat{x}$, $f_{xx} > 0$ for all $x > \hat{x}$ and $f_{xx}(\hat{x}) = 0$, (ii) $f_x(a+) > 0$ and $f_x(b-) < 0$, the following statements are true.*

1. *f_x has a unique root x^* in $(a,\, b)$ such that $x^* < \hat{x}$.*

2. *The function f is unimodal with peak at $x = x^*$.*

Proof. Proof of Lemma A.3.1.

1. Since $f_x(a) > 0$, $f_x(b) < 0$, and f_x is continuous, there exists at least one root of f_x in $(a,\, b)$. Suppose η_1, η_2 are two roots, such that $\eta_1 < \eta_2$, without loss of generality. Then, $f_x(\eta_1) = f_x(\eta_2) = 0$. By the Mean Value Theorem, there exists $u \in (\eta_1,\, \eta_2)$ such that $f_{xx}(u) = 0$. There are three possibilities, (i) $u > \hat{x}$, (ii) $u < \hat{x}$, or (iii) $u = \hat{x}$. The first two possibilities result in a contradiction, since f_{xx} has a unique root. Hence, $u = \hat{x}$. This implies $\eta_1 < \hat{x} < \eta_2$. The function $f_{xx} > 0$ for $x > \hat{x}$. Thus, $f_x(y) > f_x(\eta_2) = 0$ for all $y > \eta_2$. Taking limit as $y \to b$ we get $f_x(b-) \geq 0$ which contradicts assumption (ii) in the lemma. Thus, there exists a unique root for f_x, denoted by x^*. Suppose $x^* \geq \hat{x}$. Using assumption (ii) we get $f_x(y) > f_x(x^*) = 0$ for all $y > x^*$. Taking limit on $y \to b$, we get a contradiction. Hence $x^* < \hat{x}$.

2. Since $f_x(a+) > 0$ by assumption (ii), it follows from the part (i) of the proof that $f_x > 0$ for all $x < x^*$, $f_x < 0$ for $x > x^*$ and $f_x(x^*) = 0$. This clearly implies the function f is a *unimodal* function with its maximum attained at $x^* < \hat{x}$.

□

Proof. Proof of Theorem 2.4.1.

1. Differentiating $G(z|\xi)$ wrt ξ yields,

$$\frac{\partial G(z|\xi)}{\partial \xi} = \frac{\partial}{\partial \xi}\int_0^z g(y|\xi)\,\mathrm{d}y = \int_0^z \frac{\partial}{\partial \xi} g(y|\xi)\,\mathrm{d}y.$$

The result hence is determined by the behavior of $\frac{\partial}{\partial \xi} g(y|\xi)$ as a function of y. $\frac{\partial}{\partial \xi} g(y|\xi)$ changes sign at least once since any two density functions cross each other (and the total area under any pdf is one). If the sign change happens exactly once, and is from negative to positive, then first order stochastic dominance follows (see Lemma A.3.2). For a pdf from the exponential family, $\frac{\partial}{\partial \xi} g(y|\xi) = h(y)e^{-t(y)\cdot w(\xi)} \cdot \left[c'(\xi) - c(\xi)t(y)w'(\xi)\right]$. Since c', $w' < 0$ and $t' > 0$ (assumptions in the lemma), it follows, $\left[c'(\xi) - c(\xi)t(y)w'(\xi)\right]$ is initially negative, becomes zero at

$$t(y^*(\xi)) := \frac{c'(\xi)}{c(\xi)w'(\xi)} > 0$$

and remains positive thereafter. Note that since t is increasing and $t(0) = 0$, we have $y^*(\xi) > 0$. Lemma A.3.2 guarantees that $\int_0^z \frac{\partial}{\partial \xi} g(y|\xi)\,\mathrm{d}y \leq 0$ for all $\xi \in \Theta$ and $z \in \mathbb{R}^+$, which in turn implies first order stochastic dominance of the family of demand distributions considered in the Theorem 2.4.1.

Lemma A.3.2. *Let $f(\cdot)$ be a continuous function on $\mathbb{R}^+$ such that, $f(x) < 0$, for all $x < \hat{x}$, $f(\hat{x}) = 0$ and $f(x) > 0$, for all $x > \hat{x}$. Define $k(z) := \displaystyle\int_0^z f(y)\mathrm{d}\,y$. If $\lim_{z\to\infty} k(z) = 0$, then $k(z) \leq 0$ for all $z \in \mathbb{R}^+$.*

Proof. Proof of Lemma A.3.2. Suppose not, i.e. $\exists\, \tilde{z} \in \mathbb{R}^+$ such that $k(\tilde{z}) > 0$. Then $\tilde{z} > \hat{x}$ otherwise, $k(\tilde{z}) \leq 0$ since $f(\cdot)$ is negative until $\hat{x}$. This implies $k(\tilde{z}) \leq k(\tilde{z}) + \displaystyle\int_{\tilde{z}}^\infty f(x)\,\mathrm{d}x = \int_0^\infty f(x)\,\mathrm{d}x = k(\infty) = 0$. The first inequality follows since $f(x)$ is strictly positive for $x > \tilde{z} > \hat{x}$ and the last equality follows by the assumption in the lemma. Thus we have arrived at a contradiction. Therefore $k(z) \leq 0$ for all $z \in \mathbb{R}^+$.

□

Note that the assumptions in the above lemma are satisfied by $\frac{\partial}{\partial \xi} g(x|\xi)$, since $\int_0^\infty \frac{\partial}{\partial \xi} g(x|\xi) \mathrm{d}x = 0$.

2. The first derivative of $H_n(\cdot, \xi)$[1] is:

$$\frac{\partial}{\partial S} H_n(S, \xi) = (r - c) - (r + h - \alpha c) G(S|\xi) + \frac{r - w}{\lambda_n(\xi)} \frac{\partial}{\partial \xi} G(S|\xi).$$

Note that, $\lim_{z\to\infty} \frac{\partial G(z|\xi)}{\partial \xi} = \lim_{z\to\infty} \frac{\partial}{\partial \xi} \int_0^z g(y|\xi)\,\mathrm{d}y = \lim_{z\to\infty} \int_0^z \frac{\partial}{\partial \xi} g(y|\xi)\,\mathrm{d}y = \int_0^\infty \frac{\partial}{\partial \xi} g(y|\xi)\,\mathrm{d}y = 0$. The last equality follows since, $\int_0^\infty g(z|\xi)\,\mathrm{d}z = 1, \ \forall \xi \in \Theta \implies \frac{\partial}{\partial \xi} \int_0^\infty g(z|\xi)\,\mathrm{d}z = 0 \implies \int_0^\infty \frac{\partial}{\partial \xi} g(z|\xi)\,\mathrm{d}z = 0, \ \ \forall \xi \in \Theta$. Thus, $\dfrac{\partial}{\partial S} H_n(0, \xi) = r - c > 0$ and $\dfrac{\partial}{\partial S} H_n(\infty, \xi) = -c(1-\alpha) - h < 0$. Since $H_n(\cdot, \xi)$ is continuous, there exists at least one critical point where it equals zero. The second derivative of $H_n(\cdot, \xi)$ is

$$\frac{\partial^2}{\partial S^2} H_n(S, \xi) = -(r + h - \alpha c) \left\{ g(S|\xi) - \frac{\varpi}{\lambda_n(\xi)} \frac{\partial g(S|\xi)}{\partial \xi} \right\},$$

where $\varpi := \frac{r-w}{r+h-\alpha c} < 1$. We evaluate the term inside the brackets on the rhs for a density from the exponential family.

$$\left\{ g(S|\xi) - \frac{\varpi}{\lambda_n(\xi)} \frac{\partial g(S|\xi)}{\partial \xi} \right\} = h(S) \cdot e^{-t(S)\cdot w(\xi)} \left[c(\xi) - \frac{\varpi}{\lambda_n(\xi)} \big(c'(\xi) - c(\xi) t(S) w'(\xi) \big) \right],$$

where c', w' denote derivatives. Then, $\frac{\partial^2 H_n(S,\xi)}{\partial S^2} > 0 \iff t(S) > \frac{-\lambda_n(\xi)}{\varpi w'(\xi)} + \frac{c'(\xi)}{c(\xi) w(\xi)} \iff S > \overline{S}_n(\xi)$, where

$$\overline{S}_n(\xi) := t^{-1} \left(\frac{-\lambda_n(\xi)}{\varpi w'(\xi)} + \frac{c'(\xi)}{c(\xi) w(\xi)} \right). \tag{A.3}$$

The first inequality follows since $-w' > 0$ and $c' < 0$ by our assumption. Since $t(\cdot)$ is increasing, t^{-1} exists and is monotone. Note that, $\overline{S}_n(\xi) > 0$ for all $\xi \in \Theta$. All the conditions in Lemma A.3.1 are satisfied and Part 2 follows.

3. It follows from unimodality of $H(\cdot, \xi|\pi_n)$ that, $\frac{\partial H_n(\hat{S}_n(\xi), \xi)}{\partial S} = 0$ characterizes the unique maximizer of $H(\cdot, \xi|\pi_n)$ and $\max\{x_n, \hat{S}_n(\xi)\}$ is the maximizer of the constrained problem.

$\square$

[1] We use $H_n(S, \xi)$ and $H(S, \xi|\pi_n)$ interchangeably to simplify notation wherever necessary.

Proof. Proof of Theorem 2.4.2. Since $\hat{S}_n(\xi)$ satisfies $\dfrac{\partial}{\partial S}H_n(\hat{S}_n(\xi),\,\xi)=0$, for all $\xi\in\Theta$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial}{\partial S}H_n(\hat{S}_n(\xi),\,\xi)=\frac{\partial^2}{\partial S^2}H_n(\hat{S}_n(\xi),\,\xi)\frac{\mathrm{d}\,\hat{S}_n(\xi)}{\mathrm{d}\,\xi}+\frac{\partial^2}{\partial S\partial\xi}H_n(\hat{S}_n(\xi),\,\xi)=0. \tag{A.4}$$

We first establish $\hat{S}_n(\xi)\le\overline{S}_n(\xi)$ (defined in Equation (A.3)). Since $w(\xi)=c(\xi)=\frac{1}{\xi}$ and $h(z)=1$ in the case of newsvendor family of distributions, the definition of $\overline{S}_n(\xi)$ simplifies to $\overline{S}_n(\xi)=t^{-1}\big(\frac{\xi^2\lambda_n(\xi)}{\varpi}+\xi\big)$. Substituting this into $\frac{\partial H_n(S,\xi)}{\partial S}$ gives $\frac{\partial H_n(\overline{S}_n(\xi),\xi)}{\partial S}=(r-c)-(r+h-\alpha c)\big(1-e^{-(\frac{\xi\lambda_n(\xi)}{\varpi}+1)}\big)-\frac{r-w}{\lambda_n(\xi)}e^{-(\frac{\xi\lambda_n(\xi)}{\varpi}+1)}\big(\frac{1}{\xi}+\frac{\lambda_n(\xi)}{\varpi}\big)=-h-c(1-\alpha)-\frac{r-w}{\lambda_n(\xi)\xi}e^{-(\frac{\xi\lambda_n(\xi)}{\varpi}+1)}<0$. It then follows from unimodality of $H_n(\cdot,\xi)$ that $\hat{S}_n(\xi)<\overline{S}_n(\xi)$. From Equation (A.3) it also follows that $\frac{\partial^2\,H_n(\hat{S}_n(\xi),\xi)}{\partial S^2}<0$ for all ξ. Finally, $\frac{\partial^2\,H_n(\hat{S}_n(\xi),\xi)}{\partial S\partial\xi}=-(r+h-\alpha c)\frac{\partial G(\hat{S}_n|\xi)}{\partial\xi}+(r-w)\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\frac{\int_\xi^{\bar{\xi}}\pi_n(\eta)\mathrm{d}\eta}{\pi_n(\xi)}\right)\frac{\partial G(\hat{S}_n|\xi)}{\partial\xi}+(r-w)\frac{\int_\xi^{\bar{\xi}}\pi_n(\eta)\mathrm{d}\eta}{\pi_n(\xi)}\frac{\partial^2}{\partial\xi^2}G(\hat{S}_n|\xi)$. If $G(\cdot|\xi)$ is from the newsvendor family, $\frac{\partial^2 G(\hat{S}_n|\xi)}{\partial\xi^2}=t(\hat{S}_n)\cdot\xi^{-4}\cdot e^{\frac{-t(\hat{S}_n)}{\xi}}\big(2\xi-t(\hat{S}_n)\big)$.

To prove non-negativity of the last factor: $2\xi-t(\hat{S}_n(\xi))\ge 0\iff\hat{S}_n(\xi)\le t^{-1}(2\xi)\iff G(\hat{S}_n(\xi)|\xi)\le G(t^{-1}(2\xi)|\xi)$ for all $\xi\in\Theta$. From Equation (2.11) it follows, $G(\hat{S}_n(\xi)|\xi)\le\frac{(r-c)}{r+h-\alpha c}$ since $\frac{\partial G(\hat{S}_n(\xi)|\xi)}{\partial\xi}\le 0$. Thus, if $G(t^{-1}(2\xi)|\xi)=1-e^{-2}\ge\frac{(r-c)}{r+h-\alpha c}$, then $2\xi-t(\hat{S}_n(\xi))>0$. Therefore the condition on cost parameters along with the first order stochastic dominance of $G(S|\xi)$ implies $\frac{\partial^2}{\partial S\partial\xi}H_n(\hat{S}_n(\xi),\,\xi)$ is positive. This along with (A.4), implies $\hat{S}_n(\xi)$ is increasing. It follows that $S_n^*=\max\{\hat{S}_n(\xi),x_n\}$ is increasing as well. For any feasible menu of base-stock levels $S_n(\xi)$ offered to the retailer in period n, the corresponding payments can be determined by combining (2.8) and (A.1),

$$P_n(\xi)=(r-w)\left(S_n(\xi)-\int_0^{S_n(\xi)}G(z|\xi)\,\mathrm{d}z+\int_{\underline{\xi}}^{\xi}\int_0^{S_n(\eta)}\frac{\partial}{\partial\eta}G(z|\eta)\,\mathrm{d}z\,\mathrm{d}\eta\right)-\frac{\Pi^r_{\min}(n)}{\gamma(n)}. \tag{A.5}$$

□

Proof. Proof of Theorem 2.4.3. If in period n the supplier decides to wait an additional period before offering the menu of contracts, she raises inventory level in period n to y_n. Two scenarios are possible in period n:

1. If the sales observations in period n are censored i.e., $y_n = z_n$ we have $\lambda_{n+1}(\xi) = \frac{\overline{G}(y_n|\xi)\cdot\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \overline{G}(y_n|\eta)\cdot\pi_n(\eta)\,\mathrm{d}\eta} = \frac{\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \frac{\overline{G}(y_n|\eta)}{\overline{G}(y_n|\xi)}\cdot\pi_n(\eta)\,\mathrm{d}\eta} \leq \frac{\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \pi_n(\eta)\,\mathrm{d}\eta} = \lambda_n(\xi)$. The first equality follows from definition of $\lambda_{n+1}(\cdot)$. The inequality follows from first order stochastic dominance of demand distributions, $\overline{G}(y_n|\eta) \geq \overline{G}(y_n|\xi)$ for all $\eta \in [\xi, \overline{\xi}]$. This implies the following $\frac{\partial H_{n+1}(S,\xi)}{\partial S} \leq \frac{\partial H_n(S,\xi)}{\partial S}$ for all $S \in \mathbb{R}^+$ and $\xi \in \Theta$. $H_{n+1}(\cdot, \xi)$ is unimodal by Theorem 2.4.1. Thus $\frac{\partial H_{n+1}(\hat{S}_n(\xi),\xi)}{\partial S} \leq \frac{\partial H_n(\hat{S}_n(\xi),\xi)}{\partial S} = 0$. This implies $S^*_{n+1}(\xi) = \hat{S}_{n+1}(\xi) \leq \hat{S}_n(\xi) \leq \max\{x_n, \hat{S}_n(\xi)\} = S^*_n(\xi)$. The first equality holds since $x_{n+1} = 0$.

2. If the sales observation in period n is uncensored i.e., $z_n < y_n$ we have, $\lambda_{n+1}(\xi) = \frac{g(z_n|\xi)\cdot\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} g(z_n|\eta)\cdot\pi_n(\eta)\,\mathrm{d}\eta} = \frac{\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \frac{g(z_n|\eta)}{g(z_n|\xi)}\cdot\pi_n(\eta)\,\mathrm{d}\eta}$. Therefore, the ratio of densities $\frac{g(z_n|\eta)}{g(z_n|\xi)}$ determines the relation between $\lambda_{n+1}(\xi)$ and $\lambda_n(\xi)$. For the newsvendor family of demand distributions $\frac{\partial g(z_n|\xi)}{\partial \xi} = t'(z_n)\cdot\xi^{-3}\cdot e^{-\frac{t(z_n)}{\xi}}\cdot\big(t(z_n)-\xi\big)$. Since $t'(\cdot) > 0$,

$$\frac{\partial g(z_n|\xi)}{\partial \xi} \quad \begin{cases} \geq 0, & \text{if } \xi \leq t(z_n); \\ < 0, & \text{if } \xi > t(z_n). \end{cases}$$

   (i) For types $\xi > t(z_n)$ it follows $g(z_n|\eta) < g(z_n|\xi)$ for all $\eta > \xi$. Hence, $\lambda_{n+1}(\xi) = \frac{\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \frac{g(z_n|\eta)}{g(z_n|\xi)}\cdot\pi_n(\eta)\,\mathrm{d}\eta} \geq \frac{\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \pi_n(\eta)\,\mathrm{d}\eta} = \lambda_n(\xi)$. Therefore, $\frac{\partial H_{n+1}(S,\xi)}{\partial S} \geq \frac{\partial H_n(S,\xi)}{\partial S}$ and unimodality of $H_{n+1}(\cdot, \xi)$ and $H_n(\cdot, \xi)$ implies $\hat{S}_{n+1}(\xi) \geq \hat{S}_n(\xi)$ for all types $\xi \geq t(z_n)$. In addition, if $x_n \leq \hat{S}_n(\underline{\xi})$, then $S^*_{n+1}(\xi) \geq S^*_n(\xi)$ for all types $\xi \geq t(z_n)$.

   (ii) For types $\xi < t(z_n)$, such that $g(z_n|\xi) \leq g(z_n|\overline{\xi})$, note that $g(z_n|\eta) > g(z_n|\xi)$ for all $\eta > \xi$, since $g(z_n|\eta)$ is increasing upto $t(z_n)$ and decreasing thereafter. Therefore,

$$\lambda_{n+1}(\xi) = \frac{\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \frac{g(z_n|\eta)}{g(z_n|\xi)}\cdot\pi_n(\eta)\,\mathrm{d}\eta} \leq \frac{\pi_n(\xi)}{\int_{\underline{\xi}}^{\overline{\xi}} \pi_n(\eta)\,\mathrm{d}\eta} = \lambda_n(\xi).$$

   By arguments similar to those used in earlier part, we $\hat{S}_{n+1}(\xi) < \hat{S}_n(\xi)$.

□

Proof. Proof of Theorem 2.4.4. In the symmetric setting, $\xi \in \Theta$ is common knowledge. Under a linear price contract, the retailer pays pS for base-stock level S. Subsequently, he pays a fixed wholesale price of w per unit to satisfy demand to the extent possible. The retailer's profit if he chooses base-stock level S is $\gamma(n)\big((r-w-p)S-(r-w)\int_0^S G(z|\xi)\,\mathrm{d}z\big)$. The retailer chooses inventory level that maximizes his profit: $\frac{\partial}{\partial S}\,\gamma(n)\big(r-w-p-(r-w)G(S|\xi)\big)=0$. To coordinate the channel, inventory level must be determined using the critical fractile, $S^{fb}(\xi)=G^{-1}\big(\frac{r-c}{r+h-\alpha c}\big|\xi\big)$. To ensure the retailer chooses the coordinating inventory level $S^{fb}(\xi)$, the supplier sets $G(S|\xi)=G(S^{fb}(\xi)|\xi)=\frac{r-c}{r+h-\alpha c}$. The resulting coordinating marginal price is $p=p^{fb}:=\frac{(r-w)(h+c(1-\alpha)}{r+h-\alpha c}$. The total profit in the coordinated supply chain can be calculated using the following recursion:

$$\begin{aligned} V_n^{cs}(x_n,\pi)=\mathbb{E}_\xi\Bigg[&(r-c)-(r+h-\alpha c)\int_0^{\hat{y}_n(\xi)}G(z|\xi)\,\mathrm{d}z+\alpha\big(1-G(\hat{y}_n(\xi)|\xi)V_{n+1}^{cs}\big(0,\pi\big)\\ &+\alpha\int_0^{\hat{y}_n(\xi)}g(z|\xi)V_{n+1}^{cs}\big(\hat{y}_n(\xi)-z,\pi\big)\,\mathrm{d}z\Bigg] \end{aligned} \tag{A.6}$$

where $V_{N+1}^{cs}(x_{N+1},\pi)=0$, for all x_{N+1},π and $\hat{y}_n(\xi):=\max\{S^{fb}(\xi),x_n\}$. $\square$

Proof. Proof of Theorem 2.4.5.

1. Differentiating payment function $P(\xi)$ wrt ξ in Equation (A.5) gives

$$\begin{aligned} \frac{\mathrm{d}P(\xi)}{\mathrm{d}\xi}&=(r-w)\left(S'(\xi)-G(S(\xi)|\xi)S'(\xi)-\int_0^{S(\xi)}\frac{\partial\,G(z|\xi)}{\partial\,\xi}\,\mathrm{d}z+\int_0^{S(\xi)}\frac{\partial}{\partial\xi}G(z|\xi)\,\mathrm{d}z\right)\\ &=(r-w)S'(\xi)\left(1-G(S(\xi)|\xi)\right)\geq 0. \end{aligned}$$

   It follows from above that, $\frac{\mathrm{d}P}{\mathrm{d}S}=\frac{\frac{\mathrm{d}P(\xi)}{\mathrm{d}\xi}}{\frac{\mathrm{d}S(\xi)}{\mathrm{d}\xi}}=(r-w)\left(1-G(S(\xi)|\xi)\right)$. Consider arbitrary (x_n,π_n). Then, $\frac{\mathrm{d}\,P_n^*(S_n^*)}{\mathrm{d}\,S_n^*}=(r-w)(1-G(S_n^*(\xi|x_n,\pi_n)|\xi))\geq(r-w)(1-G(\hat{S}_n(\xi|\pi_n)|\xi))\geq(r-w)(1-G(S^{fb}(\xi|\pi_n)|\xi))=(r-w)\big(1-\frac{r-c}{r+h-\alpha c}\big)$. The first inequality follows since $\hat{S}_n(\xi)\leq S_n^*(\xi)$ and the next inequality follows since, the supplier never maintains inventory level higher than the first best level $S^{fb}(\xi)$ for any demand type ξ.

2. The marginal price the retailer of type ξ pays is $\frac{\mathrm{d}P_n^*(S_n^*)}{\mathrm{d}S_n^*} = (r-w)(1-G(S_n^*(\xi)|\xi)) \leq (r-w)(1-G(S_{n+1}^*(\xi)|\xi)) = \frac{\mathrm{d}P_{n+1}^*(S_{n+1}^*)}{\mathrm{d}S_{n+1}^*}$, using Part 1 of Theorem 2.4.3.

3. For all $\xi > t(z_n)$, $\hat{S}_{n+1}(\xi) \geq \hat{S}_n(\xi)$ from Part 2 (i) of Theorem 2.4.3. Then, $x_n \leq \hat{S}_n(\xi) \implies S_n^*(\xi) \leq S_{n+1}^*(\xi)$. Therefore, $\frac{\mathrm{d}P_n^*(S_n^*)}{\mathrm{d}S_n^*} = (r-w)(1-G(S_n^*(\xi)|\xi)) \geq (r-w)(1-G(S_{n+1}^*(\xi)|\xi)) = \frac{\mathrm{d}P_{n+1}^*(S_{n+1}^*)}{\mathrm{d}S_{n+1}^*}$.

$\square$

Proof. Proof of Theorem 2.4.6.

1. Differentiating $\Pi_n^{sr}(x_n, \pi_n)$ with respect to x_n gives

$$\frac{\mathrm{d}\,\Pi_n^{sr}(x_n, \pi_n)}{\mathrm{d}x_n} = \gamma(n) \int_{\Theta} \pi_n(\xi) \cdot \frac{\mathrm{d}}{\mathrm{d}x_n} H(S_n^*(\xi), \xi|\pi_n)\mathrm{d}\xi.$$

For any $\xi \in \Theta$, owing to unimodality of $H_n(\cdot, \xi)$,

$$\frac{\mathrm{d}}{\mathrm{d}x_n} H(\hat{S}_n(\xi) \vee x_n, \xi|\pi_n) \begin{cases} = 0, & \text{if } x_n \leq \hat{S}_n(\xi); \\ < 0, & \text{if } x_n > \hat{S}_n(\xi); \end{cases}$$

since $\hat{S}_n(\xi)$ is the unique maximizer of $H(\cdot, \xi|\pi_n)$ and is characterized by the first order condition. Thus $\frac{\mathrm{d}}{\mathrm{d}x_n} H(\hat{S}_n(\xi) \vee x_n, \xi|\pi_n) \leq 0$ for all $\xi \in \Theta \implies \frac{\mathrm{d}\,\Pi_n^{sr}(x_n, \pi_n)}{\mathrm{d}x_n} \leq 0$.

2. Consider an arbitrary period n. It is not optimal to offer contracts in period n if $\Pi_n^{sr}(x_n, \pi_n) < \Pi_n^{lr}(x_n, \pi_n) = \max_{y \geq x_n} L_n(y, \pi_n)$. To establish the result, we first construct a lower bound for $L_n(y, \pi_n)$. This lower bound is in turn used to find an upper bound on the difference, $\Pi_n^{sr}(x_n, \pi_n) - \max_{y \geq x_n} L_n(y, \pi_n)$. Then we find the range of values for x_n for which the difference is always less than zero.

$$\begin{aligned} L_n(y, \pi_n) = (w-c)y - (w+h-\alpha c) \int_0^y Q_n(z)\mathrm{d}z + \alpha \overline{Q_n}(y) V_{n+1}(0, \pi_{n+1}^c(\cdot|y)) \\ + \alpha \int_0^y q_n(z) \cdot V_{n+1}(y-z, \pi_{n+1}^e)\mathrm{d}z, \end{aligned}$$

where $Q_n(z) = \int_\Theta \pi_n(\xi) G(z|\xi)\,\mathrm{d}\xi$ is predictive demand distribution in period n. To find a lower bound on $V_{n+1}(z,\, \pi_{n+1})$, we consider a policy that produces nothing until period N nor does the supplier offer contracts to the retailer, i.e., $\tau = N+1$.

$$\begin{aligned}
V_{n+1}(z,\, \pi_{n+1}) &\geq \mathbb{E}\Big[w \min\{z,\, D_{n+1}\} - h(z - D_{n+1})^+ + \alpha\big(w \min\{x_{n+2},\, D_{n+2}\} - h(x_{n+2} - D_{n+2})^+\big) \\
&\qquad\qquad + \cdots + \alpha^{N-n+1}\big(w \min\{x_N,\, D_N\} - h(x_N - D_N)^+\big)\Big] \\
&\geq \mathbb{E}\Big[w \min\{z,\, D_{n+1}\} - h(z - D_{n+1})^+ + \alpha\big(w \min\{x_{n+2},\, D_{n+2}\} - h(x_{n+2} - D_{n+2})^+\big) \\
&\qquad\qquad + \cdots + \alpha^{N-n+1}\big(w \min\{x_N,\, D_N\} - h(x_N - D_N)^+\big)\Big] \\
&\geq -h \cdot \mathbb{E}\Big[(z - D_{n+1})^+ + \alpha(x_{n+2} - D_{n+2})^+\big) + \cdots + \alpha^{N-n+1}(x_N - D_N)^+\Big] \\
&= -h \cdot \gamma(n+1) \cdot \mathbb{E}\Big[(z - D_{n+1})^+\Big] \geq -h \cdot \gamma(n+1) \cdot z
\end{aligned}$$

Hence,

$$\begin{aligned}
L_n(y,\, \pi_n) &\geq (w-c)y - (w + h - \alpha c) \int_0^y Q_n(z)\mathrm{d}z - h\gamma(n+1)\alpha \int_0^y q_n(z) \cdot (y - z)\mathrm{d}z \\
&= (w-c)y - (w + h(1 + \alpha\gamma(n+1)) - \alpha c) \int_0^y Q_n(z)\mathrm{d}z \\
&\quad [\text{Note that } \gamma(\cdot) \text{ satisfies the following recursion: } \gamma(n) = 1 + \alpha \cdot \gamma(n+1)] \\
&= (w-c)y - (w + h\gamma(n) - \alpha c) \int_0^y Q_n(z)\mathrm{d}z \\
&=: \tilde{L}_n(y,\, \pi_n)
\end{aligned}$$

Note that $\tilde{L}_n(y,\, \pi_n)$ is concave in y, is maximized at $\tilde{y}_n(\pi_n) := Q_n^{-1}\Big(\dfrac{w-c}{w + h\gamma(n) - \alpha c}\Big) < S_n^*(\bar{\xi})$. For $x_n > S_n^*(\bar{\xi})$ it follows that

$$\begin{aligned}
&\Pi_n^{sr}(x_n,\, \pi_n) - \max_{y \geq x_n} L_n(y,\, \pi_n) \\
&\leq \gamma(n) \int_\Theta \pi_n(\xi) \cdot H_n(S^*(\xi) \vee x_n,\, \xi)\,\mathrm{d}\xi - \max_{y \geq x_n} \tilde{L}_n(y,\, \pi_n) \\
&= \gamma(n) \int_\Theta \pi_n(\xi) \cdot H_n(x_n,\, \xi)\,\mathrm{d}\xi - \tilde{L}_n(x_n,\, \pi_n) \\
&= \int_\Theta \Big\{(r-c)\gamma(n)x_n - (r + h - \alpha c)\gamma(n) \int_0^{x_n} G(z|\xi)\mathrm{d}z + \frac{(r-w)\gamma(n)}{\lambda_n(\xi)} \int_0^{x_n} \frac{\partial}{\partial \xi} G(z|\xi)\,\mathrm{d}z\Big\}\pi_n(\xi)\mathrm{d}\xi
\end{aligned}$$

$$
\begin{aligned}
&- \int_\Theta \left\{ (w-c)x_n - (w + h\gamma(n) - \alpha c) \int_0^{x_n} G(z|\xi) \right\} \pi_n(\xi)\mathrm{d}\xi \\
&= \int_\Theta \left\{ \big[(r-c) + \alpha\gamma(n+1)(r-c) - (w-c)\big]x_n - \big[\gamma(n)(r-\alpha c) - (w-\alpha c)\big] \int_0^{x_n} G(z|\xi)\mathrm{d}z \right. \\
&\left. + \frac{(r-w)\gamma(n)}{\lambda_n(\xi)} \int_0^{x_n} \frac{\partial}{\partial\xi} G(z|\xi)\,\mathrm{d}z \right\} \pi_n(\xi)\mathrm{d}\xi \\
&= \int_\Theta \left\{ \big[(r-w) + \alpha\gamma(n+1)(r-c)\big]x_n - \big[\alpha\gamma(n+1)(r-\alpha c) + (r-\alpha c) - (w-\alpha c)\big] \int_0^{x_n} G(z|\xi)\mathrm{d}z \right. \\
&\left. + \frac{(r-w)\gamma(n)}{\lambda_n(\xi)} \int_0^{x_n} \frac{\partial}{\partial\xi} G(z|\xi)\,\mathrm{d}z \right\} \pi_n(\xi)\mathrm{d}\xi \\
&= \int_\Theta \left\{ \big[(r-w) + \alpha\gamma(n+1)(r-c)\big]x_n - \big[(r-w) + \alpha\gamma(n+1)(r-\alpha c)\big] \int_0^{x_n} G(z|\xi)\mathrm{d}z \right. \\
&\left. + \frac{(r-w)\gamma(n)}{\lambda_n(\xi)} \int_0^{x_n} \frac{\partial}{\partial\xi} G(z|\xi)\,\mathrm{d}z \right\} \pi_n(\xi)\mathrm{d}\xi \\
&\le \int_\Theta \left\{ x_n - \theta_n \int_0^{x_n} G(z|\xi)\mathrm{d}z \right\} \pi_n(\xi)\mathrm{d}\xi,
\end{aligned}
$$

where $\theta_n := \frac{(r-w)+\alpha\gamma(n+1)(r-\alpha c)}{(r-w)+\alpha\gamma(n+1)(r-c)} > 1$ for all n. The first inequality follows from definition of $\tilde{L}_n$ and the last inequality is true owing to stochastic dominance of $\{G(z|\xi)\}_\xi$. The function inside the integral on the last line is concave in x_n and is negative once $x_n > U_n$ for all n and ξ. U_n is implicitly defined as the non-zero solution to $U_n = \theta_n \int_0^{U_n} G(z|\xi)\mathrm{d}\xi$. Thus the last integral is non-positive for all $x_n > K_n$ where $K_n := U_n \vee \hat{S}_n(\bar{\xi})$.

□

A.4 Analytical Examples

A.4.1 Two-Point Prior

Here, the supplier's best assessment of the market can be summarized by a two-point prior. The supplier believes that the market demand is either *high* ($\bar{\xi}$) or *low* ($\underline{\xi}$). Her initial prior is denoted by p, representing her subjective probability that demand is high in the ongoing season. The supplier updates her belief using Bayes rule described in Equation (2.1) and offers two contracts $(\overline{S}, \overline{P})$, $(\underline{S}, \underline{P})$, in some period, to screen the retailer. These contracts satisfy (IC) and (PC) constraints.

Using the approach illustrated (in §2.4.1) for the continuum-type case, we can express payments, $\underline{P}$ and $\overline{P}$, similarly in terms of $\underline{S}, \overline{S}$. The main results can be summarized in the following theorem.

Theorem A.4.1. *Assuming* $\Pi^r_{\min}(n) = 0$,

1. *any two contracts* $(\underline{S},\, \underline{P})$ *and* $(\overline{S},\, \overline{P})$ *satisfy (IC) and (PC) iff* $\underline{P}$, $\overline{P}$ *can be expressed as,* $\underline{P} = (r-w)\big[\underline{S} - \int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\big]$ *and* $\overline{P} = (r-w)\big\{\overline{S} - \int_{\underline{S}}^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z - \int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\big\}$.

2. *The supplier's incentive problem can defined as follows*

$$\begin{aligned}\Pi^{sr}_n(x_n,\, p_n) \colon =& \max\left\{\mathbb{E}_\xi\big[\Pi^{tot}_n(S(\xi)|x_n, p_n) - \Pi^r_n(S(\xi), P(\xi), \xi)\big]\middle|\overline{S}, \underline{S} \geq x_n; (IC), (PC)\right\} - cx_n\\ =& \,\gamma(n)\cdot\max\left\{H^d(\underline{S},\, \overline{S}|p_n)\,\middle|\,\overline{S}, \underline{S} \geq x_n\right\},\end{aligned}$$

$$\begin{aligned}\textit{where } H^d(\underline{S}, \overline{S}|p_n) := & \,p_n\left\{(r-c)\overline{S} - (r+h-\alpha c)\int_0^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z - (r-w)\int_0^{\underline{S}}\big(G(z|\underline{\xi}) - G(z|\overline{\xi})\big)\mathrm{d}z\right\}\\ &+ (1-p_n)\left\{(r-c)\underline{S} - (r+h-\alpha c)\int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\right\}.\end{aligned}$$

Proof. Proof of Theorem A.4.1.

1. ($\Leftarrow$) We first show (IC) and (PC) imply the conditions on $\underline{P}$, $\overline{P}$. Note that, $\Pi^r_n(S, P, \overline{\xi}) > \Pi^r_n(S, P, \underline{\xi})$ for any S, P because $G(z|\underline{\xi}) \geq G(z|\overline{\xi})$, $\forall z \geq 0$. Thus, $\Pi^r_n(\underline{S}, \underline{P}, \underline{\xi}) \leq \Pi^r_n(\underline{S}, \underline{P}, \overline{\xi}) \leq \Pi^r_n(\overline{S}, \overline{P}, \overline{\xi})$, where the last inequality follows from (IC). This reduces the (PC) to $\Pi^r_n(\underline{S}, \underline{P}, \underline{\xi}) \geq 0$. At optimality however, the supplier would make sure $\Pi^r_n(\underline{S}, \underline{P}, \underline{\xi}) = 0$ since the supplier's profit is increasing in $\underline{P}$ and the retailer's is decreasing. Therefore (PC) is replaced by $\Pi^r_n(\underline{S}, \underline{P}, \underline{\xi}) = 0$. This equation gives a solution for $\underline{P}$ in terms of $\underline{S}$ as follows:

$$\underline{P} = (r-w)\left[\underline{S} - \int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\right].$$

The (IC) constraints can be expressed as $(r-w)\big[\overline{S} - \int_0^{\overline{S}} G(z|\underline{\xi})\mathrm{d}z\big] \leq \overline{P}$ and $(r-w)\big[\overline{S} - \int_0^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z\big] - \overline{P} \geq (r-w)\big[\underline{S} - \int_0^{\underline{S}} G(z|\overline{\xi})\mathrm{d}z\big] - \underline{P} = (r-w)\big[\int_0^{\underline{S}}(G(z|\underline{\xi}) - G(z|\overline{\xi}))\mathrm{d}z\big] \Rightarrow \overline{P} \leq (r-w)\big[\overline{S} - \int_{\underline{S}}^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z - \int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\big]$. Thus, (IC) and (PC) together imply following restriction on $\overline{P}$:

$$\overline{P} \in \left[(r-w)\left\{\overline{S} - \int_0^{\overline{S}} G(z|\underline{\xi})\mathrm{d}z\right\},\, (r-w)\left\{\overline{S} - \int_{\underline{S}}^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z - \int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\right\}\right],$$

which is clearly not an empty interval. Since profits of the supplier are increasing in $\overline{P}$, at optimality:

$$\overline{P} = (r-w)\Big\{\overline{S} - \int_{\underline{S}}^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z - \int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\Big\}.$$

The converse ($\Rightarrow$) can be verified by substituting $\underline{P}$, $\overline{P}$ into (IC) and (PC).

2. The supplier's objective function in the above problem can be simplified as follows

$$\begin{aligned}
&\mathbb{E}\Big[\Pi_n^{tot}(S(\xi),\xi|x_n,p_n) - \Pi_n^r(S(\xi),P(\xi),\xi)\Big]\\
&= p_n \cdot \Big[\Pi_n^{tot}(\overline{S},\overline{\xi}|x_n,p_n) - \Pi_n^r(\overline{S},\overline{P},\overline{\xi})\Big] + (1-p_n)\cdot\Big[\Pi_n^{tot}(\underline{S},\underline{\xi}|x_n,p_n) - \overbrace{\Pi_n^r(\underline{S},\underline{P},\underline{\xi})}^{=0}\Big]\\
&= cx_n + \gamma(n)\Big[p_n\Big\{(r-c)\overline{S} - (r+h-\alpha c)\int_0^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z - (r-w)[\overline{S} - \int_0^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z] + \overline{P}\Big\}\\
&\qquad + (1-p_n)\Big\{(r-c)\underline{S} - (r+h-\alpha c)\int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\Big\}\Big]\\
&= cx_n + \gamma(n)\Big[p_n\Big\{(r-c)\overline{S} - (r+h-\alpha c)\int_0^{\overline{S}} G(z|\overline{\xi})\mathrm{d}z - (r-w)\int_0^{\underline{S}}\Big(G(z|\underline{\xi}) - G(z|\overline{\xi})\Big)\mathrm{d}z\Big\}\\
&\qquad + (1-p_n)\Big\{(r-c)\underline{S} - (r+h-\alpha c)\int_0^{\underline{S}} G(z|\underline{\xi})\mathrm{d}z\Big\}\Big].
\end{aligned}$$

□

We point out that $H^d(\underline{S},\overline{S}|p_n) = \Pi_n^{tot}(\underline{S},\overline{S}|p_n) - (r-w)p_n\int_0^{\underline{S}}\big(G(z|\underline{\xi}) - G(z|\overline{\xi})\big)\mathrm{d}z$. That is, the supplier captures all but $(r-w)p_n\int_0^{\underline{S}}\big(G(z|\underline{\xi}) - G(z|\overline{\xi})\big)\mathrm{d}z$ from the total supply chain profit. Note that, $H^d(\underline{S},\,\overline{S}|p_n)$ is additively separable in $\overline{S}$, $\underline{S}$. We denote the single-variable functions as, $H^d(\underline{S},\,\overline{S}|p_n) = H_l^d(\underline{S}|p_n) + H_h^d(\overline{S}|p_n)$. In Lemma A.4.1, we characterize the optimal menu of contracts, $\{\underline{S}_n^*,\overline{S}_n^*\}$.

Lemma A.4.1. *For exponential demand distribution with mean $\xi = \{\underline{\xi},\,\overline{\xi}\}$:*

1. *For all $\overline{S}$, $\underline{S} \geq 0$ we have,*

$$H^d(\hat{\underline{S}}_n,\hat{\overline{S}}_n|p_n) \geq H^d(\underline{S},\overline{S}|p_n),$$

where $\hat{\underline{S}}_n,\hat{\overline{S}}_n$ are the unique maximizers of $H_l^d(\underline{S}|p_n)$ and $H_h^d(\overline{S}|p_n)$, respectively.

2. *$H_h^d(\overline{S}|p_n)$ is concave in $\overline{S}$ and $H_l^d(\underline{S}|p_n)$ is unimodal in $\underline{S}$. Therefore the unique maximizers $\hat{\underline{S}}_n, \hat{\overline{S}}_n$ are characterized by the first order conditions:*

$$\hat{\overline{S}}_n = \overline{\xi} \log\left(\frac{r+h-\alpha c}{h+c(1-\alpha)}\right) \qquad \text{and} \tag{A.7}$$

$$e^{-\frac{\hat{\underline{S}}_n}{\underline{\xi}}} - e^{-\frac{\hat{\underline{S}}_n}{\overline{\xi}}} \cdot \frac{p_n(r-w)}{p_n(r-w)+(1-p_n)(r+h-\alpha c)} = \frac{(1-p_n)(h+c(1-\alpha))}{p_n(r-w)+(1-p_n)(r+h-\alpha c)}. \tag{A.8}$$

3. *$\hat{\underline{S}}_n \leq \hat{\overline{S}}_n$. Therefore, $\underline{S}_n^* = \max\{\hat{\underline{S}}_n, x_n\}$ and $\overline{S}_n = \max\{\hat{\overline{S}}_n, x_n\}$ are the optimal base-stock levels offered.*

Proof. Proof of Lemma A.4.1.

1. By definition, $H_l^d(\underline{S}|p_n) \leq H_l^d(\hat{\underline{S}}_n|p_n)$ and $H_h^d(\overline{S}|p_n) \leq H_h^d(\hat{\overline{S}}_n|p_n)$ for all $\underline{S}$, $\overline{S} \geq 0$. Adding these inequalities we get, $H^d(\underline{S}, \overline{S}|p_n) = H_l^d(\underline{S}|p_n) + H_h^d(\overline{S}|p_n) \leq H_l^d(\hat{\underline{S}}_n|p_n) + H_h^d(\hat{\overline{S}}_n|p_n) = H^d(\hat{\underline{S}}_n, \hat{\overline{S}}_n|p_n)$ and the result follows.

2. The function $H_h^d(\overline{S}|p_n)$ is concave in $\overline{S}$, therefore its maximizer is characterized by the first-order condition: $G(\hat{\overline{S}}_n|\overline{\xi}) = \dfrac{r-c}{(r+h-\alpha c)}$, which can be simplified to Equation (A.7). $H_l^d(\underline{S}|p_n)$ need not be concave in general. For exponential demand distribution, we can establish that $H_l^d(\underline{S}|p_n)$ is unimodal using Lemma A.3.1. It satisfies the conditions in that Lemma with $a = 0$, $b = \infty$ and $\hat{x} := \frac{\underline{\xi}\overline{\xi}}{\overline{\xi}-\underline{\xi}} \log\big(\frac{\overline{\xi}}{\underline{\xi}}\{1 + \frac{1-p_n}{p_n}\frac{r+h-\alpha c}{r-w}\}\big)$. Hence, the following first order condition characterizes the maximizer:

$$p_n(r-w)\big(G(\hat{\underline{S}}|\overline{\xi}) - G(\hat{\underline{S}}|\underline{\xi})\big) + (1-p_n)\big(r-c-(r+h-\alpha c)G(\hat{\underline{S}}|\underline{\xi})\big) = 0. \tag{A.9}$$

Equation (A.9) can be simplified to Equation (A.8) when demand is exponentially distributed with mean ξ.

3. In (A.9) let $LHS := \big[(1-p_n)(r+h-\alpha c) + p_n(r-w)\big]G(\hat{\underline{S}}_n|\underline{\xi}) - p_n(r-w)G(\hat{\underline{S}}_n|\overline{\xi})$. Since $G(\hat{\underline{S}}_n|\underline{\xi}) > G(\hat{\underline{S}}_n|\overline{\xi})$ by assumption, it follows: $LHS > \big[(1-p_n)(r+h-\alpha c) + p_n(r-w)\big]G(\hat{\underline{S}}_n|\overline{\xi}) - p_n(r-w)G(\hat{\underline{S}}_n|\overline{\xi}) = \big[(1-p_n)(r+h-\alpha c) + p_n(r-w) - p_n(r-w)\big]G(\hat{\underline{S}}_n|\overline{\xi}) = (1-p_n)(r+h-\alpha c)G(\hat{\underline{S}}_n|\overline{\xi})$. We have from (A.9) that, $LHS < (1-p_n)(r-c)$. This implies

$(r+h-\alpha c)G(\hat{\underline{S}}_n|\bar{\xi}) < (r-c) \Rightarrow G(\hat{\underline{S}}_n|\bar{\xi}) < \dfrac{(r-c)}{r+h-\alpha c} = G(\hat{\bar{S}}_n|\bar{\xi})$. Since $G(\cdot|\bar{\xi})$ is a distribution function, it is non-decreasing and it follows that $\hat{\bar{S}}_n > \hat{\underline{S}}_n$.

□

As in the continuous-type case, the high type retailer gets the first-best base-stock level irrespective of the supplier's belief of underlying demand. Part 3 of Lemma A.4.1 shows that the maximizers Equation (A.7) and Equation (A.8) satisfy the monotonicity constraint and thus, proves optimality of these contracts.

A.4.2 Case of Conjugate Prior

Historically, researchers have favored use of newsvendor class of demand distributions when studying unobserved lost-sales Bayesian inventory problem (Lariviere and Porteus, 1999; Bisi et al., 2011). The inverse gamma distribution is a conjugate prior for the newsvendor class, i.e., if π_n has inverse gamma distribution, Bayesian updating ensures π_{n+1} is inverse gamma with updated parameters. For the analysis in this section demand is exponentially distributed with unknown mean (denoted by ξ). The prior over Θ is a truncated inverse gamma distribution with shape and scale parameters given by, m and β, respectively. The demand distribution and the prior density are given as follows:

$$G(z|\xi) = 1 - e^{-\frac{z}{\xi}}, \ \ z \geq 0 \quad \text{and} \quad \pi_n(\xi|\, m,\, \beta) = \frac{\beta^m \cdot \xi^{-m-1} \cdot e^{\frac{-\beta}{\xi}}}{\Gamma(m) \cdot \big(F(\bar{\xi}|\, m,\, \beta) - F(\underline{\xi}|\, m,\, \beta)\big)}, \ \xi \in \Theta.$$

$\Gamma(\cdot)$ is the gamma function and $F(u|\, m,\, \beta) = \dfrac{\Gamma(m,\, \beta/u)}{\Gamma(m)}$ is the un-truncated inverse gamma distribution (with shape and scale parameters m, β, respectively), where $\Gamma(\cdot,\, \cdot)$ is the *incomplete* gamma function. If sales observation in period n is z_n, parameters of the posterior are updated as follows: $(m_n + d_n,\, \beta_n + z_n)$, where $d_n = 1$ if z_n is an uncensored observation otherwise, zero. Note that the conjugate prior relationship is *not* affected by truncation of the prior distrubition.

$$q_n(z|\, m,\, \beta) = \int_{\underline{\xi}}^{\bar{\xi}} g(z|\, \xi)\pi_n(\xi|\, m,\, \beta)\, \mathrm{d}\xi, \text{ and } Q_n(z|\, m,\, \beta) = \int_0^z q(y|\, m,\, \beta)\, \mathrm{d}y$$

are the prior predictive density and distribution, respectively. These functions can be explicitly described as follows:

$$q_n(y|m,\,\beta) = \left(\frac{m\beta^m}{(y+\beta)^{m+1}}\right)\left(\frac{F(\overline{\xi}|\;m+1,\,y+\beta) - F(\underline{\xi}|\;m+1,\,y+\beta)}{F(\overline{\xi}|\,m,\,\beta) - F(\underline{\xi}|m,\,\beta)}\right),$$
$$Q_n(z|m,\,\beta) = 1 - \left(\frac{\beta}{\beta+z}\right)^m\left(\frac{F(\overline{\xi}|\,m,\,\beta+z) - F(\underline{\xi}|\,m,\,\beta+z)}{F(\overline{\xi}|\,m,\,\beta) - F(\underline{\xi}|m,\,\beta)}\right).$$

In the above equations, if $\underline{\xi}=0$ and $\overline{\xi}=\infty$, then the fractions involving $F(\cdot)$ become unity. The optimality equations for the exponential-inverse gamma pair are as follows.

$$V_n(x_n, m_n, \beta_n) = \max\left\{\Pi_n^{lr}(x_n, m_n, \beta_n),\; \Pi_n^{sr}(x_n, m_n, \beta_n)\right\},\; n=1,\cdots,N,$$

where $V_{N+1}(x_n,\,m_n,\,\beta_n) := 0$ for all $x_n,\,m_n,\,\beta_n$ and

$$\begin{aligned}\Pi_n^{lr}(x_n, m_n, \beta_n) &= \max_{y\ge x_n} L_n(y, m_n, \beta_n)\\ &:= (w-c)y - (w+h-\alpha c)\int_0^y Q_n(z)\,\mathrm{d}z + \alpha\overline{Q}_n(y)V_{n+1}\big(0,\,m_n,\,\beta_n+y\big)\\ &\quad + \alpha\int_0^y q_n(z)V_{n+1}\big(y-z,\,m_n+1,\,\beta_n+z\big)\,\mathrm{d}z, \quad \text{for all } y\ge x_n.\end{aligned}$$

APPENDIX B

APPENDIX TO CHAPTER 2

B.1 Maximum Likelihood Function

In this section, we present closed-form expression of the maximum likelihood framework used in the paper. We present the likelihood function for the promotion period model which involves estimating utilities corresponding to three alternatives and corresponding expenditure equations. Likelihood function for the aggregate and the redemption models readily follow from this discussion, and hence are omitted here.

The likelihood of the observed data as explained by the parameters of the model (Equations 3.4 and 3.5) is given as follows.

$$\begin{aligned}
&\mathcal{L}\big(\boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\theta} \big| \{p_{ik}, y_{ik}, e_{ijk}\}\big) \\
&\quad = \prod_{k=1}^{38} \prod_{i=1}^{N} \mathbb{P}\{\mathsf{Purchase}_{ik} = 0\}^{1-p_{ik}} \cdot \big[\mathbb{P}\{\mathsf{Purchase}_{ik} = 1, \mathsf{Expenditure}_{i1k} = e_{i1k}\}^{(1-y_{ik})} \\
&\qquad\qquad \cdot \mathbb{P}\{\mathsf{Participate}_{ik} = 1, \mathsf{Expenditure}_{i2k} = e_{i2k}\}^{y_{ik}}\big]^{p_{ik}},
\end{aligned}$$

where $\{\mathsf{Purchase}_{ik}, \mathsf{Participate}_{ik}, \mathsf{Expenditure}_{ijk}\}_{\substack{i=1,\ldots,N;\ j=1,2\\ k=1,\ldots,38}} = \{p_{ik}, y_{ik}, e_{ijk}\}$ denote the observed data, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ and $\tilde{\boldsymbol{\beta}} = (\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2)$ are the coefficients in the expenditure and choice model. As in Zhang and Krishnamurthi (2004), we set the standard deviation of ξ_j, denoted as δ_j, to 1. This provides greater stability to our estimation results. $\boldsymbol{\theta} = (\theta_1, \theta_2)$ denotes the scaled version of the correlation between ν^*_{ijk} and ξ_j, which features in the bivariate logistic distribution defined over the error terms (see Equation 3.3). Given these distributional assumptions, we can compute the probabilities in above likelihood function in closed-form as follows. For details of the computations we refer the reader to Zhang and Krishnamurthi (2004).

$$\begin{aligned}
\mathbb{P}\{\mathsf{Purchase}_{ik} = 1, \mathsf{Expenditure}_{i1k} = e_{i1k}\} &= \frac{e^{V_{i1k}}}{e^{V_{i0k}} + e^{V_{i1k}} + e^{V_{i2k}}} \frac{\delta_\xi e^{\delta_\xi(\boldsymbol{\beta}_1 \mathbf{Z}_{i1k} - e_{i1k})}}{\big[1 + e^{\delta_\xi(\boldsymbol{\beta}_1 \mathbf{Z}_{i1k} - e_{i1k})}\big]^2} \\
&\quad \cdot \left[1 + \theta_1 \left(1 - \frac{e^{V_{i1t}}}{e^{V_{i0t}} + e^{V_{i1t}} + e^{V_{i2t}}}\right)\left(\frac{-1 + e^{\delta_\xi(\boldsymbol{\beta}_1 \mathbf{Z}_{i1k} - e_{i1k})}}{1 + e^{\delta_\xi(\boldsymbol{\beta}_1 \mathbf{Z}_{i1k} - e_{i1k})}}\right)\right]
\end{aligned}$$

$$\mathbb{P}\{\textsf{Participate}_{ik} = 1, \textsf{Expenditure}_{i2k} = e_{i2k}\} = \frac{e^{V_{i2k}}}{e^{V_{i0k}} + e^{V_{i1k}} + e^{V_{i2k}}} \frac{\delta_\xi e^{\delta_\xi(\boldsymbol{\beta}_2 \mathbf{Z}_{i2k} - e_{i2k})}}{\left[1 + e^{\delta_\xi(\boldsymbol{\beta}_2 \mathbf{Z}_{i2k} - e_{i2k})}\right]^2}$$
$$\cdot \left[1 + \theta_2 \left(1 - \frac{e^{V_{i2k}}}{e^{V_{i0k}} + e^{V_{i1k}} + e^{V_{i2k}}}\right)\left(\frac{-1 + e^{\delta_\xi(\boldsymbol{\beta}_2 \mathbf{Z}_{i2k} - e_{i2k})}}{1 + e^{\delta_\xi(\boldsymbol{\beta}_2 \mathbf{Z}_{i2k} - e_{i2k})}}\right)\right]$$

Taking logarithm of the likelihood function, and after some simplification we get,

$$\log \mathcal{L}(\boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\theta}, \boldsymbol{\delta_\xi}) = \sum_{k=1}^{38} \sum_{i=1}^{N} \Big\{ -\log(1 + e^{V_{i1k}} + e^{V_{i2k}}) + p_{it}(1 - y_{ik}) V_{i1k} + p_{ik} y_{ik} V_{i2k} + p_{ik} \log(\delta_\xi)$$
$$+ p_{it}(1 - y_{it}) \Psi_{i1t} + p_{it} y_{it} \Psi_{i2t} \Big\},$$

$$\text{where } \Psi_{ijk} = \delta_\xi(\boldsymbol{\beta}_j \mathbf{Z}_{ijk} - e_{ijk}) - 2\log(1 + e^{\delta_\xi(\boldsymbol{\beta}_j \mathbf{Z}_{ijk} - e_{ijk})})$$
$$+ \log\left[1 + \theta_j \left(1 - \frac{e^{V_{ijk}}}{e^{V_{i0k}} + e^{V_{i1k}} + e^{V_{i2k}}}\right)\left(\frac{-1 + e^{\delta_\xi(\boldsymbol{\beta}_j \mathbf{Z}_{ijk} - e_{ijk})}}{1 + e^{\delta_\xi(\boldsymbol{\beta}_j \mathbf{Z}_{ijk} - e_{ijk})}}\right)\right].$$

The variance-covariance matrix of the parameter estimates are computed using $-\big[\text{Hessian}\big\{\log\ \mathcal{L}\big\}\big]^{-1}$ (Wooldridge, 2002). The optimization routine was implemented in MATLAB R2013b on a i7 processor with clock rate of 3.4 GHz, and 8 GB RAM.

B.2 Endogeneity Concerns

In this section, we address three endogeneity-related concerns resulting from strategic planning of the retailer or the customer, which could potentially influence our findings.

B.2.1 Potential Targeting Issue of Gift Card Promotion E-mail

The retailer uses e-mail channel to advertise the gift card promotion to its customers. However, in the data not all customers receive the promotion e-mail. This raises the concern that the e-mail might have been targeted to a more responsive segment of customers. Such targeting, if any, would muddle the effect of promotion we observe, with inherent customer characteristics. In our extensive discussions with the retailer on this issue, we were informed that the the e-mail is sent to all customers who opened any marketing e-mail in the year prior to the promotion and depending on the e-mail category they chose. For example, customers who were unsubscribed (or have Email_cat low) do not receive the e-mail. As a proxy, the retailer used the number of customers who received an e-mail in the previous week to determine who to send the gift card e-mail to. Among these customers, only those customers who also made a purchase in the last year receive the e-mail for private events. As illustrated in Table B.1, the significant difference between the two populations seems to be the recency of marketing e-mails received in the previous year. Therefore, to overcome endogeneity concerns in our promotion period model, we include GC_email along with the Email_sent_rec and Email_cat variables. This ensures that any potential reason impacting the receipt of the e-mail is not a part of the unobserved error term, thus mitigating the endogeneity concern. To investigate the validity of this approach, we explore how well customer and communication characteristics can predict the receipt of e-mail.

Table B.1: Summary statistics for customers who received the gift card promotion e-mail and those who did not.

	GC_Email = 1		GC_Email = 0	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Purchase_rec	209.56	147.37	223.47	142.22
Purchase_freq	1.90	3.37	1.76	3.51
ln(Exp_annual)	4.49	2.89	4.53	2.85
Web_rec	130.90	143.19	160.04	142.20
Web_freq	23.31	41.03	13.24	27.70
Email_sent_rec	3.66	29.63	250.68	162.90

We run three logistic regressions to determine how well communication characteristics can predict the receipt of the e-mail. In the first model, we only include an intercept term. In the second model, we include recency and e-mail preferences of customers. In the third model, we append the second model with other customer characteristics. The results are presented in Table B.2. Since our focus here is on the prediction, we compare the pseudo–R^2 (McFadden's measure[1]) across the three models. First we note that, including communication characteristics significantly improves the model fit from the base model. More interesting is the observation that adding customer characteristics to Model 2 improves the model fit only marginally. We also note that McFadden's measure of R^2

[1]McFadden's measure of $R^2 = 1 - \frac{\ln(L)}{\ln(L_0)}$, where $\ln(L)$ is the maximized log-likelihood and $\ln(L_0)$ is the maximized log-likelihood with only the intercept term. This measure returns 0 for a model with only the intercept term.

tends to be on the lower side, and a value between 0.2–0.4 can be considered an excellent fit (see page 307 in McFadden 1978).

Table B.2: Predicting receipt of gift card promotion e-mail.

	Model 1	Model 2	Model 3
Constant	0.54002***	-0.86603***	-0.41187***
	(0.00302)	(0.01202)	(0.03410)
Email_sent_rec		-0.01084***	-0.01072***
		(0.00010)	(0.00010)
Email_cat (med)		0.67823***	0.65490***
		(0.01724)	(0.01736)
Email_cat (high)		6.08023***	6.08249***
		(0.02849)	(0.02862)
Purchase_rec			-0.00085***
			(0.00009)
Purchase_freq			-0.01573***
			(0.00322)
ln(Exp_annual)			-0.05100***
			(0.00433)
Web_rec			-0.00031***
			(0.00007)
Web_freq			0.00133***
			(0.00032)
Pseudo–R^2	~ 0	0.8013	0.8017

Note: $N = 471{,}852$
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B.2.2 Timing of Promotion

Timing of promotion is an integral part of the promotion planning process for the retailer. The fact that promotions are typically viewed as a vehicle to boost short-term sales, suggests that retailer could use it during unresponsive market conditions. This raises concerns such as, the observed effect of the promotion is exaggerated due to the prevailing market conditions. The context of our study alleviates this issue to some extent. First, gift card promotions at the retailer happen almost every month (see, for example, Figure 3.3). This ensures that there is a greater chance that the data captures different prevailing market conditions. Second, within each month, the retailer does not optimize timing of the promotion. This is pertaining to the business practices observed at the retailer. The managers informed us that over the years they have not disturbed the norm of about two gift card promotions per month.

B.2.3 Strategic Customer Response

The extensive use of promotions by retailers has trained customers to wait for a promotion to make a purchase. Such strategic behavior by customers can exaggerate the impact of the promotion. As a result, for example, the baseline sales can be dampened and promotion period sales can be boosted, which is a result of customer timing their purchase and not due to the promotion. Given that gift card promotions are run regularly by the retailer, this can be a cause for concern. However, two aspects of our data alleviate this concern. First, given the detailed customer-level data, we are able to explicitly model customer's choice to make a purchase. Such a decision model takes into consideration potential strategic response of the customer. Second, the retailer only informs customers through e-mail or website banner ads about the promotion on the morning of the promotion and the promotion typically runs for a short period of time. This could make it potentially harder for customers to plan for a purchase during the promotion. If indeed customers were strategically waiting for a gift card promotion, we would expect customers to have a larger fraction of their purchases timed during the gift card promotion. To this end we compute the fraction, $R = \frac{\#\text{Gift Cards Earned}}{\#\text{Purchases}}$ for each customer in the sample. The distribution of R in Table B.3 indicates that close to 96.5% of customers made less than 20% of their purchases resulting in a gift card.

Table B.3: Distribution of fraction of gift card purchases over total purchases per customer in the sample.

R	Frequency	Cumulative Percent
[0, 0.2]	21,007	0.965
(0.2, 0.4]	287	0.978
(0.4, 0.6]	245	0.990
(0.6, 0.8]	17	0.990
(0.8, 1]	211	1.000

The last row in Table B.3 indicates that 211 customers made all their purchases resulting in a gift card. To further check if these customers had multiple purchase instances, we tabulate the distribution of number of gift cards earned per customer (with at least one). The distribution in Table B.4 confirms that majority of customers participated in a single gift card promotion. In summary, although we cannot rule out the possibility of customers timing their purchases during a promotion, the data seems to suggest that the magnitude might be too low to influence our findings.

Table B.4: Number of gift cards earned per customer, who qualified for at least one gift card. In total, there were 1,598 gift cards earned by 1,202 customers.

# Gift Cards	Frequency	Cummulative Percent
1	962	80.03
2	159	93.26
3	49	97.34
4	12	98.34
5	8	99
6	8	99.67
8	2	99.83
9	1	99.92
10	1	100

B.3 Predictive Performance of Customer Response Model

In this section, we test the predictive power of the customer response model. The dataset is randomly sampled into two equal parts, one which is used for estimating the model (training dataset) and the other to validate the estimated model (validation dataset). For the continuous outcome (ln($\mathsf{Expenditure}$)), we compute the normalized root mean square error (NRMSE), defined as $\frac{RMSE}{\overline{\mathsf{Expenditure}}}$, where $\overline{\mathsf{Expenditure}}$ is the mean of the observed expenditures. For the discrete outcomes ($\mathsf{Purchase}$ and $\mathsf{Participate}$), we report the concordance percentage, defined as the fraction of event–no-event pairs (purchase–no-purchase or participate–no-participate conditional on purchase) such that the predicted probability of success for the event observation is higher than the predicted probability of success for the no-event observation.

The customer response models provide a direct way to predict the logarithm of the intended expenditure of customers during non-promotion periods and promotion periods (if they participate or otherwise). For measuring predictive accuracy, however, we need to compare the predicted conditional expenditure $\mathbb{E}[\mathsf{Expenditure}|\mathsf{Participate}=1]$ or $\mathbb{E}[\mathsf{Expenditure}|\mathsf{Purchase}=1]$ with the observed values of the expenditure. To this end, we first predict the expected expenditure, conditional on purchase or participation, using the closed-form expression derived in Zhang and Krishnamurthi (2004), also listed below.

$$\mathbb{E}[\ln(\mathsf{Expenditure}_{i1k})|\mathsf{Purchase}_{i1k}=1] = \ln\left(1+e^{\boldsymbol{\beta}_1 Z_{i1k}}\right) - \theta_1\left(1-\frac{e^{V_{i1k}}}{1+e^{V_{i1k}}+e^{V_{i2k}}}\right)\frac{e^{\boldsymbol{\beta}_1 Z_{i1k}}}{1+e^{\boldsymbol{\beta}_1 Z_{i1k}}}$$

$$\mathbb{E}[\ln(\mathsf{Expenditure}_{i2k})|\mathsf{Participate}_{i2k}=1] = \ln\left(1+e^{\boldsymbol{\beta}_2 Z_{i2k}}\right) - \theta_2\left(1-\frac{e^{V_{i2k}}}{1+e^{V_{i1k}}+e^{V_{i2k}}}\right)\frac{e^{\boldsymbol{\beta}_1 Z_{i2k}}}{1+e^{\boldsymbol{\beta}_1 Z_{i2k}}}$$

Note that the above formulae are illustrated for the promotion period model. The second term on the rhs of the above expression needs to be modified to account for the logistic distribution in the case of aggregate model. Next, we obtain the predicted conditional expenditure using the approach outlined on page 213, Wooldridge (2012).

1. Compute $\hat{E}^p = e^{\mathbb{E}[\ln(\mathsf{Expenditure})|\mathsf{Purchase}=1,\mathbf{X}]}$.

2. Regress E^o, which is the observed expenditure, on $\hat{E}^p$ without a constant term. Let the estimated slope parameter be denoted by $\hat{m}$.

3. The correct predicted conditional expenditure is given by $\hat{m}\hat{E}^p$.

Tables B.5 and B.6 present the predictive performance of the aggregate and promotion-period models, respectively, on the training and the validation datasets. We note that the predictive measures are consistent across the training and the validation datasets, indicating good performance of the model.

Table B.5: Predictive performance of aggregate model.

Aggregate Model	Training Set		Validation Set	
	Promotion	Non-Promotion	Promotion	Non-Promotion
NRMSE (Expenditure)	1.4823	1.8542	1.4232	1.8409
Concordance (Purchase)	0.7245	0.7187	0.7180	0.7235
N	522,702		522,703	

Table B.6: Predictive performance of promotion period model.

Promotion Period Model	Training Set		Validation Set	
	Purchase	Participate	Purchase	Participate
NRMSE (Expenditure)	1.4835	1.1849	1.5005	1.0271
Concordance	0.7668	0.6556	0.7629	0.6449
N	235,787		236,065	

B.4 Details of Propensity Score Method

The propensity score method (PSM) is a statistical approach used to estimate the effect of a treatment on outcome variable using observational data. This method accounts for the bias due to self-selection, based on observable factors. In our context, we measure the impact of redemption (treatment variable) on the total expenditure customer makes in the promotion (outcome). For each customer who qualified for a gift card, we determine their propensity to redeem their gift card using a logistic regression. We use a specification for the logistic regression, identical to the one in the redemption model § 3.6.3. Customers with similar predicted propensity to redeem are matched and the difference in their average total expenditures is computed as the effect of redemption. To ensure similar customers are matched, we restrict the difference between the propensity scores of matched customers to be less than 0.01. This restriction excludes about

As outlined in Caliendo and Kopeinig (2008), PSM makes two assumptions to ensure identifiability. First, is the assumption that the decision to redeem is only based on the observable factors included in the logistic regression. We address this to the extent possible by including variables relating to past purchasing patterns, website visit behavior, and the face value of the gift card. Second, is the assumption about common support, which states that the logistic regression should not perfectly predict the outcome for any customer. We verify that this assumption is satisfied.

B.5 Robustness Checks

B.5.1 Unobserved Customer Heterogeneity

One of the assumptions underlying our specification of the utility of purchase and participation in the promotion (see Equations 3.2, 3.5) is that the unobserved errors ν_{ijt} are i.i.d. across different observations. However, it is possible that some of the unobservable factors are correlated for observations corresponding to the same customer and independent across customers. Such unobserved customer heterogeneity can be accounted by imposing special structure on the variance–co-variance matrix of the error terms. One approach is to report robust standard errors by clustering errors at customer-level. A second approach, known as the random effects (multinomial-) logit model, imposes further restrictions on the covariance matrix. We refer the reader to § 15.8 in Wooldridge (2002) for the details of the assumptions each approach entails. Here we report estimates of the customer choice in the aggregate model (Table B.7) and in the promotion period model (multinomial logit model) below.

B.5.2 Alternate Models of Self-Selection

Self-selection models proposed in the literature differ in terms of the distributional assumptions made on the correlation between the error terms of the choice and outcome equation. Broadly, in the paper, we considered models to address self-selection issues arising from binary (aggregate and redemption models) or multinomial choice (promotion period model) . Here we consider other popular models of self-selection, such as Heckman model for binary choice (Heckman, 1979) and its extension to multinomial case, such as those proposed by Lee (1983) and Bourguignon et al. (2007). Most of these models are implemented using a two-stage methods, in which first a choice model is estimated. Estimates from this model are used to debias error terms in the outcome (expenditure) equation. Therefore, in the following we report estimation results using two-stage methods.

Aggregate Model

The Heckman model assumes that the errors ξ_{it} are i.i.d. normally distributed with mean 0 and variance σ^2, which is estimated from the data. The utility from no-purchase, U_{i0t}, is normalized to 0 and ν_{i1t} is assumed to be a standard normal distribution. The correlation between the choice and outcome is modeled using a bivariate normal distribution over ξ_{it} and ν_{i1t} with correlation ρ. We note that, the aggregate model presented in § 3.6.1 captures correlation between the negative of the error associated with the utility of purchase and expenditure equation. Therefore, we expect to see

Table B.7: Logistic model of purchase with clustered (at customer-level) standard errors and random effects specification for customer heterogeneity.

	Clustered Std. Err.		Random Effects	
	Agg. Model	Adv. Model	Agg. Model	Adv. Model
Constant	-5.28404***	-5.27685***	-5.60290***	-5.60408***
	(0.03935)	(0.03964)	(0.04453)	(0.04492)
Purchase_rec	-0.01699***	-0.01690***	-0.01252***	-0.01270***
	(0.00032)	(0.00033)	(0.00020)	(0.00020)
Purchase_rec^2	0.00004***	0.00004***	0.00003***	0.00003***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Purchase_freq	0.07558***	0.07665***	0.04233***	0.04337***
	(0.00614)	(0.00632)	(0.00215)	(0.00218)
ln(Exp_ann)	0.01836**	0.01359	-0.11053***	-0.10788***
	(0.00688)	(0.00701)	(0.00445)	(0.00450)
Web_rec	0.00148***	0.00152***	0.00212***	0.00216***
	(0.00006)	(0.00006)	(0.00006)	(0.00006)
Web_freq	0.00651***	0.00659***	0.01024***	0.01043***
	(0.00027)	(0.00028)	(0.00024)	(0.00024)
Email_cat (med)	0.10196**	0.10013**	-0.03123	-0.03351
	(0.03144)	(0.03208)	(0.02944)	(0.02992)
Email_cat (high)	0.14181***	0.13367***	0.06604***	0.04196*
	(0.02121)	(0.02138)	(0.01846)	(0.01866)
Email_sent_rec	0.00253***	0.00255***	0.00443***	0.00434***
	(0.00005)	(0.00005)	(0.00006)	(0.00006)
Promo	0.19924***	-0.01357	0.24517***	0.03830
	(0.02083)	(0.02132)	(0.02110)	(0.02177)
ln(Period_length)	1.00664***	1.00567***	1.05182***	1.05086***
	(0.00868)	(0.00874)	(0.00952)	(0.00957)
Markdown	0.18128***	0.18268***	0.19388***	0.19552***
	(0.01074)	(0.01103)	(0.01107)	(0.01129)
2014	-0.30165***	-0.28412***	-0.30028***	-0.28243***
	(0.01586)	(0.01616)	(0.01626)	(0.01658)
2015	-0.56728***	-0.55444***	-0.71711***	-0.69817***
	(0.01709)	(0.01731)	(0.01648)	(0.01679)
lnsigma			0.10025***	0.05741**
N	1,045,405	1,043,813	1,045,405	1,043,813

Standard errors in parentheses
Note: There are 21,767 customers
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

correlation coefficients of opposite signs in the two models. The results of estimation of Heckman model are presented in Table B.8.

Table B.8: Estimating aggregate and advertisement models using Heckman model (two-stage).

	Aggregate Model		Advertisement Model	
	Expenditure	Purchase	Expenditure	Purchase
Constant	4.93675***	-2.77537***	4.92016***	-2.77190***
	(0.03703)	(0.01880)	(0.03745)	(0.01902)
Purchase_rec	0.00262***	-0.00787***	0.00263***	-0.00783***
	(0.00007)	(0.00009)	(0.00007)	(0.00009)
Purchase_rec^2		0.00002***		0.00002***
		(0.00000)		(0.00000)
Purchase_freq	0.00274**	0.03948***	0.00324**	0.03980***
	(0.00102)	(0.00066)	(0.00104)	(0.00067)
ln(Exp_ann)	0.19721***	0.01031***	0.19904***	0.00780***
	(0.00343)	(0.00164)	(0.00348)	(0.00168)
Web_rec	-0.00027***	0.00061***	-0.00025***	0.00063***
	(0.00005)	(0.00002)	(0.00005)	(0.00002)
Web_freq	-0.00246***	0.00335***	-0.00248***	0.00338***
	(0.00013)	(0.00006)	(0.00013)	(0.00007)
Email_cat (med)	0.09607***	0.05785***	0.09006***	0.05807***
	(0.02422)	(0.01202)	(0.02455)	(0.01223)
Email_cat (high)	0.00035	0.07310***	-0.00832	0.06978***
	(0.01163)	(0.00704)	(0.01178)	(0.00714)
Email_sent_rec		0.00121***		0.00123***
		(0.00002)		(0.00002)
Promo	0.38852***	0.14074***	0.23047***	0.05332***
	(0.01613)	(0.00943)	(0.01775)	(0.00964)
ln(Period_length)		0.48784***		0.48749***
		(0.00440)		(0.00442)
Markdown	-0.08460***	0.07828***	-0.07729***	0.07893***
	(0.01025)	(0.00503)	(0.01038)	(0.00513)
2014	0.03363*	-0.14724***	0.03947*	-0.13681***
	(0.01522)	(0.00767)	(0.01544)	(0.00783)
2015	0.09667***	-0.27253***	0.10226***	-0.26502***
	(0.01526)	(0.00750)	(0.01547)	(0.00766)
lambda		-0.39502***		-0.39643***
		(0.01629)		(0.01640)
N	1,045,405		1,043,813	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The probability of a purchase during a promotion period is 5.9% which is 25.82% greater than the non-promotion period purchase probability. We also find a marginal advertisement effect of the promotion on customers who do not participate in the promotion. The coefficient of Promo in the $\ln(\textsf{Expenditure})$ equation suggests a strong positive effect of the promotion period on customer expenditure. The average customer expenditure during a promotion period is \$1096.41, which is 49.23% greater than the average expenditure during a non-promotion period.

Promotion Period Model

We consider three self-selection models involving multinomial choice proposed by Lee (1983); Dubin and McFadden (1984), and Dahl (2002). These two-stage methods differ in terms of the distributional assumptions made between the errors in the outcome equation (ξ_i) and the utility associated with each alternative (ν_0, ν_1, ν_2). Similar to the approach we take, Lee (1983) models joint distribution between ξ_i and ν_i^*, but does not account for potential correlation between ξ_i and the primitive error terms ν_i's. Dubin and McFadden (1984) instead model the selectivity bias term ($\mathbb{E}[\xi_i|\textsf{Purchase} = 1]$) as a linear combination of ν_i's and estimates correlation between each pair $(\xi_i, \nu_j)_{j=0,1,2}$. Dahl (2002) further relaxes the linearity assumption. Bourguignon et al. (2007) provide an extensive discussion of these methods and tests their performance on simulated data. They find that the method proposed by Dahl (2002) generally outperforms others in terms of bias correction. Here we present results from the estimation of the expenditure equations for customers who make a purchase without participation (Table B.9) and those who participate in the promotion (Table B.10).

Table B.9: Estimation of expenditure equation for customers who do not participate in the promotion.

	Lee (1983)	Dubin and McFadden (1984)	Dahl (2002)
Constant	4.49604***	3.04776***	34.34173***
	(0.09734)	(0.35231)	(10.19847)
Purchase_rec	0.00253***	0.00267***	0.00257***
	(0.00022)	(0.00022)	(0.00018)
Purchase_freq	-0.00244	-0.00090	-0.00494
	(0.00200)	(0.00215)	(0.00356)
Exp_annual	0.18971***	0.18536***	0.18002***
	(0.01000)	(0.01009)	(0.01019)
Web_rec	-0.00002	-0.00005	0.00001
	(0.00013)	(0.00013)	(0.00017)
Web_freq	-0.00136***	-0.00171***	-0.00169***
	(0.00030)	(0.00028)	(0.00035)
Email_cat (med)	0.13477*	0.12729	0.10178*
	(0.05853)	(0.06925)	(0.04664)
Email_cat (high)	-0.03908	-0.04240	-0.07042*
	(0.02851)	(0.03237)	(0.03371)
Markdown	-0.07216**	-0.08128**	-0.08083**
	(0.02594)	(0.02751)	(0.02828)
2014	0.05591	0.10119*	0.13952***
	(0.03928)	(0.04493)	(0.03945)
2015	0.09118**	0.12752**	0.16221***
	(0.03348)	(0.04039)	(0.04648)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.10: Estimation of expenditure equation for customers who participate in the promotion.

	Lee (1983)	Dubin and McFadden (1984)	Dahl (2002)
Constant	4.62953***	5.81366***	30.59510*
	(0.81997)	(1.54168)	(15.57035)
Purchase_rec	0.00093***	0.00085**	0.00104***
	(0.00027)	(0.00028)	(0.00030)
Purchase_freq	0.00654*	0.00910**	0.00254
	(0.00315)	(0.00321)	(0.00341)
ln(Exp_annual)	0.08835***	0.07582***	0.08711***
	(0.01171)	(0.01527)	(0.01292)
Web_rec	0.00009	0.00030	0.00006
	(0.00024)	(0.00024)	(0.00025)
Web_freq	-0.00147**	-0.00155**	-0.00193***
	(0.00047)	(0.00050)	(0.00053)
Email_cat (med)	0.11308	0.05271	0.07676
	(0.09488)	(0.08355)	(0.10021)
Email_cat (high)	-0.02068	-0.08645	-0.03160
	(0.04098)	(0.04943)	(0.04865)
Private	0.08234	0.18505*	0.10935
	(0.09485)	(0.08229)	(0.09288)
ln(Lowest_threshold)	0.48240***	0.60701***	0.50945***
	(0.10694)	(0.10231)	(0.10358)
Nbr_tiers	-0.09031	-0.09603	-0.08457
	(0.04936)	(0.05280)	(0.06528)
Discount	-0.03663***	-0.03936***	-0.03751***
	(0.00573)	(0.00724)	(0.00801)
Markdown	0.08986	0.12090*	0.08740
	(0.04663)	(0.05209)	(0.05540)
2014	0.25847***	0.32498***	0.27454***
	(0.05109)	(0.06370)	(0.06756)
2015	0.23580***	0.30879***	0.27139***
	(0.05701)	(0.06011)	(0.06376)

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

REFERENCES

Aquino, D. (2009). The resurgence of vendor-managed inventory: A landscape. *Gartner*.

Aviv, Y. (2001). The effect of collaborative forecasting on supply chain performance. *Management Science 47*(10), 1326–1343.

Aviv, Y. (2002). Gaining benefits from joint forecasting and replenishment processes: The case of auto-correlated demand. *Manufacturing & Service Operations Management 4*(1), 55–74.

Aviv, Y. (2004). Collaborative forecasting and its impact on supply chain performance. In D. Simchi-Levi, S. D. Wu, and Z.-J. Shen (Eds.), *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era*, pp. 393–446. Boston, MA: Springer US.

Aviv, Y. (2007). On the benefits of collaborative forecasting partnerships between retailers and manufacturers. *Management Science 53*(5), 777–794.

Azoury, K. S. (1985). Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science 31*(9), 1150–1160.

Baardman, L., M. C. Cohen, K. Panchamgam, G. Perakis, and D. Segev (2017). Scheduling promotion vehicles to boost profits. Working Paper, New York University, New York City, NY. Retrieved from https://ssrn.com/abstract=2638396orhttp://dx.doi.org/10.2139/ssrn.2638396 on September 21, 2017.

Babich, V., H. Li, P. Ritchken, and Y. Wang (2012). Contracting with asymmetric demand information in supply chains. *European Journal of Operational Research 217*(2), 333 – 341.

Bass, F. M., N. Bruce, S. Majumdar, and B. P. S. Murthi (2007). Wearout effects of different advertising themes: A dynamic bayesian model of the advertising-sales relationship. *Marketing Science 26*(2), 179–195.

Bell, D. R., J. Chiang, and V. Padmanabhan (1999). The decomposition of promotional response: An empirical generalization. *Marketing Science 18*(4), 504–526.

Bensoussan, A., B. Kadiyala, and Ö. Özer (2017). A strategic approach to collaborative inventory management. Working Paper, University of Texas at Dallas, Richardson, TX.

Bergemann, D. and J. Välimäki (2006). Information in mechanism design. In R. Blundell, W. Newey, and T. Persson (Eds.), *Proceedings of the 9th World Congress of the Econometric Society*, pp. 186–221. Cambridge University Press.

Berger, R. L. and G. Casella (2002). *Statistical Inference* (2nd ed.). Duxbury Press.

Birge, J. and B. Keskin (2015). Dynamic pricing and product differentiation with cost uncertainty and learning. Working Paper, Duke University, Durham, NC.

Bisi, A., M. Dada, and S. Tokdar (2011). A censored-data multiperiod inventory problem with newsvendor demand distributions. *Manufacturing & Service Operations Management 13*(4), 525–533.

Blackhurst, J., C. W. Craighead, and R. B. Handfield (2006). Towards supply chain collaboration: An operations audit of VMI initiatives in the electronics industry. *International Journal of Integrated Supply Management 2*(1), 91–105.

Blattberg, R. C. and R. A. Briesch (2012). Sales promotions. In Ö. Özer and R. Phillips (Eds.), *The Oxford Handbook of Pricing Management*, Chapter 24. Oxford University Press.

Bourguignon, F., M. Fournier, and M. Gurgand (2007). Selection bias corrections based on the multinomial logit model: Monte carlo comparisons. *Journal of Economic Surveys 21*(1), 174–205.

Braden, D. J. and M. Freimer (1991, November). Informational dynamics of censored observations. *Management Science 37*(11), 1390–1404.

Brinkhoff, A., Ö. Özer, and G. Sargut (2015). All you need is trust? An examination of inter-organizational supply chain projects. *Production and Operations Management 24*(2), 181–200.

Cachon, G. P. and M. Fisher (2000). Supply chain inventory management and the value of shared information. *Management Science 46*(8), 1032–1048.

Cachon, G. P. and M. A. Lariviere (2001). Contracting to assure supply: How to share demand forecasts in a supply chain. *Management Science 47*(5), 629–646.

Cachon, G. P. and R. Swinney (2009). Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science 55*(3), 497–511.

Cachon, G. P. and R. Swinney (2011). The value of fast fashion: Quick response, enhanced design, and strategic consumer behavior. *Management Science 57*(4), 778–795.

Caliendo, M. and S. Kopeinig (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys 22*(1), 31–72.

Chen, L. (2010). Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. *Operations Research 58*(2), 396–413.

Chen, L. and H. L. Lee (2009). Information sharing and order variability control under a generalized demand model. *Management Science 55*(5), 781–797.

Chen, L. and E. L. Plambeck (2008). Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing & Service Operations Management 10*(2), 236–256.

Chen, Y., S. Moorthy, and Z. J. Zhang (2005). Research noteprice discrimination after the purchase: Rebates as state-dependent discounts. *Management Science 51*(7), 1131–1140.

Chintagunta, P. K. (1993). Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science 12*(2), 184–208.

Cohen, M. C., N.-H. Z. Leung, K. Panchamgam, G. Perakis, and A. Smith (2017). The impact of linear optimization on promotion planning. *Operations Research 65*(2), 446–468.

Dahl, G. B. (2002). Mobility and the return to education: Testing a roy model with multiple markets. *Econometrica 70*(6), 2367–2420.

Dickson, P. R. and A. G. Sawyer (1990). The price knowledge and search of supermarket shoppers. *Journal of Marketing 54*(3), 42–53.

Dong, Y., M. Dresner, and Y. Yao (2014). Beyond information sharing: An empirical analysis of vendor-managed inventory. *Production and Operations Management 23*(5), 817–828.

Dubin, J. A. and D. L. McFadden (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica 52*(2), 345–362.

Ernst and Young (2013). Margin unlocked: Integrated margin management to deliver breakthrough performance in consumer products. Report.

Federal Reserve System (2010, "1" April). Electronic funds transfers; final rule. *Federal Register 75*(62), 16580–16621.

Feng, Q., G. Lai, and L. X. Lu (2015). Dynamic bargaining in a supply chain with asymmetric demand information. *Management Science 61*(2), 301–315.

Ferguson, D. (2013). How supermarkets get your data – and what they do with it. *The Guardian.*. Retreived from http://www.theguardian.com/money/2013/jun/08/supermarkets-get-your-dataonSeptember21,2017.

Ferreira, K. J., B. H. A. Lee, and D. Simchi-Levi (2016). Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management 18*(1), 69–88.

Gilpatric, S. M. (2009). Slippage in rebate programs and present-biased preferences. *Marketing Science 28*(2), 229–238.

Gönül, F. and M. Z. Shi (1998). Optimal mailing of catalogs: A new methodology using estimable structural dynamic programming models. *Management Science 44*(9), 1249–1262.

Gönül, F. F. and F. T. Hofstede (2006). How to compute optimal catalog mailing decisions. *Marketing Science 25*(1), 65–74.

Grocer, T. (2009). Northern in talks over Sainsbury's VMI deal. Retreived from http://www.thegrocer.co.uk/companies/northern-in-talks-over-sainsburys-vmi-deal/199333.article on September 21, 2017.

Gupta, S. (1988). Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing Research 25*(4), 342–355.

Ha, A. Y., S. Tong, and H. Zhang (2011). Sharing demand information in competing supply chains with production diseconomies. *Management Science 57*(3), 566–581.

Hammond, J. H. (2006). Barilla SpA (D). Harvard Business School Case 9-695-066, Harvard University, Cambridge, MA.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Kadiyala, B., Ö. Özer, and S. Şimşek (2017). Value of delayed incentive: An empirical investigation of gift card promotion. Working Paper, University of Texas at Dallas, Richardson, TX.

Kapner, S. (2013). The dirty secret of black friday 'discounts'. *The Wall Street Journal*. Retrieved from http://www.wsj.com/articles/SB10001424052702304281004579217863262940166 on September 21, 2017.

Karlin, S. and H. Scarf (1958). Inventory models of the Arrow-Harris-Marschak type with time lag. In K. Arrow, S. Karlin, and H. Scarf (Eds.), *Studies in the Mathematical Theory of Iventory and Production*. Stanford University Press.

Katariya, A. P., S. Çetinkaya, and E. Tekin (2014). Cyclic consumption and replenishment decisions for vendor-managed inventory of multisourced parts in Dell's supply chain. *Interfaces 44*(3), 300–316.

Khan, R., M. Lewis, and V. Singh (2009). Dynamic customer management and the value of one-to-one marketing. *Marketing Science 28*(6), 1063–1079.

Khouja, M., J. Pan, B. T. Ratchford, and J. Zhou (2011). Analysis of free gift card program effectiveness. *Journal of Retailing 87*(4), 444–461.

Kouvelis, P., C. Chambers, and H. Wang (2006). Supply chain management research and production and operations management: Review, trends, and opportunities. *Production and Operations Management 15*(3), 449–469.

Krishnamurthi, L. and S. P. Raj (1988). A model of brand choice and purchase quantity price sensitivities. *Marketing Science 7*(1), 1–20.

Kumar, V., X. A. Zhang, and A. Luo (2014). Modeling customer opt-in and opt-out in a permission-based marketing context. *Journal of Marketing Research 51*(4), 403–419.

Lariviere, M. A. and E. L. Porteus (1999). Stalking information: Bayesian inventory management with unobserved lost sales. *Management Science 45*(3), 346–363.

Lee, H. L., V. Padmanabhan, and S. Whang (1997). The Bullwhip effect in supply chains. *MIT Sloan Management Review 38*(3).

Lee, H. L., K. C. So, and C. S. Tang (2000). The value of information sharing in a two-level supply chain. *Management Science 46*(5), 626–643.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica 51*(2), 507–512.

Li, L. and H. Zhang (2008). Confidentiality and information sharing in supply chain coordination. *Management Science 54*(8), 1467–1481.

Long, H. (2015). This is the most requested gift for 8 years in a row. *CNNMoney*. Retieved from http://money.cnn.com/2015/08/10/investing/gift-cards-soar-in-popularity/ on September 21, 2017.

Lovejoy, W. S. (1990). Myopic policies for some inventory models with uncertain demand distributions. *Management Science 36*(6), 724–738.

Lu, Q. and S. Moorthy (2007). Coupons versus rebates. *Marketing Science 26*(1), 67–82.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Number 3. Cambridge university press.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995). *Microeconomic theory*, Volume 1. Oxford university press New York.

Mathews, R. (1995). Spartan pulls the plug on VMI. *Progressive Grocer*.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, Chapter 4. Academic Press.

McFadden, D. (1978). Quantitative methods for analyzing travel behavior of individuals: Some recent developments. In D. A. Hensher and P. R. Stopher (Eds.), *Behavioral Travel Modeling*, Chapter 13. Croom Helm London.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies 38*(2), pp. 175–208.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

Narasimhan, C. (1984). A price discrimination theory of coupons. *Marketing Science 3*(2), 128–147.

Nasiry, J. and I. Popescu (2012). Advance selling when consumers regret. *Management Science 58*(6), 1160–1177.

Neslin, S. A. (2002). *Sales Promotion*. Boston: Marketing Science Institute.

Neslin, S. A., G. A. Taylor, K. D. Grantham, and K. R. McNeil (2013). Overcoming the recency trap in customer relationship management. *Journal of the Academy of Marketing Science 41*(3), 320–337.

Oh, S. and Ö. Özer (2013). Mechanism design for capacity planning under dynamic evolutions of asymmetric demand forecasts. *Management Science 59*(4), 987–1007.

Oracle (2009, November). Peoplesoft enterprise inventory 9.1 peoplebook. Technical report.

Özer, Ö. and W. Wei (2006). Strategic commitments for an optimal capacity decision under asymmetric forecast information. *Management Science 52*(8), 1238–1257.

Özer, Ö. and Y. Zheng (2016). Markdown or everyday low price? The role of behavioral motives. *Management Science 62*(2), 326–346.

Quinn, M. (2014). Why retailers love gift cards but you shouldn't. *CBS MoneyWatch*. Retrieved from http://www.cbsnews.com/news/why-retailers-love-gift-cards-but-you-shouldnt/ on September 21, 2017.

Ren, Z. J., M. A. Cohen, T. H. Ho, and C. Terwiesch (2010). Information sharing in a long-term supply chain relationship: The role of customer review strategy. *Operations Research 58*(1), 81–93.

Roesler, A.-K. (2014). Mechanism design with endogenous information. Working Paper, Bonn Graduate School of Economics, Bonn, Germany.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rosenblum, P. (2014). Walmart's out of stock problem: Only half the story? *Forbes*. Retreived from http://www.forbes.com/sites/paularosenblum/2014/04/15/walmarts-out-of-stock-problem-only-half-the-story/ on September 21, 2017.

Sahni, N. S., D. Zou, and P. K. Chintagunta (2016). Do targeted discount offers serve as advertising? Evidence from 70 field experiments. *Management Science*. Forthcoming.

Scarf, H. (1959, 06). Bayes solutions of the statistical inventory problem. *The Annals of Mathematical Statistics 30*(2), 490–508.

Scarf, H. E. (1960). Some remarks on Bayes solutions to the inventory problem. *Naval Research Logistics Quarterly 7*(4), 591–596.

Shang, W., A. Y. Ha, and S. Tong (2016). Information sharing in a supply chain with a common retailer. *Management Science 62*(1), 245–263.

Simchi-Levi, D., P. Kaminsky, and E. Simchi-Levi (2008). *Designing & Managing the Supply Chain* (3rd ed.). New York, NY: McGraw-Hill/Irwin.

Soman, D. (2001). Effects of payment mechanism on spending behavior: The role of rehearsal and immediacy of payments. *Journal of Consumer Research 27*(4), 460–474.

Soysal, G. P. and L. Krishnamurthi (2012). Demand dynamics in the seasonal goods industry: An empirical analysis. *Marketing Science 31*(2), 293–316.

Soysal, G. P. and L. Krishnamurthi (2013). Buying and selling seasonal goods. *Kellog Insight*. Retrieved from https://insight.kellogg.northwestern.edu/article/buying_and_selling_seasonal_goods on September 21, 2017.

Stevens, L. and S. Banjo (2014). Free shipping is going to cost you more. *The Wall Street Journal*. Retrieved from https://www.wsj.com/articles/free-shipping-is-going-to-cost-you-more-1414003507 on September 21, 2017.

Su, X. (2007). Intertemporal pricing with strategic customer behavior. *Management Science 53*(5), 726–741.

Su, X. (2010). Intertemporal pricing and consumer stockpiling. *Operations Research 58*(4-part-2), 1133–1147.

Tirole, J. (2002). *The Theory of Industrial organization*. The MIT Press.

Trost, R. P. and L.-F. Lee (1984). Technical training and earnings: A polychotomous choice model with selectivity. *The Review of Economics and Statistics 66*(1), 151–156.

Van Heerde, H. J., S. Gupta, and D. R. Wittink (2003). Is 75% of the sales promotion bump due to brand switching? No, only 33% is. *Journal of Marketing Research 40*(4), 481–491.

Varian, H. R. (1980). A model of sales. *The American Economic Review 70*(4), 651–659.

Venkatesan, R. and P. W. Farris (2012). Measuring and managing returns from retailer-customized coupon campaigns. *Journal of Marketing 76*(1), 76–94.

Watson, E. (2005, August). Nestlé switches to vendor managed inventory with Tesco. Retreived from http://www.foodmanufacture.co.uk/Supply-Chain/Nestle-switches-to-vendor-managed-inventory-with-Tesco on September 21, 2017.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach* (5th ed.). South-Western, Cengage Learning.

Zhang, H., M. Nagarajan, and G. Sošić (2010). Dynamic supplier contracts under asymmetric inventory information. *Operations Research 58*(5), 1380–1397.

Zhang, H. and S. Zenios (2008). A dynamic principal-agent model with hidden information: Sequential optimality through truthful state revelation. *Operations Research 56*(3), 681–696.

Zhang, J. and L. Krishnamurthi (2004). Customizing promotions in online stores. *Marketing Science 23*(4), 561–578.

BIOGRAPHICAL SKETCH

Bharadwaj Kadiyala was born in Nidadavole, in the state of Andhra Pradesh, India. After completing high school in Hyderabad, he moved to Dhirubhai Ambani Institute of Information and Communication Technology, to pursue an undergraduate degree in Information and Communication Technology. Fascinated by courses in optimization and queueing theory, he pursued master's degrees in applied mathematics and industrial engineering at Indiana University-Purdue University Indianapolis and Bilkent University, respectively. Following that he moved to The University of Texas at Dallas to pursue doctoral studies in operations management.

CURRICULUM VITAE

Bharadwaj Kadiyala

September 6, 2017

Contact Information:

800 West Campbell Road, JSOM 3.218
Richardson, TX 75080, U.S.A.
Phone: (972) 883-6249
Email: bharadwaj.kadiyala@utdallas.edu

Educational History:

B.Tech., Information and Communication Technology, Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), India, 2008
M.S., Applied Mathematics, Purdue University, U.S.A., 2010
M.S., Industrial Engineering, Bilkent University, Turkey, 2012

Poisson Disorder Problem with Continuous Control of Costly Observations
Masters Thesis
Industrial Engineering Department, Bilkent University
Advisor: Dr. Savaş Dayanık

Estimation of Broadcast Probability in Wireless Ad-hoc Networks
Undergraduate Thesis
DA-IICT
Advisor: Dr. Sunitha V Murugan

Professional Recognitions and Honors:

Finalist, POMS College of Supply Chain Management Student Paper Competition, 2017
Recipient of the Cyrus Cantrell III UTD Phi Kappa Phi Travel Grant, UT Dallas, 2016
Inducted into Phi Kappa Phi honor society, UT Dallas, 2016
Recipient of the Betty and Gifford Johnson Travel Award, UT Dallas, 2015
First place (among 80 interns) in Intern Case Competition at Sabre Holdings, 2015
Elected by the undergraduate student body to the annual college festival (Synapse) committee, DA-IICT, 2006
All India Rank 113 in the 5th National Science Olympiad, 2003

Professional Memberships:

Institute for Operations Research and the Management Sciences (INFORMS), 2012–present
Manufacturing and Service Operations Management Society (MSOM), 2012–present
Production and Operations Management Society (POMS), 2014–present