Spring 5-15-2022

Optimization and Information Problems in Operations

Puping Jiang
Washington University in St. Louis, jiang.p@wustl.edu

Recommended Citation

WASHINGTON UNIVERSITY IN ST. LOUIS

Olin Business School

Dissertation Examination Committee:
Lingxiu Dong, Chair
Panos Kouvelis, Co-Chair
René Caldentey
Naveed Chehrazi
Jacob Feldman

Optimization and Information Problems in Operations

by

Puping Jiang

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2022
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

First, I would like to thank my dissertation committee chairs and advisers Lingxiu Dong and Panos Kouvelis for their outstanding mentorship throughout my PhD. They were always here to support and assist me by spending countless hours discussing our research. And thank you both for guiding my steps well beyond the academic life, especially your encouragement and insights at the start point of my postgraduate career. My deepest gratitude also goes to Jacob Feldman and René Caldentey for their great support and patience. My meetings with them have always been some of the best parts of my PhD memory. Every meeting was loaded with lively discussions on technical challenges and full of intellectual joy. It is hard to imagine this thesis without their sharp insights and most generous help. I would also like to express my sincere gratitude to Naveed Chehrazi for joining my dissertation committee. His insightful questions and shrewd suggestions helped me greatly improve the thesis.

Moreover, I would like to thank the WUSTL community and my fellow PhD friends for their help and all the joyful moments we have together. Also my sincere thanks go to staff members of Olin Business School's doctoral program: Jessica Hatch, Gina Poe, Shengxin Sun, Kristy Chan, and staff members of Boeing center: Evan Dalton and Andy Sample for their kind administrative support.

Last but not least, I am extremely grateful to my parents, my girlfriend, and other family members for their tireless and unconditional love, support, and encouragement throughout this journey. To them I dedicate this thesis.

Puping Jiang

Washington University in St. Louis

May 2022

Dedicated to my parents, my girlfriend, and other family members.

ABSTRACT OF THE DISSERTATION

Optimization and Information Problems in Operations

by

Puping Jiang

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2022

Professor Lingxiu Dong, Chair

Professor Panos Kouvelis, Co-Chair

The main purpose of this dissertation is to study the optimization problems and the value of information in various commercial settings, especially the emerging platform economy.

Chapter 1, "Data-Driven Asset Selling". Motivated by online asset selling marketplace business (e.g., used cars and real estate), we formulate a data-driven asset selling dynamic pricing framework which utilizes platforms' access to customers' online behavioral data. With mild assumptions on the demand model, careful characterization of the problem structure shows that the model admits some ideal properties that facilitate our regret analysis under our dynamic programming setting. Instead of studying the policy performance with a long horizon and large quantities of inventory, we study the asymptotic policy performance over a single unit of product as the demand rate grows. We propose a deterministic approximation policy (DA policy) and show that DA policy provides an upper bound for the original problem and its induced pricing policy achieves asymptotic optimality as the scale of the problem grows properly. Later we consider a dynamic pricing scenario where an idiosyncratic latent value for each asset is unknown. We propose a Thompson-Sampling-based and a MAP-

based pricing and learning policy. Since the platform is restricted in an infrequent pricing environment, within each decision epoch, an adequate amount of customer online behavior records is available. Utilizing large-sample deviation properties, we are able to conduct regret analysis on the TS and MAP policies. Finally, we use numerical experiments to show that our proposed algorithms could improve the revenue performance significantly compared with an algorithm that is currently implemented by a leading used car platform. Besides, we find that using a simple deterministic proxy of demand forecast is mostly harmless, while accurate estimation on the idiosyncratic latent value can make significant differences. Simulations also reveal that in our problem setting, the exploration step in the TS policy may not help to outperform the MAP policy. This indicates that the effectiveness of exploration highly depends on the nature of the problem, which may be of independent interest.

Chapter 2, "Cash Hedging Motivates Information Sharing in Supply Chains". Finance literature well documents that firms' cash hedging strategies heavily depend on the market conditions. Unsurprisingly, such decisions could be challenging for an upstream firm in a supply chain where the end market conditions are not transparent to him. In this paper, we study the interplay between firms' information sharing behaviors and cash hedging strategies in supply chains. First, we argue that the presence of a supplier's cash hedging decision may motivate downstream retailers' voluntary market information sharing with the supplier, since making the supplier more informed of the market conditions helps the retailer handle her risk in the wholesale price. This also forms a new reason why a supplier should consider hedging, since the cash hedging decision itself can be used as a bargaining tool during the information sharing negotiation with his retailer. Then we find for homogeneous Cournot-competing retailers, asymmetric information-sharing outcomes could emerge as an equilibrium where publicly sharing information typically will not hurt, especially, sometimes it can achieve Pareto improvement of the supply chain and consumer welfare. Finally, when a single supplier serves multiple markets, the heterogeneity across market sizes and the cor-

relation among market shocks play big roles in shaping the equilibrium. Especially in a simultaneous information-sharing game, greater market size heterogeneity and negatively correlated market shocks are more likely to result in the nonexistence of pure Nash equilibrium. When the Stackelberg sequence is introduced, greater market size heterogeneity and positively correlated market shocks are more likely to induce information sharing in the equilibrium. Furthermore, in the multi-market setting, the existence of an information-sharing channel may hurt retailers, the system as a whole, and consumer welfare.

Chapter 3, "Display Optimization under the Multinomial Logit Choice Model: Balancing Revenue and Customer Satisfaction". In this paper, we consider an assortment optimization problem in which a platform must choose pairwise disjoint sets of assortments to offer across a series of T stages. Arriving customers begin their search process in the first stage, and progress sequentially through the stages until their patience expires, at which point they make a multinomial-logit-based purchasing decision from among all products they have viewed throughout their search process. The goal is to choose the sequential displays of product offerings to maximize expected revenue. Additionally, we impose stage-specific constraints that ensure that as each customer progresses farther and farther through the T stages, there is a minimum level of "desirability" met by the collections of displayed products. We consider two related measures of desirability: purchase likelihood and expected utility derived from the offered assortments. In this way, the offered sequence of assortment must be both high earning and well-liked, which breaks from the traditional assortment setting, where customer considerations are generally not explicitly accounted for.

We show that our assortment problem of interest is strongly NP-Hard, thus ruling out the existence of a fully polynomial-time approximation scheme (FPTAS). From an algorithmic stand-point, as a warm-up, we develop a simple constant factor approximation scheme in which we carefully stitch together myopically selected assortments for each stage. Our main algorithmic result consists of a polynomial-time approximation scheme (PTAS), which

combines a handful of structural results related to the make-up of the optimal assortment sequence within an approximate dynamic programming framework.

1. Data-Driven Asset Selling

1.1 Introduction

With the fast developing business model of online asset selling platforms, e.g., Guazi (the largest used car marketplace in China), CarMax (the largest used car marketplace in the US) and Zillow (the largest real estate platform in the US), impressive growth in volume has been taking place in the last decade, and the growth is forecasted to accelerate. Real estate market is always being a significant part of economy. For used car markets, for instance, China's used car market is predicted to double the size by 2025 compared with 2020 at value worth \$306 billion.[1] According to McKinsey the US used car market was worth twice as much as new car market in 2018.[2] Especially, during the pandemic disruption, the used car market demonstrated significant robustness.[3] Due to a couple of unique features of online asset selling business, the operations of such platforms are drastically different from a traditional inventory selling business. First, each unit of asset has idiosyncratic attributes, thus the selling prices are essentially item-wise. Second, the inventory replenishment is only partially controlled by the platform mainly via acquiring assets from exogenous individual asset sellers. Besides, some specific business constraints further complicate the operations. For example, in order to keep a high inventory turnover rate, used car platforms may set targeted selling horizon for each car according to our conversations with a leading used car platform where each car typically has an on-site life length around 7 weeks after which it will

[1]Bloomberg: China Wants to Build a \$306 Billion Used-Car Market From Scratch.
[2]McKinsey: Used cars, new platforms: Accelerating sales in a digitally disrupted market.
[3]The Wall Street Journal: During Covid-19 Pandemic, the Used-Car Lot Is Hot.

be salvaged to other dealers. Also, frequent price change is unfavorable from the platform's perspective. On one hand, frequent price change is both computationally and practically (e.g., menu cost) expensive; on the other hand, frequently changing prices may encourage unideal strategic customer behaviors. In practice, a platform may update the prices at most on a weekly basis (e.g., Guazi). According to [1], around 30% of the used cars on CarMax experienced price changes during their on-site time (which is on average 14 days), but very few of them experienced price changes over three times before being sold. Although the price changes can take place more often in other large dealers, the pricing schemes are way from being regarded as frequent pricing. Therefore, the inventory dynamics and the pricing problem in the asset selling platforms is hard to be characterized by the frameworks from seasonal products pricing literature where the decision makers typically have a pool of identical items to sell and are able to change prices frequently. Unfortunately, given the significance of the asset selling business and the uniqueness of its operational model, very few work has been dedicated to formulate relevant models and this gap between literature and practice motivates our paper.

Another major business challenge for the platform is to characterize customers' valuations over assets' idiosyncratic attributes. For example, in used car business, typically a platform would conduct close inspections on each car covering over one hundred points (see [1]'s discussion on CarMax data, and Figure 1.1, screenshots from Guazi app, in Section 1.5). Those inspection details are often available for customers, but how those idiosyncratic features would translate into customers' reactions is often unclear. There could be several reasons to that, e.g., the historical data is too sparse to support an accurate estimation compared to the number of possible combinations of all features, and customers' evaluations or tastes may change over time which downgrades the relevance of data collected long time ago. So typically, the platform would have some parsimonious estimation models which capture some major features like car model, color, mileage, and some stylized rubrics on mechanical

nuances, yet still leaves room for learning customers' reactions to those idiosyncratic features (or we call *latent value*) on-the-fly for each car during price optimization. Indeed, as empirical evidence has shown (see [1]), the value of learning such car-specific latent value could be significant. In a traditional brick-and-mortar asset selling business, customer behavior data is not really informative, since the historical data is just a sequence of no-purchase before a final purchase is made and then the problem ends. Therefore learning in a non-platform setting would simply lead to a decreasing estimate over time with a markdown price trajectory (see [2]). However, in the asset-selling platform's setting, the data of customers' online behaviors before their final purchase decisions is observable from the platform's perspective, which enhances the decision maker's ability of learning an asset's latent value. This motivates our research on incorporating learning into the dynamic pricing framework.

In this paper, we build up a dynamic pricing model incorporating customers' online behavior process for an online asset selling platform, mainly motivated by the used car business. But our framework can also be adapted to other online asset selling business. Specifically based on our conversations with a leading used car platform, we have two high level modeling assumptions. First, dynamic pricing policy is designed for each car separately, in other words, we do not explicitly model the substitution effect led by price changes on other cars. This is similar to the framework in the closely related empirical work [1], while stylized way of capturing substitution effect is always possible. A side benefit of this framework is that car acquisition decision is relatively trivial. This is because on one hand, dynamic pricing on the selling side is independent of the acquisition price (which is sunk once acquired); on the other hand, given individual car seller's willingness to sell function, acquisition price optimization is a single-dimension static maximization problem trading off between the expected acquisition cost and the expected selling revenue. Therefore in this paper we only focus on the selling-side for a given car. Second, within each pricing decision epoch, the potential time delay between a customer's online behavior and offline purchase

decision is negligible, since the decision epoch is relatively long in practice, and customers typically don't have incentives to delay their visits to the offline store for their interested cars. Another way to interpret this modelling assumption is that the last step in a customer's sequential online behaviors (typically a test drive appointment) would almost surely inform the final in-store purchase decision, i.e., the customer who chooses to test drive a car will almost surely buy the car. This is because the car information provided online can be very detailed (see screenshots from Guazi app, Figure 1.1) and a customer bears both time and transportation costs to see the car, a final test drive generally should not lead to much surprise.

Our paper has the following major contributions. Our first contribution is in model formulation. Assuming log-concave individual customers' choice function in price, we build a dynamic asset pricing framework with Poisson demand process whose rate may change over a discrete horizon. Importantly, we relax the typical assumption in inventory dynamic pricing literature that in each decision epoch, at most one customer arrives at the system, and we also incorporates the volatility in demand rate. This assumption relaxation is necessary in model formulation to fit the practice reality and business interests. In reality, platforms would typically avoid implementing real-time and high-frequency dynamic pricing as we introduced before. Under the above mild assumptions, we show that this general asset selling framework admits nice structural properties and allows us to conduct regret analysis although the problem is a dynamic programming by its nature. We demonstrate both theoretically and empirically (via numerical experiments) the good performance of a simple pricing policy derived from deterministic approximation to the original problem under certain asymptotic regimes. A major advantage of constructing a model from individual utility function is to facilitate utilizing customers' online behavioral data to better understand demand pattern, which is particularly of the platform's interests. Related to this, our second contribution is to demonstrate how our model framework can utilize online behavior data to inform price

optimization and how it performs. We propose both active (Thompson-sampling-based) and passive learning (MAP-based) policies on the fly of price optimization. A typical difficulty in analyzing the performance of Thompson sampling is the complexity of characterizing the posterior belief, therefore people generally resort to weaker performance metric like Bayesian regret. Traditional regret analysis could be conducted in a few special cases where the models have ideal structures and prior beliefs are well selected in order to have nice conjugate structures. Performance analysis for MLE-type of learning strategy is also not easy since the estimation update dynamic is challenging to characterize. However in our setting, thanks to the infrequent nature of dynamic pricing, a potentially large amount of behavioral data is observable within each pricing decision epoch, consequently some large-sample statistical properties (e.g., sharper versions of Bernstein-von-Mises Theorem and Central Limit Theorem) start to take effect. Correspondingly, such large-sample properties allow us to conduct traditional regret analysis on our Thompson-sampling-based and MAP-based dynamic pricing policy. We show that our proposed Thompson-sampling-based policy and MAP-based policy both achieve asymptotic optimality at rate of $\mathcal{O}\left(\log(n)/\sqrt{n}\right)$ starting from the second decision epoch, where n is a scale factor in demand rate. Here we focus on the performance metric starting from the second decision epoch because in the first period there is no learning takes place, the performance is purely determined by the prior knowledge. Finally, we use simulations to demonstrate the potentially significant values of dynamic pricing and demand learning in our policies compared with a currently implemented policy by the platform. Numerical results reveal that using a simple deterministic proxy for demand forecast generally performs very well, but the estimation accuracy of the idiosyncratic latent value can make a significant difference. Interestingly, extensive simulations also reveal that Thompson-sampling-based policy may not outperform MAP-based policy, that is, the exploration step may not give learning strategies an edge.

Our paper calls researchers' attention to the fast-growing yet less studied business, the online asset selling business. We formulate an optimization framework that is particularly relevant to practice and propose implementation policies. Our work also highlights the opportunities in conducting rigorous policy performance analysis that is contingent to business settings, especially for learning policies when observable data scale is large.

The remainder of the paper is organized as follows. Section 1.2 reviews related literature and highlights our contributions again with the comparison to the previous work. We formulate our asset selling problem in section 1.3 and in section 1.4 we propose a deterministic approximation to the original stochastic optimization problem and show that the approximation serves as an upper bound in section 1.4.1. In section 1.4.2 we calculate the expected regret of the policy induced by the deterministic approximation. In section 1.5 we introduce how our framework can incorporate online behavioral data and conduct demand learning and price optimization via Thompson sampling and Maximum a posteriori estimation. We then show regret analysis could be done using large-sample deviation properties. We demonstrate the performance of our algorithms via numerical experiments in section 1.6 and finally we summarize our work in section 1.7.

1.2   Literature Review

Our work directly contributes the literature of asset selling. Representative work includes [1–5], etc. [2] studies a dynamic asset selling problem with Bayesian learning on demand rate, where the observable data is sequence of 0-1 purchase records and specifically the realized data is just a bunch of zeros, because once the purchase record turns out to be 1, the game ends. This is a typical offline asset selling setting, while the salient feature of our online asset selling is that a much richer data set is observable. Specifically, observing the online visit data, the platform has a good sense of potential demand process. Furthermore, besides

6

the final purchase record, the platform also has access to a sequence of customers' online behavior including clicking, adding to list, scheduling offline visit and test etc. Unlike the offline setting where the seller only observes a sequence of zero records before the game ends, which leads to a simple markdown pricing policy, online platforms are able to design more elaborate pricing policies based upon much better understandings of customers' willingness to pay. The most relevant work to our paper is [1], where the authors empirically identify significant value of car sellers' learning on cars' unobservable idiosyncratic features on an online used car marketplace, CarMax. This paper builds up a structural model under a dynamic pricing framework, which is the setting that motivates our research. Due to the technical tractability concern and pragmatic relevance, the dynamic pricing optimization is conducted on a single-car base and substitution effect is characterized via a heuristic manner. Our paper falls into the similar framework. [3] studies a dynamic asset selling problem when the seller is under debt, where a stylized aggregate demand function is assumed. Our work on the other hand, formulates the problem starting from customers' utility functions and thus is able to utilize individual level behavioral data, which directly fits practice needs. [4] designs a holistic dynamic pricing algorithm for an online truckload transportation platform which accepts truckload orders from customers and sells the orders to truck carriers. Essentially the work is a multi-unit asset selling problem and the authors propose a deterministic pricing policy derived from fluid approximation to the original stochastic optimization problem. There are a couple of salient differences between their problem and our business setting. Like many multi-product pricing papers, one critical assumption in [4] is that at most one incoming order and one carrier can arrive the marketplace in each decision epoch, however, in our business setting, real-time dynamic pricing is neither technically tractable nor managerially ideal in the sense that platforms tend to avoid confusing customers and developing strategic customers via high-frequency price changes. Furthermore, demand learning and regret analysis are not in the scope of [4]. [5] is a recent related work where the author investigates

how long a single unit of product can be sold under various structural assumptions. Our work contributes to this stream of literature by formulating our asset selling problem closely adaptive to the real business environment, and providing rigorous performance analysis and insights for the proposed policies via sharp characterization of the problem structure and extensive numerical experiments.

The second stream of related literature is dynamic pricing with inventory constraints for single and multiple products. Seminal work [6] and [7] propose deterministic approximations to the stochastic dynamic pricing problem and show the asymptotic optimality of the simple policy. [8] formulates a unified framework for both dynamic pricing and capacity allocation problems. Multi-product dynamic pricing has also been studied under more specific demand models including attrition models, paired combinatorial models, diffusion models, etc. Representative work include [9–11], etc, where some structural properties like convexity are proven for the multi-product price optimization. Rigorous performance guarantees for resolving heuristics are given by [12–14]. Especially, [12] provides the first logarithm regret for the resolving deterministic heuristic, and recent paper [14] further improve the result to be a constant. [13] shows that infrequent price update on a well-chosen subset of products is sufficient to attain decent performance. However, again this literature also generally assumes at most one customer arrives in each review period. To keep the review concise, we restrain ourselves from the vast body of literature in network revenue management.

Finally the third related stream of literature is the extensive work on dynamic pricing with demand learning. [15–19] study pricing problems under Bayesian learning settings. [20] utilizes an extended MLE method for parametric learning and pricing problems. Work on pricing with parametric learning also include [21–24], etc. [25–28] consider pricing problem with non-parametric demand learning. [29] identifies the conditions for incomplete learning to take place where the learning and optimization process would converge to a suboptimal state. [30–32] study the effect of demand model mis-specification and find that some simple

algorithms could have strong performance guarantees. [33] studies a learning and optimization pricing problem with constraints on the number of price changes. The authors verify the performance of a proposed algorithm via experiments. Our paper falls specifically into the policy category of Thompson sampling. Early work includes [34] and [35] where authors propose new metrics to measure the performance of Thompson sampling due to the difficulties in conducting typical regret analysis in a Bayesian setting. [36–39] and a series of following papers derive traditional regret analysis for Thompson sampling under some specific problem settings with well selected prior distributions. [40] provides a well-cited empirical evidence showing the superior performance of Thompson sampling over other widely used multi-armed bandit polices. Work in this literature mostly deals with problems that do not have inventory or capacity constraints. [41–43] are among the few papers consider dynamic pricing with demand learning under inventory constrained settings. Specifically, [42] applies Thompson sampling to a multi-product dynamic pricing setting and shows the asymptotic optimality under the Bayesian regret. [41] proposes a MLE-based policy for single product selling with finite inventory and compact action space. [43] improves [42] under a general demand model via a MLE-based learning and optimization strategy. Notice that the previous work on policy asymptotic performance requires scaling up in time horizon or total inventory, the performance implications on each unit of product is not the focus. In our paper, we focus on the asymptotic performance for a single unit of product (asset) instead, which requires very different insights towards the functional structures. We apply the ideas of both Thompson sampling and MAP (a natural extension of MLE with prior knowledge) to the asset selling setting where the practical reality facilitates large-sample deviation analysis, and we are able to conduct regret analysis on our algorithm. We also derive some interesting insights via various numerical comparisons.

1.3 Model

We consider a used-car platform which maximizes its expected revenue. In this paper, we focus on the single-car dynamic policies that intend to sell the car within a discrete time horizon with length T. The platform charges price p_t for each period and the decision epochs are indexed by $t \in \{1, \cdots, T-1\}$. At the end of the time horizon, the car will be salvaged to other used-car dealers at a salvage price which is denoted by p_T. The decision maker's problem is to find policy π that maximizes the expected present value:

$$\max_\pi \left\{ \mathbb{E}\left[X_1 \cdot p_1^\pi + (1 - X_1) \cdot X_2 \cdot p_2^\pi + \cdots \right. \right.$$
$$\left. \left. + \prod_{t=1}^{T-2} (1 - X_t) \cdot X_{T-1} \cdot p_{T-1}^\pi + \prod_{t=1}^{T-1} (1 - X_t) \cdot p_T \right] \right\}, \quad (1.1)$$

where $X_t \in \{0, 1\}$, $t = 1, \cdots, T-1$ are the random variables that denote whether the asset is sold in period t.

In each period potential customers come and view the car following a Poisson process with rate λ_t. We allow customer arrival rate to change over time, but within one decision epoch, we assume the rate is well approximated by a steady Poisson process. But we do allow λ_t to follow an exogenous distribution $\Lambda_t(\cdot)$ which could treated as the decision maker's future forecast. $\{\lambda_t\}_{t=1,\cdots,T_1}$ are independently distributed and each λ_t is realized as the decision maker approaches the beginning of each period t. As [44] points out, a lookahead model with future forecast typically fixes the forecast and treat it as the true future. In our framework, we do not model how the forecast itself would evolve and update over time either. Our analysis in the paper has the potential to be applied to more general scenarios, e.g. when the forecast evolution is well characterized by more sophisticated time series model (independent $\{\lambda_t\}_{t=1}^{T-1}$ is then a special case in the more general setting). An underlying assumption of time series model is that current knowledge on the time series model is treated as the ground truth

and fixed in regret analysis. Most of our results have their analogues in the more general time series framework and we will briefly mention them later. Given price p_t, each arriving customer would choose to purchase the car with probability $q_t(p_t)$ if available. Therefore the customers who choose to buy the car arrive following a Poisson process with rate $\lambda_t q_t(p_t)$. In turn, the probability that the car is sold in period t is equal to the probability that the first customer who purchases the car arrives within period t. We know that the interarrival time of Poisson process is exponentially distributed (and in our case, it follows a $Exp(\lambda_t q_t(p_t))$ distribution), so the probability that the car is sold in period t is $1 - e^{-\lambda_t q_t(p_t)}$. Let $V_t(\lambda_t)$ denote the platform's optimal expected revenue at time t given that the car has not been sold and the potential customers' arrival rate is λ_t, then we can write down the Bellman equation of the dynamic pricing problem:

$$V_t(\lambda_t) = \max_{p_t} \left\{ \left(1 - e^{-\lambda_t q_t(p_t)}\right) p_t + e^{-\lambda_t q_t(p_t)} \mathbb{E}_{\lambda_{t+1}} \left[V_{t+1}(\lambda_{t+1}) \right] \right\}, \tag{1.2}$$

with terminal condition $V_T(\lambda_T) = p_T$. We point out that here problem (1.2) is considered as the base case where there is no constraint on the price decision and the platform knows the distribution of λ_t, $t \in \{1, \cdots, T\}$ and can see the realization of λ_t for each period t before deciding the price.

We consider individual purchasing probabilities that satisfy the following assumptions:

Assumption 1.3.1 *Given* λ, $\lim_{p \to \infty} \left(1 - e^{-\lambda q(p)}\right) p = 0$.

Assumption 1.3.2 $q(\cdot) : \mathbb{R} \to [0,1]$, $q'(p) < 0$. $q(p)$ *is log-concave, i.e.,* $-q'(p)/q(p)$ *is increasing in* p.

We define the function space that contains all $q(\cdot)$ which satisfies the above assumptions as $\mathcal{Q}$. We point out that $\mathcal{Q}$ includes many commonly used demand probability functions such as linear demand $q(p) = \beta - \alpha p$, $\alpha, \beta > 0$, logistic demand $q(p) = \frac{e^{\beta - \alpha p}}{1 + e^{\beta - \alpha p}}$, $\alpha, \beta > 0$ and

exponential demand $q(p) = \beta e^{-\alpha p}$, $\alpha, \beta > 0$. Let $\pi_t(p_t; \lambda_t) = \left(1 - e^{-\lambda_t q_t(p_t)}\right) p_t + e^{-\lambda_t q_t(p_t)} \cdot$
$\mathbb{E}_{\lambda_{t+1}}\left[V_{t+1}(\lambda_{t+1})\right]$, then we have the following result:

Lemma 1.3.3 *If $q(\cdot) \in \mathcal{Q}$, then for fixed λ_t, $\pi_t(p_t; \lambda_t)$ is unimodal in p_t.*

The unimodularity of $\pi_t(p_t; \lambda_t)$[4] directly leads to the sufficiency of first-order condition to the optimal policy calculation:

Proposition 1.3.1 *Given $q(\cdot) \in \mathcal{Q}$, the optimal pricing policy $\{p_t^*\}_{t=0}^{T-1}$ is given by $\frac{\partial \pi_t(p_t^*; \lambda_t)}{\partial p_t} = 0$, for $t \in \{1, \cdots, T-1\}$.*

However, we notice that in Bellman's equation (1.2) the expectation is taken over future demand arrival rate, therefore computation is intractable for continuous forecast on future demand rate. To solve this issue, we propose a simple heuristic in the next section.

1.4 Deterministic Approximation and Heuristic

In this section, we introduce a deterministic approximation to the original problem. However, we should point out that the deterministic approximation in the typical inventory selling literature refers to the approximation on the inventory dynamics, while in our setting, we simply mean a deterministic proxy to the demand rate forecast.

1.4.1 Deterministic Approximation as Upper Bound

Suppose that for period $t \in \{1, 2, \cdots, T-1\}$, the incoming potential demand rate λ_t follows a distribution $F_t(\cdot)$. We propose the following pricing policy $\widehat{\pi}$: in any period t, for

---

[4]It is worth pointing out that the unimodularity of $\pi_t(p_t; \lambda_t)$ does not require any assumptions on the distributions of $\{\lambda_t\}_{t=1}^{T-1}$.

realized potential demand rate λ_t, price p_t is determined by solving the following dynamic programming problem:

$$\widetilde{V}_t^{DA}(\lambda_t) = \max_{p_t} \left\{ \left(1 - e^{-\lambda_t q_t(p_t)}\right) p_t + e^{-\lambda_t q_t(p_t)} \cdot \widetilde{V}_{t+1}^{DA}\left(\mathbb{E}\left[\lambda_{t+1}\right]\right) \right\}, \ \widetilde{V}_T^{DA} = p_T. \qquad \text{(DA)}$$

It is trivial to see that as the demand process degenerates into a deterministic process, the above algorithm gives the optimal price trajectory. The main result that we will show in this section is that the value function $\widetilde{V}_t^{DA}(\lambda_t)$ of simple algorithm described above actually gives an upper bound of the optimal expected value $V_t(\lambda_t)$ which is defined in (1.2).[5]

Theorem 1.4.1 *Given* $q(\cdot) \in \mathcal{Q}$, *then* $V_t(\lambda_t) \leq \widetilde{V}_t^{DA}(\lambda_t)$, $t = \{0, 1, \cdots, T-1\}$.

Proof Proof. We prove the theorem by induction. For $t = T - 1$, from definition, we have $V_{T-1}(\lambda_{T-1}) = \widetilde{V}_{T-1}^{DA}(\lambda_{T-1})$ which is because period T's salvage price p_T is exogenously given. Assume that the theorem holds for $t + 1 \in \{1, \cdots, T-1\}$, i.e., $V_{t+1}(\lambda_{t+1}) \leq \widetilde{V}_{t+1}^{DA}(\lambda_{t+1})$, we will show that $V_t(\lambda_t) \leq \widetilde{V}_t^{DA}(\lambda_t)$, which will complete the proof. From Lemma A.0.1, we know that $V_{t+1}(\lambda_{t+1})$ is increasing concave in λ_{t+1}, $t = 0, 1, \cdots, T-1$, therefore, we have

$$\mathbb{E}\left[V_{t+1}(\lambda_{t+1})\right] \leq V_{t+1}\left(\mathbb{E}[\lambda_{t+1}]\right) \leq \widetilde{V}_{t+1}^{DA}\left(\mathbb{E}[\lambda_{t+1}]\right),$$

where the first inequality is Jensen inequality and the second inequality comes from the induction assumption. Therefore, it is straightforward to see that

$$V_t(\lambda_t) = \max_{p_t} \left\{ \left(1 - e^{-\lambda_t q_t(p_t)}\right) p_t + e^{-\lambda_t q_t(p_t)} \cdot \mathbb{E}\left[V_{t+1}(\lambda_{t+1})\right] \right\}$$

$$\leq \max_{p_t} \left\{ \left(1 - e^{-\lambda_t q_t(p_t)}\right) p_t + e^{-\lambda_t q_t(p_t)} \cdot \widetilde{V}_{t+1}^{DA}\left(\mathbb{E}[\lambda_{t+1}]\right) \right\} = \widetilde{V}_t^{DA}(\lambda_t).$$

---

[5]If the demand rate forecast is modeled to evolve over time following a more sophisticated time series model mentioned in Section 1.3, $\widetilde{V}_t^{DA}(\lambda_t)$ is not necessarily concave in λ_t, which means that deterministic approximation may not serve as an upper bound of problem (1.2), yet we are still able to derive similar regret results as in Section 1.4.2 with some appropriate assumptions on the time series.

The proof is completed. ∎

1.4.2 DA-Induced Heuristic and Regret

The deterministic approximation (DA) induces a pricing policy and in this section, we would analyze the performance of this heuristic. Especially we are interested in the asymptotic behavior of the regret as the problem scales up. But before the analysis, we need to first specify the asymptotic regime.

Asymptotic Regime.

For asset selling problem, there are two dimensions in which one problem could scale up, i.e., potential demand rate and length of decision horizon. In our case, scaling up in potential demand rate (or online flow volume) is more relevant to the nature of online asset selling business, while the length of decision horizon for each unit of asset is typically set to be limited because a high turnover rate is favorable to the platforms.

Remark. In our setting, since the price action space is not bounded, scaling up in demand rate does not necessarily imply that the asset will be sold almost surely in the first period. See [5] for rigorous discussions. Our problem could be considered as the mixture case between dynamic pricing and static pricing in [5]. Numerical results in Section 1.6 show that even when the problem scale is considerably large, the expected sale duration still reaches far beyond the first epoch.

Next we formalize how the problem is scaled up in the dimension of potential demand rate by making the following definition:

Definition 1.4.1 (Scale Up in Degree γ) *We say a sequence of random variable $\{\lambda_n\}$ scales up in degree γ if $\mathbb{E}[\lambda_n] \sim n$ and $Sd(\lambda_n) \sim n^\gamma$, where $Sd(\lambda_n)$ is the standard deviation of λ_n.*

Typically we would expect to see that $0 \leq \gamma \leq 1$. We give examples for two probable scale-up degrees: $1/2$ and 1. One scenario is that the platform divides each decision epoch t (normalized to have length of 1) into n equal-length monitored time slots and the realized number of arrivals within each time slot is considered as one demand data point. The Poisson rates λ_{tni} for each time slot i with length $1/n$, $i = 1, \cdots, n$, are independent and identically distributed, then it is straightforward to find that an equivalent Poisson rate for the whole decision epoch is $\lambda_{tn} = \sum_{i=1}^{n} \lambda_{tni}/n$. When the number of slots, n, increases as the potential demand increases such that the distribution of each λ_{tni}/n keeps the same, λ_{tn} will scale up in degree $1/2$. An alternative scenario of scaling up in degree $1/2$ is that the demand flow consists of n independent and identical Poisson process with rate λ_{tni}, $i = 1, \cdots, n$. Fixing the rate of each sub-flow, λ_{tni}, the potential demand increases as the number of sub-flows increases, then λ_{tn} will also scale up in degree $1/2$. On the other hand, if we substitute the independence in the above two scenarios with perfect correlation, we would see λ_{tn} scales up in degree 1. In the context of forecast, scaling up in a degree less than 1 is saying that the platform's forecast of future demand becomes relatively more accurate as the market size grows.

Closely related to the Definition 1.4.1, we define

Definition 1.4.2 ($\mathcal{G}_\gamma$-Sequence) *Let $\{G_n(\cdot)\}$ denote the corresponding sequence of cumulative probability functions of random variable sequence $\{\lambda_n\}$ which scales up in degree γ. We say $\{G_n(\cdot)\}$ is a $\mathcal{G}_\gamma$-sequence or simply denoted as $\{G_n(\cdot)\} \in \mathcal{G}_\gamma$ if there exists a constant $\rho > 0$, such that*

$$G_n(\rho \cdot n) \sim o\left(\frac{n^{\gamma-1}}{\log(n)}\right).$$

As an example, if sequence $\{\lambda_n\}$ scales up in degree $1/2$, $G_n(\cdot)$ is sub-Gaussian and ρ is any constant such that $\rho < \mathbb{E}[\lambda_n]/n$, then $G_n(\rho \cdot n)$ decreases to zero way faster than $n^{-1/2}/\log(n)$ as n grows large, which implies that $\{G_n(\cdot)\} \in \mathcal{G}_{\frac{1}{2}}$.

To conduct the regret analysis, it is critical to understand how the value function would react to parameter changes. It is useful to define the following auxiliary function:

$$F(\lambda, A) = \max_p \{\pi(p; \lambda, A)\} = \max_p \left\{ \left(1 - e^{-\lambda q(p)}\right) p + e^{-\lambda q(p)} \cdot A \right\},$$

where $\lambda, A > 0$ and $q(\cdot) \in \mathcal{Q}$. The following lemma summarizes the sensitivity results of our auxiliary function $F(\cdot, \cdot)$, which provides important hints on the sensitivity of value function $V_t(\cdot)$.

Lemma 1.4.2 *For any $A > 0$, if $\bar{c} = \lim_{\lambda \to \infty}(-q'(p(\lambda))/q(p(\lambda)))$ exists and is positive (including infinite), then for $\lambda > \Lambda(c)$, $\frac{\partial F(\lambda, A)}{\partial \lambda}$ is upper bounded by $1/c\lambda$ where c is any number within $(0, \bar{c})$.*

Proof Proof. According to Lemma A.0.1, we know that for any $A \geq 0$:

$$0 \leq \frac{\partial F(\lambda, A)}{\partial \lambda} \leq \frac{\partial F(\lambda, 0)}{\partial \lambda} = \frac{p(\lambda) q(p(\lambda))}{1 - \lambda p(\lambda) q'(p(\lambda))}.$$

On the other hand,

$$\frac{p(\lambda) q(p(\lambda))}{1 - \lambda p(\lambda) q'(p(\lambda))} = \frac{1}{\frac{1}{p(\lambda)q(p(\lambda))} + \left(-\frac{q'(p(\lambda))}{q(p(\lambda))}\right) \cdot \lambda} \leq \frac{1}{\left(-\frac{q'(p(\lambda))}{q(p(\lambda))}\right) \cdot \lambda},$$

We know that as λ goes to infinite, $p(\lambda)$ also goes to infinite, then if

$$\bar{c} = \lim_{p \to \infty}(-q'(p(\lambda))/q(p(\lambda)))$$

16

exists and is positive, for $\lambda > \Lambda(c)$ where $\Lambda(c)$ is some positive constant depending on c, we have

$$\frac{\partial F(\lambda, A)}{\partial \lambda} \leq \frac{p(\lambda) q(p(\lambda))}{1 - \lambda p(\lambda) q'(p(\lambda))} \leq \frac{1}{c\lambda} , \ \forall \ c \in (0, \bar{c}) . \tag{1.3}$$

The proof is completed. ∎

We notice that in Lemma 1.4.2, we require $\lim_{\lambda \to \infty} (-q'(p(\lambda)))/q(p(\lambda))$ exists and positive. This *de facto* comes from the log-concavity of $q(\cdot)$ directly, because log-concavity is equivalent to that $-q'(p)/q(p)$ is increasing in p. Therefore by requiring $\lim_{\lambda \to \infty} (-q'(p(\lambda)))/q(p(\lambda)) > 0$ we do not put any further assumptions on $q(\cdot)$ other than Assumption 1.3.2. Besides the sensitivity properties of value function, it is also critical to understand the sensitivity properties of the optimal decision. Let $p(\lambda, A) = \text{argmax}_p \pi(p; \lambda, A)$, we have the following lemma:

Lemma 1.4.3 (i). $0 < \frac{\partial p(\lambda, A)}{\partial \lambda}$ and for any constant $c \in (0, \lim_{p \to \infty} \{-q'(p)/q(p)\}]$, there exists $\lambda(c)$ such that when $\lambda > \lambda(c)$, $\frac{\partial p(\lambda, A)}{\partial \lambda} < \frac{1}{c\lambda}$. (ii). $0 < \frac{\partial p(\lambda, A)}{\partial A} < 1$, for any $\lambda > 0$ and $A > 0$.

Lemma 1.4.3 will play an important role later when we conduct regret analysis. It tells us how $p(\lambda, A)$ would increase as the two parameters λ and A grow large. Especially, we know that $p(0, 0) = 0$, therefore Lemma 1.4.3 implies the following bound for $p(\lambda, A)$:

$$p(\lambda, A) < C \cdot \log(\lambda) + A, \tag{1.4}$$

where C is some positive constant independent of A.

Given the above Lemmas, we are ready to analyze the performance of the deterministic approximation. With a slight abuse of notations, we use $V_t^\pi(\lambda_{tn})$ to denote the expected present value of pricing policy $\widehat{\pi}$ given the current realized demand rate λ_{tn} and here the expectation is taken on all the future demand rate $\lambda_{\tau n}$, $\tau = t + 1, \cdots, T - 1$, based upon

current forecast. To measure the performance of our algorithm, we define the following metric:

Definition 1.4.3 *The regret of a policy π is defined as:*

$$Regret\left(\pi\right) = \mathbb{E}_{\lambda_{1n}}\left[V_1\left(\lambda_{1n}\right) - V_1^\pi\left(\lambda_{1n}\right)\right],$$

where $V_1^\pi\left(\lambda_{1n}\right)$ is the expected profit when applying pricing policy π given current demand rate λ_{1n}.

Our main results below show the asymptotic optimality in scale factor n could be achieved by DA policy:

Theorem 1.4.4 *Assume that each sequence $\{\lambda_{tn}\}$, $t = 1, \cdots, T-1$ scales up in degree γ and each corresponding distribution sequence $\{G_{tn}\left(\cdot\right)\} \in \mathcal{G}_\gamma$ and $\lim_{p\to\infty}\left\{-q'\left(p\right)/q\left(p\right)\right\} < \infty$, then*

$$0 < Regret\left(\pi_{DA}\right) < \mathcal{O}\left(\frac{\log\left(n\right)}{n^{1-\gamma}}\right).$$

We point out that in the above theorem, the condition $\lim_{p\to\infty}\left\{-q'\left(p\right)/q\left(p\right)\right\} < \infty$ is satisfied many commonly used functions in family $\mathcal{Q}$. For example, when $q\left(p\right) = \frac{e^{\beta-\alpha p}}{1+e^{\beta-\alpha p}}$, $\lim_{p\to\infty}\left\{-q'\left(p\right)/q\left(p\right)\right\} = \alpha$, when $q\left(p\right) = \beta e^{-\alpha p}$, $\lim_{p\to\infty}\left\{-q'\left(p\right)/q\left(p\right)\right\} = \alpha$. Besides, it is straightforward to check that for a series of $q_i\left(p\right)$, $i \in \mathcal{I}$, that satisfies condition $\lim_{p\to\infty}\left\{-q_i'\left(p\right)/q_i\left(p\right)\right\} < \infty$, then the multiplication $q\left(p\right) = \prod_{i\in\mathcal{I}}q_i\left(p\right)$ also satisfies $\lim_{p\to\infty}\left\{-q'\left(p\right)/q\left(p\right)\right\} < \infty$, which would be useful in the next section.

18
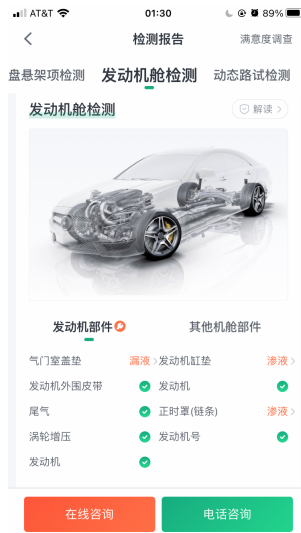
Figure 1.1.: Example of Car Profile on Guazi Used Car Platform.

1.5 Demand Learning via Online Behavioral Data

In this section, we dive deeper into customers' preferences over the asset, and discuss how the platform can better inform its pricing decisions via customers' online behavior data. We assume that the net value of car j which has been posted for t periods has form

$$u_{jt} = X_j\beta - \alpha p_{jt} + \xi_j + \epsilon_{jt}, \tag{1.5}$$

where X_{jt} is a vector of car j's observable attributes and it could be time dependent, ξ_j is car j's latent value which is unobservable from the platform's perspective, while is observable by customers and ϵ_{jt} is each customer's idiosyncratic preference shock for car j in period t with a known distribution. Figure 1.1 shows the screenshots of a used-car profile on a platform's app. The platform typically conducts a thorough inspection covering over one hundred mechanical nuances. However, when formulating an empirical model to estimate customers' reactions to a car profile, generally only a subset of major factors would be included. One reason is that the data may be too sparse to support a model that captures the full profile of the attributes. Especially, customers' preference may evolve over time, which makes the data collected long time ago less relevant. Therefore, including an unobservable latent value ξ_j besides the major factors into the empirical model and learning the latent value over the selling process is a reasonable setup. As [1] shows, under such a setting, learning the latent factor admits significant values. We assume that the effects of observable attributes β is universal for all cars which could be learned from historical data, while ξ_j is associated to each specific car j which is what the platform should learn along the dynamic pricing. Furthermore, we assume that ξ_j is supported on a compact set, i.e., $\xi_j \in [\underline{\xi}, \bar{\xi}]$. Because we are focusing on the single car scenario, without introducing confusion, we drop the j subscript hereafter.

The nature of online business enables the platform to observe each single potential customer's online behaviors including not only the final purchase decision, but also intermediate behaviors. In the used car online marketplace, typical intermediate behaviors include click, online save and appointment for test drive etc. For exposition purpose, we focus on the framework where the platform can well observe one specific type of customers' online behaviors before customers make up their purchase decisions. And we assume that whether a customer conducts such a behavior (e.g., test drive appointment) depends on the net value u_t and each customer's reservation value which is sampled from a known distribution. Then we call a customer who conducts the specific type of intermediate behavior as a *promising customer* and for each visitor, she becomes a *promising customer* with probability $q_t^1(u_t)$. Recall that the effects of observable attributes β could be estimated from historical data, in our setting, we treat them as known constants, so we can write $q_t^1(u_t)$ as $q_t^1(\xi - \alpha p_t)$, where as a generalized linear function of p_t, we assume $q_t^1(\xi - \alpha p_t) \in \mathcal{Q}$. After conducting the intermediate behavior, each *promising customer* decides to purchase the asset with probability $q_t^2(\xi - \alpha p_t)$ where $q_t^2(\cdot) \in \mathcal{Q}$. Let $q_t(\cdot)$ be the *conversion rate*. We assume that $q_t(\cdot)$ has the decomposed structure: $q_t(\xi - \alpha p_t) = q_t^1(\xi - \alpha p_t) \cdot q_t^2(\xi - \alpha p_t)$. Notice that here we made a simplification of the demand process. We assume the time delay between a customer's intermediate behavior and the final purchase decision is negligible, also we do not model any strategic behaviors. In used-car selling business, a typical decision epoch lasts for a couple of weeks, so for a popular car, we can expect most customers won't delay their final purchase decision across epochs. Especially, the recent movements of some leading used-car companies (e.g., Carvana) to advocate the online-only car selling mode make a customer's whole purchasing process much shorter than a traditional offline selling mode.

The following lemma indicates that multiplication is closed in $\mathcal{Q}$:

Lemma 1.5.1 *If $q_i(p) \in \mathcal{Q}$, $i = 1, \cdots, n$, and $q(p) = \Pi_{i=1}^n q_i(p)$, then $q(\cdot) \in \mathcal{Q}$.*

Whence we know that $q_t (\xi - \alpha p_t) \in \mathcal{Q}$. One direct extension to the above setting is that the platform is able to well observe S steps of intermediate behaviors, where $S > 1$. For example, a customer should first click the product, then add it to wish list and finally schedule a test drive. We assume that the order of these intermediate steps cannot be alternated and a latter step cannot be conducted if one earlier step is skipped. Let $q_t^s (\xi - p_t) \in \mathcal{Q}$ denote each customer's probability of conducting step s given the previous steps are conducted, then a customer's purchase probability can be written as $q_t (\xi - p_t) = \Pi_{s=0}^{S} q_t^s (\xi - p_t)$. Lemma 1.5.1 implies that $q_t (\xi - p_t) \in \mathcal{Q}$, therefore all the analysis can be easily extended to this more general setting. To ease the exposition, we use the deterministic approximation (DA) as the true optimal and conduct regret analysis. Once we do that, we will see later that it is straightforward to get the regret when the true optimal is the original problem (1.2).

1.5.1 Data Description for Parameter Learning

Before we propose the policy for learning and pricing, we need to first specify the observed data used for parameter learning. Here we denote the true latent value to learn by ξ_0 with prior distribution density $h_{\xi_0} (\cdot)$. As we indicated before, the firm can observe each customer's behaviors including both the intermediate decision and the final purchase decision. In our setting, the data that the firm would mostly utilize is the intermediate behavior observations which indicate whether visitors become *promising customers* and each observation is a Bernoulli random variable. To facilitate our analysis below, we assume the platform collects data in the following way: (i). The firm divides a single decision epoch into small monitored time slots with equal lengths such that mean number of arriving customers in each time slot equals to a predetermined constant λ_t; (ii). After one decision epoch, the platform collects the realized numbers of *promising customers* for all the time slots. Each realized observation from one slot is treated as one data point. We notice that in our frame-

work, we characterize the customer arrival process within each decision epoch as a stationary Poisson process with rate λ_{tn}, therefore evenly divided time slots have identical mean value of arriving customers. More specifically, given price p_{jt}, the number of *promising customers* of all the time slots within one decision epoch are independent and identically sampled from distribution $Pois\left(\lambda_t q_t \left(\xi_0 - \alpha p_t\right)\right)$ and the sample size is $n_t \coloneqq \lambda_{tn}/\lambda_t$. Collected data is used to estimate the unobservable value ξ_0.

1.5.2 Learning and Optimization Policies: Thompson Sampling & MAP

The firm's ability to observe *potential customers* in online marketplace is critical in our asset selling setting, because the final purchase observations would be all nos before the asset is finally sold, which cannot support effective demand learning. In this section, we propose two learning and optimization policies: an active learning policy based on Thompson sampling, and a passive learning policy based on maximum likelihood estimation. By calling an algorithm active/passive, we simply aim to distinguish algorithms with/without explorations during learning. Our active learning policy works as follows:

Algorithm. (Thompson Sampling for Dynamic Pricing)

Iterate the following steps for $t = 1, \cdots, T$:

1. *Sample Unobservable Attribute*: Sample a random parameter $\xi_t \in \Xi$ according to the posterior distribution of ξ_0 given history $\mathcal{H}_{t-1}$.

2. *Offer Price*: Solve the dynamic pricing problem and get the optimal pricing policy $\{p_\tau^*\}_{\tau=t}^{T-1}$:

$$V_t\left(\lambda_{tn}\right) = \max_p \left\{ \left(1 - e^{-\lambda_{tn}q_t(\xi_t - \alpha p_t)}\right) p_t + e^{-\lambda_{tn}q_t(\xi_t - \alpha p_t)} V_{t+1}\left(\mathbb{E}\left[\lambda_{(t+1)n}\right]\right) \right\},$$

with $V_T\left(\cdot\right) = p_T$. Then set $p_t^{TS} = p_t^*$.

3. *Update Estimation*: Observe the realized number of *potential buyers* for each time slot $X_t = \{X_{t1}, \cdots, X_{tk_t}\}$ with each observation sampled from $Pois\left(\lambda_t q_t \left(\xi_0 - \alpha p_t^{TS}\right)\right)$ and update the history $\mathcal{H}_t = \mathcal{H}_{t-1} \cup \{X_t\}$ and the posterior belief of ξ_0 based upon $\mathcal{H}_t$ via Bayes rule.

On the other hand, the passive learning policy works as follows:

Algorithm. (Maximum A Posteriori Estimation for Dynamic Pricing)

For $t = 1$:

1. *Sample Unobservable Attribute*: Sample a random parameter $\xi_1 \in \Xi$ according to the prior distribution of ξ_0.

Iterate the following steps for $t = 1, \cdots, T$:

2. *Offer Price*: Solve the dynamic pricing problem and get the optimal pricing policy $\{p_\tau^*\}_{\tau=t}^{T-1}$:

$$V_t\left(\lambda_{tn}\right) = \max_p \left\{ \left(1 - e^{-\lambda_{tn} q_t(\xi_t - \alpha p_t)}\right) p_t + e^{-\lambda_{tn} q_t(\xi_t - \alpha p_t)} V_{t+1}\left(\mathbb{E}\left[\lambda_{(t+1)n}\right]\right)\right\},$$

with $V_T\left(\cdot\right) = p_T$. Then set $p_t^{MAP} = p_t^*$.

3. *Update Estimation*: Observe the realized number of *potential buyers* for each time slot $X_t = \{X_{t1}, \cdots, X_{tk_t}\}$ with each observation sampled from $Pois\left(\lambda_t q_t \left(\xi_0 - \alpha p_t^{MAP}\right)\right)$ and update the history $\mathcal{H}_t = \mathcal{H}_{t-1} \cup \{X_t\}$ and MAP estimation of ξ_0 based upon $\mathcal{H}_t$, and let ξ_{t+1} denote the MAP estimation. $t = t + 1$.

Remark. Here we do not need to worry about the consistency of the learning, because: (i). There is no uninformative state due to the fact that $dq\left(\xi - \alpha p\right)/d\xi > 0$ (see [29]); (ii). Our policies are different from typical certainty-equivalence controls where in our setting multiple observations are available within each decision epoch.

1.5.3 Performance Analysis

A salient feature of our online pricing framework is that within each decision epoch, there could be large amount of observations, while in a typical model of sequential decision making with learning, decisions are made instantly after each new observation. Large observation sample in each decision epoch opens the gateway for us to use asymptotic behaviors of some statistics to get finer regret bounds which have not been seen in conventional settings. We start from analyze the Thompson-sampling-based policy.

Within a given decision epoch t, $t = 1, \cdots, T-1$, with price p_t, the platform observes n_t number of independent and identically distributed random draws from distribution $Pois\left(\lambda_t q\left(\xi - \alpha p_t\right)\right)$. We know that under some regularity conditions, as the i.i.d. sample size grows, the posterior distribution of ξ has asymptotic normality due to Bernstein-von-Mises Theorem. Across different decision epochs, the random draws are still independent but not identically distributed, because the prices change over time. Without loss of generality, we assume $n_1 = n_2 = \cdots = n_{T-1} = n$, then by slightly modifying the proof in [45], we still can show the asymptotic normality of the posterior distribution.

Following the notations in [45], let $(X, \mathscr{A})$ be a measurable space and $L_\xi | \mathscr{A}$, $\xi \in \Xi$, a family of probability measures, where Ξ is an open subset of $\mathbb{R}$. Let ξ be a random variable with prior distribution density $h_{\xi_0} | \mathscr{B} \cap \Xi$. Assume that λ_ξ has a finite density ρ with respect to the Lebesgue measure, which is positive on Ξ and zero on Ξ^c. Let $R_{tn,\boldsymbol{x}}$ be the posterior distribution of ξ for the sample size $t \cdot n$ given $\boldsymbol{x} \in X^{tn}$ after t decision epochs, which is defined as

$$R_{tn,\boldsymbol{x}}\left(B\right) = \frac{\int_B \left(\Pi_{\tau=1}^t \Pi_{i=1}^n l_\tau\left(x_{\tau i}, \sigma\right)\right) \rho\left(\sigma\right) d\sigma}{\int \left(\Pi_{\tau=1}^t \Pi_{i=1}^n l_\tau\left(x_{\tau i}, \sigma\right)\right) \rho\left(\sigma\right) d\sigma}, \quad B \in \mathscr{B},$$

where in our setting $l\left(\cdot,\xi\right) = \left(\lambda_\tau q_\tau\left(\xi - \alpha p_\tau\right)\right)^x e^{-\lambda_\tau q_\tau\left(\xi - \alpha p_\tau\right)}/x!$. Let $Q_{tn,\boldsymbol{x}}$ be the normal distribution centered at the maximum likelihood estimator $\xi_{tn}\left(\boldsymbol{x}\right)$ with covariance matrix $\Gamma_{tn}\left(\boldsymbol{x}\right)^{-1}$, where

$$\Gamma_{tn}\left(\boldsymbol{x}\right) = \left(\sum_{\tau=1}^{t}\sum_{i=1}^{n}\frac{\partial^2}{\partial\xi^2}f\left(x_{ti},\xi\right)\Big|_{\xi=\xi_n\left(\boldsymbol{x}\right)}\right),$$

which is positive and $f\left(x_{\tau i},\xi\right) = -\log l\left(x_{\tau i},\xi\right)$.

Define the variational distance between the measures $R_{tn,\boldsymbol{x}}$ and $Q_{tn,\boldsymbol{x}}$ as

$$d\left(R_{tn,\boldsymbol{x}},Q_{tn,\boldsymbol{x}}\right) = \sup\left\{\left|R_{tn,\boldsymbol{x}}\left(B\right) - Q_{tn,\boldsymbol{x}}\left(B\right)\right| : B \in \mathscr{B}\right\}.$$

Then we can prove the following proposition:

Proposition 1.5.1 *For any $s \geq 2$ and every compact subset K of ξ there exists a constant $c_K\left(s\right) > 0$ such that*

$$\sup_{\xi\in K} P_\xi^{tn}\left\{\boldsymbol{x} \in X^{tn} : d\left(R_{tn,\boldsymbol{x}},Q_{tn,\boldsymbol{x}}\right) > c_K\left(s\right)\cdot\left(t\cdot n\right)^{-1/2}\right\} = \mathcal{O}\left(\left(t\cdot n\right)^{-s/2}\right).$$

Proof Proof. First we prove that within each decision epoch, the regularity conditions in Appendix hold for any integer $s \geq 2$. For given epoch t and price p_t,

$$f\left(x,\xi\right) = -\log\left(\frac{\left(\lambda_t q_t\left(\xi - \alpha p_t\right)\right)^x e^{-\lambda_t q_t\left(\xi - \alpha p_t\right)}}{x!}\right)$$

$$= -x\log\left(\lambda_t q_t\left(\xi - \alpha p_t\right)\right) + \log\left(x!\right) + \lambda_t q_t\left(\xi - \alpha p_t\right),$$

and

$$f''\left(x,\xi\right) = \frac{\partial^2 f\left(x,\xi\right)}{\partial^2\xi} = \left(\frac{q_t'\left(\xi - \alpha p_t\right)}{q_t\left(\xi - \alpha p_t\right)}\right)'\cdot x + \lambda_t q_t''\left(\xi - \alpha p_t\right).$$

26

Therefore, we have

$$\left| f\left(x,\xi\right)\right| \leq \left(\left|\log\left(\lambda_t q_t\left(\xi - \alpha p_t\right)\right)\right| + x - 1\right)\cdot x + \left(x + \frac{1}{2}\right)\log\left(x\right) + \widetilde{C} < C\cdot x^2,$$

where C and $\widetilde{C}$ are some positive constants independent of x and the first inequality comes from the upper bound on $x!$: $x! \leq x^{x+\frac{1}{2}}e^{-x+1}$. In turn, for any integer $s \geq 2$, we have $\left| f\left(x,\xi\right)\right|^s < C_s\cdot x^{2s}$, where C_s is some positive constant. On the other hand, we know that given any ξ, x follows a Poisson distribution and any order of factorial moment of Poisson distribution exists, therefore given any $\xi,\sigma \in \Xi$, $\mathbb{E}_\sigma\left| f\left(\cdot,\xi\right)\right|^s < \infty$. We can apply the same argument to $\left| f''\left(x,\xi\right)\right|^s$ which is upper bounded by a linear function of x. Furthermore

$$\mathbb{E}_\xi f''\left(\cdot,\xi\right) = \left(\frac{q_t'\left(\xi - \alpha p_t\right)}{q_t\left(\xi - \alpha p_t\right)}\right)'\cdot \lambda_t q_t\left(\xi - \alpha p_t\right) + \lambda_t q_t''\left(\xi - \alpha p_t\right)$$

$$= \frac{\lambda_t}{q_t\left(\xi - \alpha p_t\right)}\cdot \left(q_t\left(\xi - \alpha p_t\right)\right)^2 > 0.$$

Then it is not difficult to see that all the regularity conditions listed in the appendix hold for any $s \geq 2$. Finally, by slightly modifying the proof in [45], we can extend the results into non-identically distributed observations across different decision epochs. ∎

Proposition 1.5.1 indicates that the probability of convergence error exceeding $c_K\left(t\cdot n\right)^{-1/2}$ diminishes faster than any polynomial order, therefore when $t\cdot n$ is large, $R_{tn,\boldsymbol{x}}\left(B\right)$ can be well bounded by $\left[Q_{tn,\boldsymbol{x}}\left(B\right) - c_K\cdot\left(tn\right)^{-1/2}, Q_{tn,\boldsymbol{x}}\left(B\right) + c_K\cdot\left(tn\right)^{-1/2}\right]$. Now that we are interested in the asymptotic regret, hereafter we will assume $Q_{tn,\boldsymbol{x}}\left(B\right) - c_K\cdot\left(tn\right)^{-1/2} < R_{tn,\boldsymbol{x}}\left(B\right) < Q_{tn,\boldsymbol{x}}\left(B\right) + c_K\cdot\left(tn\right)^{-1/2}$, for any $B \in \mathcal{B}$.

Given samples are independent and identically distributed within each decision epoch and independent distributed across different epochs, we are able to identify the consistency and asymptotic normality of the maximum likelihood estimator $\xi_{tn}\left(\boldsymbol{x}\right)$. Consistency follow

directly from the results in [24] and here we focus on the asymptotic normality and the corresponding convergence rate.

Proposition 1.5.2 *For any $s \geq 2$ and every compact subset K of ξ there exists a constant $C_K(s) > 0$ such that*

$$\sup_{\xi \in K} P_\xi^{tn} \left\{ \boldsymbol{x} \in X^{tn} : d\left(R_{tn,\boldsymbol{x}}, \widetilde{Q}_{tn,\boldsymbol{x}}\right) > C_K(s) \cdot (t \cdot n)^{-1/2} \right\} = \mathcal{O}\left((t \cdot n)^{-s/2}\right).$$

Let $V_t(\lambda_{tn}, \xi)$ be the optimal expected profit at time period t given that the current-period demand rate is λ_{tn} and the latent value is ξ. Using the function $F(\lambda, A, \xi)$ defined in Lemma A.0.9, we can write down the recursive definition for $V_t(\lambda_{tn}, \xi)$:

$$V_t(\lambda_{tn}, \xi) = \max_{p_t} \left\{ \left(1 - e^{-\lambda_{tn} q(\xi - \alpha p_t)}\right) p_t + e^{-\lambda_{tn} q(\xi - \alpha p_t)} \cdot V_{t+1}(\xi) \right\} = F(\lambda_{tn}, V_{t+1}(\xi), \xi). \tag{1.6}$$

Let $V_t^{TS}(\lambda_{tn}, \xi_t)$ be the expected profit at time period t by applying Thompson sampling pricing policy given that the current-period demand rate is λ_{tn} and the drawn sample for period t is ξ_t:

$$V_t^{TS}(\lambda_{tn}, \xi_t) = \left(1 - e^{-\lambda_{tn} q\left(\xi - \alpha p_t^{TS}(\xi_t)\right)}\right) p_t^{TS}(\xi_t)$$
$$+ e^{-\lambda_{tn} q\left(\xi - \alpha p_t^{TS}(\xi_t)\right)} \cdot \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right], \tag{1.7}$$

where we should notice that in the above the expectation on ξ_{t+1} is taken based upon the posterior belief on ξ at the beginning of period $t + 1$. Our main result gives the performance analysis for the Thompson sampling algorithm:

Theorem 1.5.2 *If $\lim_{p\to\infty}\{-q'(p)/q(p)\} < \infty$, the regret of Thompson Sampling pricing policy is given by:*

$$0 < V_1\left(\lambda_{1n}, \xi_0\right) - \mathbb{E}_{\xi_1}\left[V_1^{TS}\left(\lambda_{1n}, \xi_1\right)\right] < \frac{T \cdot \sigma_0}{\alpha} \cdot \log\left(n\right) + \mathcal{O}\left(\frac{\log\left(n\right)}{\sqrt{n}}\right),$$

$$0 < V_2\left(\lambda_{2n}, \xi_0\right) - \mathbb{E}_{\xi_2}\left[V_2^{TS}\left(\lambda_{2n}, \xi_2\right)\right] < \mathcal{O}\left(\frac{\log\left(n\right)}{\sqrt{n}}\right).$$

where $\sigma_0 = \mathbb{E}\left[\left|\xi_1 - \xi_0\right|\right]$.

We notice that in the regret starting from the first period, there is a term $\frac{T \cdot \sigma_0}{\alpha} \cdot \log(n)$ that is not decreasing as the scale factor n grows. This term is purely led by the initial misspecification on ξ_0, which is independent of whatever learning algorithm that is applied, therefore the regret starting from the second period would be more meaningful. The above calculation gives the regret of Thompson Sampling pricing policy in the deterministic approximation problem, and with some slight modifications, it is not difficult to get the regret in the original problem (1.2). To avoid confusion, here we use $V_t^*\left(\lambda_{tn}, \xi_0\right)$ to denote the true optimal profit given the current period demand rate λ_{tn} and latent value ξ_0. Then we have the following corollary:

Corollary 1.5.3 *Assume that each sequence $\{\lambda_{tn}\}$, $t = 1, \cdots, T-1$ scales up in degree γ and each corresponding distribution sequence $\{G_{tn}\left(\cdot\right)\} \in \mathcal{G}_\gamma$, then*

$$0 < \mathbb{E}_{\lambda_1}\left[V_1^*\left(\lambda_1, \xi_0\right) - \mathbb{E}_{\xi_1}\left[V_1^{TS}\left(\lambda_1, \xi_1\right)\right]\right] < \frac{T \cdot \sigma_0}{\alpha} \cdot \log\left(n\right) \cdot \mathbb{1}\left\{t = 1\right\} + \mathcal{O}\left(\frac{\log\left(n\right)}{n^{\min\{1-\gamma,1/2\}}}\right).$$

We notice that *Bernstein-von-Mises Theorem* indicates that the posterior asymptotic normal distribution is centered at MLE, it is not difficult to see that the regret analysis for the Thompson-sampling-based policy includes the analysis for the MAP-based pricing policy. Furthermore, the MAP distribution converges at the same asymptotic rate to the

true parameter as the Bayesian posterior distribution, i.e., at the rate of $1/\sqrt{n}$, therefore results similar to Theorem 1.5.2 and Corollary 1.5.3 hold for MAP.

Theorem 1.5.4 *If* $\lim_{p\to\infty}\{-q'(p)/q(p)\} < \infty$, *the regret of MAP pricing policy is given by:*

$$0 < V_1\left(\lambda_{1n},\xi_0\right) - \mathbb{E}_{\xi_1}\left[V_1^{MAP}\left(\lambda_{1n},\xi_1\right)\right] < \frac{T\cdot\sigma_0}{\alpha}\cdot\log\left(n\right) + \mathcal{O}\left(\frac{\log\left(n\right)}{\sqrt{n}}\right),$$

$$0 < V_2\left(\lambda_{2n},\xi_0\right) - \mathbb{E}_{\xi_2}\left[V_2^{MAP}\left(\lambda_{2n},\xi_2\right)\right] < \mathcal{O}\left(\frac{\log\left(n\right)}{\sqrt{n}}\right).$$

where $\sigma_0 = \mathbb{E}\left[\left|\xi_1 - \xi_0\right|\right]$.

And analogous to Corollary 1.5.3, we have

Corollary 1.5.5 *Assume that each sequence* $\{\lambda_{tn}\}$, $t = 1,\cdots,T-1$ *scales up in degree* γ *and each corresponding distribution sequence* $\{G_{tn}\left(\cdot\right)\} \in \mathcal{G}_\gamma$, *then*

$$0 < \mathbb{E}_{\lambda_1}\left[V_1^*\left(\lambda_1,\xi_0\right) - \mathbb{E}_{\xi_1}\left[V_1^{MAP}\left(\lambda_1,\xi_1\right)\right]\right] < \frac{T\cdot\sigma_0}{\alpha}\cdot\log\left(n\right)\cdot\mathbb{1}\left\{t=1\right\} + \mathcal{O}\left(\frac{\log\left(n\right)}{n^{\min\{1-\gamma,1/2\}}}\right).$$

1.6 Numerical Experiments

1.6.1 Performance of DA & The Value of Dynamic Programming

Our first numerical experiment investigates how well the deterministic approximation (DA) algorithm works. As a comparison, we also consider a heuristic which simulates the pricing algorithm conducted by a leading used car platform. The algorithm that is currently being used works in the following way: Given horizon length T, for each period t, $t = 1,\cdots,T-1$ (one period typically corresponds to one week), the platform has an estimated demand function, $\phi_t(p)$, which is the probability that the car could be sold at price p when

there are $T - t$ periods left before the car has to be salvaged at price p_T. Then the platform solves a simple optimization problem:

$$\max_p \phi_t(p) p + (1 - \phi_t(p)) p_T.$$

The solution p_t is set as the selling price at period t. For simplicity, in our simulation we assume that $\phi_t(p)$ has the correct function form, but with the future demand arrival rates to be the average forecasted rates and later when we consider the idiosyncratic learning, we assume the customers' latent value in $\phi_t(p)$ is captured by the mean of prior belief. In other words, we assume $\phi_t(p)$ has form

$$\phi_t(p) = \left(1 - e^{-\lambda_t q(p)}\right) p + e^{-\lambda_t q(p)} \left(\left(1 - e^{-\lambda_{t+1} q(p)}\right) p + \cdots + e^{-\lambda_{T-1} q(p)} p_T\right),$$

where λ_t is the observed rate in period t and $\lambda_{t+1}, \cdots, \lambda_{T-1}$ are the mean value of forecast. We have two points to make about this algorithm: First, the optimization problem that the algorithm solves actually assumes a uniform price across the horizon from period t to $T - 1$. We call this algorithm as *Uniform Approximation* (UA) algorithm. Although as time moves forward, the demand function $\phi_t(p)$ gets updated every period, which still leads to a dynamic price trajectory, the policy is not computed via dynamic programming. The algorithm itself neglects the dynamic nature of the pricing problem, which will lead to a significantly lower price trajectory compared with the true optimal. On the other hand, as we have indicated before, deterministic approximation provides an upper bound for the original problem, so DA policy would lead to a higher price trajectory than the true optimal. Second, the demand model $\phi_t(p)$ is not developed based upon customers' utilities and does not have the active idiosyncratic learning as in the Thompson sampling policy. Idiosyncratic learning has been shown to be valuable for used car marketplace by empirical work [1],

therefore with the availability of large amount of customers' online behaviors data, we would expect great opportunities in improving revenues as we will indicate by simulations later.

In our first numerical experiment, we run simulations with horizon lengths $T = 3, 5, 7$, respectively. Individual purchase probability has form $q(p) = e^{-2p} / (1 + e^{-2p})$ with salvage price $p_T = 2$. We consider two scenarios where the problem scales up in degree $1/2$ with scale n with three-point distributed forecast demand rates, $P(\lambda_t = n) = 1/2$, $P(\lambda_t = n - 3\sqrt{n}) = P(\lambda_t = n + 3\sqrt{n}) = 1/4$, and $P(\lambda_t = n) = 1/2$, $P(\lambda_t = n - \sqrt{n}) = P(\lambda_t = n + \sqrt{n}) = 1/4$. We consider scales $n = 10, 50, 100, 500, 1000$ respectively. For each given distribution, horizon length T and scale n, we run the simulation for 1200 times. The average prices and time periods at which the deal is made are summarized in Table 1.1 (with $T = 7$) and the results of revenue approximation ratios are depicted in Figure 1.2 and 1.3, where *Optimal* is the pricing policy of the original problem (1.2).

Table 1.1: Price of Deal and Time of Deal - (Price, Time)

Scale n	10	50	100	500	1000
$(n, 3\sqrt{n})$					
Algorithms					
DA	(2.184,5.940)	(2.548,4.650)	(2.807,4.215)	(3.508,3.755)	(3.849,3.569)
UA	(2.186,5.838)	(2.541,4.214)	(2.784,3.631)	(3.448,2.801)	(3.766,2.546)
OPT	(2.184,5.940)	(2.548,4.650)	(2.807,4.215)	(3.508,3.755)	(3.849,3.569)
$(n, \sqrt{n})$					
Algorithms					
DA	(2.171,5.973)	(2.552,4.650)	(2.784,4.284)	(3.513,3.685)	(3.853,3.542)
UA	(2.170,5.882)	(2.543,4.197)	(2.757,3.737)	(3.446,2.759)	(3.767,2.545)
OPT	(2.171,5.973)	(2.552,4.650)	(2.784,4.284)	(3.513,3.685)	(3.853,3.542)

Notes. $T = 7$, $p_T = 2$. In (Price, Time), Price and Time are the simulated expected price and time at which the deal is made.

According to some summary statistics from the used car platform we are working with, the platform typically sets the time horizon to be 7 weeks (and conducts weekly price adjust-

ment). On average, the deal prices are around 20% higher than the salvage prices and the time each car stays on the platform is around 3 weeks. Therefore, as we can see from Table 1.1, when the problem scale is at 50, the parameters and corresponding results may be most relevant to the real business environment. Our simulation results show that both DA and UA work pretty well and especially, as shown by Figure 1.2 and 1.3, the performance of DA policy almost totally achieves the optimal policy (they are almost not distinguishable from each other on the graph). As the time horizon increases and the problem scale grows, the performance of UA depreciates, while our proposed algorithm DA performs quite robustly.
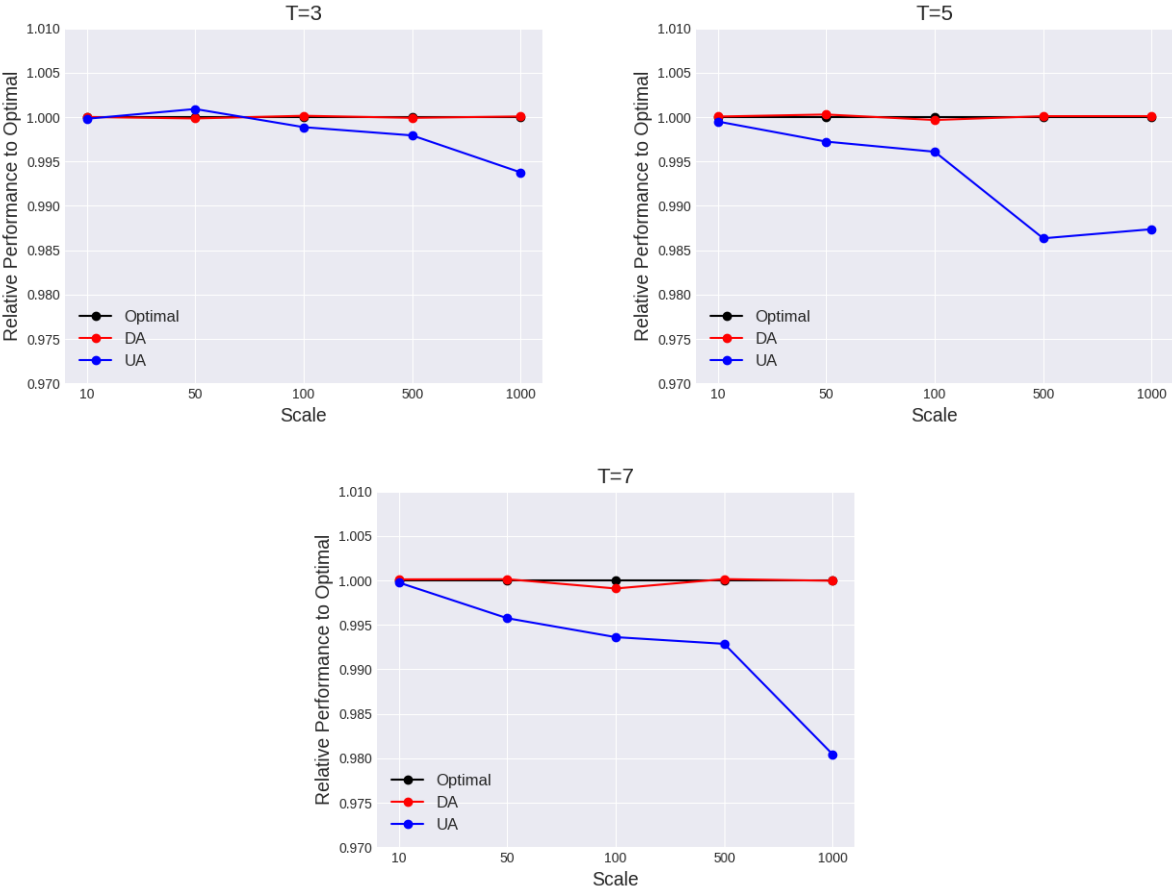
Figure 1.2.: (Color online) Performance Comparison of Dynamic Pricing Algorithms $(n(\pm 3\sqrt{n}))$
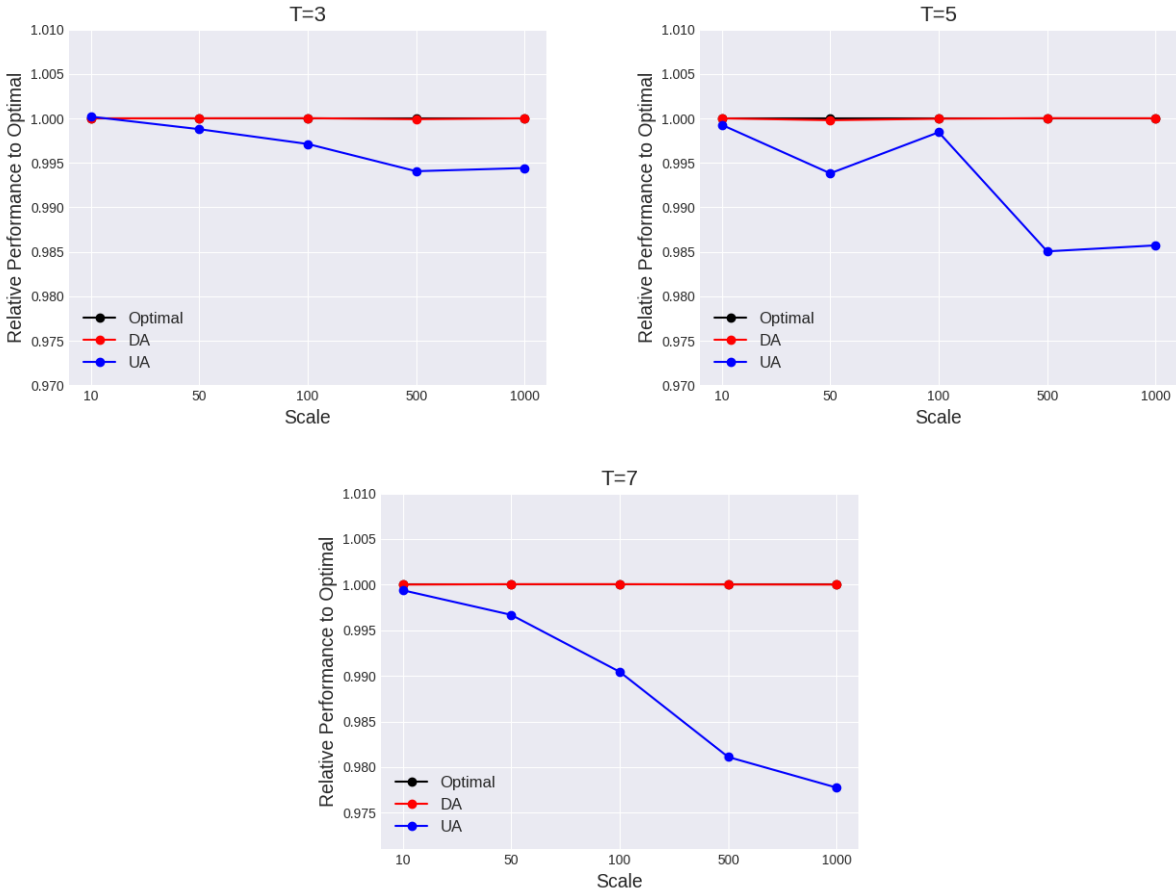
Figure 1.3.: (Color online) Performance Comparison of Dynamic Pricing Algorithms $(n(\pm\sqrt{n}))$

In the simulations we run, on average DA achieves revenue around $1\% - 2\%$ higher than UA. We notice that typically used car platforms' profit rate per car is around 10%,[6] therefore such revenue increase could imply a considerable increase in profit.

The major takeaway from the simulation results is that neglecting the randomness in the demand rate forecast is generally harmless once the dynamic formulation and the average trend of demand rate are correctly characterized.

---

[6]See [1] and CarMax Reports First Quarter Fiscal 2021 Results.

1.6.2 Performance of TS, MAP & The Value of Learning

Our second numerical experiment investigates the performance of *Thompson Sampling* (TS) and *Maximum A Posteriori* algorithms and demonstrates the value of idiosyncratic learning. As our previous experiment has already shown the performance of deterministic approximation and the corresponding revenue gap is small, in our second experiment, we assume that the forecasted demand arrival rates are constants. In TS, the posterior distribution is simulated via MCMC method and we take 1000 iterations before each posterior sampling. In our simulations, MAP runs much faster than TS, which is because optimization can be conducted much more efficiently than MCMC. For comparison, we consider the algorithm using the mean value of posterior distribution as the latent value input instead of sampling from the posterior distribution. We call this algorithm as *Bayesian Mean* (BM) algorithm. In implementation, BM algorithm requires taking average of a big sample of random draws from posterior distribution. We also simulate the performance of the dynamic programming algorithm using the mean value of the prior belief on the latent value as input, i.e., *Dynamic Programming with No Learning* (DP_NL) algorithm. Finally, we simulate the algorithm that is most similar to the platform's current running one, *Uniform with No Learning* (U_NL) algorithm, i.e., applying uniform price optimization using the mean value of prior distribution on latent value (i.e., without idiosyncratic learning) as we described before. In the experiment, we have $\lambda_1 = \cdots = \lambda_{T-1} = 1$, ground truth latent value ξ is sampled from standard normal distribution $N(0,1)$, prior belief over ξ has no bias, which is also a standard normal distribution. Each arrival's purchasing behavior consists of two stages: First with probability $e^{\xi-p}/(1 + e^{\xi-p})$, the arrival becomes a *promising customer*. This stage corresponds to customers' online behaviors that indicate their strong interests in the item. Then with probability $e^{\xi-p}/(1 + e^{\xi-p})$, each *promising customer* decides to purchase the item. Each algorithm is simulated for 360 times. The simulation details for $T = 7$

are summarized in Table 1.2 and Table 1.3, where Table 1.2 summarizes the performance of algorithms from period 2 to the end, while Table 1.3 summarizes the whole horizon. The reason that we make this distinction is that in the first period when we do not have historical data, learning in the algorithm won't make a difference, the revenue loss is largely driven by the prior misspecification in the first period. All simulated revenue approximation ratios are shown in Figure 1.4 (start from period 2) and Figure 1.5 (whole horizon), where *Optimal* is the optimal pricing policy of (1.2) solved by a clairvoyant decision maker who knows the true latent value ξ.

Table 1.2: Price of Deal and Time of Deal - (Price, Time), (Start from Period 2)

Scale	10	50	100	500	1000
Algorithms					
TS	(2.255,5.836)	(2.563,5.178)	(2.855,4.756)	(3.389,4.333)	(3.719,4.183)
BM	(2.246,5.911)	(2.541,5.256)	(2.889,4.758)	(3.383,4.453)	(3.728,4.236)
MLE	(2.254,5.853)	(2.581,5.222)	(2.905,4.814)	(3.425,4.494)	(3.735,4.347)
DP-NL	(2.230,5.608)	(2.474,4.925)	(2.697,4.464)	(3.171,4.361)	(3.550,4.064)
U-NL	(2.221,5.581)	(2.440,4.764)	(2.645,4.192)	(3.104,3.914)	(3.412,3.692)
OPT	(2.283,6.086)	(2.621,5.319)	(2.942,4.842)	(3.454,4.572)	(3.735,4.372)

Notes. $T = 7$, $p_T = 2$. In (Price, Time), Price and Time are the simulated expected price and time at which the deal is made.

Table 1.3: Price of Deal and Time of Deal - (Price, Time)

Scale	10	50	100	500	1000
Algorithms					
TS	(2.189,4.808)	(2.534,3.481)	(2.712,3.214)	(3.289,2.561)	(3.659,2.200)
BM	(2.191,4.797)	(2.536,3.450)	(2.711,3.189)	(3.309,2.542)	(3.699,2.139)
MAP	(2.19,4.819)	(2.539,3.442)	(2.735,3.217)	(3.327,2.594)	(3.699,2.211)
DP-NL	(2.193,4.714)	(2.530,3.394)	(2.700,3.258)	(3.270,2.800)	(3.606,2.578)
U-NL	(2.182,4.706)	(2.485,3.142)	(2.651,2.892)	(3.187,2.208)	(3.489,1.975)
OPT	(2.244,5.031)	(2.649,4.058)	(2.910,3.692)	(3.503,3.100)	(3.866,2.747)

Notes. $T = 7$, $p_T = 2$. In (Price, Time), Price and Time are the simulated expected price and time at which the deal is made.
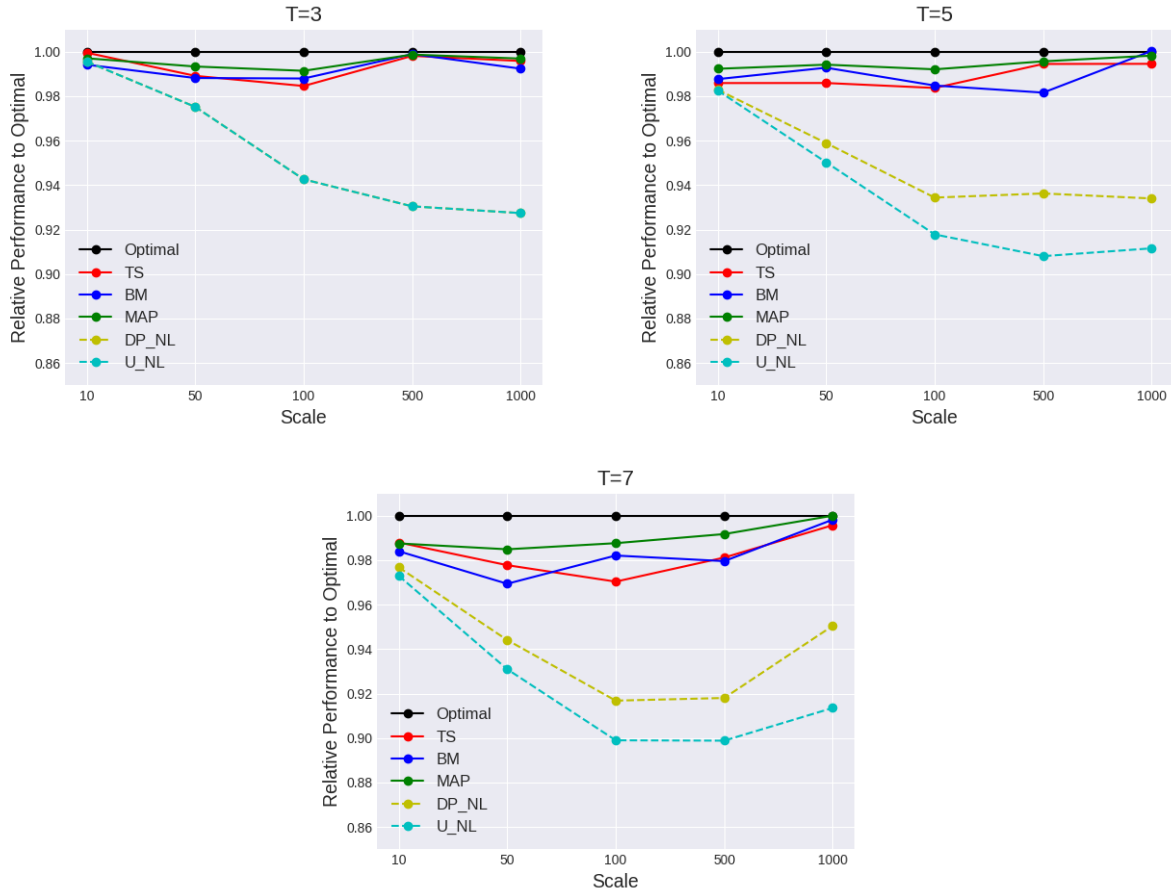
Figure 1.4.: (Color online) Performance Comparison of Dynamic Pricing Algorithms (Start from Period 2)

As we can see from the results summary, TS, BM and MAP all achieve very high approximation ratios and their simulated performances are close. But in most simulated cases, MAP appears to slightly outperform the other algorithms. This implies that the sophistication in the exploration step of the active learning strategies does not necessarily bring the edge over the passive learning strategies. The necessity of implementing active learning strategies should depend on the specific problem settings. Although our problem scales up in the demand rate for fixed horizon length, as a robustness check, we also conduct simulations for longer horizon and fixed demand rate, see Table 1.4 and Figure 1.6, 1.7. Again, we find in most cases, MAP slightly outperforms the other strategies.
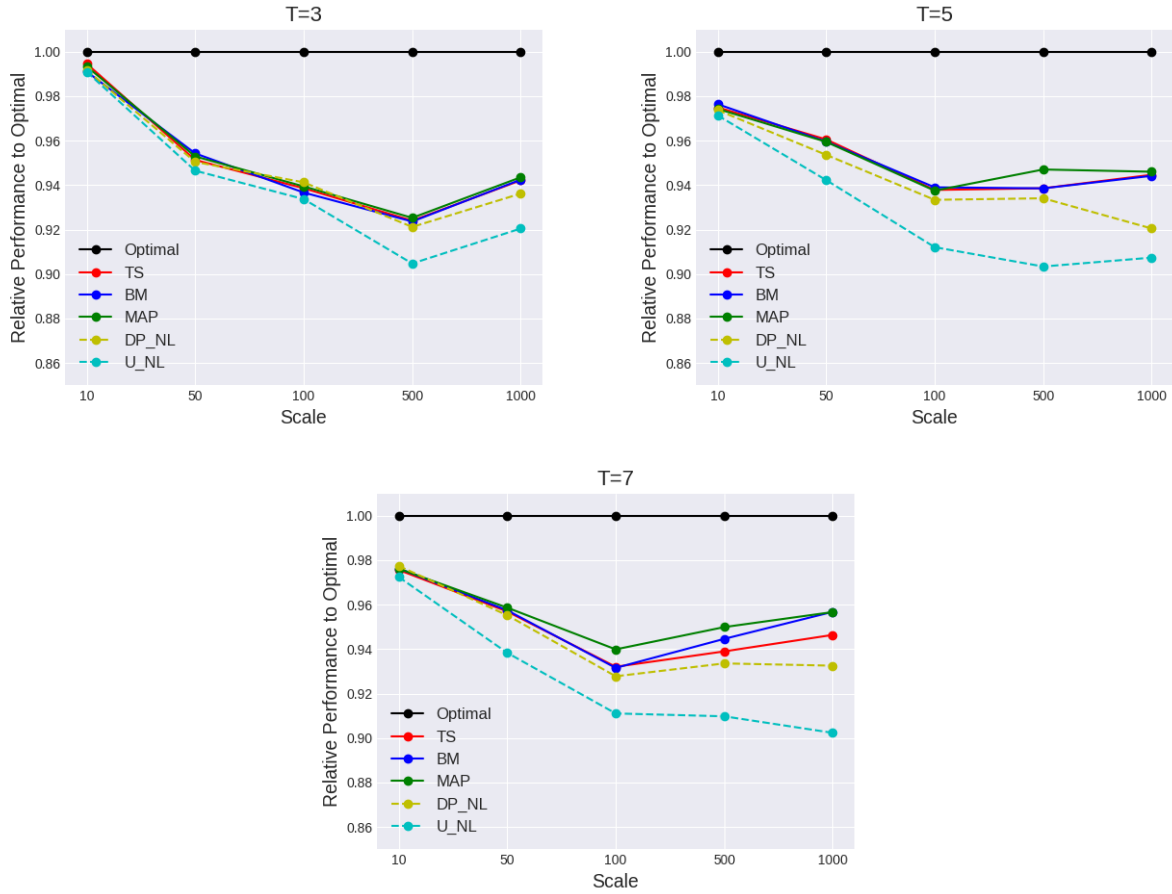
Figure 1.5.: (Color online) Performance Comparison of Dynamic Pricing Algorithms (Start from Period 1)

Finally, compared to the numerical results in Section 1.6.1, our simulation results in Section 1.6.2 indicate that the estimation accuracy of the idiosyncratic latent value can make a more significant difference. This observation provides some meaningful insights to the managers that correctly characterizing the average demand dynamics can generally lead to a decent performance, while more effort should be invested into accurately understanding the idiosyncratic value of each unit of asset.

Table 1.4: Price of Deal and Time of Deal - (Price, Time)

Horizon Length	10	15	20	25	30
$\xi_0 \sim N(0,1)$					
Algorithms					
TS	(2.199,8.197)	(2.262,11.547)	(2.307,14.931)	(2.339,18.411)	(2.359,21.903)
BM	(2.200,8.228)	(2.266,11.528)	(2.314,14.989)	(2.347,18.458)	(2.357,21.747)
MAP	(2.197,8.222)	(2.265,11.536)	(2.314,14.908)	(2.358,18.639)	(2.379,21.828)
DP-NL	(2.196,8.119)	(2.249,11.339)	(2.296,14.586)	(2.332,17.836)	(2.333,21.300)
U-NL	(2.196,8.006)	(2.239,11.067)	(2.283,14.200)	(2.311,17.261)	(2.307,20.642)
OPT	(2.222,8.625)	(2.327,12.525)	(2.411,15.892)	(2.449,19.703)	(2.459,23.392)
$\xi_0 \sim N(0,2)$					
Algorithms					
TS	(2.256,7.231)	(2.330,10.217)	(2.355,13.706)	(2.396,16.894)	(2.450,18.314)
BM	(2.262,7.136)	(2.320,10.189)	(2.343,13.639)	(2.385,16.858)	(2.474,18.514)
MAP	(2.270,7.208)	(2.359,10.333)	(2.388,13.719)	(2.416,17.367)	(2.491,18.897)
DP-NL	(2.246,7.039)	(2.306,9.939)	(2.329,13.075)	(2.334,16.314)	(2.428,17.281)
U-NL	(2.235,6.936)	(2.289,9.744)	(2.306,12.717)	(2.306,15.917)	(2.378,16.997)
OPT	(2.557,8.217)	(2.643,11.681)	(2.644,15.922)	(2.686,19.989)	(2.801,22.436)

Notes. $Scale = 5$, $p_T = 2$. In (Price, Time), Price and Time are the simulated expected price and time at which the deal is made.

1.7 Conclusion

Despite the fast-growing business of online asset selling including used car and real estate marketplaces, limited amount of research has been devoted to developing models that are closely adaptive to the real business environments. Our paper aims to fill this gap via formulating a dynamic asset selling framework which can be extended into a data-driven setting. We propose implementable algorithms that are proven to achieve asymptotic optimality as the demand rates scale up. As most of the existing literature investigates asymptotic performance of high-frequency dynamic pricing policies in the regime where the time horizon or initial inventory grows large, due to the nature of our problem, we study the asymptotic
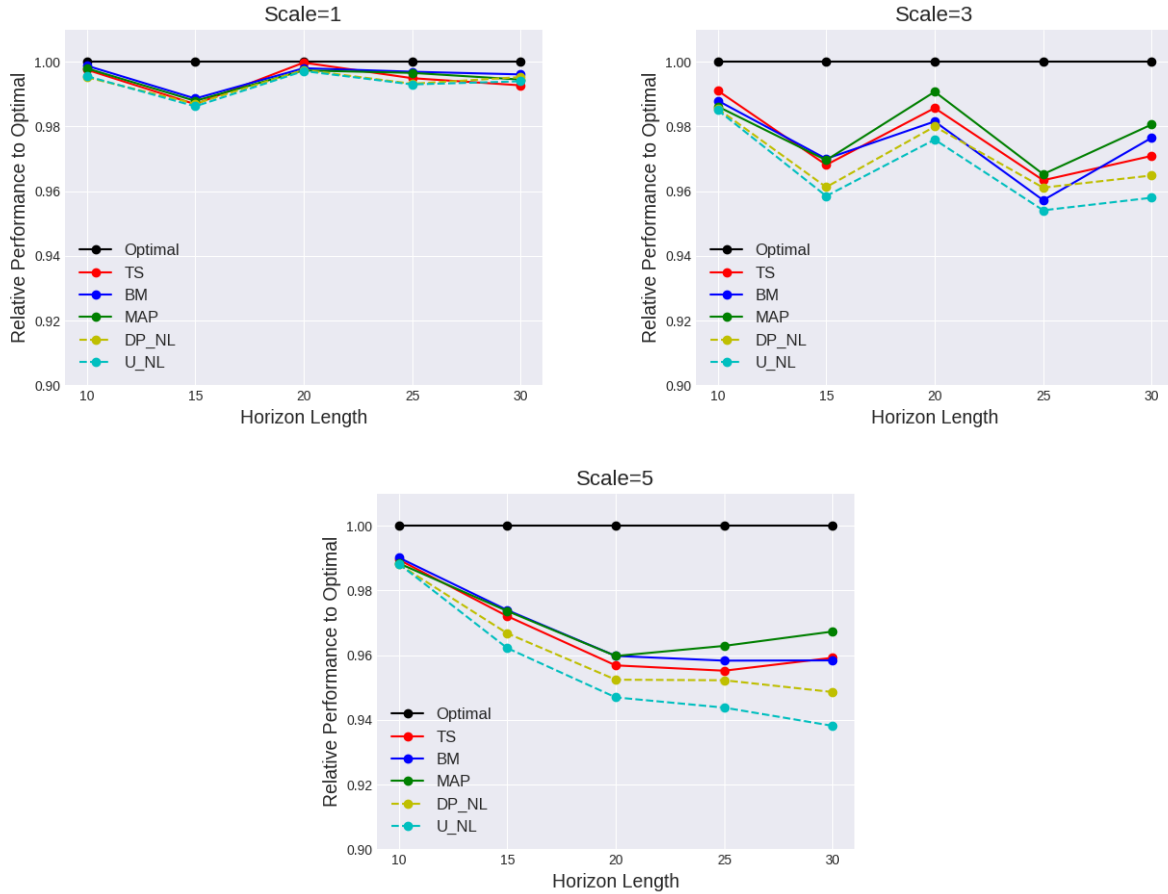
Figure 1.6.: (Color online) Performance Comparison of Dynamic Pricing Algorithms

performance of infrequent dynamic pricing policies on a single unit of asset over a fixed time horizon instead. Our model well fits some business restrictions encountered in real practice including infrequent price change and volatility in demand patterns. Especially, our model is build upon individual customer's utility function, which naturally facilitates utilizing online behavioral data for demand learning. We prove that our general asset selling framework admits ideal mathematics properties and allows us to conduct regret analysis for various policies under a dynamic programming setting. Finally, we use numerical experiments to show that our proposed algorithms can potentially improve the revenue performance significantly compared with an algorithm that is currently implemented by a leading used car platform.
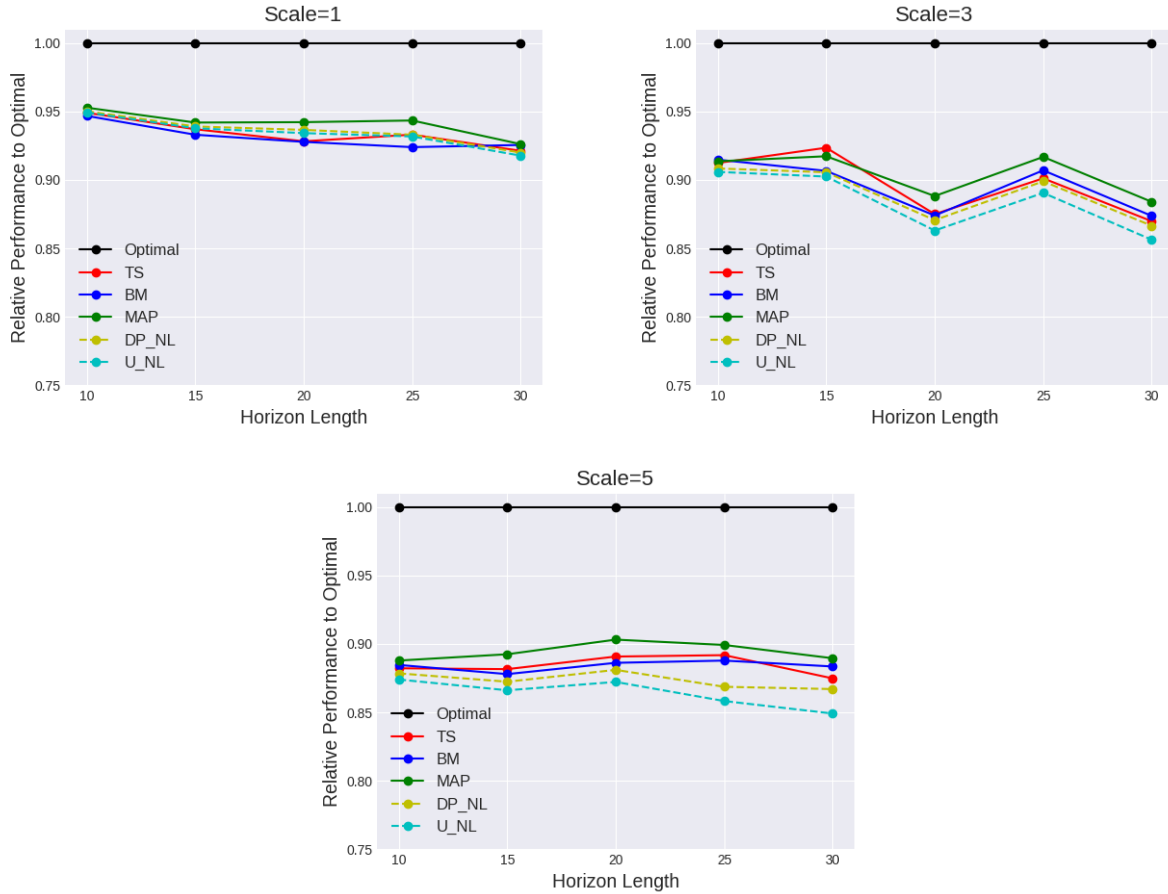
Figure 1.7.: (Color online) Performance Comparison of Dynamic Pricing Algorithms

Besides, we find that using simple deterministic proxy of demand forecast is mostly harmless, while accurate estimation on the idiosyncratic latent value may make more significant differences. Simulations also reveal that in our problem setting, the exploration step in an active learning policy may not help outperform a passive learning policy. This indicates that the effectiveness of active learning highly depends on the nature of the problem, which may be of independent interest.

2. Cash Hedging Motivates Information Sharing in Supply Chains

2.1   Introduction

Some major operational and financial activities of a firm are largely driven by the firm's understanding of the market conditions. Therefore, it can become quite challenging for a firm to carry out those activities under an information asymmetric environment, like a supply chain where the upstream supplier is not fully aware of the end market conditions. Supply chain experts argue that market demand information sharing among retailers and their suppliers improve supply chain efficiency, and under appropriate contracts can result in benefits for all involved parties ([46]). Erroneous understanding of market demand conditions amplified through retail forecast errors propagate upstream to drive chaotic ordering, which drives resource inefficiencies, increased inventories, or drastic shortages. The so-called "bullwhip effects" of informational inefficiencies in the chain are well understood ([47]), but the information sharing practices encounter significant barriers in their implementation and are not as frequently observed in practice ([48], [49]).

The information sharing literature in supply chain systems with vertical relationships among retailers and suppliers, and in the presence of horizontal competition among retailers, argues that voluntary market demand information sharing among retailers and their supplier (vertical information sharing) within such systems is hard to achieve as an equilibrium ([50] and [51]). Such information sharing has two effects: "direct effects" due to the changes in subsequent decisions of the parties (retailer and supplier) involved in the information sharing, and "indirect (leakage) effects" due to changes in decisions by other competing firms (retailers

in our context) based on inferences of information from actions of the informed parties. The direct effect benefits the supplier, who may appropriate extra value in the chain through contract pricing using such information but disadvantages the retailer. The leakage effect also hurts the retailer in their downstream competition, and thus under both effects, it is unlikely for the retailer to voluntarily reveal market demand information in a non-cooperative game equilibrium ([50]).

The usual studied setting that confers the above observations is a two-level supply chain with one upstream firm (supplier, "he") and multiple downstream firms (retailers, "she") horizontally competing in a Cournot competition by selling a homogeneous product at a constant marginal cost. The supplier supplies the retailer in a Stackelberg fashion via a wholesale price contract. Within this setting, the private market demand information, when revealed to the supplier, allows him to appropriate value through the wholesale price (direct effect), and by observing the supplier prices, the competing retailers engage in more aggressive Cournot competition (leakage effect). The limitation of this setting is that it does not offer any opportunities for the retailer to gain from sharing the market demand information through subsequent actions of the supplier. For example, the supplier may use the market demand information to undertake operational actions justified in the presence of strong market signals, such as operational improvement efforts that lower costs or financial hedging contracts that hedge relevant cost risks. When the supplier takes such actions, he alters his cost structure, and he may pass some of the benefits to the downstream players. These trickled down to the retailers' benefits may offer incentives to them for sharing their demand information.

In our paper, we advance the argument that vertical market demand information sharing may happen when such information leads to supplier operational actions altering its cost structure in a way beneficial to the downstream players (retailers). We start with a bilateral supply chain of a supplier and a retailer, with the supplier supplying in a Stackelberg

43

fashion the retailer via a wholesale price contract. The supplier is financing his operational improvements through his own cash flows, but these cash flows are volatile. The supplier invests the realized cash outcome to increase production efficiency. The retailer faces price-sensitive random demand, with the market potential her private information. The retailer decides to reveal her information and orders from the supplier. There are various scenarios to be modeled where the supplier controls actions that depend on the market demand information, and when executed, alter his cost structure in ways that may benefit the retailer. For example, the supplier decides on operational improvement investments depending on realized internal cash flow and revealed market potential. Such decisions can lower supplier costs, with some of the savings passed on through the wholesale price to the retailer. Alternatively, the supplier may engage in financial hedging in the presence of operational improvement investments convex on the invested cash flow, with such hedging justified in the presence of strong market potential ([52]). In the remaining of the paper, we are going to model the second such scenario. It is practical, easy to execute, and has appealing analytical tractability. Our model allows us to enhance our understanding of how operational actions, information-sharing policies, and financial decisions may interact within a supply chain setting.

When we investigate the vertical information-sharing problem by taking the supplier's cash flow hedging decision process into consideration, the value of building up such an information-sharing channel can be justified under some conditions. The supplier's hedging policy determines the wholesale cost risk the retailer faces, while the end market demand information allows the supplier to make better hedging decisions on his own interest. However, the resulting cost outcome under supplier hedging may benefit the retailer as well. In certain circumstances, the indirect benefits of the retailer from the supplier's hedging turn out to be enough to overcome the disadvantageous direct effect of information revelation, and result in the retailer's voluntary information sharing. On the other hand, our finding

also provides a new reason why a firm should consider hedging. The presence of the hedging option can be used to help achieve agreement on information sharing during the supplier's negotiation with the retailer.

We extend the basic bilateral supply chain into the case where there are multiple homogeneous Cournot-competing retailers, thus bringing up the opportunity for leakage effects. The classic analysis of this setting (see [50]) substantiated the disadvantageous direct and leakage effects for the retailers, and lead to a symmetric equilibrium of no-voluntary private market information by the retailers. However, when the information allows cost control actions by the supplier, as in engaging in financial hedging, the information-sharing game ends up in an asymmetric equilibrium even though the competing retailers are identical. Furthermore, we are able to show that building up an information-sharing channel generally won't backfire and sometimes can Pareto improve the system and consumer welfare.

Finally, we study the information-sharing game in a supply chain where we have independent heterogeneous retailers, i.e., are located in different markets of differing potential market sizes. While in the absence of supplier cost control actions, such setting would have led to two separate and independent vertical channels, in the presence of such actions (e.g., our modeled financial hedging that depends on the total market), the two vertical channels are interlinked. We find that the heterogeneity in market sizes, and the correlation between market shocks, play significant roles in shaping vertical information sharing equilibrium. Contrary to the single market competition setting, although the supplier can never do worse by building up an information-sharing channel, information sharing may not only hurt the retailer but also the system as a whole and consumer welfare. While in our two previous settings, the supply chain efficiency is improved with information sharing, in the heterogeneous independent market system having more informed agents does not necessarily mean a more efficient system. In these cases, whether information-sharing channel benefits the system really depends on the structure of the game (i.e., market sizes and correlation of market

shocks). We also point out that the simultaneous information-sharing game of retailers from different markets may not admit a pure Nash equilibrium, which suggests that in practice, a sequential game process may be necessary to guarantee a stable outcome.

In contrast to the previous literature that proved the direct and leakage effects are disadvantageous for retailers in vertical information sharing with their common supplier, our paper argues that supplier's cost control actions based upon revealed market information may provide indirect benefits to retailers to counter these disadvantageous effects. Our modeling of financial cash flow hedging by the supplier, with cash flow investments driving operational cost improvements, analytically supports this argument. Interestingly, vertical information sharing may take place under plausible conditions, and in a counter-intuitive result, the Nash equilibrium may be asymmetric even for the case of two identical retailers. In our paper, we gain an understanding of the previously unexplored interaction between the two major operational policies: informational operations and operational cost investments (e.g., cash flow allocation and/or financial hedging). The financial policies of the supplier on operational improvements are dependent on accurate market information, and their execution may provide benefits to downstream retailers. Thus, in an indirect fashion, the supplier's cost control actions provide the hidden incentives for retailers' voluntary information sharing.

2.2 Literature Review

Our paper revisits the issue of incentives for firms to share information vertically in the presence of horizontal competition. The setting we consider is that of a supplier serving many competing retailers that have private information about the market demand. The supplier employs an operational action that depends on market information (e.g., a financial hedge of volatile cash flows in support of operational improvement investments) as an indirect way to offer positive incentives to their counterparties (e.g., retailers) to reveal their private

information. Therefore, the paper touches upon the literature of vertical information sharing in the presence of horizontal competition, information sharing under oligopoly, and financial hedging in support of operational actions.

Early economic work dealt with information sharing in an oligopoly. The classic references in this literature are [53–59]. These papers are concerned with whether a firm has incentives to share its private information with competing firms in an oligopoly. All of them study horizontal information sharing without addressing issues of vertical information sharing. A typical message of this work is that information sharing is unlikely to emerge as an equilibrium in Cournot competition but may take place in Bertrand competition. Our paper studies vertical information sharing in the presence of horizontal Cournot competition. While it models both the direct and leakage effects for these environments, it brings up an alternative operational viewpoint in the operation of the vertical channel. The retailers' private information is often enacting operational actions of the supplier that indirectly benefit the retailers. This hidden incentive has important implications for the resulting information sharing equilibrium under certain conditions.

The closest literature to our work is vertical information sharing in the context of horizontal competition. In particular when the studied structure is that of a single supplier and multiple retailers. For the case of demand uncertainty, [50] shows that both the direct effect and the leakage effect are disadvantageous for the Cournot competing retailers to voluntarily share information with the supplier. Furthermore, vertical information sharing reduces social benefits. [51] considers the stylized setting of a supplier serving two competing retailers who sell differentiated goods. The paper shows that the supplier's optimal strategy is independent of the type of downstream competition, Cournot or Bernard, and that no voluntary information sharing takes place. Motivated by these results, our work argues on the need to model the implications of the revealed information on the supplier's operational actions

that alter his cost structure, and then through the vertical channel contracting mechanism creating benefits for the downstream retailers.

Later work on vertical information sharing with horizontal competition brought up aspects of confidentiality ([60,61]) and multiple competing supply chains ([62–65]). Our work considers a single supplier supplying horizontal competitors and thus does not deal with competing supply chains. Furthermore, we do not study signaling issues involved in confidential or voluntary disclose of information cases. Any shared information by the retailers is publicly disclosed as in [50].

Most of the supply chain management literature works within a stylized supplier-retailer bilateral supply chain, with the retailer modeled as a newsvendor in the presence of demand uncertainty and the supplier responding to orders according to a contract. Authors model information asymmetry issues via the retailer having private information about the demand distribution. Information sharing studies in this literature either quantify the benefits of information sharing by reduced "bullwhip" inefficiencies in ordering and service levels, or address issues of contracting. The contracting work deals with double-marginalization and offering incentives for sharing demand information. Our work departs substantially from this literature as we are concerned with issues of vertical information sharing in the presence of horizontal competition. The retailer competition is with market clearing quantities in an uncertain size market. We are not concerned with the short-term market matching issues (excess inventories or shortages) of a newsvendor retailer.

The value of information sharing in supply chains has been extensively studied by both theoretical work (see [48,66–68]), and empirical work (see [69,70]). It is challenging to align supply chain agents' interests towards information sharing, with quite a few papers studying associated contract issues. [67] study the contract design problem to induce credible information sharing along the bilateral supply chain. [71] and [72] show that revenue/profit sharing and buyback contracts could better align supply chains and achieve vertical information

sharing. [73] design contracts that prevent information sharing. The non-monotone profits of supply chain agents in forecast accuracy has been widely documented by work including [74–78], etc. As we mentioned before, the "majority of these papers use a serial system isolated from horizontal competition" (see [50]) and emphasize short-term market mismatch gains through reduced inventories and shortages. As our focus has shifted to better understanding how operational actions of the supplier, in response to the revealed information, play in environments of information leakage due to retailer competition, we have suppressed these effects. This way the model focuses on the main issue we study and is tractable.

Another related literature comes from the financial hedging area. Our paper is motivated by observations in the corporate finance literature that volatility in a firm's cash flows may compromise its ability to invest in operational improvement efforts. Our supplier may be concerned about volatile cash flows for a multiplicity of reasons: material prices, production yields, exchange rates, if a foreign supplier, and demand uncertainties. As pointed out in [79], the supplier engaging in financial hedging better deploys operational improvement investment opportunities, and this way lowers expected costs and risks. Some empirical evidence in support of such practices appears in [80] and [81]. [82] present empirical evidence that cash hedging policies in support of operational actions may be more effective than cost hedging of independent factors (materials, exchange rates, etc.). Following the spirit of [79] and [83], in our paper, the supplier uses internal funds to finance investments in cost-reducing operational improvements. Addressing horizontal competition but not in a supply chain context, [52] studies the cash hedging game with Cournot competing homogeneous retailers and shows that asymmetric information sharing outcome may emerge as Nash equilibrium in spite of the symmetric model setting. [84] studies a general duopolistic risk exposure game and characterizes the conditions under which the competing firms are willing to expose themselves to risks. The paper does not specify a business context, yet the framework built in the paper is closely related to the horizontal hedging game in [52]. [85] study the cash

hedging game in a bilateral supply chain but with fully symmetric information. It shows that the correlation between agents' cash flows, market size and volatility play critical roles in determining the hedging equilibrium. To the best of our knowledge, there is no work so far that investigates the interplay between informational policies (e.g., vertical information sharing) and financial hedging in supply chains. Importantly, we identify a new reason why a supplier should keep the option of hedging and call the awareness of his retailers to the presence of his hedging lever, which is to garner more market information. Our paper captures the hidden incentive for the retailer to share private information with the supplier, when this information drives the right operational action of the supplier. Under certain conditions, these new incentives overcome direct and leakage effects of vertical information sharing for the retailer, and lead to voluntary information sharing equilibria. However, the resulting equilibria may be surprising, with an asymmetric equilibrium potentially emerging in a setting with multiple homogeneous retailers, and for heterogeneous separate retailers' correlation among their market shocks shapes the equilibrium.

We organize the remainder of this paper as follows. Section 2.3 introduces our basic bilateral supply chain model. Then we conduct equilibrium analysis and identify the conditions for voluntary information sharing. And we explain the main intuitions via a profit decomposition method. Section 2.4 extends the bilateral supply chain into a two-stage supply chain where there are multiple homogeneous Cournot-competing retailers. Section 2.5 studies the information-sharing game when the single supplier serves two independent but heterogeneous markets. And finally, we conclude in Section 2.6 with a summary of our main results and insights.

2.3 A Bilateral Supply Chain

Consider a simple supply chain, there is a supplier "he" with marginal production cost $c = C(\cdot)$ which is a convex decreasing function of investment, i.e., $C'(\cdot) < 0$ and $C''(\cdot) > 0$. The cost investment can include production technology upgrading and capacity planning, etc. The investment comes from an internal random cash flow ξ, $\xi \sim F_\xi(\cdot)$. The supplier can choose whether to hedge the cash flow, if he hedges, then the realized cash flow is $\bar{\xi} = \mathbb{E}[\xi]$. Here we make two major assumptions. First, the supplier does not have access to external finance, i.e., the supplier solely relies on his internal finance for investment. Second, the marginal return to investment is higher than the opportunity cost of funds, which implies that the realized cash flow will be fully invested into the cost improvement. Therefore, the corresponding realized marginal cost is $c = C(\xi)$ if the supplier does not hedge and $\bar{c} = C(\bar{\xi})$ if the supplier hedges. See [52] for more discussions on these two assumptions.

The supplier sells products via a wholesale price contract to the retailer. The retailer "she" is a monopoly in a linear-demand market with demand function $p = a - bq$, where p is retailing price and q is demand. The intercept a reflects the potential market size, $a \sim F_a(\cdot)$ and a is supported on $[\underline{a}, \bar{a}]$. To ensure a positive production quantity, we assume that for all realizations of potential market size a, cash flow and corresponding cost c, $a > c$. In this production step after the cost realization, we assume both the supplier and the retailer are not financially constrained. This is because unlike the cost investment which is generally a longer-term decision and has to be committed at the beginning, cost incurred during production process is short-term and is easier to be financed (also see [52] for the discussions on this). The retailer can observe the market condition a and has the option to share this information with the supplier before the realization of a. Following the majority of the information sharing literature, information sharing in our paper refers to a pre-committed

agreement on market information disclose. Like cost investment, information sharing is also a long-term decision that should take place before the production process starts.

$C(\cdot)$, $F_\xi(\cdot)$ and $F_a(\cdot)$ are public known. Then the timeline of events is as follows:

1. The retailer decides whether to disclose the market condition with her supplier.

2. Market condition a is realized. Upon getting the information of a (if the retailer shares a), the supplier chooses whether to hedge the cash flow ξ.

3. After the realization of ξ, the supplier invests ξ into cost improvement and then decide the wholesale price.

4. The retailer chooses the quantity to buy from the supplier. And finally profits are realized.

2.3.1 Analysis

To analyze the equilibrium, we first derive the equilibrium under given information-sharing policy and then discuss retailer's information-sharing decision via comparison.

The Retailer Shares Information.

For given realization of a and wholesale price w, the retailer's problem is:

$$\max_q \pi_R(q; a) = \max_q (p - w) \cdot q = \max_q (a - bq - w)q,$$

which leads to $q^* = (a - w)/2b$ and $\pi_R^*(q^*; a) = (a - w)^2/4b$. Given the retailer's best response, if the supplier does not hedge, then for any realizations of a and c, the supplier's problem is

$$\max_w \pi_S(w; a, c) = \max_w (w - c) \cdot q^* = \max_w \frac{1}{2b}(w - c)(a - w),$$

which leads to $w^* = (a + c)/2$ and $\pi_S(w^*; a, c) = (a - c)^2/8b$. Therefore, in the equilibrium, $q^* = (a - w^*)/2b = (a - c)/4b$ and $\pi_R(q^*; a) = (a - c)^2/16b$. So for the given a, the supplier's expected profit is

$$\pi_S^{nh*}(a) = \mathbb{E}_\xi[\pi_S^*(w^*; a, c)] = \mathbb{E}_\xi\left[\frac{(a - c)^2}{8b}\right] = \frac{1}{8b}\left(a^2 - 2a\mathbb{E}_\xi[c] + \mathbb{E}_\xi[c^2]\right).$$

On the other hand, if the supplier hedges, then for any realization of a, the supplier's problem is

$$\max_w \pi_S(w; a, \bar{c}) = \max_w \frac{1}{2b}(w - \bar{c})(a - w),$$

which leads to $w^* = (a + \bar{c})/2$ and $\pi_S(w^*; a, \bar{c}) = (a - \bar{c})^2/8b$. Therefore, in the equilibrium, $q^* = (a - w^*)/2b = (a - \bar{c})/4b$ and $\pi_R(q^*; a) = (a - \bar{c})^2/16b$. So for the given a, the supplier's expected profit is

$$\pi_S^{h*}(a) = \mathbb{E}_\xi[\pi_S^*(w^*; a, \bar{c})] = \frac{(a - \bar{c})^2}{8b}.$$

Now we get

$$\pi_S^{h*}(a) - \pi_S^{nh*}(a) = \frac{1}{8b}\left(2(\mathbb{E}_\xi[c] - \bar{c})a - (\mathbb{E}_\xi[c^2] - \bar{c}^2)\right),$$

therefore, the supplier hedges if and only if

$$a \geq \frac{\mathbb{E}_\xi[c^2] - \bar{c}^2}{2(\mathbb{E}_\xi[c] - \bar{c})} \triangleq t, \tag{2.1}$$

which is consistent with the result in [85]. We assume that $t \in [\underline{a}, \bar{a}]$. Based on this, we can get the expected optimal profit of the retailer as follows,

$$\pi_R^{s*} = \mathbb{E}_\xi\left[\mathbb{E}_a\left[\frac{(a - w^*)^2}{4b}\right]\right] = \mathbb{E}_\xi\left[\int_{\underline{a}}^t \frac{(a - c)^2}{16b}dF(a) + \int_t^{\bar{a}} \frac{(a - \bar{c})^2}{16b}dF(a)\right]$$

$$= \frac{1}{16b}\left(\mathbb{E}_a[a^2] - 2\left(\mathbb{E}_\xi[c]\int_{\underline{a}}^t a\, dF(a) + \bar{c}\int_t^{\bar{a}} a\, dF(a)\right) + \mathbb{E}_\xi[c^2] F(t) + \bar{c}^2 \bar{F}(t)\right). \tag{2.2}$$

The Retailer Does Not Share Information.

For the given wholesale price w, the retailer's problem is the same as before, so we still have that $q^* = (a - w)/2b$ and $\pi_R(q^*; a) = (a - w)^2/4b$. Therefore, for any given a and c, the profit of the supplier is $\pi_S(w; a, c) = (w - c) \cdot q^* = (w - c)(a - w)/2b$. However, the supplier does not know the realized value of a, so for a given c, the supplier's problem is

$$\max_w \mathbb{E}_a\left[\pi_S(w; a, c)\right] = \max_w \mathbb{E}_a\left[\frac{(w - c)(a - w)}{2b}\right] = \max_w \frac{(w - c)(\mathbb{E}_a[a] - w)}{2b},$$

which leads to that $w^* = (\mathbb{E}_a[a] + c)/2$ and $\pi_S(w^*; c) = (\mathbb{E}_a[a] - c)^2/8b$. So if the supplier does not hedge, his expected profit is

$$\pi_S^{nh*} = \mathbb{E}_\xi[\pi_S^*(w^*; c)] = \mathbb{E}_\xi\left[\frac{(\mathbb{E}_a[a] - c)^2}{8b}\right] = \frac{1}{8b}\left(\mathbb{E}_a[a]^2 - 2\mathbb{E}_a[a]\mathbb{E}_\xi[c] + \mathbb{E}_\xi\left[c^2\right]\right).$$

On the other hand, the hedged supplier's expected profit is

$$\pi_S^{h*} = \mathbb{E}_\xi[\pi_S^*(w^*; \bar{c})] = \mathbb{E}_\xi\left[\frac{(\mathbb{E}_a[a] - \bar{c})^2}{8b}\right] = \frac{1}{8b}\left(\mathbb{E}_a\left[a^2\right] - 2\mathbb{E}_a[a]\bar{c} + \bar{c}^2\right).$$

Now we get

$$\pi_S^{h*} - \pi_S^{nh*} = \frac{1}{8b}\left(2\left(\mathbb{E}_\xi[c] - \bar{c}\right)\mathbb{E}_a[a] - \left(\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2\right)\right),$$

therefore, the supplier hedges if and only if

$$\mathbb{E}_a[a] \geq \frac{\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2}{2\left(\mathbb{E}_\xi[c] - \bar{c}\right)} = t. \tag{2.3}$$

54

So when $\mathbb{E}_a[a] \geq t$, the supplier will hedge and $w^* = (\mathbb{E}_a[a] + \bar{c})/2$. The corresponding expected profit of the retailer is

$$\pi_R^{ns*} = \mathbb{E}_a \left[\frac{\left(a - \frac{\mathbb{E}_a[a] + \bar{c}}{2} \right)^2}{4b} \right] = \frac{1}{16b} \left(4\mathbb{E}_a\left[a^2\right] - 3\mathbb{E}_a[a]^2 - 2\mathbb{E}_a[a]\bar{c} + \bar{c}^2 \right). \quad (2.4)$$

When $\mathbb{E}_a[a] < t$, the supplier will not hedge and $w^* = (\mathbb{E}_a[a] + c)/2$. The corresponding expected profit of the retailer is

$$\pi_R^{ns*} = \mathbb{E}_\xi \left[\mathbb{E}_a \left[\frac{\left(a - \frac{\mathbb{E}_a[a] + c}{2} \right)^2}{4b} \right] \right] = \frac{1}{16b} \left(4\mathbb{E}_a\left[a^2\right] - 3\mathbb{E}_a[a]^2 - 2\mathbb{E}_a[a]\mathbb{E}_\xi[c] + \mathbb{E}_\xi[c^2] \right).$$

$$(2.5)$$

2.3.2 Is Voluntary Information Sharing Possible?

By now, we have two assumptions on the parameters: (i), For any realizations of a and c, we have $a > c$; (ii), $t \in [\underline{a}, \bar{a}]$. We will check that under these two assumptions, whether voluntary information sharing of the retailer is possible, i.e., $\pi_R^{s*} > \pi_R^{ns*}$. The following proposition characterizes the conditions under which voluntary information sharing takes place:

Proposition 2.3.1 *Voluntary information sharing takes place if and only if*

$$\begin{cases} 2 \left(\mathbb{E}_\xi[c] - \bar{c} \right) \int_{\underline{a}}^t F(a)\, da > 3Var(a) \ , \ \textit{if } \mathbb{E}_a[a] \geq t \\ 2 \left(\mathbb{E}_\xi[c] - \bar{c} \right) \int_t^{\bar{a}} \bar{F}(a)\, da > 3Var(a) \ , \ \textit{if } \mathbb{E}_a[a] < t. \end{cases} \quad (2.6)$$

Conditions in Proposition 2.3.1 involve three terms: $\mathbb{E}_\xi[c] - \bar{c}$ measures the effectiveness of cost reduction effect of cash hedging; $\int_{\underline{a}}^t F(a)\, da$ measures the likelihood of information

sharing to change the cash hedging decision and we call it as *information factor of hedging*; $Var(a)$ measures the market volatility. As we can see, there are typically three requirements that make (2.6) hold: (i). The cash hedging effect of unit cost reduction, $\mathbb{E}_\xi[c] - \bar{c}$, should be significant, ; (ii). Average market size $\mathbb{E}[a]$ is close to the cash hedging threshold t, i.e., $\int_{\underline{a}}^t F(a)\,da$ or $\int_t^{\bar{a}} F(a)\,da$ is big enough; (iii). Some but limited market volatility $Var(a)$.

The following corollary gives simpler necessary conditions of voluntary information sharing:

Corollary 2.3.1 *The following condition is necessary for voluntary information sharing to take place:*

$$\mathbb{E}_\xi[c] - \bar{c} > \frac{3Var(a)}{2\int_{\underline{a}}^{\mathbb{E}[a]} F(a)\,da} \ , \ \ if\ t \le \mathbb{E}[a] \ and \ \mathbb{E}_\xi[c] - \bar{c} > \frac{3Var(a)}{2\int_{\mathbb{E}[a]}^{\bar{a}} F(a)\,da} \ , \ \ if\ t > \mathbb{E}[a].$$

If potential market size, a, follows a symmetric distribution, the following distribution-free condition is necessary for voluntary information sharing:

$$\mathbb{E}_\xi[c] - \bar{c} > 3\sigma_0,$$

where σ_0 is the standard deviation of random variable a.

Now we use an example to help better demonstrate the structural requirements of (2.6).

Example 1 *(Scale up in β) Assume that the production cost function has form $c_\beta(\cdot) = \beta \cdot c(\cdot)$, where $\beta > 0$ is a constant and $c(\cdot)$ is a convex decreasing function. The potential market size has form $a_\beta = a + (\beta - 1) \cdot t$, where a is a random variable following distribution $F(\cdot)$ with support on $[\underline{a}, \bar{a}]$ and $t = \frac{\mathbb{E}_\xi[c^2] - \bar{c}^2}{2(\mathbb{E}_\xi[c] - \bar{c})}$. Additionally, we assume $\mathbb{E}[a] \ge t > \underline{a}$ and for any realization of c, we have $c < \underline{a}$. We notice that such assumption can be compatible if and only if $t > \max_\xi \{c(\xi)\}$. For example, $c(\xi) = (1 - \xi)^{3/2}$, which is convex decreasing and*

$\xi \sim U([0,1])$, then $t \approx 1.34 > 1 = \max_\xi \{c(\xi)\}$. We denote the distribution of a_β by $F_\beta(\cdot)$. Now condition (2.6) is written as

$$2\left(\mathbb{E}_\xi[c_\beta] - \bar{c}_\beta\right) \int_{\underline{a}+(\beta-1)t}^{\beta t} F_\beta(a_\beta)\, da_\beta > 3Var(a_\beta),$$

which is equivalent to

$$2\beta \cdot \left(\mathbb{E}_\xi[c] - \bar{c}\right) \int_{\underline{a}}^{t} F(a)\, da > 3Var(a),$$

therefore, condition (2.6) holds if and only if

$$\beta > \frac{3Var(a)}{2\left(\mathbb{E}_\xi[c] - \bar{c}\right) \cdot \int_{\underline{a}}^{t} F(a)\, da},$$

that is, when the scale-up parameter β is greater than some threshold, voluntary information sharing would take place. Specifically, in our example, we fix the volatility of market size and apply translation on the mean market size, which keeps terms $\int_{\underline{a}+(\beta-1)t}^{\beta t} F_\beta(a_\beta)\, da_\beta$ and $Var(a_\beta)$ constant, while the scaling-up increases the magnitude of cost reduction effect.

The implication of our equilibrium analysis has two folds. First, we show that the presence of cost risk from the supplier could motivate the retailer's voluntary information sharing even in the very simple bilateral supply chain. Second, it is straightforward to see that the supplier can never do worse by having more information in our game, therefore we identify a new reason why a firm would like to hedge. The presence of the hedge lever itself could help a supplier garner more information about the end market. In other words, a supplier's cash hedging strategy could be used as a natural bargain tool during the information sharing negotiation with the retailer.

2.3.3 Insightful Decomposition

Retailer's equilibrium profit difference between sharing and not sharing information is the compound product of retailer's information sharing behavior and supplier's cash hedging decisions. We conduct the following decomposition to demonstrate the intuitions of the value of information sharing:

$$\pi_R^{s*} - \pi_R^{ns*} = \overbrace{\left(\pi_R^{s*} - \widetilde{\pi}_R^s\right)}^{\Delta_1} - \overbrace{\left(\pi_R^{ns*} - \widetilde{\pi}_R^{ns}\right)}^{\Delta_2} + \overbrace{\left(\widetilde{\pi}_R^s - \widetilde{\pi}_R^{ns}\right)}^{\Delta_3}, \tag{2.7}$$

where $\widetilde{\pi}_R^s$ and $\widetilde{\pi}_R^{ns}$ are the retailer's equilibrium profits when sharing and not sharing information with the supplier not having cash hedging lever respectively. Therefore, Δ_1 and Δ_2 represent the effects of supplier's hedging decision on retailer's surplus with and without retailer's information sharing respectively. The last term, Δ_3, represents retailer's pure effect of information sharing when the supplier does not have hedging decisions. Terms Δ_1 and Δ_2 are the effects of hedging decision under given information sharing scheme, therefore $\Delta_1 - \Delta_2$ represents the *information rent on hedging effect*. Following the idea of profit function decomposition in [85], we can further decompose Δ_1 and Δ_2 into the *cost-reduction effect of hedging* and the *flexibility loss of hedging* respectively.

$$\Delta_1 = \mathbb{E}_\xi \left[\int_t^{\bar{a}} \left(\frac{(a - \bar{c})^2}{16b} - \frac{(a - c)^2}{16b} \right) dF(a) \right]$$

$$= \underbrace{\int_t^{\bar{a}} \frac{2\left(\mathbb{E}_\xi[c] - \bar{c}\right)a - \left(\mathbb{E}_\xi[c^2] - \bar{c}^2\right)}{16b} dF(a)}_{\text{cost reduction effect of hedging} > 0} \underbrace{- \frac{Var(c)}{16b} \cdot \overline{F}(t)}_{\text{flexibility loss of hedging} < 0}. \tag{2.8}$$

$$
\Delta_2 =
\begin{cases}
\mathbb{E}_\xi \left[\int_{\underline{a}}^{t} \left(\dfrac{\left(a - \frac{\mathbb{E}_a[a]+c}{2}\right)^2}{4b} - \dfrac{\left(a - \frac{\mathbb{E}_a[a]+\bar{c}}{2}\right)^2}{4b} \right) dF(a) \right], & \text{if } \mathbb{E}_a[a] \geq t, \\[3em]
0, & \text{if } \mathbb{E}_a[a] < t.
\end{cases}
$$

$$
=
\begin{cases}
\underbrace{\int_{\underline{a}}^{t} \dfrac{(\mathbb{E}_\xi[c] - \bar{c}) \left(((2a - \mathbb{E}_a[a]) - \mathbb{E}_\xi[c]) + ((2a - \mathbb{E}_a[a]) - \bar{c}) \right)}{16b} dF(a)}_{\text{cost reduction effect of hedging} > 0} \\
\qquad \underbrace{-\dfrac{Var(c)}{16b} \cdot F(t)}_{\text{flexibility loss of hedging} < 0}, & \text{if } \mathbb{E}_a[a] \geq t, \\[2em]
0, & \text{if } \mathbb{E}_a[a] < t.
\end{cases}
$$

$$(2.9)$$

In the above decomposition, the *cost reduction effect of hedging* is the expected profit difference between the retailer who has the hedging decision and the retailer who cannot hedge but also has to determine wholesale price before the realization of cash flow ξ given that the realized market size is big (here "big" refers to the market size that is bigger than hedging threshold t, and the expected profit here is in terms of the unnormalized conditional market size distribution). On the other hand, the *flexibility loss of hedging* is the expected profit difference between the retailer who cannot hedge and has to determine wholesale price before the realization of cash flow and the retailer who cannot hedge but is able to determine wholesale price after the realization of cash flow given that the realized market size is big. Correspondingly, we refer to the difference between the *cost reduction effect of hedging* in Δ_1 and Δ_2 as the *information rent on cost reduction effect of hedging* and the difference

between the *flexibility loss of hedging* in Δ_1 and Δ_2 as the *information rent on flexibility loss of hedging*. On the other hand, for Δ_3, we have

$$
\begin{aligned}
\Delta_3 &= \mathbb{E}_\xi \left[\int_{\underline{a}}^{\bar{a}} \left(\frac{(a-c)^2}{16b} - \frac{\left(a - \frac{\mathbb{E}_a[a]+c}{2}\right)^2}{4b} \right) dF(a) \right] \\
&= \int_{\underline{a}}^{\bar{a}} \left(\frac{(a - \mathbb{E}_\xi[c])^2}{16b} - \frac{((2a - \mathbb{E}_a[a]) - \mathbb{E}_\xi[c])^2}{16b} \right) dF(a) \\
&= \frac{\mathbb{E}_a\left[a^2 - (2a - \mathbb{E}_a[a])^2\right]}{16b} = \underbrace{-\frac{3Var(a)}{16b}}_{\text{wholesale cost of information sharing} < 0} .
\end{aligned} \tag{2.10}
$$

We have several observations from the above calculations (2.8), (2.9) and (2.10). First, we notice that given information sharing outcome, the value of hedging decision represented by Δ_1 and Δ_2 involve the distributions of both potential market size a and cash flow ξ and the cost function $c(\xi)$. On the other hand, the value of information sharing represented by Δ_3 only involves the volatility of the market, which is independent of the specific market distribution form and the firm's cost details. Besides, we notice that Δ_3 is always a loss for the retailer because additional market information gives the supplier more flexibility in setting the wholesale price in order to exploit more profits from the retailer. So we call Δ_3 as the *wholesale cost of information sharing*. Second, there are two positive terms in the above decomposition that may incentivize the retailer to share information, i.e., the *cost reduction effect of hedging* given information is shared and the negative *flexibility loss of hedging* given no information is shared (if $\mathbb{E}_a[a] \geq t$). If the market size a could be well characterized by a symmetric distribution, then we have clearer intuition on what drives the retailer's voluntary information sharing: When $\mathbb{E}_a[a] < t$, only in the scenario where the retailer shares information, the cost reduction effect of hedging and the flexibility loss of hedging are nontrivial, and it is straightforward to see that it is the *information rent on cost reduction effect of hedging* that may incentivize the retailer's information sharing. When

$\mathbb{E}_a[a] \geq t$ and the distribution of a is symmetric, for the *information rent on cost reduction effect of hedging*, we have

$$\int_t^{\bar{a}} \frac{(\mathbb{E}_\xi[c] - \bar{c})((a - \mathbb{E}_\xi[c]) + (a - \bar{c}))}{16b} dF(a)$$
$$- \int_{\underline{a}}^t \frac{(\mathbb{E}_\xi[c] - \bar{c})(((2a - \mathbb{E}_a[a]) - \mathbb{E}_\xi[c]) + ((2a - \mathbb{E}_a[a]) - \bar{c}))}{16b} dF(a)$$
$$\geq \int_t^{\bar{a}} \frac{(\mathbb{E}_\xi[c] - \bar{c})((t - \mathbb{E}_\xi[c]) + (t - \bar{c}))}{16b} dF(a)$$
$$- \int_{\underline{a}}^t \frac{(\mathbb{E}_\xi[c] - \bar{c})((t - \mathbb{E}_\xi[c]) + (t - \bar{c}))}{16b} dF(a) = 0.$$

The second inequality is because of the condition $\mathbb{E}_a[a] \geq t$, and the last equality is because of the symmetry of market size distribution. Therefore, the cost reduction effect of hedging is always magnified by additional information, which encourages the retailer to share the information. On the other hand, for the *information rent on flexibility loss of hedging*, we have

$$-\frac{Var(c)}{16b} \cdot \bar{F}(t) + \frac{Var(c)}{16b} \cdot F(t) = \frac{Var(c)}{16b} \cdot (2F(t) - 1) \leq \frac{Var(c)}{16b} \cdot (2F(\mathbb{E}_a[a]) - 1) = 0.$$

Therefore, the flexibility loss of hedging is also magnified by additional information which discourages the retailer from sharing the information. So in summary, when the *information rent on cost reduction effect of hedging* outweighs the *information rent on flexibility loss of hedging* and the *wholesale cost of information sharing*, the retailer would share the information with the supplier voluntarily.

As we have indicated on one hand, information sharing facilitates better cash hedging policies that helps to enhance system efficiency. On the other hand, information sharing increases the inefficiency by more severe double marginalization. It is possible to design surplus division mechanism to better align the supply chain. While simple information

sharing compensation contract cannot work in the traditional vertical information sharing setting of a bilateral supply chain (see [50]), we characterize the conditions of the existence of such contract by the following proposition.

Proposition 2.3.2 *Information sharing compensation contracts that can Pareto improve the supply chain exist if and only if*

$$\begin{cases} 2\left(\mathbb{E}_\xi\left[c\right]-\bar{c}\right)\displaystyle\int_{\underline{a}}^{t}F\left(a\right)da > Var\left(a\right) \ , \ \ if\ \mathbb{E}_a\left[a\right]\geq t \\ 2\left(\mathbb{E}_\xi\left[c\right]-\bar{c}\right)\displaystyle\int_{t}^{\bar{a}}\bar{F}\left(a\right)da > Var\left(a\right) \ , \ \ if\ \mathbb{E}_a\left[a\right]< t. \end{cases}$$

We notice that the conditions in the above proposition are strictly weaker than the conditions in Proposition 2.3.1, which indicates that when information sharing is not favorable for the retailer, there may still be an opportunity to induce information sharing and Pareto improve system performance by implementing compensation mechanism for information between the supplier and the retailer.

2.4 Competing Retailers with Demand Uncertainty Information

We now consider a two-tier supply chain with one supplier and n retailers engaged in Cournot competition. And denote the set of n retailers by $N = \{1, 2, \cdots, n\}$. Following the framework of [50], we assume the downstream demand curve is $p = a + \theta - Q$, where $\theta \sim N\left(0, \sigma_0^2\right)$. Each retailer can observe a private signal Y_i about θ and $Y_i = \theta + \epsilon_i$, where $\epsilon_i \sim N\left(0, \sigma^2\right)$. [1] Here σ_0 and σ are common knowledge. To avoid the pathological cases, here we assume that a is large enough compared with σ_0 and any realization of production cost

---

[1]To facilitate our analysis, we assume normally distributed market size and signals, which may violate our assumption $a > c$ in the previous section. The justification could be that both the supplier and the retailer commit to serving the market once their expected earning is positive (which can be guaranteed if the mean value of market size is big enough compared to the cost) even if in some extreme cases they may lose money. Such a relaxation does not change the main insights of our results.

c such that the probability of $a + \theta < c$ is negligible. Given θ, Y_i, $i \in N$, are independently distributed. Similar to the case above, the game consists of four stages: (i). Each retailer decides whether to publicly disclose her private market signal[2]; (ii). The private signals are realized and the supplier observes the signals, then determines whether to hedge the cash flow or not; (iii). The cash flow is realized, and in turn the production cost c is realized, then the supplier determines the wholesale price w; (iv). After knowing the wholesale price w, the retailers determine their order quantities q_i, $i \in N$, and then the profits are realized.

Remark. (Observable Information) *Like [50] and [51], for simplicity, our model assumes all the retailers can observe the publicly shared market information and the wholesale price. Retailers do not need to know the realized production cost. However, as we will see later, the supplier's equilibrium wholesale price decision is a linear combination of the realized production cost and the shared information. Since the shared information is public knowledge, all the retailers can perfectly infer the realized cost. Importantly, the wholesale price decision does not need to serve as a signal of competitors' private information (see [60]). We conjecture that information sharing with confidentiality should also lead to more information sharing in the equilibrium as indicated by [60]. However, information sharing with confidentiality would significantly complicate our analysis, while it is irrelevant to our main insights derived from the bilateral supply chain model.*

We regulate that if there are multiple equilibriums in the information-sharing game, we apply Pareto refinement rule.

Let K, $K \subseteq N$, denote the set of retailers who disclose their private signals and the supplier chooses to acquire their information. Here $|K| = k$. Using Bayesian rule, the post-distribution of θ given sample Y_j, $j \in K$, is $N(\mu_1(k), \sigma_1^2(k))$, where $\sigma_1^2(k) = \left(\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2/k} \right)^{-1}$ and $\mu_1(k) = \sigma_1^2(k) \sum_{j \in K} Y_j / \sigma^2 = \frac{1}{k+s} \sum_{j \in K} Y_j$, where $s = \sigma^2/\sigma_0^2$. Let $\Pi_S(k)$ denote the

---

[2]As pointed out in [60], [50] actually assumes a public information revelation. In this paper, we also study the problem under this benchmark. As shown in [60], signaling issues arise in the information-sharing game with confidentiality, which may give the retailers more incentive to share information.

supplier's expected equilibrium profit when there are k out of n retailers share information. Let $\Pi_R^S(k)$ and $\Pi_R^N(k)$ denote the expected equilibrium profit of retailer who shares and does not share information given that k out of $n-1$ other retailers share information respectively.

2.4.1 Equilibrium Analysis

We refer readers to Appendix A and C for the technical details of analysis. The following equations give the marginal utility of sharing information given $k-1$ out of the rest $n-1$ retailers share information:

$$
\Pi_R^S(k) - \Pi_R^N(k-1)
$$

$$
= \frac{1}{4(n+1)^2}\left[-\frac{\sigma_0^4 \sigma^2 \left(\left((2n+1)k - n\right)\sigma_0^2 + 2n\sigma^2 \right)\left(\left((2n+3)k + n\right)\sigma_0^2 + 2(n+2)\sigma^2 \right)}{\left(k\sigma_0^2 + \sigma^2\right)\left((k-1)\sigma_0^2 + \sigma^2\right)\left((n+k)\sigma_0^2 + 2\sigma^2\right)^2} \right.
$$

$$
\left. + 2(\mathbb{E}_\xi[c] - \bar{c})\left(\int_{-\infty}^{t-a} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right) \right], \quad k > 1,
$$

where $\sigma_k^2 = \frac{k\sigma_0^2}{k + \sigma^2/\sigma_0^2}$ and $\Phi(\cdot)$ is the cdf of standard normal distribution. And marginal utility of information sharing when none of the other retailers share information has a slightly different form:

$$
\Pi_R^S(1) - \Pi_R^N(0)
$$

$$
= \frac{1}{4(n+1)^2}\left(-\left(4\left(B_2^0\right)^2 - \left(A_1^1\right)^2\right)\left(\sigma_0^2 + \sigma^2\right) + 2\left(\mathbb{E}_\xi[c] - \bar{c}\right)\int_{-\infty}^{|t-a|} \Phi\left(\frac{x}{\sigma_1}\right) dx \right).
$$

We notice that all retailers hold information is an equilibrium if and only if $\Pi_R^S(1) - \Pi_R^N(0) \leq 0$. All retailers disclose information is an equilibrium if and only if $\Pi_R^S(n) - \Pi_R^N(n-1) > 0$. k out of n retailers share information is an equilibrium if and only if

$\Pi_R^S(k) - \Pi_R^N(k-1) > 0$ and $\Pi_R^S(k+1) - \Pi_R^N(k) \leq 0$, for $k = 2, \cdots, n-1$. One of the above three conditions must hold, therefore we have

Proposition 2.4.1 *The information-sharing game must admit a pure Nash equilibrium.*

Given the existence of the equilibrium, the following corollary answers how the system's equilibrium performance changes after building up the information-sharing channel.

Corollary 2.4.1 *If no-information-sharing is not a Nash equilibrium, i.e., $\Pi_R^S(1) > \Pi_R^N(0)$, there must exist an information sharing equilibrium that improves the whole system upon the no-information-sharing outcome in Pareto sense.*

Corollary 2.4.1 indicates that building up an information-sharing channel may not only generate nontrivial equilibrium outcomes, but can also benefit every agent in the system.

Next, we use the following two-retailer example to demonstrate the possibility of asymmetric equilibrium outcome.

Example 2 *(Scale up in β) We construct our example based upon Example 1 where the production cost function has form $c_\beta(\cdot) = \beta \cdot c(\cdot)$, where $\beta > 0$ is a constant and $c(\xi) = (1-\xi)^{3/2}$, which is convex decreasing and $\xi \sim U([0,1])$. The potential market size has form $a_\beta = a + (\beta - 1) \cdot t + \theta$, where $a = t - 1$, $t = \frac{\mathbb{E}_\xi[c^2] - \bar{c}^2}{2(\mathbb{E}_\xi[c] - \bar{c})} \approx 1.34$, $\theta \in N(0, \sigma_0^2)$ with $\sigma_0 = 1$ and signal accuracy $\sigma = 0.5$.* [3] *We know that the largest realization of $c_\beta(\cdot)$ is β, while $\mathbb{E}[a_\beta] \approx 1.34 \cdot \beta - 1$, therefore when β is larger than 15 or so, the largest realized production cost is at least 3 standard deviations smaller than market size, which implies that it is well justified to neglect the pathological cases where realized production cost exceeds the market size. According to our calculation above, given the example setting, we have that*

$$\Pi_R^S(1) - \Pi_R^N(0) = -2.873 + 0.118 \cdot \beta \ , \ \Pi_R^S(2) - \Pi_R^N(1) = -0.711 + 0.021 \cdot \beta,$$

---

[3]As we mentioned in the model setup, one caveat is that the support of the market size distribution is infinite which may lead to the pathological case where production cost exceeds the market size in the linear demand model, however, we argue that when the scale factor β is big enough, it is legit to neglect the pathological possibility and keep the setting of normal distributions.

both profit differences are linear in the scale factor β. As shown in Figure 2.1, we see that the information sharing equilibrium outcome transfers from (N, N) to (S, N) (or (N, S)) to (S, S) as β grows. For fixed market volatility, when scale factor β is small, the value of
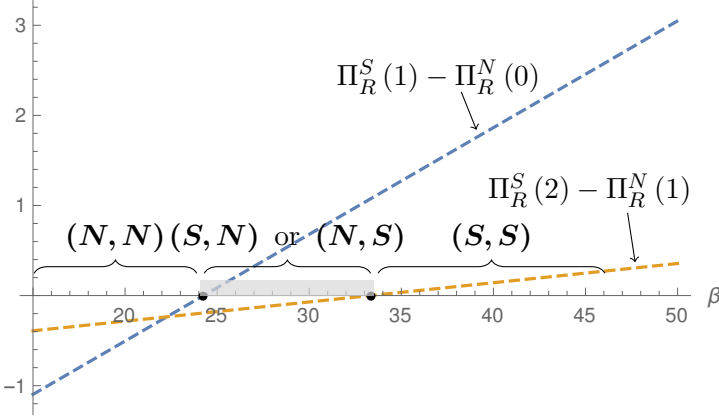
Figure 2.1.: Information Sharing Equilibrium Outcome

information is relatively salient, which implies that the wholesale price loss due to information sharing outweighs the potential cash hedging benefits. As β grows, the magnitude of production cost becomes large, and correspondingly the benefits of cash hedging led by information sharing starts to exceed the wholesale price loss, then more agents turn to sharing information.

Marginal Utility Decomposition and Distribution.

In this section, we show that how the marginal utility changes with one additional retailer sharing information could be decomposed following the decomposition idea introduced in Section 2.3.3. That helps us understand how the potential benefits of more information are

distributed among supply chain agents. We can rewrite the marginal utility difference of the retailer who is thinking about sharing information as follows:

$$\Pi_R^S(k) - \Pi_R^N(k-1) = \frac{1}{4(n+1)^2} \left(\Delta_I^{SN}(k) + \Delta_H(k) \right),$$

where

$$\Delta_H(k) = 2(\mathbb{E}_\xi[c] - \bar{c}) \left(\int_{-\infty}^{t-a} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right),$$

$$\Delta_I^{SN}(k) = -\frac{\sigma_0^4 \sigma^2 \left(\left((2n+1)k - n \right)\sigma_0^2 + 2n\sigma^2 \right) \left(\left((2n+3)k + n \right)\sigma_0^2 + 2(n+2)\sigma^2 \right)}{\left(k\sigma_0^2 + \sigma^2 \right) \left((k-1)\sigma_0^2 + \sigma^2 \right) \left((n+k)\sigma_0^2 + 2\sigma^2 \right)^2} < 0.$$

$\Delta_H(k)$ characterizes the *information rent on cost reduction effect of hedging* within which the term $\left(\int_{-\infty}^{t-a} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right)$ is the extension of the notion *information factor of hedging* introduced in the simple bilateral model. It is not difficult to show that $\Delta_H(k) > 0$. $\Delta_I^{SN}(k)$ on the other hand captures the *wholesale cost of information sharing* for the marginal retailer who changes her position on disclosing her information and it is the summation of the *direct effect* and *indirect (leakage) effect* of information sharing in [50]. For the retailers who have already shared their information, we have

$$\Pi_R^S(k) - \Pi_R^S(k-1) = \frac{1}{4(n+1)^2} \left(\Delta_I^{SS}(k) + \Delta_H(k) \right),$$

where $\Delta_I^{SS}(k)$ denotes the benefits of receiving additional information for the retailer who has already shared her information, which is always positive,

$$\Delta_I^{SS}(k) = \frac{\sigma_0^4 \sigma^2}{(k\sigma_0^2 + \sigma^2)((k-1)\sigma_0^2 + \sigma^2)} > 0.$$

And for the retailers who do not share information, we have

$$\Pi_R^N(k) - \Pi_R^N(k-1) = \frac{1}{4(n+1)^2}\left(\Delta_I^{NN}(k) + \Delta_H(k)\right)$$

$$= \frac{1}{4(n+1)^2}\left(\Delta_I^{SS}(k) + \left(\Delta_I^{NN}(k) - \Delta_I^{SS}(k)\right) + \Delta_H(k)\right),$$

where $\Delta_I^{NN}(k)$ denotes the net benefit for the retailer who never shares information. [4] For the supplier, we have a similar decomposition

$$\Pi_S(k) - \Pi_S(k-1) = \frac{n}{4(n+1)}\left(\Delta_I^S(k) + \Delta_H(k)\right),$$

where $\Delta_I^S(k)$ is the for the supplier

$$\Delta_I^S(k) = \frac{\sigma_0^4\sigma^2}{(k\sigma_0^2 + \sigma^2)((k-1)\sigma_0^2 + \sigma^2)}.$$

We find that the *information rent on cost reduction effect of hedging* has the same value among all retailers which is always $\frac{1}{n(n+1)}$ of the supplier's value. However, the *wholesale cost of information sharing* differs. Specially, $\Delta_I^{SS}(k)$ is always positive and it has the same form as $\Delta_I^S(k)$. This indicates that more information from other retailers help the retailers who have already shared information before and the supplier in the similar way (except for a factor $\frac{1}{n(n+1)}$), since the market information they possess always stays the same. $\Delta_I^{SN}(k)$ is always negative, because the retailer does not gain new knowledge of the market but forgoes her information advantage to the supplier and other competing retailers. And finally, $\Delta_I^{NN}(k)$ is indefinite. On one hand, those retailers who never share information gain better knowledge of the market from a newly shared piece of information as other retailers and the supplier (captured by $\Delta_I^{SS}(k)$), but on the other hand, $\Delta_I^{NN}(k) - \Delta_I^{SS} < 0$ represents

---

[4]$\Delta_I^{NN}(k) - \Delta_I^{SS}(k) = -\frac{4(n+1)^2\sigma_0^6\sigma^2\left(\left(2k^2(k+1)+2k(k+1)n+n^2\right)\sigma_0^6 + 2(k(4k+3)+2(k+1)n)\sigma_0^4\sigma^2 + (5+10k+2n)\sigma_0^2\sigma^4 + 4\sigma^6\right)}{\left((k-1)\sigma_0^2+\sigma^2\right)\left((n+k)\sigma_0^2+2\sigma^2\right)^2\left(k\sigma_0^2+\sigma^2\right)\left((n+k+1)\sigma_0^2+2\sigma^2\right)^2}.$

the profit loss due to the enhanced ability of other retailers and the supplier to infer her private information. $\Delta_I^{NN}(k)$ reflects the involved interplay between those two effects. But of course, this profit loss is not as much as the *wholesale cost* undergone by the retailer who directly shares her information. The above insights are summarized by ordering the various rents as follows:

Corollary 2.4.2 $\Delta_I^{SN}(k) < \Delta_I^{NN}(k) < \Delta_I^{SS}(k) \; \big(= \Delta_I^{S}(k) \big).$

And obviously, it is the supplier who always garners the majority of the benefit from the shared information.

Consumer Welfare.

Finally, we investigate the impact on customer surplus and social surplus when voluntary information sharing takes place. Customer surplus when k out of n retailers share information is defined as $CS(k) = \frac{1}{2}\mathbb{E}[Q^2]$. where Q is the total production quantity. Let $\Pi_{SC}(k)$ denote the total supply chain surplus when k out of n retailers share information, then the corresponding total social surplus could be defined as $W(k) = \Pi_{SC}(k) + CS(k)$ (see [50]). Unlike what documented in the previous literature, in our framework, the consumer surplus may increase in a non-trivial equilibrium (i.e., effective information sharing takes place),

Proposition 2.4.2 *When no-information-sharing is not a Nash equilibrium, there must exist an information sharing equilibrium where consumer surplus strictly improves upon the no-information-sharing outcome.*

Proposition 2.4.2 indicates that building up an information-sharing channel typically won't hurt and may lead to a strict win-win-win outcome.

2.5 Two Markets

In this section, we investigate the case when the supplier serves retailers selling into two separate markets, how the supplier's cash hedging behavior interacts with the retailers' information sharing incentives. The supplier's hedging is what links the two vertical channels for the markets. We index the two markets by m ($m \in \{1,2\}$) and for market m, the retailer m faces linear demand $p_m = a_m + \theta_m - q_m$, where for tractability, we assume the two markets follow two-point distribution:

$$P\left(\theta_1 = \delta, \theta_2 = \delta\right) = P\left(\theta_1 = -\delta, \theta_2 = -\delta\right) = \alpha,$$
$$P\left(\theta_1 = \delta, \theta_2 = -\delta\right) = P\left(\theta_1 = -\delta, \theta_2 = \delta\right) = \frac{1}{2} - \alpha,$$

and $\delta, \delta > 0$ and $\alpha \in \left[0, \frac{1}{2}\right]$. The correlation coefficient between the two markets is $4\alpha - 1$. Because we are interested in the effect of differentiated average market size, we do not consider the difference in market volatility between the two markets. Moreover, to focus on the tension generated from the two markets' difference and facilitate our analysis, we assume each retailer m can observe accurate condition θ_m in her own market.

We assume that the supplier can do price discrimination, i.e., provides different wholesale prices for different markets. There are potentially four different information sharing scenarios, we use the pair (m_1, m_2), $m_1, m_2 \in \{S, N\}$ to denote each scenario where S stands for information sharing and N stands for no information sharing.

Remark. **(Observable Information)** *Here we do not need the information sharing to be public, which means the shared information from a retailer will not be disclosed to the other. Actually, since the two retailers are in separate markets, and both have full knowledge of their own markets, there is also no need to know the other's market condition during the production process.*

2.5.1 Equilibrium Analysis

For the analysis details, please refer to Appendix B. We need to highlight that in this two-market game the supplier's hedging decision is driven by his expected belief over the expected average market size between the two markets upon the available information. In effect the market heterogeneity is not fully reflected by the supplier's hedging decision. We find that when $\frac{a_1+a_2}{2} \ge t+\delta$ ($\frac{a_1+a_2}{2} \le t-\delta$), the supplier would always (never) hedge, i.e., information sharing has no impact on the supplier's hedging decisions. We only focus on the cases where information sharing has impact on supplier's hedging, $2(t-\delta) < a_1 + a_2 < 2(t+\delta)$. Calculations give the following results:

**Proposition 2.5.1** *Nash equilibrium under Pareto refinement is summarized by Table 2.1.*

Table 2.1: Information Sharing Equilibrium Summary (w. Pareto Refinement)

	$A2:$ $2(t-\delta) < a_1 + a_2 <$ $2(t-2\alpha\delta)$	$A1:$ $2(t-2\alpha\delta) \le$ $a_1 + a_2 < 2t$	$B1:$ $2t \le a_1 + a_2 <$ $2(t+2\alpha\delta)$	$B2:$ $2(t+2\alpha\delta) \le$ $a_1 + a_2 < 2(t+\delta)$
(S,S)	$a_1, a_2 > \max\Big\{t -$ $\delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big), t -$ $\delta\Big(1-\frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)\Big\}$	$t - \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big) <$ $a_1, a_2 <$ $t + \delta\Big(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	$t - \delta\Big(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar c}\Big) <$ $a_1, a_2 <$ $t + \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	$a_1, a_2 < \min\Big\{t +$ $\delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big), t +$ $\delta\Big(1-\frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)\Big\}$
(S,N)	—	$a_1 >$ $t - \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$ $a_2 \ge$ $t + \delta\Big(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	$a_1 <$ $t + \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$ $a_2 \le$ $t - \delta\Big(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	—
(N,S)	—	$a_1 \ge$ $t + \delta\Big(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$ $a_2 >$ $t - \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	$a_1 \le$ $t - \delta\Big(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$ $a_2 <$ $t + \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	—
(N,N)	$a_1, a_2 \le$ $t - \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	$a_1, a_2 \le$ $t - \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	$a_1, a_2 \ge$ $t + \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$	$a_1, a_2 \ge$ $t + \delta\Big(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar c}\Big)$

The fact that the supplier's hedging decision somewhat neglects the heterogeneity across the markets implies that supply chain agents' preference over the hedging outcome may not be

aligned, which results in a new incentive for the retailers to disclose market information. Figure 2.2 depicts the equilibrium outcomes under different correlation parameter α and volatility δ pairs. Two high-level observations on the equilibrium outcomes are that

(a) $\alpha = 0.2$, $\delta = 0.5$

(b) $\alpha = 0.2$, $\delta = 1$

(c) $\alpha = 0.4$, $\delta = 0.5$

(d) $\alpha = 0.4$, $\delta = 1$
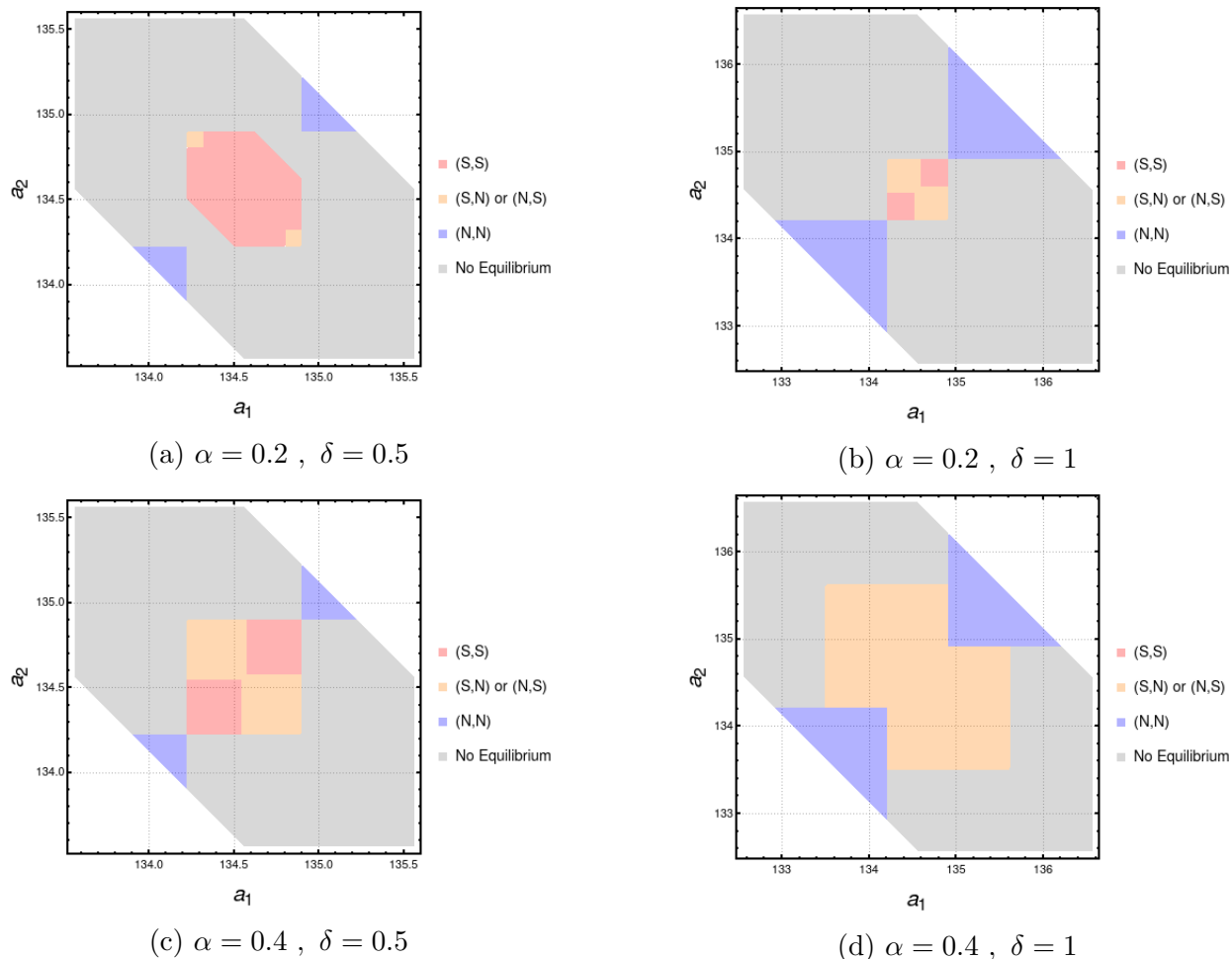
Note: The blank regions denote the cases where $a_1 + a_2 \geq 2(t + \delta)$ or $a_1 + a_2 \leq 2(t - \delta)$.

Figure 2.2.: Equilibrium Outcomes under Different Parameters ($c_\beta(\xi) = \beta \cdot (1 - \xi)^{3/2}$, $\xi \sim U[0, 1]$, $\beta = 100$)

- Information sharing can emerge as an equilibrium when both retailers' markets have moderate sizes. When both markets are either rather large or rather small, no information sharing will take place in the equilibrium. This is intuitive because the supplier's hedging decision is driven by the average size of the two markets. If both markets

72

have similar but extreme sizes (very large or very small), information sharing would have a rather limited influence on the supplier's hedging policy. Therefore, sharing information would mainly incur wholesale price exploitation from the supplier, which would discourage both retailers from sharing information.

- When the two markets have distinctly different sizes and are more negatively correlated, it is more likely that the information-sharing game admits no pure Nash equilibrium. The intuition is that for rather heterogeneous markets (in expected size and correlation), as we showed in equilibrium analysis (see Appendix B), retailer 1 would tend to follow (deviate from) retailer 2's information-sharing strategy, while retailer 2 would tend to deviate from (follow) retailer 1's information-sharing strategy, which can end up with no pure Nash equilibrium.

2.5.2   Does Information-Sharing Channel Improve Efficiency?

In this section, we try to answer the question of whether building up an information-sharing channel can help improve the system's efficiency. To do so, we compare the total system welfare under information-sharing outcomes (S, S) and (S, N) with the benchmark outcome (N, N) respectively. It is straightforward to see that the supplier can never do worse after building up the information sharing. The reason is the same as before because the supplier can simply choose to neglect any collected information and act as if in the scenario when there is no information sharing. But what is unclear under the duopolistic setting is that how the potential inefficiency introduced by the interactions between the two retailers would influence their welfare. We find that the information-sharing channel can backfire both for the retailers and even the whole system. Specifically, we have the following results:

Proposition 2.5.2 *The conditions under which building up information-sharing channel can improve system efficiency are summarized in Table 2.2. The conditions under which building up information-sharing channel can benefit the two retailers are summarized in Table 2.3.*

For technical details, please refer to Appendix B. The following gives a numerical example

Table 2.2: Conditions for Improved System Efficiency under Given Information Sharing Outcome

	$A2:$ $2(t-\delta) < a_1+a_2 <$ $2(t-2\alpha\delta)$	$A1:$ $2(t-2\alpha\delta) \le$ $a_1+a_2 < 2t$	$B1:$ $2t \le a_1+a_2 <$ $2(t+2\alpha\delta)$	$B2:$ $2(t+2\alpha\delta) \le$ $a_1+a_2 < 2(t+\delta)$
(S,S)	$a_1+a_2 >$ $2t-2\delta+\frac{\delta^2}{3\alpha(\mathbb{E}_\xi[c]-\bar{c})}$	$a_1+a_2 >$ $2t-2\delta+\frac{\delta^2}{3\alpha(\mathbb{E}_\xi[c]-\bar{c})}$	$a_1+a_2 <$ $2t+2\delta-\frac{\delta^2}{3\alpha(\mathbb{E}_\xi[c]-\bar{c})}$	$a_1+a_2 <$ $2t+2\delta-\frac{\delta^2}{3\alpha(\mathbb{E}_\xi[c]-\bar{c})}$
(S,N)	$-$	$a_1+a_2 > 2t-4\alpha\delta+$ $\frac{2\delta^2(1-4\alpha+8\alpha^2)}{3(\mathbb{E}_\xi[c]-\bar{c})}$	$a_1+a_2 < 2t+4\alpha\delta-$ $\frac{2\delta^2(1-4\alpha+8\alpha^2)}{3(\mathbb{E}_\xi[c]-\bar{c})}$	$-$
(N,S)	$-$	$a_1+a_2 > 2t-4\alpha\delta+$ $\frac{2\delta^2(1-4\alpha+8\alpha^2)}{3(\mathbb{E}_\xi[c]-\bar{c})}$	$a_1+a_2 < 2t+4\alpha\delta-$ $\frac{2\delta^2(1-4\alpha+8\alpha^2)}{3(\mathbb{E}_\xi[c]-\bar{c})}$	$-$

Table 2.3: Conditions for Improved Profits for Retailer 1 and 2 under Given Information Sharing Outcome

	$A2:$ $2(t-\delta) < a_1+a_2 <$ $2(t-2\alpha\delta)$	$A1:$ $2(t-2\alpha\delta) \le a_1+a_2 <$ $2t$	$B1:$ $2t \le a_1+a_2 <$ $2(t+2\alpha\delta)$	$B2:$ $2(t+2\alpha\delta) \le$ $a_1+a_2 < 2(t+\delta)$
(S,S)	$a_1 >$ $t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$	$a_1 >$ $t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$	$a_1 <$ $t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$	$a_1 <$ $t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$
	$a_2 >$ $t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$	$a_2 >$ $t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$	$a_2 <$ $t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$	$a_2 <$ $t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$
(S,N)	$-$	$a_1 > t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 < t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$-$
	$-$	$a_2 > t+$ $(1-4\alpha)\delta\left(1-\frac{3(1-4\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_2 < t-$ $(1-4\alpha)\delta\left(1-\frac{3(1-4\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$-$
(N,S)	$-$	$a_1 > t+$ $(1-4\alpha)\delta\left(1-\frac{3(1-4\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 < t-$ $(1-4\alpha)\delta\left(1-\frac{3(1-4\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$-$
	$-$	$a_2 > t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_2 < t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$-$

Note: For a given outcome $((S,S), (S,N)$ or $(N,S))$ and a total market size region $(A_1, A_2, B_1$ or $B_2)$, the upper box summarizes the conditions that the retailer 1 gets better off and the lower box summarizes the conditions that the retailer 2 gets better off.

where effective information sharing emerges in the equilibrium, while both retailers and the whole system get worse off.

Example 3 *(Negative Impact of Information Sharing) Similar as Example 1, we assume production cost function has form $c_\beta(\cdot) = \beta \cdot c(\cdot)$, where $\beta > 0$ is a scale-up factor and $c(\xi) = (1-\xi)^{3/2}$, $\xi \sim U([0,1])$. Let $\beta = 100$, then $t_\beta = \frac{\mathbb{E}_\xi[c_\beta^2] - \bar{c}_\beta^2}{2(\mathbb{E}_\xi[c_\beta] - \bar{c}_\beta)} = \beta \cdot \frac{\mathbb{E}_\xi[c^2] - \bar{c}^2}{2(\mathbb{E}_\xi[c] - \bar{c})} \approx 134$. Consider two markets with identical mean market sizes $a_1 = a_2 = 134.53$ and market shock $\delta = 1$. The two markets are negatively correlated with $\alpha = 0.03$. Then $269.01 = 2(t_\beta - 2\alpha\delta) < a_1 + a_2 < 2t_\beta = 269.13$ and $134.21 = t_\beta - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c_\beta] - \bar{c}_\beta}\right) < a_1, a_2 < t_\beta + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c_\beta] - \bar{c}_\beta}\right) = 135.41$, which means that (S, S) is a Nash equilibrium (see Table 2.1). On the other hand, we have $a_1 + a_2 < 2t_\beta - 2\delta + \frac{\delta^2}{3\alpha(\mathbb{E}_\xi[c_\beta] - \bar{c}_\beta)} = 269.52$, $a_1 < t_\beta - \delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c_\beta] - \bar{c}_\beta)}\right) = 144.33$ and $a_2 < t_\beta - \delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c_\beta] - \bar{c}_\beta)}\right) = 144.33$. From Table 2.2 and Table 2.3 we know that under the equilibrium (S, S), both the retailers and the whole system get worse off compared with the no-information-sharing outcome (N, N).*

One direct observation from Table 2.2 is that an information-sharing channel can improve the system efficiency when the average market size $(a_1 + a_2)/2$ is in a moderate range. On the other hand, each retailer's trade-off is between the cost of giving the supplier information advantage in wholesale price exploitation and the benefit of shaping the supplier's hedging decisions into a more favorable manner. However, the heterogeneity between the two retailers leads to their quite different preferences over the supplier's hedging policy. The result of such a difference is that they may fail to shape the supplier's hedging policy into their most ideal one, or manage to induce the right hedging policy but at a high cost of forgoing private information. Especially, when the two markets are highly negatively correlated (i.e., α is close to 0) and $\mathbb{E}_\xi[c] - \bar{c} > 12\delta$, we can find from Table 2.1 that in Case $(A2)$ and $(B2)$, there exist nontrivial regions where (S, S) emerges as the equilibrium. However, from Table 2.2, we can see that the region for (S, S) to benefit the system would not exist when α is close

to 0, which means that information-sharing channel induces nontrivial sharing decisions but hurts the whole system. This is true even when the two markets have the same average market size, which means that such inefficiency introduced by information-sharing channel can hurt both retailers at the same time as well as the system.

Furthermore, it is straightforward to see that when the retailers can directly observe the market realization and are monopolists in their own markets, the expected consumer surplus introduced in Section 2.4 in each market is equal to one half of the corresponding retailer's expected equilibrium profit, i.e., $CS_m = \frac{1}{2}\mathbb{E}[Q^2] = \frac{1}{2}\pi_{Rm}$. Therefore, our observations on retailers' profits under different information-sharing outcomes still hold on consumer surplus. We summarize the above results in the following proposition:

Proposition 2.5.3 *When the supplier serves multiple markets, the existence of information-sharing channel may hurt (all) the retailers, consumer surplus in (all) the market(s) and the whole system.*

This result presents some quite different insights from Corollary 2.4.1 and [50]. When a single supplier serves multiple markets, nontrivial information sharing may still emerge in equilibriums in our cash-hedging-information-sharing game, but there may not exist a simple information compensation scheme that achieves win-win-win outcome for the system. The above are still true when Stackelberg game is introduced as below.

2.5.3 Stackelberg Information-Sharing Game

The potential nonexistence of Nash equilibrium in the simultaneous information-sharing game implies that in practice a Stackelberg game sequence may need to be introduced. Calculations give the following results:

Proposition 2.5.4 *The equilibrium outcomes under the Stackelberg game are summarized by Table 2.4 and Table 2.5. The surplus improvement conditions keep the same as Proposition 2.5.2.*

Table 2.4: Stackelberg Information Sharing Equilibrium Summary

	$A2:$ $2(t-\delta) < a_1+a_2 <$ $2(t-2\alpha\delta)$	$A1:$ $2(t-2\alpha\delta) \le a_1+a_2 <$ $2t$	$B1:$ $2t \le a_1+a_2 <$ $2(t+2\alpha\delta)$	$B2:$ $2(t+2\alpha\delta) \le$ $a_1+a_2 < 2(t+\delta)$
	$\mathbb{E}_\xi[c] - \bar{c} \ge (12\alpha + \frac{3}{2})\delta$			
(S,S)	$a_1 >$ $t-\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 >$ $t-\delta\left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 >$ $t-\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 < t-\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ or $a_1 < t+\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t-\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) \le$ $a_2 < t+\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 <$ $t+\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 \ge t+\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ or $a_1 > t-\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t-\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) <$ $a_2 < t+\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 <$ $t+\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 <$ $t+\delta\left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$
(S,N)	–	$a_2 \ge t+\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_2 \le t-\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	–
(N,S)	–	$a_1 \ge t+\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t-\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) \le$ $a_2 < t+\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 \le t-\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t-\delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) <$ $a_2 < t+\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	–
(N,N)	$a_1 \le$ $t-\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ or $a_2 \le$ $t-\delta\left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 \le$ $t-\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 < t-\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 \ge$ $t+\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 \ge t+\delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 \ge$ $t+\delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ or $a_2 \ge$ $t+\delta\left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$

For details, we refer readers to the Stackelberg discussions in Appendix C. We notice that in our framework, the interactions between the two retailers only take place in the information sharing step, therefore under a given information sharing outcome, retailers' optimal profits keep the same forms as their counterparts under the simultaneous information-sharing game. This is why Table 2.2 and 2.3 still hold under the Stackelberg game.

Figure 2.3 is the analog to Figure 2.2 with the same production cost structure and parameters under the Stackelberg game. Table 2.4 and 2.5 suggest that information sharing still emerges in the equilibrium when the average market size is moderate and similar to

Table 2.5: Stackelberg Information Sharing Equilibrium Summary

	A2: $2(t-\delta) < a_1+a_2 < 2(t-2\alpha\delta)$	A1: $2(t-2\alpha\delta) \leq a_1+a_2 < 2t$	B1: $2t \leq a_1+a_2 < 2(t+2\alpha\delta)$	B2: $2(t+2\alpha\delta) \leq a_1+a_2 < 2(t+\delta)$
	$\mathbb{E}_\xi[c]-\bar{c} < \left(12\alpha+\frac{3}{2}\right)\delta$			
(S,S)	$a_1 > t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 > t-\delta\left(1-\frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 > t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 < t+\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 < t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 > t-\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 < t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 < t+\delta\left(1-\frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$
(S,N)	–	$a_1 > t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t+\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) \leq$ $a_2 \leq t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ or $\mathbb{E}_\xi[c]-\bar{c} > 12\alpha\delta$ $a_2 > t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 < t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) \leq$ $a_2 \leq t-\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ or $\mathbb{E}_\xi[c]-\bar{c} > 12\alpha\delta$ $a_2 < t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	–
(N,S)	–	$\mathbb{E}_\xi[c]-\bar{c} \leq 12\alpha\delta$ $a_2 > t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$\mathbb{E}_\xi[c]-\bar{c} \leq 12\alpha\delta$ $a_2 < t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	–
(N,N)	$a_1 \leq t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ or $a_2 \leq t-\delta\left(1-\frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 \leq$ $t-\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 < t+\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ or $a_1 \leq t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t+\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) \leq$ $a_2 \leq t-\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 \geq$ $t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ $a_2 > t-\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ or $a_1 \geq t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ $t+\delta\left(1-\frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) \leq$ $a_2 \leq t-\delta\left(1-\frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$	$a_1 \geq$ $t+\delta\left(1-\frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ or $a_2 \geq t+\delta\left(1-\frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$

the simultaneous game, effective information sharing is more likely to happen when the two markets are more positively correlated.

Specifically, we have the following observations:

- Different from the simultaneous game, complete $((S,S))$ or partial $((S,N), (N,S))$ information-sharing outcomes would emerge as equilibrium outcomes when the difference between the two markets' average sizes is large. The insights follow the similar logic introduced in simultaneous game analysis: for highly heterogeneous retailers, they are more likely to have opposite interests towards supplier's hedging decision, which gives them the incentive to conduct more active information-sharing policies in order to twist supplier's hedging decisions into a more favorable manner or to offset the neg-

(a) $\alpha = 0.2$, $\delta = 0.5$

(b) $\alpha = 0.2$, $\delta = 1$

(c) $\alpha = 0.4$, $\delta = 0.5$

(d) $\alpha = 0.4$, $\delta = 1$
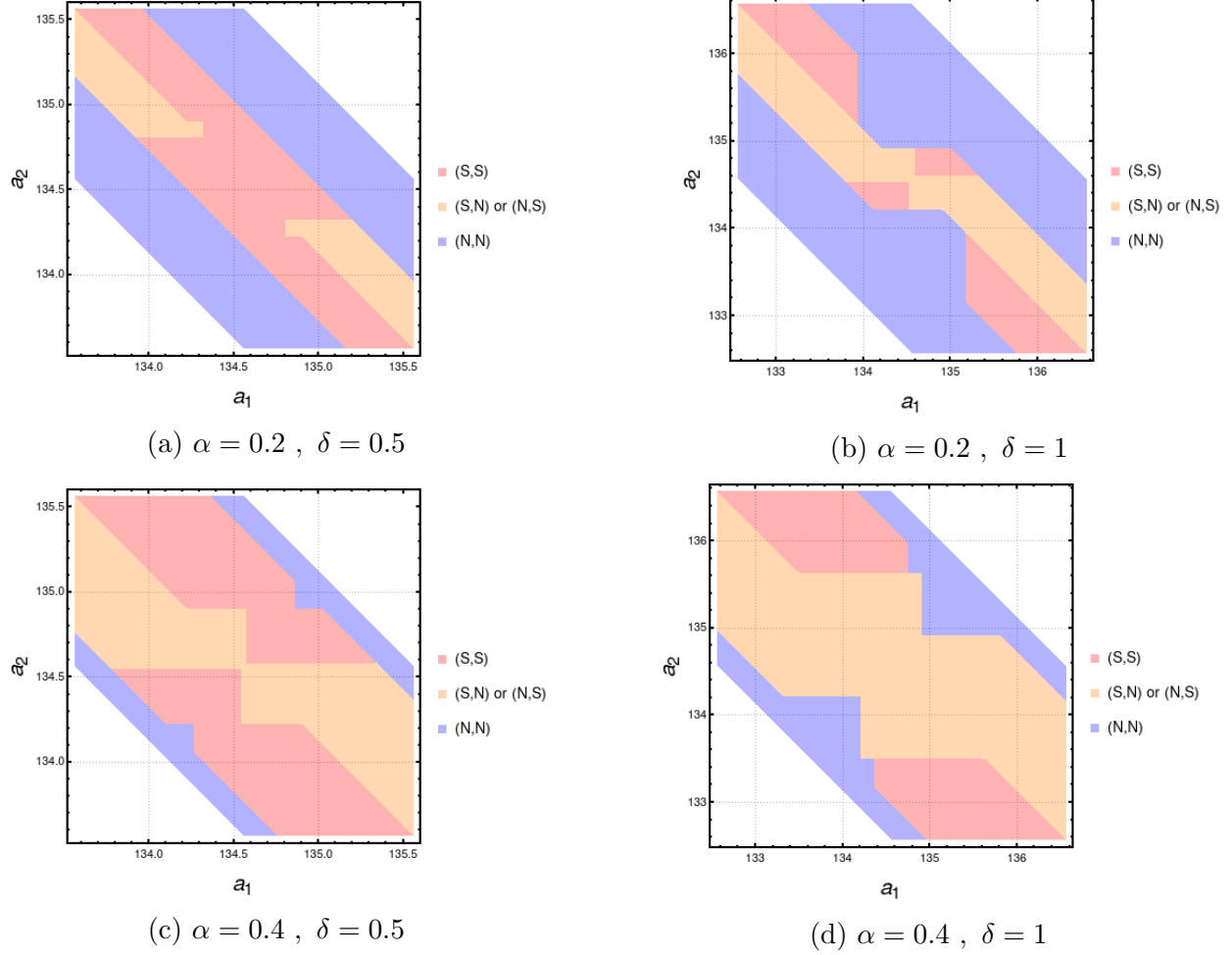
Figure 2.3.: Stackelberg Equilibrium Outcomes under Different Parameters ($c_\beta(\xi) = \beta \cdot (1-\xi)^{3/2}$, $\xi \sim U[0,1]$, $\beta = 100$)

ative impact of the competing retailer's information-sharing effort in shaping supplier's hedging decision.

- Information sharing is more likely to take place in the equilibrium when the two markets are more positively correlated. The intuition is that the supplier's cash hedging decision is determined by the average size of the two markets. If the two markets are highly negatively correlated, the market shocks would largely cancel out each other (which is especially true as we assumed that the shock magnitudes of the two markets are the

same) which means that the average size between the two markets under any realized market conditions would be similar to the expected average market size when there is no information sharing. Therefore, neither of the retailers would have enough incentive to disclose information.

Similar to the simultaneous game analysis, we are also interested in the performance implications of implementing an information-sharing channel. We end this section by the following numerical example which shows that even when Stackelberg sequence is introduced, both the retailers and the whole system can get worse off. Especially, retailer 1 as the Stackelberg leader still can get worse off.

Example 4 *(Negative Impact of Information Sharing under Stackelberg Game) We keep the parameter settings the same as Example 3 except that $a_1 = 134.25$ and $a_2 = 134.80$. Then $\mathbb{E}_\xi\left[c_\beta\right] - \bar{c}_\beta = 4.64 > (12\alpha + 3/2)\,\delta = 1.86$, $269.01 = 2\left(t_\beta - 2\alpha\delta\right) < a_1 + a2 < 2t = 269.13$ and $a_1 < t_\beta + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi\left[c_\beta\right] - \bar{c}_\beta}\right) = 135.41$, $134.21 = t_\beta - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi\left[c_\beta\right] - \bar{c}_\beta}\right) < a_2 < t_\beta + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi\left[c_\beta\right] - \bar{c}_\beta}\right) = 135.41$, which means that (S, S) is a Nash equilibrium (see Table 2.4 and 2.5). On the other hand, we have $a_1 + a_2 < 2t_\beta - 2\delta + \frac{\delta^2}{3\alpha\left(\mathbb{E}_\xi\left[c_\beta\right] - \bar{c}_\beta\right)} = 269.52$, $a_1 < t_\beta - \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi\left[c_\beta\right] - \bar{c}_\beta\right)}\right) = 144.33$ and $a_2 < t_\beta - \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi\left[c_\beta\right] - \bar{c}_\beta\right)}\right) = 144.33$. From Table 2.2 and Table 2.3 we know that under the equilibrium (S, S), both the retailers and the whole system get worse off compared with the no-information-sharing outcome (N, N).*

Example 4 shows a scenario where the whole system and both retailers get worse off simultaneous, which typically only happens in the marginal parameter regions. It is more common to end up in a case where one of the retailer gets worse off in the equilibrium. For example, when $\mathbb{E}_\xi\left[c\right] - \bar{c} \geq \left(12\alpha + \frac{3}{2}\right)\delta$, $2\left(t - 2\alpha\delta\right) \leq a_1 + a_2 < 2t$ and $a_2 \geq t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi\left[c\right] - \bar{c}}\right)$, according to Table 2.4 and 2.5, the Stackelberg equilibrium is (S, N). Then from Table 2.3, we know that if only a_2 is big enough, we can find an a_1 such that $a_1 \leq t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi\left[c\right] - \bar{c}}\right)$,

which is the case where the Stackelberg leader, retailer 1, gets worse off in the equilibrium, (S, N).

In summary, our analysis on the two-market model consolidates our insights derived from the single-market model before. That is, on one hand, the supplier's cash hedging decision can provide incentive for retailers' voluntary information sharing. And we need to highlight that the mechanism of how the incentive takes place relies on the supply chain structure. On the other hand, clearly the supplier can never do worse by having more information, we identify a new reason why the supplier should consider cash hedging, which is to help the negotiation with downstream buyers on market information sharing.

2.6   Conclusion

Informational operations and financial risk management are two critical parts of business operations. Extensive research has been done in both areas, while little is known about the implications of the interplay between the two operational levers. Our work is the first paper that aims to understand such interaction via a game theoretical framework.

Classical literature indicates that effective horizontal information sharing and vertical market information sharing on top of the widely used wholesale price contracts in Cournot competition are unlikely to achieve. We first show that even in the simple bilateral supply chain linked via wholesale price contract, the supplier's cash hedging decision can induce the retailer's voluntary information sharing. We formulate a profit decomposition framework to explain the intuition of the results. We identify three driving forces, i.e., *information rent on cost reduction effect of hedging*, which is beneficial to the retailer; *information rent on flexibility loss of hedging* and *wholesale cost of information sharing*, both of which are unfavorable to the retailer. Therefore when the first benefit outweighs the latter two costs, voluntary information sharing would take place. This is more likely to happen when expected

cost reduction after cash hedging is significant, expected market size is around some threshold such that the supplier's cash hedging policy is largely driven by the realization of market size, and the volatility of market size is limited. We also identify this as a new reason why a supplier should consider hedging and utilize the presence of his hedging option to achieve information sharing agreement with his retailers. Such intuition continues to apply in the general setting where there are multiple Cournot-competing retailers. Two major findings for the multiple homogeneous retailers setting are: asymmetric information outcomes can emerge as an equilibrium even though all the retailers are ex-ante homogeneous; building up a public information-sharing channel generally won't hurt and sometimes can Pareto improve the system and the consumer welfare.

However, we later find that some of the insights are dependent on the supply chain structure, and in particular, multiple separate markets with heterogeneous retailers. We discuss the situation where a single supplier serves two separate but correlated markets. We find that the supplier's hedging decision is dependent on the average size of the two markets. We identify that the heterogeneity across the average market sizes and the random shock correlation between the two markets play significant roles in shaping the equilibrium outcomes. Specially, we find that when market size heterogeneity is more significant, and the market shock correlation is more negative, the information-sharing game is more likely to admit no pure Nash equilibrium. For this case, the Stackelberg sequence may be introduced to guarantee a stable outcome. Besides, we find that having a public information-sharing channel may not be beneficial for retailers or even the system as a whole, while the supplier can never be worse off. This is because the retailers in heterogeneous markets tend to have different interests in supplier's hedging decisions. Therefore the competition between different markets may introduce significant inefficiency, which could lead to unfavorable information-sharing outcomes from retailers' standpoint. This implies that although voluntary information shar-

ing can still emerge as equilibrium in such a system, there may not exist a simple lump-sum compensation that achieves a win-win outcome.

Our paper serves as a first try to investigate firms' information policy and financial risk management policy along the supply chain in a unified framework, and it generates managerial insights that reveal highly nontrivial interactions between the two operational tools. Some insights in this paper remain to be supported and verified by future empirical study, but our work opens the door to this intersection area that got little attention and understanding from operations academia before.

3. Display Optimization under the Multinomial Logit Choice Model: Balancing Revenue and Customer Satisfaction

3.1 Introduction

The assortment optimization problem has come to be one of the well-studied problems in the field of revenue management. At a high level, this now classic problem considers a retailer or platform who is tasked with selecting a revenue-maximizing subset (assortment) of products to display to arriving customers. Over the years, this standard problem blueprint has been adapted and reshaped in many ways so as to capture a multitude of retailing settings ranging from traditional brick and mortar shelf offerings to product recommendation displays in e-commerce settings [86, Chap. 5].

As noted above, the focus for the many flavors of the assortment problem considered in the literature has almost exclusively been to uncover profitable assortments, irrespective of the fact that the revenue-maximizing assortment could be undesirable to the vast majority of customers. [1] To the best of our knowledge, the work of [87] is the first paper to consider an assortment setting where the goal is to choose an assortment that garners a large revenue, while also delivering high expected utility to arriving customers. The intent of such a framework is to incorporate customer considerations into the assortment planning decision, rather than focusing exclusively on revenue. This effort, for example, could serve to increase the chance for repeat customers, as customers who witness an assortment from which they draw a high utility are more likely to return to the platform or store for future purchases.

---

[1]For example, it is easy to construct instances of the assortment problem where the optimal assortment consists of a few high-revenue products that are purchased infrequently.

In this paper, we adapt the revenue-utility trade-off framework of [87] to a setting where the retailer must choose assortments to offer over T stages. Each arriving customer is assumed to sequentially progress through the stages, stopping her search at some stage $t \in [T]$ [2]. Then, from amongst all the products she has viewed across stages $1, \ldots t$, she is assumed to make a multinomial-logit-based purchasing choice. Similar to the existing works of [88] and [89], who both utilize a near-identical framework to capture choice behavior in e-commerce applications, our goal is also to broadly model an e-commerce setting where customers sequentially browse pages of displayed product recommendations. The goal of the retailer is to choose assortments to offer across each stage with the intention to maximize expected revenue. However, the twist in our setting is that we enforce stage-specific constraints meant to ensure that the assortments offered meet a minimum level of "desirability". In Section 3.1.1, we formalize both the structure of these constraints, as well as the notion of a desirable assortment. Finally, we consider the problem extension with additional cardinality constraints. With modifications of the previous PTAS, we construct a new PTAS to find a proxy that meets the cardinality constraints while slightly violating the desirability constraints.

3.1.1 Problem formulation

We consider a platform that has access to n items, indexed by the set $[n] = \{1, \ldots, n\}$, where the revenue earned from selling a single unit of item $i \in [n]$ is denoted r_i. These items are offered across T stages (or pages) of product displays, where we use $A_t \subseteq [n]$ to denote the set of item displayed in stage $t \in [T]$. Each item can be offered in at most one stage, and so the assortments $A_1, \ldots, A_T$ must be pairwise disjoint. We use $\mathcal{A} = (A_1, \ldots, A_T)$ to denote any such pairwise disjoint sequence of assortments.

---

[2]Throughout the paper, for $x \in \mathbb{Z}_+$, we use the notation $[x] = \{1, \ldots, x\}$ and $[x]_0 = \{0, 1, \ldots, x\}$

Purchasing dynamics. Beginning at the first stage, arriving customers are assumed to browse the stages sequentially, stopping at stage $t \in [T]$ with probability λ_t. These stopping probabilities will henceforth be referred to as stage weights. Each customer's consideration set consists of only the products she has viewed in this initial browsing phase. So, for example, under the offered sequence $\mathcal{A} = (A_1, \ldots, A_T)$, a customer who browses only the first two stages has a consideration set of $A_1 \cup A_2$. A type-t customer is called of patience level t, and customers are categorized into K classes based upon their patience levels ($K \le T$). A class-k group consists of customers of patience levels $[t_k : t_{k+1}) = \{t_k, t_k + 1, \cdots, t_{k+1} - 1\}$, $k \in [K]$, $t_1 = 1$ and $t_{K+1} = T + 1$. Finally, from among the products considered, the customer makes a purchasing decision according to a multinomial logit (MNL) choice model ([90]). Accordingly, we use w_i to denote the weight associated with item $i \in [n]$ and $w(S) = \sum_{i \in S} w_i$ to denote the total weight of assortment $S \subseteq [n]$. As dictated by the MNL model, a customer with consideration set $S \subseteq [n]$ selects item $i \in S$ with probability $\frac{w_i}{1 + w(S)}$.

The assortment problem. We begin by noting that, under the purchasing dynamics described above, any sequence of pairwise disjoint assortments $\mathcal{A} = (A_1, \ldots, A_T)$ can be equivalently represented as a nested sequence of consideration sets $\mathcal{S} = (S_1, \ldots, S_T)$, where $S_t = \biguplus_{\tau \le t} A_\tau$. This vantage point turns out to be a more convenient way to express the decisions of the platform, and hence for the remainder of the paper, we focus on selecting a sequence of assortments $\mathcal{S} = (S_1, \ldots, S_T)$ that satisfy $S_1 \subseteq \ldots \subseteq S_T$. In other words, we assume that the stage-t assortment decision is to select the subset of products viewed by customers who stop browsing at stage t, rather than the choice of the set of items to offer at stage t. Given this translated view of the platform's decision, the expected revenue earned from offering the sequence of assortments $\mathcal{S} = (S_1, \ldots, S_T)$ is given by

$$\mathcal{R}(\mathcal{S}) = \sum_{t \in [T]} \lambda_t \cdot \left(\sum_{i \in S_t} \frac{\rho_i}{1 + w\left(S_t\right)} \right),$$

86

where $\rho_i = r_i w_i$. Additionally, let $cs\,(w(S_t))$ denote the "customer satisfaction" of assortment S_t, and we have the following two assumption on $cs\,(\cdot)$,

Assumption 3.1.1 $cs'\,(\cdot) \geq 0$.

Assumption 3.1.2 $cs\,(c \cdot x) \geq c \cdot cs\,(x)$, $\forall 0 \leq c \leq 1$. $cs\,(c \cdot x) \leq c \cdot cs\,(x)$, $\forall c \geq 1$.

The chosen sequence of assortments $\mathcal{S}$ must abide by class-dependent lower bounds of the expected "customer satisfaction" $\alpha_1, \cdots, \alpha_K$, i.e., $\sum_{t \in [t_k:t_{k+1})} \lambda_t \cdot cs\,(w(S_t)) \geq \alpha_k$, $k \in [K]$, here for exposition's simplicity, we drop the weight summation $\sum_{i \in [t_k:t_{k+1}]} \lambda_i$ in the denominator when we calculate the expectation. We place no restrictions on the sequence of lower bounds $\alpha_1, \cdots, \alpha_K$, nor do we place any restrictions on the stage weights $\lambda_1, \ldots, \lambda_T$. Moving forward, we use $\mathcal{F}_K = \{(S_1, \ldots, S_T) : S_t \subseteq [n], \sum_{t \in [t_k:t_{k+1})} \lambda_t \cdot cs\,(w(S_t)) \geq \alpha_k \; \forall \; k \in [K], S_1 \subseteq \ldots \subseteq S_T\}$ to denote all feasible sequences of assortments. Combining everything, our assortment problem of interest can be formulated as follows

$$\max_{\mathcal{S} \in \mathcal{F}_K} \mathcal{R}(\mathcal{S}), \qquad\qquad \text{(DISPLAY-OPT-}K\text{)}$$

whose optimal solution is denoted as $\mathcal{S}^* = (S_1^*, \ldots, S_T^*)$.

Distinguishing feature. The main differentiating element of our setting is the addition of the customer satisfaction constraints for each class $k \in [K]$. There are two main forms of $cs\,(\cdot)$ function with practical interpretations, which help highlight our intention to consider a framework where the platform wishes to balance its need to offer profitable assortments (those with high expected revenue) with the dueling desire to offer assortments that are "well-liked". These two forms are formalized below, where in both cases, we view our problem framework from the e-commerce-based lens discussed previously.

1. <u>Purchase likelihood:</u> $cs\,(w(S)) = \frac{w(S)}{1+w(S)}$. The platform would like to ensure that, as customers progress farther through the displayed set of products, a minimal threshold

is met for the likelihood that a purchase is made. This idea can be captured by enacting a lower bound α_k on the purchase probability of each customer class. More precisely, we enforce that $\sum_{t \in [t_k : t_{k+1})} \lambda_t \cdot \frac{w(S_t)}{1 + w(S_t)} \geq \alpha_k$ for each class $k \in [K]$.

2. <u>Expected utility:</u> $cs\,(w(S)) = \ln(1 + w(S)) + \gamma$. Our second formulation views these constraints through the lens of customer utilities. Specifically, under MNL preferences, the expected utility derived from an assortment S is $\ln(1 + w(S)) + \gamma$, where γ is the Euler-Mascheroni constant ([87]). Hence, constraints of the form $\sum_{t \in [t_k : t_{k+1})} \lambda_t \cdot (\ln(1 + w(S_t)) + \gamma) \geq \alpha_k$ enforce a lower bound on the expected utility of class-k customers.

Note that in the special case where $K = T$, i.e., each class k only contains one type of customers, the above two scenarios of constraints can be formulated into the same stage-dependent form: $w\,(S_t) \geq W_t, t \in [T]$, where W_t can be assumed to be monotone in t without loss of generality.

3.1.2 Contributions

Below, we provide a high-level overview of our two algorithmic results.

The PTAS. Our main algorithmic contribution is the development of a polynomial time approximation scheme (PTAS) for DISPLAY-OPT-K. More formally, for any fixed $\epsilon > 0$, we develop a polynomial-time algorithm that returns a sequence of assortments whose expected revenue is at least $(1 - \epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$. Our PTAS turns out to be a best-case result in light of the fact that we also show DISPLAY-OPT-K to be strongly NP-Hard, thus ruling out the possibility of an FPTAS.

Our approach begins by partitioning the items based on their respective weights, grouping items whose weights differ by at most a factor of $1 + \epsilon$. This partitioning of the items into

so-called weight classes, allows for a more granular view of any assortment as the set of items offered from each weight class. With this notion in mind, we show how to construct a "proxy" optimal assortment $\hat{\mathcal{S}} \in \mathcal{F}$, whose expected revenue satisfies $\mathcal{R}(\hat{\mathcal{S}}) \geq (1-\epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$, and more importantly, whose sequence of assortments $\hat{S}_1, \ldots, \hat{S}_T$ span the weight classes in a succinct way. Specifically, under $\hat{\mathcal{S}}$, the set of items offered from each weight class are either (i) added sequentially from largest to smallest ρ-value or (ii) match those added by $\mathcal{S}^*$ up to the first $\frac{1}{\epsilon}$ items, after which the left-over items from the particular class are again sequentially added by descending ρ-value. Furthermore, we construct $\hat{S}$ so that in each stage, the number of classes assigned this second categorization is never more than $\frac{2}{\epsilon}$. In this way, we establish that, in any stage, there are never more than $\frac{2}{\epsilon^2}$ products that disobey the ρ-order of particular class. Exploiting this special structure, we develop a dynamic program that approximately recovers $\hat{\mathcal{S}}$, incurring only an ϵ-loss in expected revenue in the process.

3.1.3   Related Literature

In what follows, we detail the past work that most closely resembles ours. With this summary, we hope to highlight the fact that our problem setting is a natural extension of quite a few well-studied frameworks. However, to the best of our knowledge, we see no straightforward way to directly apply the ideas in the papers summarized below to our setting.

The MNL-based assortment problem.   There are a multitude of papers that consider a single stage variant of our problem, which has come to be known as the assortment optimization problem. When the offered assortment is unconstrained, the seminal result of [91] shows that the optimal assortment consists of all products priced above a certain threshold. [92] consider a cardinality constrained variant of the assortment problem, which they show admits an optimal algorithm whose running time scales quadratically in the number of products. [93]

show that the addition of a simple knapsack constraint renders the MNL-based assortment problem NP-Hard. Nonetheless, they provide a fully polynomial time approximation scheme (FPTAS) for this problem, which extends to settings where choice is governed by a mixed-MNL model with $O(1)$ customer segments. Finally, to the best of our knowledge, [87] are the first to consider a variant of the assortment problem where the sole objective is not merely to maximize expected revenue, but also has customer-based utility considerations. Unlike our work, where this trade-off is capture through a set of constraints, [87] model this trade-off in the objective function, which has both expected revenue and expected utility terms. They provide a parametric linear-programming-based approach to solve their problem optimally.

Assortment over time. The original assortment over time problem, as conceived by [94], considers a variant of our problem under a general choice model, where the weight constraints are replaced with constraints enforcing that at most one product can be added in each stage. Hence, when choice is governed by an MNL model, their objective function precisely matches ours, as seen in DISPLAY-OPT-K. Additionally, [94] impose that the stage weights are uniform, i.e. $\lambda_1 = \ldots = \lambda_T$. In this setting, the authors provide a $\frac{1}{2\alpha}$-approximation, which assumes black-box access to a $\frac{1}{\alpha}$-approximation for the single-stage cardinality constrained assortment problem under the presumed choice model. [89] extend this problem landscape by allowing for any collection of stage weights $\{\lambda_t\}_{t \in [t]}$ that follow a new better than used in expectation (NBUE) distribution, while also enforcing a general cardinality constraint that allows up to C to products to be added in each stage. In this more general setting, the authors provide a $\frac{\pi^2}{6\alpha}$-approximation, where the $\frac{1}{\alpha}$ term has the same interpretation as noted previously. Finally, [88] provide a polynomial time approximation scheme (PTAS) to this extended version of the assortment over time problem for arbitrary stage weights, however, their approach only caters to the setting when choice is governed by an MNL model.

The incremental knapsack problem. The incremental knapsack problem considers a T-stage setting identical to ours, except that the objective function in DISPLAY-OPT-K is replaced with $\mathcal{R}(S) = \sum_{t \in [T]} \lambda_t \cdot (\sum_{i \in S_t} r_i)$, and the inequality in the weight constraints are flipped so that a sequence of assortments $S_1 \subseteq S_2 \subseteq \ldots \subseteq S_T$ is feasible only if $w(S_t) \leq W_t$ for each stage $t \in [T]$. When the stage-weights satisfy $\lambda_1 = \ldots = \lambda_T$, the problem is referred to as the stage-invariant incremental knapsack problem. Surprisingly, [95] show that the simpler stage-invariant version of the problem is strongly NP-Hard, even when $r_i = w_i$ for each item $i \in [n]$. The authors go on to provide a constant factor guarantee for the general problem, as well as a PTAS when $T = O(\sqrt{\log n})$ for the stage-invariant version. Subsequently, [96] developed a PTAS for the stage-invariant version when there are no restrictions on the input parameters. The PTAS of [97] applies to the most general form of the incremental knapsack problem, but it requires that $T = O(1)$. Most recently, [98] provide the first PTAS for the general problem whose running time in polynomial in the input, and [99] provide a polynomial time $(\frac{1}{2} - \epsilon)$-approximation for the generalized incremental knapsack, where the rewards are stage-dependent. Also quite recently, [100] provide an FPTAS for a variant of the incremental knapsack problem, where, among other differences, each item is endowed with a deadline stage, indicating the latest stage where the particular can be introduced.

3.2 A Polynomial-Time Approximation Scheme $(K = T)$

In this section, we present a PTAS for DISPLAY-OPT-K when $K = T$, which represents a best-case algorithmic result in light of the fact that DISPLAY-OPT-K is strongly NP-Hard, as shown in Appendix C.2.1.[3] The exact nature of this PTAS is formalized in the

---

[3]In Appendix C.1, we also provide a strongly-polynomial $1/2-$approximation to the scenario when $K = T$, which may be of independent interest.

following theorem, whose proof unfolds over the remainder of this section. All proofs for this section can be found in Appendix C.2.

Theorem 3.2.1 *For any $\epsilon > 0$, there is an algorithm that returns a sequence of assortments $\mathcal{S} \in \mathcal{F}_T$ with expected revenue $\mathcal{R}(\mathcal{S}) \geq (1 - \epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$, whose running time is $O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right)$, where $|\mathcal{I}|$ denotes the size of the input.*

3.2.1 Preliminaries

This section serves to introduce key pieces of notations that will dramatically simplify the exposition of the PTAS. For ease of exposition, we assume throughout the remainder of the paper that $\frac{1}{\epsilon}$ in an integer. Also, we denote the maximal and minimal weights as $w_{\max} = \max_{i \in [n]} w_i$ and $w_{\min} = \min_{i \in [n]} w_i$ respectively.

The weight classes. Let $\mathcal{C}_q = \{i \in [n] : w_{\min} \cdot (1 + \epsilon)^q \leq w_i < w_{\min} \cdot (1 + \epsilon)^{q+1}\}$ for $q \in [\mathcal{Q}]_0$, where $\mathcal{Q} = \lceil \log_{1+\epsilon}(\frac{w_{\max}}{w_{\min}}) \rceil$, denote the set of items in weight class q. Clearly, the collection of all weight classes $\{\mathcal{C}_q\}_{q \in [\mathcal{Q}]_0}$ represents a partitioning of the products. For assortment $S \subseteq [n]$ and $q \in [\mathcal{Q}]_0$, we let $S_q = S \cap \mathcal{C}_q$ denote the class-q products offered from S. Based on the make-up of S_q, we give class-q one of the following three labels:

- <u>Exhausted classes:</u> Let $Q_E(S) = \{q \in [\mathcal{Q}]_0 : |S_q| = |\mathcal{C}_q|\}$ denote the set of weight classes for which all products are included in the assortment S.

- <u>Active classes:</u> Let $Q_A(S) = \{q \in [\mathcal{Q}]_0 : 1 \leq |S_q| < |\mathcal{C}_q|\}$ denote the collection of weight classes that are "in-use", but not exhausted.

- <u>Empty classes:</u> Let $Q_\emptyset(S) = \{q \in [\mathcal{Q}]_0 : |S_q| = 0\}$ denote the set of classes from which no products have been offered.

Additionally, for assortment $S \subseteq [n]$ and $k \in [|S|]_0$, we let $S[k]$ denote the k highest ρ-valued items in S. For the most part, we will use $\mathcal{C}_q[k]$ to refer to the k items with the largest ρ-value from class q.

$\frac{1}{\epsilon}$-**capped-class-q assortments.** For class $q \in [\mathcal{Q}]_0$ and $k \in [|\mathcal{C}_q|]_0$, let $S_q^*(k) \subseteq \mathcal{C}_q$ denote the assortment that contains the "first" k products added from class q under $\mathcal{S}^*$. We assume that products offered in earlier stages are added before those in later stages, and within a stage, products are added in decreasing ρ-order. A $\frac{1}{\epsilon}$-capped-class-q assortment, or simply a "capped" class-q assortment for short, is defined as

$$
\mathcal{C}_q^*(k) = \begin{cases} S_q^*(k), & \text{if } k \leq \frac{1}{\epsilon} \\ S_q^*\left(\frac{1}{\epsilon}\right) \cup \left(\mathcal{C}_q \setminus S_q^*\left(\frac{1}{\epsilon}\right)\right)\left[k - \frac{1}{\epsilon}\right] & \text{otherwise.} \end{cases}
$$

Rounded total weight. We make-use of a "rounded" total weight of assortment $S \subseteq [n]$, defined as $\hat{w}(S) = \sum_{q \in [\mathcal{Q}]_0} |S_q| \cdot w_{\min} \cdot (1 + \epsilon)^q$. In other words, $\hat{w}(S)$ is the total weight of S if the weights of the products in each class q were rounded down to $w_{\min} \cdot (1 + \epsilon)^q$. We clearly have that $\hat{w}(S) \leq w(S) \leq (1 + \epsilon) \cdot \hat{w}(S)$, since the weights in each class differ by at most a factor of $1 + \epsilon$.

3.2.2 Constructing the proxy assortment

In this section, we show how to construct a proxy optimal sequence of assortment $\hat{\mathcal{S}} \in \mathcal{F}_T$, whose expected revenue satisfies $\mathcal{R}(\hat{\mathcal{S}}) \geq (1 - O(\epsilon)) \cdot \mathcal{R}(\mathcal{S}^*)$. We will construct $\hat{\mathcal{S}}$ from basic building blocks, endowing it with a special structure that allows for its approximate recovery via an efficient dynamic-programming-based approach. The main building block that we utilize to construct $\hat{\mathcal{S}}$ is what we refer to as a "fill event", which is summarized below, and formally presented in Algorithm 1.

Fill event. Consider an arbitrary stage $t \in [T]$ and an assortment $S \subseteq [n]$ that does not satisfy the stage-t weight constraint with respect to the total rounded weight, i.e. $\hat{w}(S) < W_t$. A fill event consists of the following steps. We begin by picking the lowest indexed active class $q_{\min} = \min\{q \in [\mathcal{Q}]_0 : q \in Q_A(\tilde{S})\}$ and adding products from $\mathcal{C}_{q_{\min}} \setminus \tilde{S}$ to $\tilde{S}$ in decreasing order of ρ-value until either $\hat{w}(\tilde{S}) \geq W_t$ (this check is with respect to the rounded total weight), or class $q_{\min}$ is exhausted. We repeat these steps until either (i) a feasible assortment is returned or (ii) there are fewer than $\frac{1}{\epsilon}$ active classes. The following lemma details the key properties held by the assortment returned after a fill event, which concern its total weight or its number of active stages.

Lemma 3.2.2 *Consider an arbitrary stage $t \in [T]$ and an assortment $S \subseteq [n]$ whose total rounded weight satisfies $\hat{w}(S) < W_t$. For assortment $\tilde{S} = \text{FILL}(S,t)$, we have that $w(\tilde{S}) \leq (1 + 3\epsilon) \cdot W_t$ if $\hat{w}(\tilde{S}) \geq W_t$. Otherwise, we have that $|Q_A(\tilde{S})| \leq \frac{1}{\epsilon}$.*

Constructing the proxy sequence $\hat{\mathcal{S}}$. To build $\hat{\mathcal{S}}$, we sequentially construct three candidate sequences of assortments that are loosely based on the make-up of $\mathcal{S}^*$. For the first two, it is important to note that feasibility is checked with respect to the total rounded weight.

- **Candidate 1:** The first candidate sequence of assortment is given by $\mathcal{S}^{(1)}$, whose contents are built as follows. For each stage $t \in [T]$ and class $q \in [\mathcal{Q}]_0$, we set

$$
S_{t,q}^{(1)} = \begin{cases} \mathcal{C}_q[k_{t,q}^*], & \text{if } k_{t,q}^* \leq \frac{1}{\epsilon} \\ \mathcal{C}_q[\lceil (1+\epsilon) \cdot k_{t,q}^* \rceil], & \text{otherwise,} \end{cases}
$$

where $k_{t,q}^* = |S_{t,q}^*|$ denotes the number of products offered from class q in stage t under $\mathcal{S}^*$. If $\hat{w}(\mathcal{S}_t^{(1)}) \geq W_t$ for each stage $t \in [T]$, then set $\hat{\mathcal{S}} = \mathcal{S}^{(1)}$. Otherwise, we move to Candidate 2, which builds on top of $\mathcal{S}^{(1)}$ using fill events.

- **Candidate 2:** The second candidate sequence of assortment $\mathcal{S}^{(2)}$ is built recursively as follows. For each stage $t \in [T]$, we set

$$
S_t^{(2)} = \begin{cases} S_t^{\text{temp}}, & \text{if } \hat{w}(S_t^{\text{temp}}) \geq W_t \\[2mm] \text{FILL}(S_t^{\text{temp}}, t), & \text{otherwise,} \end{cases}
$$

where $S_t^{\text{temp}} = S_t^{(1)} \cup S_{t-1}^{(2)}$ (assuming that $S_0^{(2)} = \emptyset$). Again, if $\hat{w}(\mathcal{S}_t^{(2)}) \geq W_t$ for each stage $t \in [T]$, then set $\hat{\mathcal{S}} = \mathcal{S}^{(2)}$. Otherwise, we move to Candidate 3.

- **Candidate 3:** Let $T^< = \{t \in [T] : \hat{w}(S_t^{(2)}) < W_t\}$ denote the stages for which $\mathcal{S}^{(2)}$ remains infeasible with respect to the weight constraints, and let $Q_A^< = \biguplus_{t \in T^<} Q_A(S_t^{(2)})$ give the active classes during these stages. The third final and candidate solution $\mathcal{S}^{(3)}$ is defined as follows:

$$
S_t^{(3)} = \left(\biguplus_{q \notin Q_A^<} S_{t,q}^{(2)} \right) \cup \left(\biguplus_{q \in Q_A^<} C_q^*(k_{t,q}^{(2)}) \right),
$$

where $k_{t,q}^{(2)} = |S_{t,q}^{(2)}|$ denotes the number of products offered from class q in stage t under $\mathcal{S}^{(2)}$. Finally, we set $\hat{\mathcal{S}} = \mathcal{S}^{(3)}$ if this third candidate is indeed reached.

The efficacy of $\hat{\mathcal{S}}$ as a proxy. The following lemma shows that $\hat{\mathcal{S}}$ is indeed a good proxy for $\mathcal{S}^*$. We prove the result by showing its validity for each of the three candidate assortments described above, conditioned that the particular candidate is in fact set to be $\hat{\mathcal{S}}$.

Lemma 3.2.3 *Let $\hat{\mathcal{S}}$ be the proxy sequence of assortments defined above. Then, $\hat{\mathcal{S}} \in \mathcal{F}$ and $\mathcal{R}(\hat{\mathcal{S}}) \geq (1 - 10\epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$.*

Algorithm 1 fill event

1: **procedure** $\text{FILL}(S, t)$
2:     $\tilde{S} \leftarrow S$
3:     **while** $|Q_A(\tilde{S})| > \frac{1}{\epsilon}$ **do**
4:         $q_{\min} \leftarrow \min\{q : q \in Q_A(\tilde{S})\}$
5:         $A \leftarrow \mathcal{C}_{q_{\min}} \setminus \tilde{S}$
6:         **for** $i \in A$ **do** [Iterate from largest to smallest ρ-value]
7:             $\tilde{S} = \tilde{S} \cup \{i\}$
8:             **if** $\hat{w}(\tilde{S}) \geq W_t$ **then**
9:                 **return** $\tilde{S}$
10:             **end if**
11:         **end for**
12:     **end while**
13:     **return** $\tilde{S}$
14: **end procedure**

3.2.3 The approximation scheme

In this section, we present a dynamic program that can be used to recover a sequence of assortments whose expected revenue is within an $O(\epsilon)$-factor of that of $\hat{\mathcal{S}}$. To do so, we begin by constructing a universe of assortments $\mathcal{U}$ that contains $\hat{S}_t$ for any stage $t \in [T]$. We then show how to fold this universe of assortments into a dynamic programming formulation of the problem, which is guaranteed to return an assortment whose expected revenue is at least as large as that of $\hat{S}$. Unfortunately, due to the exponential size of $\mathcal{U}$, the running time of this dynamic program is not polynomial. As such, we set about applying two updates to $\mathcal{U}$, which together yield an alternative universe of polynomially-many assortments $\mathcal{U}_{\text{small}}$. Furthermore, we show that replacing $\mathcal{U}_{\text{small}}$ with $\mathcal{U}$ in the aforementioned dynamic program only degrades its performance by at most an $O(\epsilon)$-factor.

Constructing $\mathcal{U}$. Central to our approach is the notion that, for any stage $t \in [T]$ and class $q \in [\mathcal{Q}]_0$, we must have either

- $\hat{S}_{t,q} = \mathcal{C}_q[k]$ for some $k \in [|\mathcal{C}_q|]_0$, or

- $\hat{S}_{t,q} = \mathcal{C}_q^*(k)$ for some $k \in [|\mathcal{C}_q|]_0$

based on the make-up of the three candidates outlined in Section 3.2.2. Given this structure, we can guess $\hat{\mathcal{S}}_t$ exactly for any stage $t \in [T]$ by enumerating over all combinations of the following three parameters:

(i) A subset of classes $Q_{\text{CAP}} \subseteq [\mathcal{Q}]_0$, intended to contain the capped class-q assortments.

(ii) For each Q_{CAP}, a collection of assortments $\{A_q\}_{q \in Q_{\text{CAP}}}$ such that $A_q \subseteq \mathcal{C}_q$ and $|A_q| \leq \frac{1}{\epsilon}$.

(iii) For each pair $(Q_{\text{CAP}}, \{A_q\}_{q \in Q_{\text{CAP}}})$, a collection of utilization vectors

$$\mathcal{K} = \{K = (k_0, \ldots, k_{\mathcal{Q}}) \in [|\mathcal{C}_0|]_0 \times \ldots \times [|\mathcal{C}_{\mathcal{Q}}|]_0 : k_q \in \{|A_q|, \frac{1}{\epsilon}+1, \ldots, |\mathcal{C}_q|\} \, \forall \, q \in Q_{\text{CAP}}\},$$

each of which indicates the total number of products added from class $q \in [\mathcal{Q}]_0$.

A parameter triplet $(Q_{\text{CAP}}, \{A_q\}_{q \in Q_{\text{CAP}}}, K)$ satisfying conditions (i)-(iii) translates into the assortment S with

$$
S_q = \begin{cases}
\mathcal{C}_q[k_q], & \text{if } q \notin Q_{\text{CAP}} \\
A_q, & \text{if } q \in Q_{\text{CAP}}, k_q = |A_q| \\
A_q \cup (\mathcal{C}_q \setminus A_q) [k_q - |A_q|] & \text{otherwise.}
\end{cases}
$$

We build $\mathcal{U}$ by enumerating over all such assortments corresponding to valid parameter triplets. While $\mathcal{U}$ clearly contains an exponential number of assortments, it is easy to see that $\hat{\mathcal{S}}_t \in \mathcal{U}$ for any stage $t \in [T]$.

The dynamic program. The state space of our dynamic program will be a stage $t \in [T]$, and an assortment $S_{t-1} \in \mathcal{U}$ indicating the set of products offered in stage $t-1$ (we assume $S_0 = \emptyset$). The value functions $\mathcal{V}_{\mathcal{U}}(t, S_{t-1})$ indicate the maximum expected revenue that can be accrued from stage $t, \ldots, T$ given that the stage t assortment is S_{t-1}. Formally, the recursion is as follows:

$$
\mathcal{V}_{\mathcal{U}}(t, S_{t-1}) = \max_{\substack{S_t \in \mathcal{U}: \\ w(S_t) \geq W_t, \\ S_{t-1} \subseteq S_t}} \left\{ \lambda_t \cdot \frac{1}{1 + w(S_t)} \cdot \sum_{i \in S_t} \rho_i + \mathcal{V}_{\mathcal{U}}(t+1, S_t) \right\}, \tag{3.1}
$$

with base case $\mathcal{V}_{\mathcal{U}}(T+1, \cdot) = 0$. Moreover, if there is no stage-t assortment that is feasible in a particular state, we set the corresponding value function to be negative infinity. Since the sequence of states

$$
(0, \emptyset) \rightarrow (1, \hat{\mathcal{S}}_1) \rightarrow \cdots \rightarrow (T, \hat{\mathcal{S}}_1)
$$

is feasible, we get that $\mathcal{V}_{\mathcal{U}}(0, \emptyset) \geq \mathcal{R}(\hat{\mathcal{S}})$. As such, this dynamic program will produce a sequence of assortments $\mathcal{S} \in \mathcal{F}$, whose expected revenue is at least as large as that of $\hat{\mathcal{S}}$,

albeit in a running time that is clearly scales with the size of $\mathcal{U}$, and hence is exponential in the input. In the sequel, we set about reducing the size of the universe of assortments to choose from, while at the same time sacrificing little in terms of the efficacy of the dynamic program in (3.1).

Step 1: bounding the number of capped assortments. In what follows, we argue that it is sufficient to focus on a universe of assortments built exactly like $\mathcal{U}$, except we update condition (i) to consider only $Q_{\text{CAP}} \subseteq [\mathcal{Q}]_0$ such that $|Q_{\text{CAP}}| \leq \frac{2}{\epsilon}$. To establish this result, we upper bound the total number of capped assortments under $\hat{S}_t$ by $\frac{2}{\epsilon}$, for any stage $t \in [T]$. This result is formally stated in the following lemma, where it is critical to recall that $\hat{\mathcal{S}}$ only uses capped assortments if $\hat{\mathcal{S}} = \mathcal{S}^{(3)}$. Moreover, in this case, we have that $\hat{S}_{t,q}$ is a capped assortment if and only if $q \in \{q \in Q_A^{<} : k_{t,q}^{(2)} > 0\}$. Hence bounding the number of capped assortment under $\hat{\mathcal{S}}$ in a particular stage requires only bounding the size of this latter set.

Lemma 3.2.4 *For any stage $t \in [T]$, we have that*

$$\left| \left\{ q \in Q_A^{<} : k_{t,q}^{(2)} > 0 \right\} \right| \leq \frac{2}{\epsilon},$$

where $Q_A^{<}$ is as defined is the construction of $\mathcal{S}^{(3)}$.

Step 2: rounding the utilization vectors. Fixing the parameter pair $(Q_{\text{CAP}}, \{A_q\}_{q \in Q_{\text{CAP}}})$ throughout this discussion, we propose the following "up-rounded" version of the collection of utilization vectors $\mathcal{K}$. Specifically, for any utilization vector $K \in \mathcal{K}$, we construct its rounded counterpart $K^{\uparrow} = (k_0^{\uparrow}, \ldots, k_{\mathcal{Q}}^{\uparrow})$ as follows, where for ease of notation, we use $S \subseteq [n]$ to denote the assortment corresponding to the parameter triplet $(Q_{\text{CAP}}, \{A_q\}_{q \in Q_{\text{CAP}}}, K)$. Also, let $q_{\max}(K) = \max\{q \in [\mathcal{Q}]_0 : k_q > 0\}$ denote the highest indexed non-empty class and let $L = \lceil \log_{1+\epsilon}(\frac{n}{\epsilon}) \rceil$. It is worth noting that this up-rounding

99

scheme closely mirrors the one proposed in [98], and for this reason, we omit proof details that would require directly replicating the arguments of this paper.

- For class $q < [q_{\max}(K) - L + 1]^+$, we set $k_q^{\uparrow} = |\mathcal{C}_q|$.

- For class $q \geq [q_{\max}(K) - L + 1]^+$ such that $k_q \leq \frac{1}{\epsilon}$ we set $k_q^{\uparrow} = k_q$.

- For class $q \geq [q_{\max}(K) - L + 1]^+$ such that $k_q > \frac{1}{\epsilon}$, we let $w_q^{\uparrow}$ be an over-estimate of $w(S_q)$ by an additive factor of at most $\text{Power}_2[(\frac{\epsilon}{L} \cdot w(S))]$, where the operator $\text{Power}_2[\cdot]$ rounds up its input to the nearest power of 2. Specifically, we define $w_q^{\uparrow} = \mu_q \cdot \text{Power}_2[\frac{\epsilon}{L} \cdot w(S)]$ for the unique integer $\mu_q \geq 1$ satisfying

$$(\mu_q - 1) \cdot \text{Power}_2 \left[\frac{\epsilon}{L} \cdot w(S) \right] < w(S_q) \leq \mu_q \cdot \text{Power}_2 \left[\frac{\epsilon}{L} \cdot w(S) \right]. \qquad (3.2)$$

Next, if $q \in Q_{\text{CAP}}$. we set

$$k_q^{\uparrow} = \max \left\{ k \in [|\mathcal{C}_q|]_0 : w(A_q \cup (\mathcal{C}_q \setminus A_q)[k - |A_q|]) \leq w_q^{\uparrow} \right\}.$$

Otherwise, if $q \notin Q_{\text{CAP}}$, we set

$$k_q^{\uparrow} = \max \left\{ k \in [|\mathcal{C}_q|]_0 : w(\mathcal{C}_q[k]) \leq w_q^{\uparrow} \right\}.$$

The following lemma establishes two crucial properties of this rounding scheme, the first of which formalizes the notion that we are indeed up-rounding the utilization vectors, and the second bounds the total weight added by this rounding.

Lemma 3.2.5 *The up-rounding scheme outlined above maintains the following two properties:*

(i) Monotonicity: Consider any pair of parameter triplets $(Q_{\text{CAP}}, \{A_q\}_{q \in Q_{\text{CAP}}}, K)$ and $(Q_{\text{CAP}}^+, \{A_q^+\}_{q \in Q_{\text{CAP}}^+}, K^+)$ respectively corresponding to assortments S and S^+ that sat-

isfy $S \subseteq S^+$. Then, the assortments $S^\uparrow$ and $S^{+\uparrow}$ corresponding to the up-rounded triplets $(Q_{\mathrm{CAP}}, \{A_q\}_{q \in Q_{\mathrm{CAP}}}, K^\uparrow)$ and $(Q_{\mathrm{CAP}}^+, \{A_q^+\}_{q \in Q_{\mathrm{CAP}}^+}, K^{+\uparrow})$ must satisfy $S^\uparrow \subseteq S^{+\uparrow}$.

(ii) *Weight added: For any $K \in \mathcal{K}$, let S and $S^\uparrow$ denote the assortments represented by $(Q_{\mathrm{CAP}}, \{A_q\}_{q \in Q_{\mathrm{CAP}}}, K)$ and $(Q_{\mathrm{CAP}}, \{A_q\}_{q \in Q_{\mathrm{CAP}}}, K^\uparrow)$ respectively. Then we have that $w(S) \leq w(S^\uparrow) \leq (1 + 3\epsilon) \cdot w(S)$.*

Constructing $\mathcal{U}_{\mathrm{small}}$. We build $\mathcal{U}_{\mathrm{small}}$ by adding all assortments corresponding to every combination of the following updated version of the original three parameters:

(i) A subset of classes $Q_{\mathrm{CAP}} \in \{Q \subseteq [\mathcal{Q}]_0 : |Q| \leq \frac{2}{\epsilon}\}$.

(ii) For each Q_{CAP}, a collection of assortments $\{A_q\}_{q \in Q_{\mathrm{CAP}}}$ such that $A_q \subseteq \mathcal{C}_q$ and $|A_q| \leq \frac{1}{\epsilon}$.

(iii) For each pair $(Q_{\mathrm{CAP}}, \{A_q\}_{q \in Q_{\mathrm{CAP}}})$, a collection of utilization vectors $\mathcal{K}^\uparrow = \{K^\uparrow : K \in \mathcal{K}\}$.

The following lemma reveals that $\mathcal{U}_{\mathrm{small}}$ is indeed considerable smaller than $\mathcal{U}$, however it is important to note that due to the up-rounding scheme, we are no longer guaranteed to have that $\hat{S}_t \in \mathcal{U}_{\mathrm{small}}$ for any stage $t \in [T]$.

Lemma 3.2.6 $|\mathcal{U}_{\mathrm{small}}| = O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right)$.

The approximate dynamic program. Our final step is to replace $\mathcal{U}$ with $\mathcal{U}_{\mathrm{small}}$ in (3.1), thus yielding the following updated recursion:

$$\mathcal{V}_{\mathcal{U}_{\mathrm{small}}}(t, S_{t-1}) = \max_{\substack{S_t \in \mathcal{U}_{\mathrm{small}}: \\ w(S_t) \geq W_t, \\ S_{t-1} \subseteq S_t}} \left\{ \lambda_t \cdot \frac{1}{1 + w(S_t)} \cdot \sum_{i \in S_t} \rho_i + \mathcal{V}_{\mathcal{U}_{\mathrm{small}}}(t+1, S_t) \right\}, \qquad (3.3)$$

with base case $\mathcal{V}_{\mathcal{U}_{\mathrm{small}}}(T+1, \cdot) = 0$. Again, if there is no stage-t assortment that is feasible in a particular state, we set the corresponding value function to be negative infinity. Given

Lemma 3.2.6, we can compute the value function for any state (t, S_{t-1}) by simply enumerating over all assortments $S_t \in \mathcal{U}_{\text{small}}$, and hence all value functions can be computed in a running time of

$$O(|\mathcal{U}_{\text{small}}|^2 \cdot T) = O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right),$$

which matches the running time specified in Theorem 3.2.1. Our final result of this section bounds the revenue lose incurred by replacing $\mathcal{U}$ with $\mathcal{U}_{\text{small}}$ as seen in (3.3).

Lemma 3.2.7 $\mathcal{V}_{\mathcal{U}_{\text{small}}}(0, \emptyset) \geq (1 - 3\epsilon) \cdot \mathcal{R}(\hat{\mathcal{S}})$.

3.3 A Polynomial-Time Approximation Scheme ($K = 1$)

In this section, we present a PTAS for DISPLAY-OPT-K when $K = 1$.

The approximate dynamic program. Consider the following recursion:

$$\mathcal{V}_{\mathcal{U}_{\text{small}}}(t, r, S_{t-1}) = \max_{\substack{S_t \in \mathcal{U}_{\text{small}}: \\ S_{t-1} \subseteq S_t}} \left\{\lambda_t \cdot cs(S_t) + \mathcal{V}_{\mathcal{U}_{\text{small}}}\left(t + 1, \lceil r - \lambda_t \cdot \mathcal{R}(S_t)\rceil_{1+\delta}, S_t\right)\right\}, \qquad (3.4)$$

where $\delta = \frac{14\epsilon}{T+1}$, with base case

$$\mathcal{V}_{\mathcal{U}_{\text{small}}}(T + 1, r, \cdot) = \begin{cases} 0, & \text{if } r \leq 0, \\ -\infty, & \text{if } r > 0. \end{cases}$$

Given Lemma 3.2.6, we can compute the value function for any state (t, r, S_{t-1}) by simply enumerating over all possible rounded revenue values and assortments $S_t \in \mathcal{U}_{\text{small}}$, and hence all value functions can be computed in a running time of $O\left(\left(|\mathcal{U}_{\text{small}}| \cdot \log_{1+\delta}(r_{\max}/r_{\min} w_{\min})\right)^2 \cdot T\right) = O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right)$. Finally, the performance of the above PTAS is summarized as follows,

Theorem 3.3.1 *For any $\epsilon > 0$, there is an algorithm that returns a sequence of assortments $\mathcal{S} \in \mathcal{F}_1$ with expected revenue $\mathcal{R}(\mathcal{S}) \geq (1 - \epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$, whose running time is $O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right)$, where $|\mathcal{I}|$ denotes the size of the input.*

3.4 A Polynomial-Time Approximation Scheme for General K $(1 < K < T)$

Given our algorithms proposed in Section 3.2 and 3.3, we are now ready to introduce our PTAS for the general case of DISPLAY-OPT-K when $1 < K < T$. The algorithm below is a combination of the two algorithms above.

The approximate dynamic program. For every group $k \in [K]$ customers, we can write down the following inner-group recursion for $t \in [t_k : t_{k+1} - 1)$:

$$\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}(t, r, S_{t-1}, S_{t_{k+1}-1})$$

$$= \max_{\substack{S_t \in \mathcal{U}_{\text{small}} \cap S_{t_{k+1}-1}: \\ S_{t-1} \subseteq S_t}} \left\{ \lambda_t \cdot cs\left(S_t\right) + \mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}\left(t+1, \lceil r - \lambda_t \cdot \mathcal{R}(S_t)\rceil_{1+\delta}, S_t, S_{t_{k+1}-1}\right) \right\},$$

$$(3.5)$$

with base case

$$\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}(t_{k+1} - 1, r, S_{t_{k+1}-1}, S_{t_{k+1}-1}) = \begin{cases} 0, & \text{if } r \leq 0, \\ -\infty, & \text{if } r > 0. \end{cases}$$

Let $\mathcal{T}\left(\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}\left(t_k, r, S_{t_k-1}, S_{t_{k+1}-1}\right)\right)$ be the induced sequence of assortments $\mathcal{S}^{(k)} = \left(S_{t_k}, \cdots, S_{t_{k+1}-1}\right)$ via computing $\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}\left(t_k, r, S_{t_k-1}, S_{t_{k+1}-1}\right)$. Similar to Section 3.3, for given S_{t_k-1} and $S_{t_{k+1}-1}$, the running time of computing an inner-group recursion $\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}\left(t_k, r, S_{t_k-1}, S_{t_{k+1}-1}\right)$ takes $O\left(\left(|\mathcal{U}_{\text{small}}| \cdot \log_{1+\delta}\left(r_{max}/r_{min}w_{\min}\right)\right)^2 \cdot (t_{k+1} - t_k - 1)\right)$

We define $\mathcal{R}^{(k)}\left(\mathcal{S}^{(k)}\right) = \sum_{t=t_k}^{t_{k+1}-1} \lambda_i \cdot \left(\sum_{i \in S_t} \frac{\rho_i}{1+w(S_t)}\right)$, and $\mathcal{R}^{(k)}(\emptyset) = -\infty$. Then we have the outer-group recursion for break periods $t_1, \cdots, t_K$,

$$\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-outer}(t_k, S_{t_k-1})$$

$$= \max_{\substack{r \in \mathsf{Dom}^\delta \\ S_{t_{k+1}-1} \in \mathcal{U}_{\text{small}}: \\ S_{t_{k+1}-1} \supseteq S_{t_k-1} \\ \mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}\left(t_k, r, S_{t_k-1}, S_{t_{k+1}-1}\right) \geq \alpha_k}} \left\{ \mathcal{R}^{(k)}\left(\mathcal{T}\left(\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-inner}\left(t_k, r, S_{t_k-1}, S_{t_{k+1}-1}\right)\right)\right) \right.$$

$$\left. + \mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k+1)-outer}(t_{k+1}, S_{t_{k+1}-1}) \right\}, \quad (3.6)$$

where $\mathsf{Dom}^\delta = \{0\} \cup \left\{(1+\delta)^k : \underline{k} \leq k \leq \overline{k}\right\}$, where $(1+\delta)^{\underline{k}} = \left\lfloor r_{\min} w_{\min} / \left((1+\delta)^T \cdot (1+w_{\min})\right)\right\rfloor_{1+\delta}$ and $(1+\delta)^{\overline{k}} = \left\lfloor r_{\max} / (1+\delta)^T \right\rfloor_{1+\delta}$. Base case is $\mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-outer}(t_{K+1}, \cdot) = \mathcal{V}_{\mathcal{U}_{\text{small}}}^{(k)-outer}(T+1, \cdot) = 0$.

Given Lemma 3.2.6, we can compute the value function for any state (t, r, S_{t-1}) by simply enumerating over all possible rounded revenue values and assortments $S_t \in \mathcal{U}_{\text{small}}$, and hence all value functions in the outer-group recursion can be computed in a running time of $O\left(\left(|\mathcal{U}_{\text{small}}| \cdot \log_{1+\delta}(r_{\max}/r_{\min} w_{\min})\right)^2 \cdot K\right)$. Therefore in total we can fully solve the dynamic programming (3.5) and (3.6) in running time of

$$\sum_{k \in [K]} O\left(\left(|\mathcal{U}_{\text{small}}| \cdot \log_{1+\delta}(r_{\max}/r_{\min} w_{\min})\right)^2 \cdot (t_{k+1} - t_k - 1)\right)$$

$$+ O\left(\left(|\mathcal{U}_{\text{small}}| \cdot \log_{1+\delta}(r_{\max}/r_{\min} w_{\min})\right)^2 \cdot K\right)$$

$$= O\left(\left(|\mathcal{U}_{\text{small}}| \cdot \log_{1+\delta}(r_{\max}/r_{\min} w_{\min})\right)^2 \cdot T\right)$$

$$= O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right).$$

Combining the proof for Theorem 3.2.1 and Theorem 3.3.1, one can show that

Theorem 3.4.1 *For any $\epsilon > 0$, there is an algorithm that returns a sequence of assortments $\mathcal{S} \in \mathcal{F}_K$ with expected revenue $\mathcal{R}(\mathcal{S}) \geq (1 - \epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$, whose running time is $O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right)$, where $|\mathcal{I}|$ denotes the size of the input.*

3.5 Cardinality Constraints with $T = O(1)$

In many commercial settings, a natural business constraint is the cardinality constraint. Imagine an online product display setting where the number of products shown on each page (stage) is limited. The cardinality-constrained problem is formulated as follows

$$\max_{\mathcal{S} \in \mathcal{F}_K^{Card}} \mathcal{R}(\mathcal{S}), \qquad \text{(DISPLAY-OPT-}K\text{-CARD)}$$

where $\mathcal{F}_K^{Card} = \mathcal{F}_K \cap \{(S_1, \ldots, S_T) : S_t \subseteq [n], |S_t \setminus S_{t-1}| \leq C\}$. Our algorithm works when the cardinality constraints are stage-dependent, but here for exposition simplicity, we assume the cardinality limits for all stages are the same, i.e., C.

Inspired by the algorithm before, we are able to give a PTAS that violates the customer satisfaction constraints by ϵ when $T = O(1)$. Given any assortment S, let permutation $P(S)$ rank product set S according to $\{\rho_i\}_{i \in S}$ in the descending order.

For the true optimal assortment $\mathcal{S}^*$, we first construct a proxy $\hat{\mathcal{S}} = \left(\hat{S}_1, \cdots, \hat{S}_T\right)$ as Section 3.2.2 does, then we construct a proxy $\hat{\mathcal{S}}^{\text{small}} = \left(\hat{S}_1^{\text{small}}, \cdots, \hat{S}_T^{\text{small}}\right) \in \bigtimes_{t=1}^T \mathcal{U}_{\text{small}}^{(t)}$ to approximate $\hat{\mathcal{S}}$. The definition of candidate space $\mathcal{U}_{\text{small}}^{(t)}$ will be introduced shortly. We construct $\hat{\mathcal{S}}$ in a way similar to **Candidate 1** with a slight modification. For each stage $t \in [T]$ and class $q \in [\mathcal{Q}]_0$, we set

$$\hat{S}_{t,q} = \begin{cases} \mathcal{C}_q[k_{t,q}^*], & \text{if } k_{t,q}^* \leq \frac{1}{\epsilon} \\ \mathcal{C}_q[\lceil (1 - \epsilon) \cdot k_{t,q}^* \rceil], & \text{otherwise.} \end{cases}$$

The efficacy proof of $\hat{\mathcal{S}}$ as a proxy is given by the following lemma.

Lemma 3.5.1 *Let $\hat{\mathcal{S}}$ be the proxy sequence of assortments defined above. Then, $\hat{S}_1 \subseteq \cdots \subseteq \hat{S}_T$, $\left| \hat{S}_t \setminus \hat{S}_{t-1} \right| \leq C$, $\sum_{t \in [t_k : t_{k+1})} \lambda_t \cdot cs \left(w \left(\hat{S}_t \right) \right) > (1 - 2\epsilon) \cdot \alpha_k$, for $t \in [T]$, $k \in [K]$. And $\mathcal{R}(\hat{\mathcal{S}}) \geq (1 - 2\epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$.*

Then we consider constructing a candidate assortment space $\mathcal{U}_{\text{small}}^{(t)}$, $t \in [T]$. The building block of the candidate assortment space comes from a modification of $\mathcal{U}_{\text{small}}$. We define space $\mathcal{U}_{\text{small}}^{(t)}$ for stage t by parameter sets $\left(K_1^{\downarrow}, \cdots, K_t^{\downarrow}; l_{t1}, \cdots, l_{tt} \right)$. Here $K_\tau^{\downarrow} = \left(k_{\tau 0}^{\downarrow}, \cdots, k_{\tau Q}^{\downarrow} \right)$ is the rounded utilization vector for products **added** in stage $\tau \in [t]$. On the other hand, $l_{t\tau}$ is the number that records how many products from class $[q_{\max}(K_{\tau-1}) - L + 1]^+$ to $[q_{\max}(K_\tau) - L]^+$ are added according to the descending ρ order in stage $t \in [T]$, i.e.,

$$P \left(\biguplus_{q = [q_{\max}(K_{\tau-1}) - L + 1]^+}^{[q_{\max}(K_\tau) - L]^+} \left(\mathcal{C}_q \setminus \mathcal{C}_q \left[k_{(t-1)q}^{\Sigma} \right] \right) \right) [l_{t\tau}],$$

where we define $K_t^{\Sigma} = \left(k_{t0}^{\Sigma}, \cdots, k_{tQ}^{\Sigma} \right)$ and $k_{tq}^{\Sigma} = \sum_{\tau=1}^{t} k_{\tau q}$. Then a parameter set $\left(K_1^{\downarrow}, \cdots, K_t^{\downarrow}; l_{t1}, \cdots, l_{tt} \right)$ is translated into the assortment S_t with

$$S_t = \left(\biguplus_{q \in [Q]} \mathcal{C}_q \left[k_{tq}^{\Sigma} \right] \right) \bigcup \left(\biguplus_{\tau=1}^{t} P \left(\biguplus_{q = [q_{\max}(K_{\tau-1}) - L + 1]^+}^{[q_{\max}(K_\tau) - L]^+} \left(\mathcal{C}_q \setminus \mathcal{C}_q \left[k_{(t-1)q}^{\Sigma} \right] \right) \right) [l_{t\tau}] \right).$$

Let $K_t = (k_{t1}, \cdots, k_{tQ})$ denote the utilization vector of $\left(\hat{S}_t \setminus \hat{S}_{t-1} \right)$. Then consider a rounded utilization vector $K_\tau^{\downarrow} = \left(k_{\tau 0}^{\downarrow}, \cdots, k_{\tau Q}^{\downarrow} \right)$ which is defined in an analogous way as (3.2) with some modifications as follows,

- For class $q < [q_{\max}(K_\tau) - L + 1]^+$, we set $k_{\tau q}^{\downarrow} = 0$.

- For class $q \geq [q_{\max}(K_\tau) - L + 1]^+$ such that $k_{\tau q} \leq \frac{1}{\epsilon}$ we set $k_{\tau q}^{\downarrow} = k_{\tau q}$.

106

- For class $q \geq [q_{\max}(K_\tau) - L + 1]^+$ such that $k_{\tau q} > \frac{1}{\epsilon}$, we define $w_q^\downarrow = (\mu_q - 1) \cdot$ $\mathrm{Power}_2 \left[\frac{\epsilon}{L} \cdot w(\hat{S}_\tau \setminus \hat{S}_{\tau-1}) \right]$, where $\mu_q \geq 1$ is the unique integer such that

$$(\mu_q - 1) \cdot \mathrm{Power}_2 \left[\frac{\epsilon}{L} \cdot w(\hat{S}_\tau \setminus \hat{S}_{\tau-1}) \right] < w(\hat{S}_{\tau,q} \setminus \hat{S}_{\tau-1,q}) \leq \mu_q \cdot \mathrm{Power}_2 \left[\frac{\epsilon}{L} \cdot w(\hat{S}_\tau \setminus \hat{S}_{\tau-1}) \right].$$

On the other hand, define $\rho_q^\downarrow = (\xi_q - 1) \cdot \mathrm{Power}_2 \left[\frac{\epsilon}{L} \cdot \rho(\hat{S}_\tau \setminus \hat{S}_{\tau-1}) \right]$, where $\xi_q \geq 1$ is the unique integer such that

$$(\xi_q - 1) \cdot \mathrm{Power}_2 \left[\frac{\epsilon}{L} \cdot \rho(\hat{S}_\tau \setminus \hat{S}_{\tau-1}) \right] < \rho(\hat{S}_{\tau,q} \setminus \hat{S}_{\tau-1,q}) \leq \xi_q \cdot \mathrm{Power}_2 \left[\frac{\epsilon}{L} \cdot \rho(\hat{S}_\tau \setminus \hat{S}_{\tau-1}) \right].$$

Then $k_{\tau q}^\downarrow = \min \left\{ k \in [|\mathcal{C}_q|]_0 : w((\mathcal{C}_q \setminus \hat{S}_{\tau-1,q})[k]) \geq w_q^\downarrow, \ \rho((\mathcal{C}_q \setminus \hat{S}_{\tau-1,q})[k]) \geq \rho_q^\downarrow \right\}.$

And let

$$l_{t\tau} = \left| \biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\hat{S}_{t,q} \setminus \hat{S}_{\tau-1,q} \right) \right|, \quad \tau \in [t].$$

The efficacy of proxy $\hat{\mathcal{S}}^{\mathrm{small}}$ is summarized by the following lemma.

Lemma 3.5.2 *Let $\hat{\mathcal{S}}^{\mathrm{small}}$ be the proxy sequence of assortments defined above. Then, $\hat{S}_1^{\mathrm{small}} \subseteq \cdots \subseteq \hat{S}_T^{\mathrm{small}}$, $\left| \hat{S}_t^{\mathrm{small}} \setminus \hat{S}_{t-1}^{\mathrm{small}} \right| \leq C$, $\sum_{t \in [t_k : t_{k+1})} \lambda_t \cdot cs \left(w \left(\hat{S}_t^{\mathrm{small}} \right) \right) > (1 - 5\epsilon) \cdot \alpha_k$, for $t \in [T]$, $k \in [K]$. And $\mathcal{R}(\hat{\mathcal{S}}^{\mathrm{small}}) \geq (1 - 2(T+3) \cdot \epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$.*

In each stage t, we can enumerate all possible parameter sets $\left(K_1^\downarrow, \cdots, K_t^\downarrow; l_{t1}, \cdots, l_{tt} \right)$ in $\mathcal{U}_{\mathrm{small}}^{(t)}$ to find $\hat{S}_t^{\mathrm{small}}$. The total number of possible combinations is upper bounded by $\left| \mathcal{U}_{\mathrm{small}}^{(t)} \right| = O\left(\left(\ln \left(\frac{n w_{\max}}{w_{\min}} \right) \cdot \ln \left(\frac{n \rho_{\max}}{\rho_{\min}} \right) \cdot 2^{O(L/\epsilon)} \right)^t \cdot n^t \right) < O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{T}{\epsilon^2})} \right) = O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})} \right).$

Finally, to find a proxy of $\mathcal{S}^*$ we can write down the recursions like (3.5) and (3.6). The only differences are: (i) Substitute $\mathcal{U}_{\mathrm{small}}$ with $\mathcal{U}_{\mathrm{small}}^{(t)}$ for each stage t; (ii) Add cardinality constraints $|S_t \setminus S_{t-1}| \leq C$ for each stage t; (iii) For given approximation error ϵ, substitute the satisfaction constraints α_k with $(1 - \epsilon) \cdot \alpha_k$, $k \in [K]$.

Define

$$\mathcal{F}_K^{Card-\epsilon} = \left\{ (S_1, \ldots, S_T) : S_t \subseteq [n], S_1 \subseteq \ldots \subseteq S_T, |S_t \setminus S_{t-1}| \leq C, \ \forall \ t \in [T], \right.$$

$$\left. \sum_{t \in [t_k:t_{k+1})} \lambda_t \cdot cs\left(w(S_t)\right) \geq (1-\epsilon) \cdot \alpha_k \ \forall \ k \in [K] \right\}.$$

One can show that

Theorem 3.5.3 *For any $\epsilon > 0$, there is an algorithm that returns a sequence of assortments $\mathcal{S} = (S_1, \cdots, S_T) \in \mathcal{F}_K^{Card-\epsilon}$ with expected revenue $\mathcal{R}(\mathcal{S}) \geq (1-\epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$, whose running time is $O\left(|\mathcal{I}|^{O(1)} \cdot n^{O(\frac{1}{\epsilon^2})}\right)$, where $|\mathcal{I}|$ denotes the size of the input.*

3.6 Numerical Experiments

In this section, we investigate the empirical performance of our algorithm under cardinality constraints. As a benchmark, we compare the performance of our algorithm against the product framing algorithm, *NEST Algorithm*, proposed by [89]. Our goal is to understand how good the revenue performance of our algorithm is, given that we are guaranteeing additional customer satisfaction performance which is hopefully higher than the solution given by pure revenue-maximizing algorithms. All experiments were conducted on a standard laptop with 16× 2.30GHz Intel Core i7 CPUs and 32GB of RAM. The algorithms were implemented using Matlab.

3.6.1 Modified Algorithm

To achieve better computational efficiency, we implement our algorithm in a modified way. The high-level idea of our PTAS is first constructing a relatively small candidate assortment space, then solving a dynamic programming over stages via backward induction

and exhaustive enumeration among the candidate space. The way we construct the candidate space is basically by categorizing products into weight classes according to their rounded weights, then enumerating enough representative combinations of product numbers for each class and picking those numbers of products following the descending $\rho-$order within each corresponding class. In our modified heuristic, we follow the high-level idea of the PTAS, but construct the candidate space in a simpler way. One major change to improve the computational efficiency is to make L independent of the problem instance, especially, we let $L = 2$ or 3. For example, when $L = 2$, there are three weight classes: $C_{q_{\max}}$, $C_{q_{\max}-1}$ and $\mathcal{I} \setminus (C_{q_{\max}} \cup C_{q_{\max}-1})$. For stage t, we consider all combinations of utilization triplet (k_1, k_2, k_3) such that $k_1 + k_2 + k_3 \leq t \cdot C$. Each triplet fully characterizes one assortment that picks k_1, k_2 and k_3 products following the descending $\rho-$order from weight classes $C_{q_{\max}}$, $C_{q_{\max}-1}$ and $\mathcal{I} \setminus (C_{q_{\max}} \cup C_{q_{\max}-1})$ respectively. Then stage t's candidate space $\mathcal{U}^{(t)}$ consists of all assortments that can be characterized by a feasible utilization triplet. Finally, we implement the dynamic programming formulated in Section 3.5.

3.6.2 Experiments with Expedia Search Data

In our first numerical experiment, we use the public data set of hotel search and booking records from Expedia platform ([101]). We follow the exact fitting method as [102].

Experimental Setup

We randomly choose 50 search queries from Expedia's Site 5 data, where each search query contains all search results of a customer's single hotel search. According to the statistics summary Table C.1 from [102], the average number of search results of a search query is around 25. We treat all search results under a search query as the total item set $\mathcal{I} = [n]$ where each product's revenue, i.e., hotel booking price, is given by the data, and MNL purchase

weight is fitted by the parametric model. Then we run our algorithm comparison for each selected query. We consider display products over three consecutive pages, i.e., $T = 3$; we assume there is only one customer division, i.e., $K = 1$; and customer type distribution vector $(\lambda_1, \lambda_2, \lambda_3)$ is uniformly distributed. We consider two scenarios of cardinality constraints, $C = 5$ and $C = 10$. Under a given scenario, we first run the *NEST* algorithm, and denote the solution by $\mathcal{S}^{NEST} = \left(S_1^{NEST}, S_2^{NEST}, S_3^{NEST}\right)$. Then for customer satisfaction constraints $\alpha_\beta = \beta \cdot cs\left(\mathcal{S}^{NEST}\right)$, $\beta \in \{0, 1.025, 1.05, 1.075, 1.1\}$, we run our modified algorithm respectively, and denote the solution derived from our algorithm by $\hat{\mathcal{S}}_\beta = \left(\hat{S}_{\beta 1}, \hat{S}_{\beta 2}, \hat{S}_{\beta 3}\right)$. Performance is measured by two metrics, the revenue ratio $\mathcal{R}\left(\hat{\mathcal{S}}_\beta\right)/\mathcal{R}\left(\mathcal{S}^{NEST}\right)$ and the customer satisfaction ratio $cs\left(\hat{\mathcal{S}}_\beta\right)/cs\left(\mathcal{S}^{NEST}\right)$. For each given factor β, we average the revenue ratio and the customer satisfaction ratio over the outputs of the search queries. Notice that for $\beta > 0$, our algorithm is not guaranteed to generate a feasible solution. When we average the performance metrics across the outputs of the search queries, we only account for search queries that generate feasible solutions. Table 3.1 summarizes the performance comparisons between our algorithm and the *NEST* algorithm, where $f_{feasible}$ represents the proportion of search queries that generate feasible solutions by our algorithm. Table 3.2 summarizes the running times of the tested algorithms (our algorithm is denoted by ADP). Since our algorithm's complexity does not depend on the constraint factor β, the average running time of our algorithm is measured upon the unconstrained problem instances.

Discussions

The high-level observation is that although the *NEST* algorithm runs much faster than our algorithm, under some instance settings, our algorithm attains a significant customer satisfaction improvement without sacrificing much revenue. This is especially true for a small cardinality constraint, C. It is interesting that our algorithm can achieve improvements in

| Algorithm Parameters | | | | $\left(\mathcal{R}\left(\hat{\mathcal{S}}_\beta\right)/\mathcal{R}\left(\mathcal{S}^{NEST}\right), cs\left(\hat{\mathcal{S}}_\beta\right)/cs\left(\mathcal{S}^{NEST}\right)\right)$ (%) | | | | | |
| | | | | $f_{feasible}$ (%) | | | | | |
C	L	ϵ	δ	$\beta = 0$	$\beta = 1$	$\beta = 1.025$	$\beta = 1.050$	$\beta = 1.075$	$\beta = 1.100$
5	2	0.6	0.2	(95.35, 107.88) 100	(99.59, 113.27) 76	(98.70, 114.93) 68	(99.00, 116.44) 60	(98.74, 118.02) 52	(99.00, 119.60) 44
5	2	0.6	0.6	(95.89, 110.09) 100	(99.91, 115.35) 78	(99.82, 117.04) 70	(100.26, 117.86) 66	(100.17, 119.13) 60	(100.34, 121.29) 50
5	2	0.6	1	(93.57, 108.95) 100	(97.74, 115.96) 72	(97.67, 117.34) 66	(97.76, 119.29) 58	(98.19, 120.23) 54	(98.18, 121.52) 48
5	2	0.8	0.2	(100.27, 109.64) 100	(102.23, 112.79) 86	(102.28, 113.10) 84	(102.88, 115.30) 68	(102.78, 116.13) 62	(102.28, 118.68) 46
5	2	1	0.2	(103.35, 111.09) 100	(104.00, 112.40) 92	(104.44, 113.43) 84	(105.11, 115.11) 72	(105.35, 116.46) 62	(105.61, 117.62) 54
5	3	0.6	0.2	(101.16, 111.40) 100	(101.33, 112.72) 92	(101.01, 114.59) 80	(100.82, 115.14) 76	(100.72, 116.79) 66	(100.62, 117.90) 60
5	3	0.8	0.2	(101.29, 112.52) 100	(101.39, 113.66) 92	(101.47, 114.89) 84	(101.48, 115.39) 80	(102.18, 116.99) 68	(102.22, 119.38) 54
5	3	1	0.2	(103.53, 110.66) 100	(103.73, 112.27) 88	(104.00, 114.02) 76	(104.04, 114.58) 72	(103.60, 115.66) 64	(103.23, 117.17) 54
10	2	0.6	0.2	(81.43, 89.47) 100	(96.17, 108.02) 40	(98.88, 111.21) 28	(101.74, 114.94) 20	(102.25, 115.94) 18	(103.04, 116.82) 16
10	2	0.8	0.2	(90.65, 97.71) 100	(99.95, 107.31) 56	(101.03, 110.65) 32	(100.82, 112.37) 26	(102.50, 114.05) 20	(104.50, 118.80) 10
10	2	1	0.2	(96.05, 101.84) 100	(100.77, 107.51) 66	(101.63, 110.52) 46	(102.70, 112.96) 34	(102.01, 115.74) 24	(103.55, 119.25) 16

Table 3.1: Performance Comparisons

both revenue and customer satisfaction for small C, given that the objective function of our algorithm is not revenue maximizing. The explanation is as follows. The *NEST* algorithm achieves the revenue upper bound for one of the stages, meaning that our algorithm can

Algorithm Parameters				Average Running Time (sec.)	
C	L	ϵ	δ	ADP	NEST
5	2	0.6	0.2	1.777	0.018
5	2	0.6	0.6	0.996	0.017
5	2	0.6	1	1.045	0.019
5	2	0.8	0.2	1.742	0.017
5	2	1	0.2	1.875	0.017
5	3	0.6	0.2	15.053	0.023
5	3	0.8	0.2	13.601	0.019
5	3	1	0.2	13.858	0.022
10	2	0.6	0.2	8.692	0.023
10	2	0.8	0.2	9.784	0.020
10	2	1	0.2	10.742	0.018

Table 3.2: Running Times

never attain a higher revenue at that particular stage. However, the *NEST* algorithm uses a greedy strategy to construct other stages, which may significantly harm the overall revenue performance. Recall that the average number of search results for each search query is around 25. When $C = 10$, $T = 3$, the overall revenue performance is likely to be only determined by the first one or the first two stages (i.e., not adding more products in later stages), in which case the *NEST* algorithm does not need to resort to the greedy construction much, the revenue performance, in turn, is more likely to be good. Another interesting observation is that larger ϵ may lead to better revenue and customer satisfaction performance. This is because, in our algorithm modification, we set $L = 2$ or 3, then when ϵ is small, the heaviest L weight classes may not contain enough products. Especially, they may exclude the "good" products that have small weights but large ρ values. But still surprisingly, even when we have ϵ to be 1, our algorithm can achieve impressive performance compared with the *NEST* algorithm.

3.6.3   Experiments with Synthetic Data

We also conduct the comparisons between our algorithm and the *NEST* algorithm using synthetic data.

Generative Model

We randomly generate 99 products whose MNL weights take value from $[1 \cdot scale, 2.5 \cdot scale]$, where we consider three scales of the problem with $scale = 0.01, 0.1, 1$. Specifically, for each interval among $[1 \cdot scale, 1.5 \cdot scale]$, $[1.5 \cdot scale, 2 \cdot scale]$, and $[2 \cdot scale, 2.5 \cdot scale]$, we uniformly generate 33 products. Especially, with $scale = 0.01$, the MNL weights are roughly at the same scale as the fitted results based upon the Expedia data. As a comparison, we consider another scale with $scale = 0.1$. We also randomly generate revenue for each product, taking value from $[100, 250]$. Specifically, to mimic the pricing effect on customers' willingness to pay, we uniformly generate revenues from $[200, 250]$ for the products with MNL weights picked from $[1 \cdot scale, 1.5 \cdot scale]$; revenues from $[150, 200]$ for the products with MNL weights picked from $[1.5 \cdot scale, 2 \cdot scale]$; and revenues from $[100, 150]$ for the products with MNL weights picked from $[2 \cdot scale, 2.5 \cdot scale]$. We still let $T = 3$, and consider three cardinality scenarios with $C = 10$, $C = 15$, and $C = 20$. After extensive trials, to well balance the performance and running time, we fix $\epsilon = 0.6$ and $\delta = 0.4$. Like before, for each problem instance, we measure the comparisons by the revenue ratio and the customer satisfaction ratio. For each given setup parameters combination $(scale, C)$, we generate 20 instances and average the two ratio metrics across the instances where our algorithm generates feasible solutions. Table 3.3 summarizes the performance comparisons between our algorithm and the *NEST* algorithm, and Table 3.4 summarizes the running times of the tested algorithms.

Algorithm Parameters					$\left(\mathcal{R}\left(\hat{\mathcal{S}}_\beta\right)/\mathcal{R}\left(\mathcal{S}^{NEST}\right), cs\left(\hat{\mathcal{S}}_\beta\right)/cs\left(\mathcal{S}^{NEST}\right)\right)$ (%) $f_{feasible}$ (%)			
C	L	ϵ	δ	$scale$	$\beta = 0$	$\beta = 1$	$\beta = 1.1$	$\beta = 1.2$
10	2	0.6	0.4	0.01	$(101.36, 123.37)$ 100	$(101.36, 123.37)$ 100	$(101.36, 123.37)$ 100	$(103.36, 126.67)$ 65
10	2	0.6	0.4	0.1	$(86.09, 115.56)$ 100	$(86.09, 115.56)$ 100	$(85.36, 115.82)$ 100	$(84.86, 121.43)$ 20
15	2	0.6	0.4	0.01	$(98.02, 122.72)$ 100	$(98.02, 122.72)$ 100	$(98.02, 122.72)$ 100	$(98.90, 127.15)$ 60
15	2	0.6	0.4	0.1	$(85.67, 111.38)$ 100	$(85.67, 111.38)$ 100	$(81.63, 112.82)$ 95	$-$
20	2	0.6	0.4	0.01	$(100.02, 120.81)$ 100	$(100.02, 120.81)$ 100	$(100.02, 120.81)$ 100	$(99.60, 125.17)$ 60
20	2	0.6	0.4	0.1	$(81.75, 109.70)$ 100	$(81.75, 109.70)$ 100	$(84.34, 114.57)$ 30	$-$

Table 3.3: Performance Comparisons

Algorithm Parameters					Average Running Time (sec.)	
C	L	ϵ	δ	$scale$	ADP	NEST
10	2	0.6	0.4	0.01	12.894	0.025
10	2	0.6	0.4	0.1	7.367	0.023
15	2	0.6	0.4	0.01	50.526	0.024
15	2	0.6	0.4	0.1	25.976	0.022
20	2	0.6	0.4	0.01	140.327	0.031
20	2	0.6	0.4	0.1	68.920	0.023

Table 3.4: Running Times

Discussions

The high-level observation is like before. Our algorithm may significantly improve customer satisfaction while keeping a decent revenue performance under some instance settings. Our algorithm performs very well compared with the *NEST* algorithm in problem instances that are at a similar scale as the practical setting ($scale = 0.01$). However, our algorithm's revenue performance may lose more when the MNL weights of products become

larger ($scale = 0.1$). There are two reasons for this phenomenon. On the one hand, when all products' weights are very small, the total weight of picked products does not matter much. The revenue performance is mainly determined by the total ρ value of the assortment. Therefore the optimal revenue-maximizing assortment tends to pick as many products as possible following the descending ρ−order. This simple structure is aligned with the ρ−order picking idea of our algorithm (especially when the weight class division is coarse, i.e., ϵ is large). On the other hand, when the MNL weights are large, the optimal assortment tends to pick the top revenue products without including too many products. Therefore, the cardinality constraint is less likely to be binding for later stages (consider the extreme case when the weights tend to be infinite, the optimal revenue-maximizing solution is to pick the product with the highest revenue in the first stage and not add any products in later stages). As we explained before, this is when the *NEST* algorithm will perform very well in terms of revenue.

3.7 Future Work

Our work leaves a few unanswered questions. First, one natural question of DISPLAY-OPT-K-CARD is whether there is an efficient algorithm to find a proxy that strictly meets both the satisfaction constraints and the cardinality constraints. A second open question is whether an FPTAS exists for the setting where $T = O(1)$, as our proof that DISPLAY-OPT-K is strongly NP-Hard requires the use of $\Omega(n)$ stages, and hence does not rule out such a result.

REFERENCES

[1] Guofang Huang, Hong Luo, and Jing Xia. Invest in information or wing it? a model of dynamic pricing with seller learning. Management Science, 2018.

[2] Robin Mason and Juuso Välimäki. Learning about the arrival of sales. Journal of Economic Theory, 146(4):1699–1711, 2011.

[3] Omar Besbes, Dan A Iancu, and Nikolaos Trichakis. Dynamic pricing under debt: Spiraling distortions and efficiency losses. Management Science, 64(10):4572–4589, 2018.

[4] Yufeng Cao, Anton Kleywegt, and He Wang. Dynamic pricing for truckload transportation marketplaces. Available at SSRN 3700227, 2020.

[5] Arnoud V den Boer. How long does it take to sell a product? Available at SSRN 3943319, 2021.

[6] Guillermo Gallego and Garrett Van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. Management science, 40(8):999–1020, 1994.

[7] Guillermo Gallego and Garrett Van Ryzin. A multiproduct dynamic pricing problem and its applications to network yield management. Operations research, 45(1):24–41, 1997.

[8] Constantinos Maglaras and Joern Meissner. Dynamic pricing strategies for multiproduct revenue management problems. Manufacturing & Service Operations Management, 8(2):136–148, 2006.

[9] Lingxiu Dong, Panos Kouvelis, and Zhongjun Tian. Dynamic pricing and inventory control of substitute products. Manufacturing & Service Operations Management, 11(2):317–339, 2009.

[10] Yalçın Akçay, Harihara Prasad Natarajan, and Susan H Xu. Joint dynamic pricing of multiple perishable products under consumer choice. Management Science, 56(8):1345–1361, 2010.

[11] Philipp W Keller, Retsef Levi, and Georgia Perakis. Efficient formulations for pricing under attraction demand models. Mathematical Programming, 145(1-2):223–261, 2014.

[12] Stefanus Jasin. Reoptimization and self-adjusting price control for network revenue management. Operations Research, 62(5):1168–1178, 2014.

[13] Qi Chen, Stefanus Jasin, and Izak Duenyas. Real-time dynamic pricing with minimal and flexible price adjustment. Management Science, 62(8):2437–2455, 2016.

[14] Yining Wang and He Wang. Constant regret re-solving heuristics for price-based revenue management. arXiv preprint arXiv:2009.02861, 2020.

[15] Yossi Aviv and Amit Pazgal. Pricing of short life-cycle products through active learning. Under revision for Management Science, 2002.

[16] Yossi Aviv and Amit Pazgal. A partially observed markov decision process for dynamic pricing. Management science, 51(9):1400–1416, 2005.

[17] Victor F Araman and René Caldentey. Dynamic pricing for nonperishable products with demand learning. Operations research, 57(5):1169–1188, 2009.

[18] Vivek F Farias and Benjamin Van Roy. Dynamic pricing with a prior on market response. Operations Research, 58(1):16–29, 2010.

[19] J Michael Harrison, N Bora Keskin, and Assaf Zeevi. Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. Management Science, 58(3):570–586, 2012.

[20] Arnoud V den Boer. Dynamic pricing with multiple products and partially specified demand distribution. Mathematics of operations research, 39(3):863–888, 2014.

[21] Josef Broder and Paat Rusmevichientong. Dynamic pricing under a general parametric choice model. Operations Research, 60(4):965–980, 2012.

[22] Arnoud V den Boer and Bert Zwart. Simultaneously learning and optimizing using controlled variance pricing. Management science, 60(3):770–783, 2014.

[23] N Bora Keskin and Assaf Zeevi. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. Operations Research, 62(5):1142–1167, 2014.

[24] Mengzhenyu Zhang, Hyun-Soo Ahn, and Joline Uichanco. Data-driven pricing for a new product. Available at SSRN 3545574, 2020.

[25] Paat Rusmevichientong, Benjamin Van Roy, and Peter W Glynn. A nonparametric approach to multiproduct pricing. Operations Research, 54(1):82–98, 2006.

[26] Omar Besbes and Assaf Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. Operations Research, 57(6):1407–1420, 2009.

[27] Omar Besbes and Assaf Zeevi. Blind network revenue management. Operations research, 60(6):1537–1550, 2012.

[28] Zizhuo Wang, Shiming Deng, and Yinyu Ye. Close the gaps: A learning-while-doing algorithm for single-product revenue management problems. Operations Research, 62(2):318–331, 2014.

[29] N Bora Keskin and Assaf Zeevi. On incomplete learning and certainty-equivalence control. Operations Research, 66(4):1136–1167, 2018.

[30] Omar Besbes and Assaf Zeevi. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. Management Science, 61(4):723–739, 2015.

[31] Mila Nambiar, David Simchi-Levi, and He Wang. Dynamic learning and pricing with model misspecification. Management Science, 65(11):4980–5000, 2019.

[32] Maxime C Cohen, Georgia Perakis, and Robert S Pindyck. A simple rule for pricing with limited knowledge of demand. Management Science, 2020.

[33] Wang Chi Cheung, David Simchi-Levi, and He Wang. Dynamic pricing and demand learning with limited price experimentation. Operations Research, 65(6):1722–1731, 2017.

[34] Daniel Russo and Benjamin Van Roy. Learning to optimize via posterior sampling. Mathematics of Operations Research, 39(4):1221–1243, 2014.

[35] Daniel Russo and Benjamin Van Roy. An information-theoretic analysis of thompson sampling. The Journal of Machine Learning Research, 17(1):2442–2471, 2016.

[36] Shipra Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In Conference on learning theory, pages 39–1. JMLR Workshop and Conference Proceedings, 2012.

[37] Shipra Agrawal, Vashist Avadhanula, Vineet Goyal, and Assaf Zeevi. Thompson sampling for the mnl-bandit. arXiv preprint arXiv:1706.00977, 2017.

[38] Min-hwan Oh and Garud Iyengar. Thompson sampling for multinomial logit contextual bandits. In Advances in Neural Information Processing Systems, pages 3151–3161, 2019.

[39] Nima Hamidi and Mohsen Bayati. On worst-case regret of linear thompson sampling. arXiv preprint arXiv:2006.06790, 2020.

[40] Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. Advances in neural information processing systems, 24:2249–2257, 2011.

[41] Arnoud V den Boer and Bert Zwart. Dynamic pricing and learning with finite inventories. Operations research, 63(4):965–978, 2015.

[42] Kris Johnson Ferreira, David Simchi-Levi, and He Wang. Online network revenue management using thompson sampling. Operations research, 66(6):1586–1602, 2018.

[43] Qi Chen, Stefanus Jasin, and Izak Duenyas. Joint learning and optimization of multi-product pricing with finite resource capacity and unknown demand parameters. Operations Research, 69(2):560–573, 2021.

[44] Warren B Powell. A unified framework for stochastic optimization. European Journal of Operational Research, 275(3):795–821, 2019.

[45] C Hipp and R Michel. On the bernstein-v. mises approximation of posterior distributions. The Annals of Statistics, pages 972–980, 1976.

[46] Hau L Lee, Venkata Padmanabhan, and Seungjin Whang. Information distortion in a supply chain: The bullwhip effect. Management Science, 43(4):546–558, 1997.

[47] Yossi Sheffi. The resilient enterprise: overcoming vulnerability for competitive advantage. Zone Books, 2007.

[48] Hau L Lee and Seungjin Whang. Information sharing in a supply chain. International Journal of Manufacturing Technology and Management, 1(1):79–93, 2000.

[49] Yossi Sheffi. The power of resilience: How the best companies manage the unexpected. MIT Press, 2015.

[50] Lode Li. Information sharing in a supply chain with horizontal competition. Management Science, 48(9):1196–1212, 2002.

[51] Hongtao Zhang. Vertical information exchange in a supply chain with duopoly retailers. Production and Operations Management, 11(4):531–546, 2002.

[52] Tim Adam, Sudipto Dasgupta, and Sheridan Titman. Financial constraints, competition, and hedging in industry equilibrium. The Journal of Finance, 62(5):2445–2473, 2007.

[53] William Novshek and Hugo Sonnenschein. Fulfilled expectations cournot duopoly with information acquisition and release. The Bell Journal of Economics, pages 214–218, 1982.

[54] Richard N Clarke. Collusion and the incentives for information sharing. The Bell Journal of Economics, pages 383–394, 1983.

[55] Xavier Vives. Duopoly information equilibrium: Cournot and bertrand. Journal of Economic Theory, 34(1):71–94, 1984.

[56] Lode Li. Cournot oligopoly with information sharing. The RAND Journal of Economics, pages 521–536, 1985.

[57] Esther Gal-Or. Information sharing in oligopoly. Econometrica: Journal of the Econometric Society, pages 329–343, 1985.

[58] Esther Gal-Or. Information transmission—cournot and bertrand equilibria. The Review of Economic Studies, 53(1):85–92, 1986.

[59] Michael Raith. A general model of information sharing in oligopoly. Journal of Economic Theory, 71(1):260–288, 1996.

[60] Lode Li and Hongtao Zhang. Confidentiality and information sharing in supply chain coordination. Management Science, 54(8):1467–1481, 2008.

[61] Esther Gal-Or, Tansev Geylani, and Anthony J Dukes. Information sharing in a channel with partially informed retailers. Marketing Science, 27(4):642–658, 2008.

[62] Albert Y Ha and Shilu Tong. Contracting and information sharing under supply chain competition. Management Science, 54(4):701–715, 2008.

[63] Albert Y Ha, Shilu Tong, and Hongtao Zhang. Sharing demand information in competing supply chains with production diseconomies. Management Science, 57(3):566–581, 2011.

[64] Noam Shamir and Hyoduk Shin. Public forecast information sharing in a market with competing supply chains. Management Science, 62(10):2994–3022, 2016.

[65] Albert Y Ha, Quan Tian, and Shilu Tong. Information sharing in competing supply chains with production cost reduction. Manufacturing & Service Operations Management, 19(2):246–262, 2017.

[66] Hau L Lee, Kut C So, and Christopher S Tang. The value of information sharing in a two-level supply chain. Management Science, 46(5):626–643, 2000.

[67] Gérard P Cachon and Martin A Lariviere. Contracting to assure supply: How to share demand forecasts in a supply chain. Management Science, 47(5):629–646, 2001.

[68] Vishal Gaur, Avi Giloni, and Sridhar Seshadri. Information sharing in a supply chain under arma demand. Management Science, 51(6):961–969, 2005.

[69] Özalp Özer, Yanchong Zheng, and Kay-Yut Chen. Trust in forecast information sharing. Management Science, 57(6):1111–1137, 2011.

[70] Ruomeng Cui, Gad Allon, Achal Bassamboo, and Jan A Van Mieghem. Information sharing in supply chains: An empirical and theoretical valuation. Management Science, 61(11):2803–2824, 2015.

[71] Chuan He, Johan Marklund, and Thomas Vossen. Research note—vertical information sharing in a volatile market. Marketing Science, 27(3):513–530, 2008.

[72] Guangwen Kong, Sampath Rajagopalan, and Hao Zhang. Revenue sharing and information leakage in a supply chain. Management Science, 59(3):556–572, 2013.

[73] Yiwei Chen and Özalp Özer. Supply chain contracts that prevent information leakage. Management Science, 65(12):5619–5650, 2019.

[74] Julia Miyaoka and Warren H Hausman. How improved forecasts can degrade decentralized supply chains. Manufacturing & Service Operations Management, 10(3):547–562, 2008.

[75] Terry A Taylor and Wenqiang Xiao. Does a manufacturer benefit from selling to a better-forecasting retailer? Management Science, 56(9):1584–1598, 2010.

[76] Ling-Chieh Kung and Ying-Ju Chen. Impact of reseller's and sales agent's forecasting accuracy in a multilayer supply chain. Naval Research Logistics (NRL), 61(3):207–222, 2014.

[77] Thunyarat Amornpetchkul, Izak Duenyas, and Özge Şahin. Mechanisms to induce buyer forecasting: Do suppliers always benefit from better forecasting? Production and Operations Management, 24(11):1724–1749, 2015.

[78] Meng Li and Nicholas C Petruzzi. Demand uncertainty reduction in decentralized supply chains. Production and Operations Management, 26(1):156–161, 2017.

[79] Kenneth A Froot, David S Scharfstein, and Jeremy C Stein. Risk management: Coordinating corporate investment and financing policies. The Journal of Finance, 48(5):1629–1658, 1993.

[80] Shehzad L Mian. Evidence on corporate hedging policy. Journal of Financial and quantitative Analysis, pages 419–439, 1996.

[81] Robert CW Fok, Carolyn Carroll, and Ming C Chiou. Determinants of corporate hedging and derivatives: A revisit. Journal of Economics and Business, 49(6):569–585, 1997.

[82] Panos Kouvelis and Danko Turcic. Supporting operations with financial hedging: Cash hedging vs. cost hedging in an automotive industry. Production and Operations Management, 30(3):738–749, 2021.

[83] Kenneth A Froot, David S Scharfstein, and Jeremy C Stein. A framework for risk management. Journal of Applied Corporate Finance, 7(3):22–33, 1994.

[84] Alexei Alexandrov. When should firms expose themselves to risk? Management Science, 61(12):3001–3008, 2015.

[85] Panos Kouvelis, Xiaole Wu, and Yixuan Xiao. Cash hedging in a supply chain. Management Science, 65(8):3928–3947, 2019.

[86] Guillermo Gallego, Huseyin Topaloglu, et al. Revenue management and pricing analytics, volume 209. Springer, 2019.

[87] Mika Sumida, Guillermo Gallego, Paat Rusmevichientong, Huseyin Topaloglu, and James Davis. Revenue-utility tradeoff in assortment optimization under the multinomial logit model with totally unimodular constraints. Technical report, Tech. rep., Cornell University, Ithaca, NY, 2019.

[88] Ali Aouad and Danny Segev. Display optimization for vertically differentiated locations under multinomial logit preferences. Management Science, 2020.

[89] Guillermo Gallego, Anran Li, Van-Anh Truong, and Xinshang Wang. Approximation algorithms for product framing and pricing. Operations Research, 68(1):134–160, 2020.

[90] R. D. Luce. Individual choice behavior: A theoretical analysis. Frontiers in Econometrics, 2:105–142, 1959.

[91] Kalyan Talluri and Garrett Van Ryzin. Revenue management under a general discrete choice model of consumer behavior. Management Science, 50(1):15–33, 2004.

[92] Paat Rusmevichientong, Zuo-Jun Max Shen, and David B. Shmoys. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. Operations Research, 58(6):1666–1680, 2010.

[93] Antoine Désir, Vineet Goyal, and Jiawei Zhang. Near-optimal algorithms for capacity constrained assortment optimization. Available at SSRN, 2543309, 2014.

[94] James M Davis, Huseyin Topaloglu, and David P Williamson. Assortment optimization over time. Operations Research Letters, 43(6):608–611, 2015.

[95] Daniel Bienstock, Jay Sethuraman, and Chun Ye. Approximation algorithms for the incremental knapsack problem via disjunctive programming. arXiv preprint arXiv:1311.4563, 2013.

[96] Yuri Faenza and Igor Malinovic. A ptas for the time-invariant incremental knapsack problem. In International Symposium on Combinatorial Optimization, pages 157–169. Springer, 2018.

[97] Federico Della Croce, Ulrich Pferschy, and Rosario Scatamacchia. On approximating the incremental knapsack problem. Discrete Applied Mathematics, 264:26–42, 2019.

[98] Ali Aouad and Danny Segev. An approximate dynamic programming approach to the incremental knapsack problem. arXiv preprint arXiv:2010.07633, 2020.

[99] Yuri Faenza, Danny Segev, and Lingyi Zhang. Approximation algorithms for the generalized incremental knapsack problem. arXiv preprint arXiv:2009.07248, 2020.

[100] Zuguang Gao, John R Birge, and Varun Gupta. Approximation schemes for multi-period binary knapsack problems. In International Computer Science Symposium in Russia, pages 131–146. Springer, 2021.

[101] Kaggle. Personalize expedia hotel searches—icdm 2013, 2013. https://www.kaggle.com/c/expedia-personalized-sort/data.

[102] Jacob Feldman and Danny Segev. The multinomial logit model with sequential offerings: Algorithmic frameworks for product recommendation displays. Operations Research, 2022.

[103] David G Luenberger and Yinyu Ye. Linear and Nonlinear Programming Fourth Edition. Springer, 2016.

[104] Valentin Vladimirovich Petrov. Sums of independent random variables, volume 82. Springer Science & Business Media, 2012.

[105] Michael R Garey and David S Johnson. Computers and intractability. A Guide to the, 1979.

APPENDICES

A. Appendix for Chapter 1

Appendix: Proofs

Proof Proof for Lemma 1.3.3. We notice that the objective function could be written as $\phi(p)(p - A) + A$ which after variable substitution $\widetilde{p} = p - A$ could be further transformed into $\widetilde{\phi}(\widetilde{p})\widetilde{p} + A$ where $\widetilde{\phi}(\widetilde{p}) = \left(1 - e^{-\lambda \widetilde{q}(\widetilde{p})}\right)$ and $\widetilde{q}(\widetilde{p}) = q\left(\widetilde{p} + A\right) = q(p)$. Whence $q'(p)/q(p)$ is decreasing in p is equivalent to $\widetilde{q}'(\widetilde{p})/\widetilde{q}(\widetilde{p})$ is decreasing in $\widetilde{p}$. In other words, $q(\cdot) \in \mathcal{Q}$ is equivalent to $\widetilde{q}(\cdot) \in \mathcal{Q}$. Therefore, to prove Lemma 1.3.3, it is equivalent to rove the following statement: If $q(\cdot) \in \mathcal{Q}$, then for fixed λ, $\pi(p; \lambda) = \left(1 - e^{-\lambda q(p)}\right) p$ is unimodal in p.

From Assumption 1.3.1, we know that optimal solution is finite. Thereby to prove the unimodularity of $\pi(\cdot; \lambda)$, it is equivalent to showing that all the points where first-order condition holds have to be local maximum (i.e., negative second-order derivative), which implies that there is only one local maximum, i.e., global maximum. We can take the first-order and second-order derivatives of $\pi(p; \lambda)$:

$$
\begin{cases}
\pi'(p; \lambda) = 1 + e^{-\lambda q(p)}\left(-1 + \lambda p q'(p)\right), & \text{(A.1)} \\
\pi''(p; \lambda) = \lambda e^{-\lambda q(p)}\left(2q'(p) - \lambda p \left(q'(p)\right)^2 + p q''(p)\right). & \text{(A.2)}
\end{cases}
$$

We need to prove that when (A.1) is zero, (A.2) is negative. When first-order condition holds, we have $p = -\left(e^{\lambda q(p)} - 1\right)/\lambda q'(p)$. With this condition, proving (A.2) is negative is equivalent to proving

$$
\left(e^{\lambda q(p)} + 1\right) q'(p) < \frac{\left(e^{\lambda q(p)} - 1\right) q''(p)}{\lambda q'(p)}.
$$

Because we know that $\lambda > 0$ and $q'(p) < 0$ (from Assumption 1.3.2), it is equivalent to proving

$$\lambda \frac{\left(e^{\lambda q(p)} + 1\right)}{\left(e^{\lambda q(p)} - 1\right)} > \frac{q''(p)}{(q'(p))^2}.$$

We multiply both sides by $q(p)$ and then minus 1:

$$\lambda q(p) \frac{\left(e^{\lambda q(p)} + 1\right)}{\left(e^{\lambda q(p)} - 1\right)} - 1 > \frac{q''(p)q(p) - (q'(p))^2}{(q'(p))^2}.$$

Assumption 1.3.2 implies that $q''(p)q(p) - (q'(p))^2 < 0$, therefore the right-hand side of the above inequality is negative. On the other hand, the left-hand side is always positive. To see this, we have

$$\lambda q(p) \frac{\left(e^{\lambda q(p)} + 1\right)}{\left(e^{\lambda q(p)} - 1\right)} - 1 > 0 \Leftrightarrow (\lambda q(p) + 1) + e^{\lambda q(p)}(\lambda q(p) - 1) > 0.$$

Define function $f(x) = (x + 1) + e^x(x - 1)$, we have $f'(x) = 1 + e^x x > 0$ for $x > 0$ and $f(0) = 0$, therefore the right-hand of the above is true.

The proof is completed. ∎

Lemma A.0.1 *Define*

$$F(\lambda, A) = \max_p \left\{ \pi(p; \lambda, A) \right\} = \max_p \left\{ \left(1 - e^{-\lambda q(p)}\right) p + e^{-\lambda q(p)} \cdot A \right\},$$

where $\lambda, A > 0$ and $q(\cdot) \in \mathcal{Q}$, then (i). $p^*(\lambda, A) \geq A$ and $\lim_{\lambda \to \infty} p^*(\lambda, A) = \infty$, $\lim_{A \to \infty} p^*(\lambda, A) = \infty$, (ii). $F(\lambda, A)$ is increasing concave in λ and $\lim_{\lambda \to \infty} F(\lambda, A) = \infty$, (iii). $F(\lambda, A)$ is increasing convex in A (which implies $\lim_{A \to \infty} F(\lambda, A) = \infty$) and (iv). $\frac{\partial^2 F(\lambda, A)}{\partial \lambda \partial A} < 0$, i.e., $F(\lambda, A)$ is submodular.

Proof Proof. We notice that the objective function could be written as $\phi(p)(p - A) + A$ which after variable substitution $\widetilde{p} = p - A$ could be further transformed into $\widetilde{\phi}(\widetilde{p})\widetilde{p} + A$

[126]

where $\widetilde{\phi}(\widetilde{p}) = \left(1 - e^{-\lambda \widetilde{q}(\widetilde{p})}\right)$ and $\widetilde{q}(\widetilde{p}) = q\,(\widetilde{p} + A) = q(p)$. It is straightforward to see that the optimal price $p^* > A$ which is the first part of the lemma. The limiting behaviors of p^* in part (i) are also straightforward.

Now we look at part (ii) of the lemma. Let

$$\widetilde{F}\left(\lambda\right) = \max_{p} \left\{\phi(p)p\right\},$$

where $\phi(p) = \left(1 - e^{-\lambda q(p)}\right)$ and $q(\cdot)$ is log-concave, then $\widetilde{F}(\lambda)$ is increasing concave in λ.

The following theorem from [103] is useful (see [103] Ch.11.7):

Theorem A.0.2 *(Sensitivity Theorem) Let f, $\boldsymbol{h} \in C^2$ and consider the family of problems*

$$\max_{\boldsymbol{x}} \ f(\boldsymbol{x})$$

$$s.t. \ \boldsymbol{h}\left(\boldsymbol{x}\right) = \boldsymbol{c}. \tag{A.3}$$

Suppose for $\boldsymbol{c} = \boldsymbol{0}$ there is a local solution $\boldsymbol{x}^*$ that is a regular point and that, together with its associated Lagrange multiplier vector $\boldsymbol{\lambda}$, satisfies the second-order sufficiency conditions for a strict local maximum. Then for every $\boldsymbol{c} \in E^m$ in a region containing $\boldsymbol{0}$ there is an $\boldsymbol{x}(\boldsymbol{c})$ depending continuously on $\boldsymbol{c}$, such that $\boldsymbol{x}(\boldsymbol{0}) = \boldsymbol{x}^*$ and such that $\boldsymbol{x}(\boldsymbol{c})$ is a local maximum of (A.3). *Furthermore,*

$$\nabla_{\boldsymbol{c}} f\left(\boldsymbol{x}(\boldsymbol{c})\right)|_{\boldsymbol{c}=\boldsymbol{0}} = \boldsymbol{\lambda}.$$

Define $z = \lambda q(p)$, then we can write our optimization problem in the similar form to (A.3):

$$\widetilde{F}(\lambda + \delta_\lambda) = \max_{p,z} \ \left(1 - e^{-z}\right)p$$

$$s.t. \ z/q\left(p\right) - \lambda = \delta_\lambda.$$

To avoid confusion of notations, we use μ to denote the Lagrange multiplier for the above problem. It is not difficult to show that at solution (p^*, z^*), the second-order sufficiency conditions for a strict local maximum are satisfied (see [103] Ch.11 for the definition of second-order sufficiency conditions). From Theorem A.0.2, we know that $\nabla_{\delta_\lambda} \widetilde{F} (\lambda + \delta_\lambda) |_{\delta_\lambda = 0} = \mu(\lambda)$, therefore to prove $\widetilde{F}(\lambda)$ is increasing concave in λ, it is equivalent to proving that $\mu(\lambda) > 0$ and $\mu'(\lambda) < 0$. We can write down the Karush-Kuhn-Tucker conditions at the optimal solution (p^*, z^*) with $\delta_\lambda = 0$ (with slight abuse of notations, we drop $*$ from the superscript of optimal solutions):

$$
\begin{cases}
\left(1 - e^{-z}\right) + \mu \cdot \dfrac{z q'(p)}{q^2(p)} = 0, \\[2mm]
e^{-z} p - \mu \cdot \dfrac{1}{q(p)} = 0, \\[2mm]
\dfrac{z}{q(p)} - \lambda = 0,
\end{cases}
\tag{A.4}
$$

which leads to

$$
\mu(\lambda) = \frac{p(\lambda) q\left(p(\lambda)\right)}{1 - \lambda p(\lambda) q'(p(\lambda))}.
\tag{A.5}
$$

It is straightforward to see that the optimal price cannot be negative, and we also have $q'(\cdot) > 0$ from assumption, therefore we have shown that $\mu > 0$. Next we prove that $\mu(\cdot)$ is a decreasing function. From the above expression of μ, we have

$$
\begin{aligned}
\frac{d\mu(\lambda)}{d\lambda} = {} & \frac{1}{\left(1 - \lambda p(\lambda) q'\left(p(\lambda)\right)\right)^2} \times \\
& \left(p(\lambda) p'(\lambda) q'(p(\lambda)) \left[1 - \lambda p(\lambda) q'\left(p(\lambda)\right)\right] \right. \\
& \left. + q(p(\lambda)) \left\{ p^2(\lambda) q'(p(\lambda)) + p'(\lambda) \left[1 + \lambda p^2(\lambda) q''(p(\lambda))\right] \right\} \right),
\end{aligned}
\tag{A.6}
$$

[128]

and we want to show that the above derivative is negative, which is equivalent to show that the numerator of the above result is negative. After algebra manipulation, we can write the numerator as

$$\left\{ p'(\lambda) \left[p(\lambda) q'(p(\lambda)) + q(p(\lambda)) \right] + p^2(\lambda) q(p(\lambda)) q'(p(\lambda)) \right\}$$
$$+ \lambda p^2(\lambda) p'(\lambda) \left[q(p(\lambda)) q''(p(\lambda)) - (q'(p(\lambda)))^2 \right]. \tag{A.7}$$

Given that we have $q'(\cdot) < 0$ and $q(p)q''(p) - (q'(p))^2 < 0$ (equivalent to that $q'(p)/q(p)$ is decreasing in p), to prove (A.7) is negative, it is sufficient to have $p'(\lambda) > 0$ and $p(\lambda)q'(p(\lambda)) + q(p(\lambda)) < 0$. Proving the two lemmas below will complete the proof.

Lemma A.0.3 *Optimal price is nondecreasing in the potential demand rate, i.e., $p'(\lambda) > 0$.*

Proof Proof. It is sufficient to prove that near the optimal solution $p^*(\lambda)$, the objective function $\pi(p, \lambda) = \phi(p)p$ is supermodular in (p, λ). Indeed, we have

$$\frac{\partial^2 \pi(p, \lambda)}{\partial p \partial \lambda} = e^{-\lambda q(p)} \left(q(p) + p \left(1 - \lambda q(p) \right) q'(p) \right),$$

and on the other hand, first-order condition gives us

$$1 - e^{-\lambda q(p)} \left(1 - p \lambda q'(p) \right) = 0.$$

Therefore, along the curve of $p^*(\lambda)$, we have

$$\frac{\partial^2 \pi(p, \lambda)}{\partial p \partial \lambda} = e^{-\lambda q(p)} \left(p + q(p) e^{\lambda q(p)} \right) > 0,$$

and because of continuity, the cross derivative of $\pi(p, \lambda)$ is in turn positive around the solution $p^*(\lambda)$. Therefore, $p^*(\lambda)$ is increasing in λ. ∎

[129]

Lemma A.0.4 *For optimal price $p(\lambda)$, we have $p(\lambda)q'(p(\lambda)) + q(p(\lambda)) < 0$.*

Proof Proof. From first-order condition, we have

$$1 - e^{-\lambda q(p(\lambda))} \left(1 - p(\lambda)\lambda q'(p(\lambda))\right) = 0,$$

which leads to $p(\lambda) = \frac{1 - e^{\lambda q(p(\lambda))}}{\lambda q'(p(\lambda))}$. Therefore,

$$p(\lambda)q'(p(\lambda)) + q(p(\lambda)) < 0 \;\Leftrightarrow\; \lambda q(p(\lambda)) < e^{\lambda q(p(\lambda))} - 1.$$

The latter inequality holds because $e^x - 1 > x$ for $x \in \mathbb{R}$. The proof is completed. ■

Now we have completed the proof for the second part of Lemma A.0.1, then we show that $F(\lambda, A)$ is increasing convex in A. To do that, we write the optimization problem as:

$$F(\lambda, A + \delta_A) = \max_{p,z} \; \left(1 - e^{-\lambda q(p)}\right) p + e^{-\lambda q(p)} z,$$

$$s.t. \; z - A = \delta_A.$$

Similar as before, from Theorem A.0.2, we know that $\nabla_{\delta_A} F\left(\lambda, A + \delta_A\right)|_{\delta_A = 0} = \mu\left(A\right)$, therefore to prove $F(\lambda, A)$ is increasing convex in A, it is equivalent to proving that $\mu(A) > 0$ and $\mu'(A) > 0$. Karush-Kuhn-Tucker conditions at the optimal solution (p^*, z^*) with $\delta_A = 0$ gives us

$$\mu\left(A\right) = e^{-\lambda q(p(A))} > 0, \tag{A.8}$$

whence we have already shown that $F\left(\lambda, A\right)$ is increasing in A. On the other hand, given the expression of $\mu\left(A\right)$, we have

$$\frac{d\mu\left(A\right)}{dA} = -e^{-\lambda q(p(A))} \lambda q'\left(p\left(A\right)\right) p'\left(A\right).$$

To show that $d\mu(A)/dA > 0$, we only need to show that $p'(A) > 0$. The following lemma completes the proof:

Lemma A.0.5 *Optimal price is nondecreasing in the future value, i.e., $p'(A) > 0$.*

Proof Proof. Similar to the proof for Lemma A.0.3, it is sufficient to prove that near the optimal solution $p^*(A)$, the objective function $\pi(p, A) = \left(1 - e^{-\lambda q(p)}\right) p + e^{-\lambda q(p)} A$ is supermodular in (p, A). Indeed, we have

$$\frac{\partial^2 \pi(p, A)}{\partial p \partial A} = -e^{-\lambda q(p)} \lambda q'(p) > 0,$$

Therefore, $p(A)$ is increasing in A. ∎

Finally we prove that $\frac{\partial^2 F(\lambda, A)}{\partial \lambda \partial A} < 0$. Similar to the method we use in A.4, we get

$$\frac{\partial F(\lambda, A)}{\partial \lambda} = \mu(\lambda, A) = \frac{(p(A) - A)q(p(A))}{1 - \lambda(p(A) - A)q'(p(A))},$$

whence

$$
\begin{aligned}
\frac{\partial^2 F(\lambda, A)}{\partial \lambda \partial A} &= \frac{\partial \mu(\lambda, A)}{\partial A} \\
&= \frac{1}{(1 - \lambda(p(A) - A)q'(p(A)))^2} \cdot \Big((p(A) - A)p'(A)q'(p(A))(1 - \lambda(p(A) - A)q'(p(A))) \\
&\quad - q(p(A))\left(1 - p'(A)\left(1 + \lambda(p(A) - A)^2 q''(p(A))\right)\right) \Big).
\end{aligned}
$$

[131]

To prove the above is negative, we only need to show that the numerator of the above is negative.

$$(p(A) - A)\, p'(A)q'(p(A))\, (1 - \lambda\, (p(A) - A)\, q'(p(A)))$$
$$- q(p(A))\, \left(1 - p'(A)\, \left(1 + \lambda\, (p(A) - A)^2\, q''(p(A))\right)\right)$$
$$= (p(A) - A)\, p'(A)q'(p(A))\, (1 - \lambda\, (p(A) - A)\, q'(p(A))) - q(p(A))\, (1 - p'(A))$$
$$+ \lambda\, (p(A) - A)^2\, p'(A)q(p(A))q''(p(A))$$
$$\leq (p(A) - A)\, p'(A)q'(p(A))\, (1 - \lambda\, (p(A) - A)\, q'(p(A)))$$
$$- q(p(A))\, (1 - p'(A)) + \lambda\, (p(A) - A)^2\, p'(A)(q'(p(A)))^2$$
$$= q(p(A))\, (p'(A) - 1) + (p(A) - A)\, p'(A)q'(p(A)), \tag{A.9}$$

where the inequality is because $p(A) > A$, $p'(A) > 0$ (Lemma 1.4.3) and $q(\cdot)$ is log-concave. So it is sufficient to show that (A.9) is negative. According to the definition of $p(A)$, we can write down the first-order condition which involves $p(A)$:

$$e^{-\lambda q(p(A))} \left(e^{\lambda q(p(A))} - 1 + (p(A) - A)\lambda q'(p(A))\right) = 0,$$

which leads to

$$p(A) - A = \frac{1 - e^{\lambda q(p(A))}}{\lambda q'(p(A))}. \tag{A.10}$$

Use the above formula in (A.9), we need to prove the following expression is negative

$$q(p(A))\, (p'(A) - 1) - \frac{(e^{\lambda q(p(A))} - 1)p'(A)}{\lambda} = -q(p(A)) + \frac{p'(A)}{\lambda} \left(1 + \lambda q(p(A)) - e^{\lambda q(p(A))}\right). \tag{A.11}$$

Because we have $1 + x < e^x$, (A.11) is indeed negative. Therefore we have shown that $F(\lambda, A)$ is submodular.

Now, we finished the proof for Lemma A.0.1. ∎

Lemma A.0.6 $|F(\lambda, A) - F(\lambda, B)| < |A - B|$.

Proof Proof for Lemma A.0.6: From (A.8) we know that

$$\frac{dF(\lambda, A)}{dA} = e^{-\lambda q(p(A))} < 1,$$

then the lemma follows immediately. ∎

Proof Proof for Lemma 1.4.3: Because of the unimodularity of $\pi(p; \lambda, A)$, $p(\lambda, A)$ is the solution to

$$\frac{d\pi(p; \lambda, A)}{dp} = e^{-\lambda q(p)}\left(e^{\lambda q(p)} - 1 + (p - A)\lambda q'(p)\right) = 0,$$

which leads to

$$p(\lambda, A) = \frac{1 - e^{\lambda q(p(\lambda, A))} + A\lambda q'(p(\lambda, A))}{\lambda q'(p(\lambda, A))}.$$

We take derivative of both sides in λ

$$\frac{\partial p(\lambda, A)}{\partial \lambda} = -\frac{1 + e^{\lambda q(p(\lambda, A))}\left(\lambda q(p(\lambda, A)) - 1\right)}{\lambda^2 q'(p(\lambda, A))}$$
$$+ \frac{\partial p(\lambda, A)}{\partial \lambda} \cdot \left(-e^{\lambda q(p(\lambda, A))} + \frac{\left(e^{\lambda q(p(\lambda, A))} - 1\right) q''(p(\lambda, A))}{\lambda\left(q'(p(\lambda, A))\right)^2}\right),$$

which leads to

$$\frac{\partial p(\lambda, A)}{\partial \lambda} = -\frac{\left(1 + e^{\lambda q(p(\lambda, A))}\left(\lambda q(p(\lambda, A)) - 1\right)\right) q'(p(\lambda, A))}{\lambda\left(\left(1 + e^{\lambda q(p(\lambda, A))}\right)\lambda q'(p(\lambda, A))^2 - \left(e^{\lambda q(p(\lambda, A))} - 1\right) q''(p(\lambda, A))\right)} > 0. \quad \text{(A.12)}$$

On the other hand, we know that $q(p)$ is log-concave, i.e., $q(p)q''(p) \leq q'(p)^2$, which is equivalent to $q''(p) \leq q'(p)^2/q(p)$ (because $q(p) > 0$). To prove $\frac{\partial p(\lambda, A)}{\partial \lambda} > 0$, we need to show the denominator of (A.12) is positive. Because $e^{\lambda q(p)} - 1 > 0$, it is sufficient to prove that

$$\left(1 + e^{\lambda q(p(\lambda, A))}\right)\lambda q'(p(\lambda, A))^2 - \left(e^{\lambda q(p(\lambda, A))} - 1\right) q'(p(\lambda, A))^2/q(p(\lambda, A)) > 0,$$

[133]

which is equivalent to

$$\left(1 + e^{\lambda q(p(\lambda, A))}\right) \lambda q(p) - \left(e^{\lambda q(p(\lambda, A))} - 1\right) > 0.$$

It is easy to check that the above holds, because $(1 + e^x)x - (e^x - 1) \geq 0$ for $x \geq 0$. Therefore

$$
\begin{aligned}
0 &< \frac{\partial p(\lambda, A)}{\partial \lambda} \\
&\leq -\frac{\left(1 + e^{\lambda q(p(\lambda, A))}\left(\lambda q(p(\lambda, A)) - 1\right)\right) q'(p(\lambda, A))}{\lambda\left(\left(1 + e^{\lambda q(p(\lambda, A))}\right) \lambda q'(p(\lambda, A))^2 - \left(e^{\lambda q(p(\lambda, A))} - 1\right) q'(p(\lambda, A))^2/q(p(\lambda, A))\right)} \\
&= \frac{1}{\lambda\left(1 + \lambda q(p(\lambda, A)) / \left(1 + e^{\lambda q(p(\lambda, A))}\left(\lambda q(p(\lambda, A)) - 1\right)\right)\right)\left(-q'(p(\lambda, A))/q(p(\lambda, A))\right)} \\
&< \frac{1}{\lambda\left(-q'(p(\lambda, A))/q(p(\lambda, A))\right)},
\end{aligned}
$$

where the last second inequality is due to $1 + e^x(x - 1) > 0$ for $x > 0$. Especially, when λ becomes large, which in turn implies a large $p(\lambda, A)$ (obviously), the derivative $\partial p(\lambda, A)/\partial \lambda$ should be upper bounded by $1/c\lambda$, where c is any absolute constant within interval $(0, \lim_{p \to \infty}\{-q'(p)/q(p)\}]$.

We take derivative of both sides in A

$$\frac{\partial p(\lambda, A)}{\partial A} = 1 + \frac{\partial p(\lambda, A)}{\partial A}\left(-e^{\lambda q(p(\lambda, A))} + \frac{\left(e^{\lambda q(p(\lambda, A))} - 1\right) q''(p(\lambda, A))}{\lambda(q'(p(\lambda, A)))^2}\right),$$

which leads to

$$\frac{\partial p(\lambda, A)}{\partial A} = \frac{\lambda q'(p(\lambda, A))^2}{\left(1 + e^{\lambda q(p(\lambda, A))}\right) \lambda(q'(p(\lambda, A)))^2 - \left(e^{\lambda q(p(\lambda, A))} - 1\right) q''(p(\lambda, A))}. \tag{A.13}$$

[134]

Therefore, from (A.13), we have

$$\frac{\partial p\left(\lambda, A\right)}{\partial A} \leq \frac{\lambda q'(p(\lambda, A))^2}{\left(1 + e^{\lambda q(p(\lambda,A))}\right)\lambda q'(p(\lambda, A))^2 - \left(e^{\lambda q(p(\lambda,A))} - 1\right)q'(p(\lambda, A))^2/q(p(\lambda, A))}$$

$$= \frac{\lambda q(p)}{\left(1 + e^{\lambda q(p)}\right)\lambda q(p) - \left(e^{\lambda q(p)} - 1\right)}$$

$$\leq \max_x \left\{\frac{x}{\left(1 + e^x\right)x - \left(e^x - 1\right)}\right\} = 1.$$

In conclusion, we have shown that $0 < \frac{\partial p(\lambda, A)}{\partial A} < 1$. ∎

Lemma A.0.7 *For $A > B$, we have $\left(1 - e^{-\lambda q(p(\lambda,A))}\right)p\left(\lambda, A\right) \leq \left(1 - e^{-\lambda q(p(\lambda,B))}\right)p\left(\lambda, B\right)$.*

Proof Proof. We prove by contradiction. Suppose that for $A > B$, $\left(1 - e^{-\lambda q(p(\lambda,A))}\right)p\left(\lambda, A\right) > \left(1 - e^{-\lambda q(p(\lambda,B))}\right)p\left(\lambda, B\right)$, according to Lemma 1.4.3, $p\left(\lambda, A\right) > p\left(\lambda, B\right)$, then

$$\left(1 - e^{-\lambda q(p(\lambda,A))}\right)p\left(\lambda, A\right) + e^{-\lambda q(p(\lambda,A))}B > \left(1 - e^{-\lambda q(p(\lambda,B))}\right)p\left(\lambda, B\right) + e^{-\lambda q(p(\lambda,A))}B$$

$$> \left(1 - e^{-\lambda q(p(\lambda,B))}\right)p\left(\lambda, B\right) + e^{-\lambda q(p(\lambda,B))}B,$$

where the first inequality comes from the proof assumption and the second inequality is because $e^{-\lambda q(p)}$ is an increasing function of p. The above inequality contradicts the definition of $p\left(\lambda, B\right)$, i.e., $p\left(\lambda, B\right) = \mathrm{argmax}_p\left\{\left(1 - e^{-\lambda q(p)}\right)p + e^{-\lambda q(p)}B\right\}$. Therefore, we have proven the lemma. ∎

Lemma A.0.8 *If $F\left(\cdot\right)$ is an increasing concave function, $\{\lambda_n\}$ scales up in degree $\gamma \geq 0$, then*

$$\lim_{n\to\infty}\frac{F\left(\mathbb{E}\left[\lambda_n\right]\right) - \mathbb{E}\left[F\left(\lambda_n\right)\right]}{F\left(\mathbb{E}\left[\lambda_n\right]\right)} = \mathcal{O}\left(\frac{1}{n^\gamma}\right),$$

and if for some $c > 0$, $F'(\lambda) < \frac{1}{c\lambda}$, and $\{G_n(\cdot)\} \in \mathcal{G}_\gamma$, where $\lambda_n \sim G_n(\cdot)$, then

$$\lim_{n \to \infty} F(\mathbb{E}[\lambda_n]) - \mathbb{E}[F(\lambda_n)] = \mathcal{O}\left(\frac{1}{n^\gamma}\right).$$

Proof Proof: We have

$$F(\mathbb{E}[\lambda_n]) - \mathbb{E}[F(\lambda_n)] = \int_{\mathbb{R}^+} (F(\mathbb{E}[\lambda_n]) - F(x)) \, dG_n(x)$$

$$= \int_0^{\mathbb{E}[\lambda_n]} (F(\mathbb{E}[\lambda_n]) - F(x)) \, dG_n(x) + \int_{\mathbb{E}[\lambda_n]}^\infty (F(\mathbb{E}[\lambda_n]) - F(x)) \, dG_n(x)$$

$$\leq \int_0^{\mathbb{E}[\lambda_n]} (F(\mathbb{E}[\lambda_n]) - F(x)) \, dG_n(x),$$

where the last inequality is because $F(\cdot)$ is an increasing function. Therefore,

$$\frac{F(\mathbb{E}[\lambda_n]) - \mathbb{E}[F(\lambda_n)]}{F(\mathbb{E}[\lambda_n])} \leq \frac{\int_0^{\mathbb{E}[\lambda_n]} (F(\mathbb{E}[\lambda_n]) - F(x)) \, dG_n(x)}{F(\mathbb{E}[\lambda_n])}$$

$$= \int_0^{\mathbb{E}[\lambda_n]} \frac{\int_x^{\mathbb{E}[\lambda_n]} F'(\lambda) \, d\lambda}{\int_0^{\mathbb{E}[\lambda_n]} F'(\lambda) \, d\lambda} dG_n(x) \leq \int_0^{\mathbb{E}[\lambda_n]} \frac{\mathbb{E}[\lambda_n] - x}{\mathbb{E}[\lambda_n]} dG_n(x) \sim \frac{\sigma_n}{2\mathbb{E}[\lambda_n]},$$

[136]

where the second inequality is because $F'(\cdot)$ is a decreasing function ($F(\cdot)$ is a concave function). When $G_n(\cdot)$ is a symmetric distribution, the last approximation can be derived from Hölder's inequality. When $F'(\lambda) \leq 1/c\lambda$ for some $c > 0$, then

$$
\begin{aligned}
F(\mathbb{E}[\lambda_n]) - \mathbb{E}[F(\lambda_n)] &\leq \int_0^{\mathbb{E}[\lambda_n]} (F(\mathbb{E}[\lambda_n]) - F(x))\, dG_n(x) \\
&= \int_0^{\tau(n)} (F(\mathbb{E}[\lambda_n]) - F(x))\, dG_n(x) + \int_{\tau(n)}^{\mathbb{E}[\lambda_n]} \int_x^{\mathbb{E}[\lambda_n]} F'(\lambda)\, d\lambda dG_n(x) \\
&\leq F(\mathbb{E}[\lambda_n]) \cdot G_n(\tau(n)) + \int_{\tau(n)}^{\mathbb{E}[\lambda_n]} \frac{\mathbb{E}[\lambda_n] - x}{cx}\, dG_n(x) \\
&\leq F(\mathbb{E}[\lambda_n]) \cdot G_n(\tau(n)) + \frac{1}{c\tau(n)} \int_{\tau(n)}^{\mathbb{E}[\lambda_n]} (\mathbb{E}[\lambda_n] - x)\, dG_n(x) \\
&\leq \left(\frac{1}{c} \log(\mathbb{E}[\lambda_n]) + F(0)\right) \cdot G_n(\tau(n)) + \frac{\sigma_n}{c\tau(n)},
\end{aligned}
$$

where $\tau(n)$ is any positive function of n, the first term in the second inequality is because $F(\cdot)$ is an increasing function, the second term in the second inequality and the first term in the last inequality are because $F'(\lambda) \leq 1/c\lambda$ and the second term in the last inequality is due to Hölder's inequality. The above inequalities lead to

$$
F(\mathbb{E}[\lambda_n]) - \mathbb{E}[F(\lambda_n)] \leq \inf_{\tau(\cdot)} \left\{ \left(\frac{1}{c} \log(\mathbb{E}[\lambda_n]) + F(0)\right) \cdot G_n(\tau(n)) + \frac{\sigma_n}{c\tau(n)} \right\}.
$$

It is straightforward to see that the proof of the lemma is completed. $\blacksquare$

Proposition A.0.1 *Assume that each sequence $\{\lambda_{tn}\}$, $t = 1, \cdots, T-1$ scales up in degree γ and each corresponding distribution sequence $\{G_{tn}(\cdot)\} \in \mathcal{G}_\gamma$, then $\widetilde{V}_t^{\pi_{DA}}(\mathbb{E}[\lambda_{tn}]) - \mathbb{E}[V_t(\lambda_{tn})] \sim \mathcal{O}((T-t)/n^{1-\gamma})$, for $t = 1, \cdots, T-1$.*

Proof Proof for Proposition A.0.1: We know that $\widetilde{V}_T^{DA} = V_T = p_T$, according to Lemma 1.4.2 and Lemma A.0.8, there exist constants $C_{T-1} > 0$ and $N_{T-1} > 0$, such that for $n_{T-1} > N_{T-1}$,

$$
\begin{aligned}
&\widetilde{V}_{T-1}^{DA}\left(\mathbb{E}\left[\lambda_{(T-1)n}\right]\right) - \mathbb{E}\left[V_{T-1}\left(\lambda_{(T-1)n}\right)\right] = V_{T-1}\left(\mathbb{E}\left[\lambda_{(T-1)n}\right]\right) - \mathbb{E}\left[V_{T-1}\left(\lambda_{(T-1)n}\right)\right] \\
&\leq C_{T-1} \cdot \frac{\sigma_{(T-1)n}}{\mathbb{E}\left[\lambda_{(T-1)n}\right]} \sim \mathcal{O}\left(\frac{1}{n_{T-1}^{1-\gamma}}\right).
\end{aligned}
$$

From Lemma A.0.6, we know that

$$
\begin{aligned}
0 &\leq \widetilde{V}_{T-2}^{DA}\left(\lambda_{(T-2)n}\right) - V_{T-2}\left(\lambda_{(T-2)n}\right) \\
&\leq \widetilde{V}_{T-1}^{DA}\left(\mathbb{E}\left[\lambda_{(T-1)n}\right]\right) - \mathbb{E}\left[V_{T-1}\left(\lambda_{(T-1)n}\right)\right] \leq C_{T-1} \cdot \frac{\sigma_{(T-1)n}}{\mathbb{E}\left[\lambda_{(T-1)n}\right]},
\end{aligned}
$$

whence

$$
\begin{aligned}
0 &\leq \mathbb{E}\left[\widetilde{V}_{T-2}^{DA}\left(\lambda_{(T-2)n}\right)\right] - \mathbb{E}\left[V_{T-2}\left(\lambda_{(T-2)n}\right)\right] \\
&\leq \widetilde{V}_{T-1}^{DA}\left(\mathbb{E}\left[\lambda_{(T-1)n}\right]\right) - \mathbb{E}\left[V_{T-1}\left(\lambda_{(T-1)n}\right)\right] \leq C_{T-1} \cdot \frac{\sigma_{(T-1)n}}{\mathbb{E}\left[\lambda_{(T-1)n}\right]}.
\end{aligned}
$$

According to Lemma 1.4.2 and Lemma A.0.8 again, we have constants $C_{T-2} > 0$ and $N_{T-2} > 0$ where N_{T-2} is independent of N_{T-1} such that for $n_{T-2} > \max\{N_{T-2}, N_{T-1}\}$,

$$
\widetilde{V}_{T-2}^{DA}\left(\mathbb{E}\left[\lambda_{(T-2)n}\right]\right) - \mathbb{E}\left[\widetilde{V}_{T-2}^{DA}\left(\lambda_{(T-2)n}\right)\right] \leq C_{T-2} \cdot \frac{\sigma_{(T-2)n}}{\mathbb{E}\left[\lambda_{(T-2)n}\right]} \sim \mathcal{O}\left(\frac{1}{n_{T-2}^{1-\gamma}}\right).
$$

Here we should highlight that the reason that N_{T-2} could be independent of N_{T-1} is because to utilize the result in Lemma A.0.8, it is sufficient to have $\left(\widetilde{V}_{T-2}^{DA}(\cdot)\right)'$ be upper bounded by

[138]

$1/c\lambda$ for some constant c, and the latter is guaranteed by Lemma 1.4.2 which is independent of the next period. The above two inequalities leads to

$$\widetilde{V}_{T-2}^{DA}\left(\mathbb{E}\left[\lambda_{(T-2)n}\right]\right) - \mathbb{E}\left[V_{T-2}\left(\lambda_{(T-2)n}\right)\right]$$
$$= \left(\widetilde{V}_{T-2}^{DA}\left(\mathbb{E}\left[\lambda_{(T-2)n}\right]\right) - \mathbb{E}\left[\widetilde{V}_{T-2}^{DA}\left(\lambda_{(T-2)n}\right)\right]\right)$$
$$+ \left(\mathbb{E}\left[\widetilde{V}_{T-2}^{DA}\left(\lambda_{(T-2)n}\right)\right] - \mathbb{E}\left[V_{T-2}\left(\lambda_{(T-2)n}\right)\right]\right)$$
$$\le C_{T-2} \cdot \frac{\sigma_{(T-2)n}}{\mathbb{E}\left[\lambda_{(T-2)n}\right]} + C_{T-1} \cdot \frac{\sigma_{(T-1)n}}{\mathbb{E}\left[\lambda_{(T-1)n}\right]} \sim \mathcal{O}\left(\frac{1}{n_{T-2}^{1-\gamma}} + \frac{1}{n_{T-1}^{1-\gamma}}\right).$$

Following the similar approach, for $t = \{1, \cdots, T-1\}$, there exist constants $C_t > 0$ and $N_t > 0$ such that for $n_t > \max\{N_t, \cdots, N_{T-1}\}$,

$$\widetilde{V}_t^{DA}\left(\mathbb{E}\left[\lambda_{tn}\right]\right) - \mathbb{E}\left[V_t\left(\lambda_{tn}\right)\right] \sim \mathcal{O}\left(\sum_{\tau=t}^{T-1}\left(\frac{1}{n_\tau^{1-\gamma}}\right)\right),$$

specifically, when we let $n_t = n_{t+1} = \cdots = n_{T-1} = n > \max\{N_t, \cdots, N_{T-1}\}$, we have

$$\widetilde{V}_t^{DA}\left(\mathbb{E}\left[\lambda_{tn}\right]\right) - \mathbb{E}\left[V_t\left(\lambda_{tn}\right)\right] \sim \mathcal{O}\left(\frac{T-t}{n^{1-\gamma}}\right).$$

The proof is completed. ∎

Proof Proof for Theorem 1.4.4: From Lemma 1.4.3, for time period t, we can generate an upper bound on the difference of the price induced by $\widetilde{V}_t^{DA}\left(\lambda_{tn}\right)$, p_t^{DA}, and the price induced by $V_t\left(\lambda_{tn}\right)$, p_t:

$$0 < p_t^{DA} - p_t \le \widetilde{V}_{t+1}^{DA}\left(\mathbb{E}\left[\lambda_{(t+1)n}\right]\right) - \mathbb{E}\left[V_{t+1}\left(\lambda_{(t+1)n}\right)\right] \sim \mathcal{O}\left(\frac{T-t-1}{n^{1-\gamma}}\right). \qquad (A.14)$$

In this proof, we give the bound on the regret ratio, $\left(V_t\left(\lambda_{tn}\right) - V_t^{DA}\left(\lambda_{tn}\right)\right)/V_t\left(\lambda_{tn}\right)$. The bound on the regret $\left(V_t\left(\lambda_{tn}\right) - V_t^{DA}\left(\lambda_{tn}\right)\right)$ could be constructed following the sim-

ilar manner in Theorem 1.5.2. It turns out that the regret ratio and the regret share the same asymptotic rate, which is because the denominator in the regret ratio, $V_t(\lambda_{tn})$, increases rather slowly in both T and n. We first look at the regret. Notice that $\mathbb{E}_{\lambda_{(T-1)n}}\left[V_{T-1}\left(\lambda_{(T-1)n}\right)\right] - \mathbb{E}_{\lambda_{(T-1)n}}\left[V_{T-1}\left(\lambda_{(T-1)n}\right)\right] = 0$, we assume that

$$
\left(\mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}\left(\lambda_{(t+1)n}\right)\right] - \mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}^{DA}\left(\lambda_{(t+1)n}\right)\right]\right)
$$
$$
\leq \mathcal{O}\left((T-t-1)(T-t-2)\log(n)/n^{1-\gamma}\right),
$$

then

$$
\begin{aligned}
V_t(\lambda_{tn}) - V_t^{DA}(\lambda_{tn}) &= \left(\left(1 - e^{-\lambda_{tn}q(p_t)}\right)p_t + e^{-\lambda_{tn}q(p_t)}\mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}\left(\lambda_{(t+1)n}\right)\right]\right) \\
&\quad - \left(\left(1 - e^{-\lambda_{tn}q\left(p_t^{DA}\right)}\right)p_t^{DA} + e^{-\lambda_{tn}q\left(p_t^{DA}\right)}\mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}^{DA}\left(\lambda_{(t+1)n}\right)\right]\right) \\
&= \left(1 - e^{-\lambda_{tn}q\left(p_t^{DA}\right)}\right)\left(p_t - p_t^{DA}\right) \\
&\quad + e^{-\lambda_{tn}q\left(p_t^{DA}\right)}\left(\mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}\left(\lambda_{(t+1)n}\right)\right] - \mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}^{DA}\left(\lambda_{(t+1)n}\right)\right]\right) \\
&\quad + \left(e^{-\lambda_{tn}q\left(p_t^{DA}\right)} - e^{-\lambda_{tn}q(p_t)}\right)\left(p_t - \mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}\left(\lambda_{(t+1)n}\right)\right]\right) \\
&\leq \left(\mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}\left(\lambda_{(t+1)n}\right)\right] - \mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}^{DA}\left(\lambda_{(t+1)n}\right)\right]\right) \\
&\quad + \left(e^{-\lambda_{tn}q(\widetilde{p}_t)}\lambda_{tn}q\left(\widetilde{p}_t\right)\left(-\frac{q'\left(\widetilde{p}_t\right)}{q\left(\widetilde{p}_t\right)}\right)\right)\left(p_t^{DA} - p_t\right)\cdot\left(p_t - \mathbb{E}_{\lambda_{(t+1)n}}\left[V_{t+1}\left(\lambda_{(t+1)n}\right)\right]\right) \\
&\leq \mathcal{O}\left(\frac{(T-t-1)(T-t-2)\log(n)}{n^{1-\gamma}}\right) + \mathcal{O}\left(\log(n)\right)\cdot\left(p_t^{DA} - p_t\right) \\
&\leq \mathcal{O}\left(\frac{(T-t)(T-t-1)\log(n)}{n^{1-\gamma}}\right),
\end{aligned} \tag{A.15}
$$

where in the first inequality, $\widetilde{p}_t$ is some value in $[p_t, p_t^{DA}]$. The second inequality comes from the facts that $xe^{-x} < 1$, $\lim_{p\to\infty} -q'(p)/q(p) < \infty$ and (1.4). We notice that all of the above inequalities are actually independent of the realization of λ_{tn}, therefore we have

[140]

$\left(\mathbb{E}_{\lambda_{tn}} \left[V_t \left(\lambda_{tn} \right) \right] - \mathbb{E}_{\lambda_{tn}} \left[V_t^{DA} \left(\lambda_{tn} \right) \right] \right) \leq \mathcal{O} \left(\left(T - t \right) \left(T - t - 1 \right) \log \left(n \right) / n^{1-\gamma} \right)$. According to the proof by induction, we can conclude that

$$Regret \left(\pi_{DA} \right) < \mathcal{O} \left(\frac{T^2 log \left(n \right)}{n^{1-\gamma}} \right).$$

The proof is completed. ∎

Lemma A.0.9 *Define*

$$F \left(\lambda, A, \xi \right) = \max_p \left\{ \pi \left(p; \xi, A \right) \right\} = \max_p \left\{ \left(1 - e^{-\lambda q(\xi - \alpha p)} \right) p + e^{-\lambda q(\xi - \alpha p)} A \right\},$$

where $q(\xi - \alpha p) \in \mathcal{Q}$, then (i). $\lim_{\xi \to \infty} p^* \left(\xi, A \right) = \infty$, (ii). $F(\lambda, A, \xi)$ is increasing convex in ξ, especially $0 \leq \frac{\partial F(\lambda, A, \xi)}{\partial \xi} \leq 1/\alpha$ and $\lim_{\xi \to \infty} F(\lambda, A, \xi) = \infty$, (iii). , $F(\lambda, A, \xi)$ is submodular, i.e., $\frac{\partial^2 F(\lambda, A, \xi)}{\partial \xi \partial A} < 0$.

Proof Proof for Lemma A.0.9. The idea of the proof is similar to A.0.1. First of all, similar as before, we have $F(\xi, A) = \left(1 - e^{-\lambda q(\xi - \alpha p)} \right) \left(p - A \right) + A$. Therefore, to investigate the $\partial F \left(\xi, A \right) / \partial \xi$, it is sufficient to focus on $\widetilde{F} \left(\xi, A \right) = \left(1 - e^{-\lambda q(\xi - \alpha p)} \right) \left(p - A \right)$. Let $z = \xi - \alpha p$, then we can write the problem as

$$\widetilde{F} \left(\xi + \delta_\xi, A \right) = \max_{p, z} \left(1 - e^{-\lambda q(z)} \right) p$$

$$s.t. \ z + \alpha p - \xi = \delta_\xi.$$

According to Theorem A.0.2, we have

$$\frac{\partial \widetilde{F} \left(\xi, A \right)}{\partial \xi} = \nabla_{\delta_\xi} \widetilde{F} \left(\xi + \delta_\xi, A \right) |_{\delta_\xi = 0} = \mu \left(\xi \right),$$

[141]

where $\mu(\xi)$ is the Lagrangian multiplier of the above optimization problem and the Lagrangian has form $L(p, z; \mu) = \left(1 - e^{-\lambda q(z)}\right) p - \mu(z + \alpha p - \xi)$. We can write down the Karush-Kuhn-Tucker conditions:

$$
\begin{cases}
\left(1 - e^{-\lambda q(z)}\right) - \alpha\mu = 0, \\
e^{-\lambda q(z)}(p - A)\lambda q'(z) - \mu = 0, \\
z + \alpha p - \xi = 0,
\end{cases}
\tag{A.16}
$$

which leads to

$$
\mu(\xi) = \frac{(p(\xi) - A)\lambda q'(\xi - \alpha p(\xi))}{1 + (p(\xi) - A)\alpha\lambda q'(\xi - \alpha p(\xi))}.
\tag{A.17}
$$

It is straightforward to see that $\mu(\xi) > 0$, because $q'(\cdot) > 0$ and in the equilibrium $p(\xi) > A$. On the other hand,

$$
\mu(\xi) = \frac{(p(\xi) - A)\lambda q'(\xi - \alpha p(\xi))}{1 + (p(\xi) - A)\alpha\lambda q'(\xi - \alpha p(\xi))} = \frac{1}{\alpha} \cdot \frac{(p(\xi) - A)\lambda q'(\xi - \alpha p(\xi))}{1/\alpha + (p(\xi) - A)\lambda q'(\xi - \alpha p(\xi))} \le \frac{1}{\alpha}.
$$

Therefore, we have shown that $\widetilde{F}(\xi, A)$ is increasing in ξ with $\mu'(\xi) < 1/\alpha$. Given the expression of $\mu(\xi)$, we have

$$
\mu'(\xi) = \frac{\lambda p'(\xi) q'(\xi - \alpha p(\xi)) - \lambda(p(\xi) - A)(\alpha p'(\xi) - 1)q''(\xi - \alpha p(\xi))}{(1 + \alpha\lambda(p(\xi) - A)q'(\xi - \alpha p(\xi)))^2}.
$$

We would like to prove $\mu'(\xi) > 0$, which is equivalent to proving the numerator of the above is positive. Based upon (A.16) and (A.17), we can get the expressions for $q(\xi)$ and $q'(\xi)$ in the equilibrium, whence after algebraic manipulation, the numerator of the above can be written as

$$
\frac{e^{\lambda q(\xi - \alpha p(\xi))}\lambda^2 q'(\xi - \alpha p(\xi))^3}{\alpha\left((1 + e^{\lambda q(\xi - \alpha p(\xi))})\lambda(q'(\xi - \alpha p(\xi)))^2 - (e^{\lambda q(\xi - \alpha p(\xi))} - 1)q''(\xi - \alpha p(\xi))\right)}.
$$

[142]

Because $q'(\cdot) > 0$, to prove the above is positive, it is equivalent to proving the denominator of the above is positive. Because $q(\cdot)$ is log-concave, for the denominator of the above, we have

$$\left(1 + e^{\lambda q(\xi - \alpha p(\xi))}\right) \lambda \left(q'\left(\xi - \alpha p\left(\xi\right)\right)\right)^2 - \left(e^{\lambda q(\xi - \alpha p(\xi))} - 1\right) q''\left(\xi - \alpha p\left(\xi\right)\right)$$

$$\geq \left(1 + e^{\lambda q(\xi - \alpha p(\xi))}\right) \lambda \left(q'\left(\xi - \alpha p\left(\xi\right)\right)\right)^2 - \left(e^{\lambda q(\xi - \alpha p(\xi))} - 1\right) \frac{\left(q'\left(\xi - \alpha p\left(\xi\right)\right)\right)^2}{q\left(\xi - \alpha p\left(\xi\right)\right)}$$

$$= \frac{\left(q'\left(\xi - \alpha p\left(\xi\right)\right)\right)^2}{q\left(\xi - \alpha p\left(\xi\right)\right)} \cdot \left(\left(1 + e^{\lambda q(\xi - \alpha p(\xi))}\right) \lambda q\left(\xi - \alpha p\left(\xi\right)\right) - \left(e^{\lambda q(\xi - \alpha p(\xi))} - 1\right)\right) \geq 0.$$

The first inequality is because $e^{\lambda q(\xi - \alpha p(\xi))} > 1$ and the last inequality is due to $(1 + e^x) x > e^x - 1$, for $x \geq 0$. Therefore, we have shown that $\widetilde{F}(\xi, A)$ (or equivalently $F(\xi, A)$) is increasing convex in ξ.

On the other hand, we can write (A.17) as

$$\mu(A) = \frac{(p(A) - A) \lambda q'\left(\xi - \alpha p(A)\right)}{1 + (p(A) - A) \alpha \lambda q'\left(\xi - \alpha p(A)\right)}.$$

Therefore we have

$$\frac{\partial^2 \widetilde{F}(\xi, A)}{\partial \xi \partial A} = \frac{\lambda \left((p'(A) - 1) q'\left(\xi - \alpha p(A)\right) - \alpha \left(p(A) - A\right) p'(A) q''\left(\xi - \alpha p(A)\right)\right)}{\left(1 + \alpha \lambda \left(p(A) - A\right) q'\left(\xi - \alpha p(A)\right)\right)^2}.$$

We would like to prove that the above is negative, which is equivalent to proving the numerator is negative. Using the similar techniques that we have seen before, we can write the above numerator as

$$\frac{e^{\lambda q(\xi - \alpha p(A))} \lambda \left(q'\left(\xi - \alpha p(A)\right)\right)^3}{-\left(1 + e^{\lambda q(\xi - \alpha p(A))}\right) \lambda \left(q'\left(\xi - \alpha p(A)\right)\right)^2 + \left(e^{\lambda q(\xi - \alpha p(A))} - 1\right) q''\left(\xi - \alpha p(A)\right)}.$$

[143]

To prove the above is negative, it is further equivalent to proving the denominator of the above is negative. Actually, we have

$$- \left(1 + e^{\lambda q(\xi - \alpha p(A))}\right) \lambda \left(q'\left(\xi - \alpha p(A)\right)\right)^2 + \left(e^{\lambda q(\xi - \alpha p(A))} - 1\right) q''\left(\xi - \alpha p(A)\right)$$

$$\leq - \left(1 + e^{\lambda q(\xi - \alpha p(A))}\right) \lambda \left(q'\left(\xi - \alpha p(A)\right)\right)^2 + \left(e^{\lambda q(\xi - \alpha p(A))} - 1\right) \cdot \frac{\left(q'\left(\xi - \alpha p(A)\right)\right)^2}{q\left(\xi - \alpha p(A)\right)}$$

$$= - \frac{\left(q'\left(\xi - \alpha p(A)\right)\right)^2}{q\left(\xi - \alpha p(A)\right)} \cdot \left(\left(1 + e^{\lambda q(\xi - \alpha p(A))}\right) \lambda q\left(\xi - \alpha p(A)\right) - \left(e^{\lambda q(\xi - \alpha p(A))} - 1\right)\right) \leq 0,$$

where the first inequality is because of the log-concavity of $q(\cdot)$ and $\left(e^{\lambda q(\xi - \alpha p(A))} - 1\right) > 0$ and the last inequality is due to $(1 + e^x)x > e^x - 1$. Therefore, we have shown the submodularity of $\widetilde{F}(\xi, A)$ (or equivalently $F(\xi, A)$). ∎

Proof Proof for Proposition 1.5.2: Let $L_{tn}(\xi) = \frac{1}{t} \sum_{\tau=1}^{t} \left(\frac{1}{n} \sum_{i=1}^{n} f_\tau(x_{\tau i}|\xi)\right)$ and $L_t(\xi) = \frac{1}{t} \left(\sum_{\tau=1}^{t} \mathbb{E}_{\xi_0} f_\tau(x|\xi_0)\right)$, then according to the definition of $\xi_{tn}(\boldsymbol{x})$ and ξ_0, we have $L'_{tn}(\xi_{tn}) = 0$ and $L'_t(\xi_0) = 0$, therefore for some ξ_1 between ξ_{tn} and ξ_0 such that

$$\sqrt{tn}\left(\xi_{tn} - \xi_0\right) = -\frac{\sqrt{tn} L'_{tn}(\xi)}{L''_{tn}(\xi_1)}.$$

For the numerator, we have

$$\sqrt{tn} L'_{tn}(\xi_0) = \sqrt{tn} \left(\frac{1}{t} \sum_{\tau=1}^{t} \left(\frac{1}{n} \sum_{i=1}^{n} f'_\tau(x_{\tau i}|\xi_0)\right) - 0\right)$$

$$= \sqrt{tn} \left(\frac{1}{t} \sum_{\tau=1}^{t} \left(\frac{1}{n} \sum_{i=1}^{n} f'_\tau(x_{\tau i}|\xi_0)\right) - \frac{1}{t} \sum_{\tau=1}^{t} \mathbb{E}_{\xi_0} f'_\tau(x|\xi_0)\right) \to N\left(0, \frac{1}{t} \sum_{\tau=1}^{t} Var_{\xi_0}\left(f'_\tau(x|\xi_0)\right)\right),$$

where the last convergence comes from Lyapunov's central limit theorem. Duo to the similar argument in the proof of Proposition 1.5.1, arbitrary order of moment of $f'_\tau(x|\xi_0)$ exists (because arbitrary order of moment of Poisson distribution exists and $f'_\tau(x|\xi_0)$ can be bounded

[144]

by some linear function of x). Then it is not difficult to verify that the Lyapunov's condition required for Lyapunov's central limit theorem holds here. Furthermore, we can bound the convergence error using Local Limit Theorem (see Appendix, also see [104] Chapter VII). Let $Y_{\tau i} = f'_\tau (x_{\tau i}|\xi_0) / \sqrt{Var_{\xi_0} (f'_\tau (x|\xi_0))}$, $\tau = 1, \cdots, t$, $i = 1, \cdots, n$. And let $g_\tau (\cdot)$ be the probability density function of $Y_{\tau i}$. Then it is straightforward to verify that $\mathbb{E}|Y_{\tau i}|^3 < \infty$, $\mathbb{E}Y_{\tau i} = 0$, $\mathbb{E}Y_{\tau i}^2 = 1$ and there exists some constant C_τ such that $g_\tau (\cdot) \leq C_\tau$. Let $g_{\tau n} (x)$ be the density of the random variable $\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_{\tau i}$, then by Local Limit Theorem,

$$\left| g_{\tau n} (x) - \phi (x) \right| \leq \frac{A\beta_{3\tau}^3 \max (1, C_\tau^5)}{\sqrt{n} (1 + |x|^2)}, \tag{A.18}$$

where $\phi (\cdot)$ is the pdf of standard normal distribution, $\beta_{3\tau} = \mathbb{E}|Y_{\tau i}|^3$ and A is a constant. Let $\widehat{g}_{\tau n} (\cdot)$ be the probability density function of $\frac{1}{\sqrt{n}} \sum_{i=1}^n f'_\tau (x_{\tau i}|\xi_0)$, then by (A.18), we have

$$\left| \widehat{g}_{\tau n} (x) - \widehat{\phi} (x) \right| \leq \frac{A\beta_{3\tau}^3 \sqrt{V_\tau} \max (1, C_\tau^5)}{\sqrt{n} (V_\tau + |x|^2)}, \tag{A.19}$$

where $\widehat{\phi} (\cdot)$ is the pdf of normal distribution $N (0, V_\tau)$ and $V_\tau = Var_{\xi_0} (f'_\tau (x|\xi_0))$. Then let $\widetilde{g}_{tn} (\cdot)$ denote the pdf of $\sqrt{tn} L'_{tn} (\xi_0) = \frac{1}{\sqrt{tn}} \sum_{\tau=1}^t \sum_{i=1}^n f'_\tau (x_{\tau i}|\xi_0)$, then it is not difficult to see that

$$\left| \widetilde{g}_{\tau n} (x) - \widetilde{\phi} (x) \right| \leq \frac{\widetilde{A}_t}{\sqrt{tn} \left(\widetilde{V}_t + |x|^2 \right)}, \tag{A.20}$$

where $\widetilde{\phi} (\cdot)$ is the pdf of normal distribution $N \left(0, \widetilde{V}_t \right)$ and $\widetilde{V}_t = \frac{1}{t} \sum_{\tau=1}^t Var_{\xi_0} (f'_\tau (x|\xi_0))$.

On the other hand, for the denominator $L''_{tn} (\xi_1)$, by the Law of Large Numbers, we have

$$L''_{tn} (\xi_1) \to L''_{tn} (\xi_0) \to \frac{1}{t} \sum_{\tau=1}^t \mathbb{E}_{\xi_0} f''_\tau (x|\xi_0) = -I_t (\xi_0) < 0.$$

[145]

Therefore, by Slutsky's Theorem, we have

$$-\frac{\sqrt{tn}L'_{tn}(\xi)}{L''_{tn}(\xi_1)} \xrightarrow{d} N\left(0, \frac{\frac{1}{t}\sum_{\tau=1}^{t} Var_{\xi_0}(f'_\tau(x|\xi_0))}{(I(\xi_0))^2}\right),$$

then because

$$\frac{1}{t}\sum_{\tau=1}^{t} Var_{\xi_0}(f'_\tau(x|\xi_0)) = \frac{1}{t}\sum_{\tau=1}^{t}\left(\mathbb{E}_{\xi_0}(f'_\tau(x|\xi_0))^2 - (\mathbb{E}_{\xi_0}f'_\tau(x|\xi_0))^2\right)$$

$$= \frac{1}{t}\sum_{\tau=1}^{t}\mathbb{E}_{\xi_0}(f'_\tau(x|\xi_0))^2 = I_t(\xi_0),$$

we have

$$-\frac{\sqrt{tn}L'_{tn}(\xi)}{L''_{tn}(\xi_1)} \xrightarrow{d} N\left(0, \frac{1}{I_t(\xi_0)}\right).$$

Similar to what we did on the numerator, we can use Local Limit Theorem to bound the degeneration rate of the denominator. But because the degeneration rate of the denominator is on a higher order compared with the convergence rate of the numerator, when we bound the asymptotic normality behavior of $-\sqrt{tn}L'_{tn}(\xi)/L''_{tn}(\xi_1)$, we can neglect the effect from the denominator. Let $g^*_{tn}(\cdot)$ be the pdf of $\xi_{tn}(\boldsymbol{x})$, then from (A.18), we have

$$\left|g^*_{tn}(x) - \phi^*_{tn}(x)\right| \leq \frac{\widetilde{C}}{\sqrt{tn}(1+|x|^2)}, \tag{A.21}$$

where $\phi^*_{tn}(\cdot)$ is the pdf of distribution $N\left(\xi_0, (tnI_t(\xi_0))^{-1}\right)$ and $\widetilde{C}$ is some constant independent of t and n.

$$Q_{tn,\boldsymbol{x}}(B) = \int_{\mathbb{R}} P(\xi \in B|\xi_{tn}(\boldsymbol{x}) = \delta) g^*_{tn}(\delta) d\delta = \int_B \left(\int_{\mathbb{R}} Q'_{tn,\boldsymbol{x}}(\xi|\xi_{tn}(\boldsymbol{x}) = \delta) g^*_{tn}(\delta) d\delta\right) d\xi$$

$$= \int_B \left(\int_{\mathbb{R}} Q'_{tn,\boldsymbol{x}}(\xi|\xi_{tn}(\boldsymbol{x}) = \delta) \phi^*_{tn}(\delta) d\delta\right) d\xi$$

$$+ \int_{\mathbb{R}} P(\xi \in B|\xi_{tn}(\boldsymbol{x}) = \delta)(g^*_{tn}(\delta) - \phi^*_{tn}(\delta)) d\delta,$$

[146]

where the first term is the probability function of $N\left(\xi_0, \Gamma_{tn}\left(\boldsymbol{x}\right)^{-1} + \left(tnI_t\left(\xi_0\right)\right)^{-1}\right)$ which we denote by $\widetilde{Q}_{tn,\boldsymbol{x}}\left(B\right)$. We notice that the variance the normal distribution $\widetilde{Q}_{tn,\boldsymbol{x}}\left(\cdot\right)$ converges to zero in the rate of $1/tn$. For the second term, we have

$$\left|\int_{\mathbb{R}} P\left(\xi \in B|\xi_{tn}\left(\boldsymbol{x}\right) = \delta\right)\left(g_{tn}^*\left(\delta\right) - \phi_{tn}^*\left(\delta\right)\right)d\delta\right|$$
$$\leq \int_{\mathbb{R}} P\left(\xi \in B|\xi_{tn}\left(\boldsymbol{x}\right) = \delta\right)\left|g_{tn}^*\left(\delta\right) - \phi_{tn}^*\left(\delta\right)\right|d\delta$$
$$\leq \int_{\mathbb{R}} \left|g_{tn}^*\left(\delta\right) - \phi_{tn}^*\left(\delta\right)\right|d\delta \overset{\text{(A.21)}}{\leq} \int_{\mathbb{R}} \frac{\widetilde{C}}{\sqrt{tn}\left(1 + |\delta|^2\right)}d\delta = \frac{\widetilde{C}'}{\sqrt{tn}},$$

where $\widetilde{C}'$ is a constant. We notice that the second term also diminishes in the order of $1/\sqrt{tn}$ which is similar to the convergence error between $R_{tn,\boldsymbol{x}}$ and $Q_{tn,\boldsymbol{x}}$. Combined with Proposition 1.5.2, there exists a constant $\widetilde{C}_K$ such that

$$\sup_{\xi \in K} P_\xi^{tn}\left\{\boldsymbol{x} \in X^{tn} : d\left(R_{tn,\boldsymbol{x}}, \widetilde{Q}_{tn,\boldsymbol{x}}\right) > C_K\left(s\right)\cdot\left(t\cdot n\right)^{-1/2}\right\} = \mathcal{O}\left(\left(t\cdot n\right)^{-s/2}\right). \qquad \text{(A.22)}$$

$\blacksquare$

Proposition A.0.2 $\left|V_t\left(\lambda_{tn}, \xi\right) - V_t\left(\lambda_{tn}, \xi_0\right)\right| \leq \left(T - t\right)\left|\xi - \xi_0\right|/\alpha.$

Proof Proof for Proposition A.0.2: We know that $V_T\left(\cdot\right) = p_T$, according to Lemma A.0.9,

$$\left|V_{T-1}\left(\lambda_{(T-1)n}, \xi\right) - V_{T-1}\left(\lambda_{(T-1)n}, \xi_0\right)\right| \leq \frac{\left|\xi - \xi_0\right|}{\alpha}.$$

Assume that

$$\left|V_{t+1}\left(\lambda_{(t+1)n}, \xi\right) - V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right)\right| \leq \frac{\left(T - t - 1\right)\left|\xi - \xi_0\right|}{\alpha}.$$

From Lemma A.0.6, we know that

$$
\left| V_t \left(\lambda_{tn}, \xi \right) - F \left(\lambda_{tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right), \xi \right) \right|
$$
$$
= \left| F \left(\lambda_{tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi \right), \xi \right) - F \left(\lambda_{tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right), \xi \right) \right|
$$
$$
\leq \left| V_{t+1} \left(\lambda_{(t+1)n}, \xi \right) - V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right) \right| \leq \frac{(T - t - 1) \left| \xi - \xi_0 \right|}{\alpha},
$$

on the other hand, again by Lemma A.0.9,

$$
\left| F \left(\lambda_{tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right), \xi \right) - V_t \left(\lambda_{tn}, \xi_0 \right) \right|
$$
$$
= \left| F \left(\lambda_{tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right), \xi \right) - F \left(\lambda_{tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right), \xi_0 \right) \right| \leq \frac{\left| \xi - \xi_0 \right|}{\alpha}.
$$

Therefore,

$$
\left| V_t \left(\lambda_{tn}, \xi \right) - V_t \left(\lambda_{tn}, \xi_0 \right) \right|
$$
$$
\leq \left| V_t \left(\lambda_{tn}, \xi \right) - F \left(\lambda_{(tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right), \xi \right) \right| + \left| F \left(\lambda_{tn}, V_{t+1} \left(\lambda_{(t+1)n}, \xi_0 \right), \xi \right) - V_t \left(\lambda_{tn}, \xi_0 \right) \right|
$$
$$
\leq \frac{(T - t) \left| \xi - \xi_0 \right|}{\alpha}.
$$

Therefore using proof by induction, we have shown that for $t = \{1, \cdots, T - 1\}$,

$$
\left| V_t \left(\lambda_{tn}, \xi \right) - V_t \left(\lambda_{tn}, \xi_0 \right) \right| \leq \frac{(T - t) \left| \xi - \xi_0 \right|}{\alpha}.
$$

■

[148]

Proof Proof for Theorem 1.5.2: From Lemma 1.4.3, for time period t, we can bound the difference between the optimal price p_t and Thompson sampling price $p_t^{TS}(\xi_t)$ given the random draw ξ_t (we will simply write as p_t^{TS} when no confusion is created):

$$\left| p_t^{TS}(\xi_t) - p_t \right| \leq \left| V_{t+1}\left(\lambda_{(t+1)n}, \xi_t\right) - V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) \right| \leq \frac{(T-t-1)|\xi_t - \xi_0|}{\alpha}. \qquad \text{(A.23)}$$

When $\lim_{n\to\infty} \lambda_{tn} q_t (\xi_0 - \alpha p_t) < \infty$, $\xi_t > \xi_0$, which implies $p_t^{TS} > p_t$ Similar to the proof of Theorem 1.4.4, we have

$$
\begin{aligned}
0 &\leq V_t\left(\lambda_{tn}, \xi_0\right) - V_t^{TS}\left(\lambda_{tn}, \xi_t\right) = \left(1 - e^{-\lambda_{tn} q_t(\xi_0 - \alpha p_t)}\right) p_t - \left(1 - e^{-\lambda_{tn} q_t\left(\xi_0 - \alpha p_t^{TS}\right)}\right) p_t^{TS} \\
&\quad + e^{-\lambda_{tn} q_t(\xi_0 - \alpha p_t)} V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - e^{-\lambda_{tn} q_t\left(\xi_0 - \alpha p_t^{TS}\right)} \cdot \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right] \\
&= \left(1 - e^{-\lambda_{tn} q_t\left(\xi_0 - \alpha p_t^{TS}\right)}\right)\left(p_t - p_t^{TS}\right) \\
&\quad + e^{-\lambda_{tn} q_t\left(\xi_0 - \alpha p_t^{TS}\right)}\left(V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right]\right) \\
&\quad + \left(e^{-\lambda_{tn} q_t\left(\xi_0 - \alpha p_t^{TS}\right)} - e^{-\lambda_{tn} q_t(\xi_0 - \alpha p_t)}\right)\left(p_t - V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right)\right) \\
&\leq \left(V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right]\right) \\
&\quad + \left(e^{-\lambda_{tn} q_t(\xi_0 - \alpha \widetilde{p})} \lambda_{tn} q_t\left(\xi_0 - \alpha \widetilde{p}\right)\left(-\frac{q_t'\left(\xi_0 - \alpha \widetilde{p}\right)}{q_t\left(\xi_0 - \alpha \widetilde{p}\right)}\right)\right)\left(p_t^{TS} - p_t\right) \cdot \left(p_t - V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right)\right) \\
&\leq \log(n)\left(p_t^{TS} - p_t\right) + \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right] \\
&\leq \frac{(T-t-1)\log(n)}{\alpha} \cdot |\xi_0 - \xi_{tn}| + \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right]. \qquad \text{(A.24)}
\end{aligned}
$$

[149]

On the other hand, when $\xi_{tn} \leq \xi_0$, $p_t^{TS} \leq p_t$,

$$0 \leq V_t\left(\lambda_{tn}, \xi_0\right) - V_t^{TS}\left(\lambda_{tn}, \xi_{tn}\right) = \left(1 - e^{-\lambda_{tn}q_t\left(\xi_0 - \alpha p_t^{TS}\right)}\right)\left(p_t - p_t^{TS}\right)$$

$$+ e^{-\lambda_{tn}q_t(\xi_0 - \alpha p_t)}\left(V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right]\right)$$

$$+ \left(e^{-\lambda_{tn}q_t(\xi_0 - \alpha p_t)} - e^{-\lambda_{tn}q_t\left(\xi_0 - \alpha p_t^{TS}\right)}\right)\left(\mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right] - p_t\right)$$

$$\leq \left(p_t - p_t^{TS}\right) + \left(V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right]\right)$$

$$\leq \frac{(T - t - 1)\left|\xi_0 - \xi_{tn}\right|}{\alpha} + \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) - V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right], \qquad \text{(A.25)}$$

where the second inequality is because $p_t \geq p_t^{TS}$ implies $e^{-\lambda_{tn}q_t(\xi_0 - \alpha p_t)} \geq e^{-\lambda_{tn}q_t\left(\xi_0 - \alpha p_t^{TS}\right)}$ and $p_t > V_{t+1}\left(\lambda_{(t+1)n}, \xi_0\right) \geq \mathbb{E}_{\xi_{t+1}}\left[V_{t+1}^{TS}\left(\lambda_{(t+1)n}, \xi_{t+1}\right)\right]$, which makes the third term in the first equality to be negative. So based on (A.24) and (A.25), we have the following bound on the expected regret:

$$0 \leq \mathbb{E}_{\xi_t}\left[V_t\left(\lambda_{tn}, \xi_0\right) - V_t^{TS}\left(\lambda_{tn}, \xi_t\right)\right] \leq \frac{\log(n)}{\alpha} \cdot \sum_{\tau=t}^{T-1}(T - \tau - 1) \cdot \mathbb{E}_{\xi_\tau}\left|\xi_0 - \xi_\tau\right|. \qquad \text{(A.26)}$$

Now we calculate the expected learning error on ξ_0,

$$\mathbb{E}_{\xi_{tn}^{TS}}\left|\xi_{tn}^{TS} - \xi_0\right| = \int_{\underline{\xi}}^{\bar{\xi}}\left|\xi_{tn}^{TS} - \xi_0\right| R'_{tn,\boldsymbol{x}}\left(\xi_t^{TS}\right) d\xi_{tn}^{TS}$$

$$= \int_{\underline{\xi}}^{\bar{\xi}}\left|\xi_{tn}^{TS} - \xi_0\right| \widetilde{Q}'_{tn,\boldsymbol{x}}\left(\xi_t^{TS}\right) d\xi_{tn}^{TS} + \int_{\underline{\xi}}^{\bar{\xi}}\left|\xi_{tn}^{TS} - \xi_0\right|\left(R'_{tn,\boldsymbol{x}}\left(\xi_{tn}^{TS}\right) - \widetilde{Q}'_{tn,\boldsymbol{x}}\left(\xi_t^{TS}\right)\right) d\xi_{tn}^{TS}$$

$$\leq \frac{C_1}{\sqrt{tn}} + \left(\bar{\xi} - \underline{\xi}\right) \cdot \int_{\underline{\xi}}^{\bar{\xi}}\left|R'_{tn,\boldsymbol{x}}\left(\xi_{tn}^{TS}\right) - \widetilde{Q}'_{tn,\boldsymbol{x}}\left(\xi_t^{TS}\right)\right| d\xi_{tn}^{TS}$$

$$\leq \frac{C_1}{\sqrt{tn}} + \frac{C_2}{\sqrt{tn}} = \mathcal{O}\left(\frac{1}{\sqrt{tn}}\right),$$

where the first inequality comes from the property of normal distribution and the second inequality comes from Proposition 1.5.2. The above result indicates that the expected learning

[150]

error decreases at rate $1/\sqrt{tn}$ as data accumulates over time t and potential demand rate n. Therefore, we can write (A.26) as

$$0 < \mathbb{E}_{\xi_t}\left[V_t\left(\lambda_{tn}, \xi_0\right) - V_t^{TS}\left(\lambda_{tn}, \xi_t\right)\right] < \mathcal{O}\left(\log\left(n\right) \cdot \int_{\tau=t}^{T} \frac{T-t}{\sqrt{tn}} d\tau\right) , \text{ for } t = 2, \cdots, T-1.$$
(A.27)

We should notice that (A.27) holds for $t > 1$. This is because upon making the initial price decision in period 1, there is no data available to learn the latent value ξ_0, which makes the regret of the first period do depend upon the prior distribution over ξ_0, $h_{\xi_0}\left(\cdot\right)$ and such a regret can never be avoided by any heuristics. Therefore, in summary we have

$$0 < V_1\left(\lambda_{1n}, \xi_0\right) - \mathbb{E}_{\xi_1}\left[V_1^{TS}\left(\lambda_{1n}, \xi_1\right)\right] < \frac{T \cdot \sigma_0}{\alpha} \cdot \log\left(n\right) + \mathcal{O}\left(\frac{T^{3/2} \cdot \log\left(n\right)}{\sqrt{n}}\right),$$
(A.28)

$$0 < V_2\left(\lambda_{2n}, \xi_0\right) - \mathbb{E}_{\xi_2}\left[V_2^{TS}\left(\lambda_{2n}, \xi_2\right)\right] < \mathcal{O}\left(\frac{T^{3/2} \cdot \log\left(n\right)}{\sqrt{n}}\right),$$
(A.29)

where $\sigma_0 = \mathbb{E}\left[\left|\xi_1 - \xi_0\right|\right]$, $\xi_1 \sim h_{\xi_0}\left(\cdot\right)$.

∎

Proof Proof for Corollary 1.5.3. To avoid confusion, let p_t^* denote the true optimal price derived from solving original optimization problem (1.2). From (A.14) and (A.23), we can get that

$$\left|p_t^{TS}\left(\xi_t\right) - p_t^*\right| = \left|p_t^{TS}\left(\xi_t\right) - p_t + p_t - p_t^*\right| \leq \left|p_t^{TS}\left(\xi_t\right) - p_t\right| + \left|p_t\left(\xi_t\right) - p_t^*\right|$$
$$\leq \mathcal{O}\left(\max\left\{\frac{T-t}{\sqrt{tn}}, \frac{T-t}{n^{1-\gamma}}\right\}\right) , \ t > 1,$$

and for $t = 1$, we have

$$\left|p_1^{TS}\left(\xi_1\right) - p_1^*\right| \leq \frac{T \cdot \sigma_0}{\alpha} + \mathcal{O}\left(\max\left\{\frac{T}{\sqrt{n}}, \frac{T}{n^{1-\gamma}}\right\}\right).$$

Then apply the similar analysis in Theorem 1.4.4, we can show that

$$0 < V_t^* \left(\lambda_{tn}, \xi_0 \right) - \mathbb{E}_{\xi_t} \left[V_t^{TS} \left(\lambda_{tn}, \xi_t \right) \right] < \mathcal{O} \left(\max \left\{ \frac{(T-t)^2}{n^{1-\gamma}}, \frac{(T-t)^{3/2}}{\sqrt{n}} \right\} \cdot \log(n) \right).$$

Especially, for $t = 1$, we need to include the error led by initial belief bias, and we get

$$0 < V_1^* \left(\lambda_{1n}, \xi_0 \right) - \mathbb{E}_{\xi_1} \left[V_1^{TS} \left(\lambda_{1n}, \xi_1 \right) \right]$$

$$< \frac{T \cdot \sigma_0}{\alpha} \cdot \log(n) \cdot \mathbb{1} \left\{ t = 1 \right\} + \mathcal{O} \left(\max \left\{ \frac{T^2}{n^{1-\gamma}}, \frac{T^{3/2}}{\sqrt{n}} \right\} \cdot \log(n) \right).$$

Take expectation over the demand rate, we complete the proof. ∎

Proof   Proof for Theorem 1.5.4. Since the proof is already included in the proof for Theorem 1.5.2, we omit it here. ∎

Proof   Proof for Corollary 1.5.5. Since the proof mirrors the proof for Corollary 1.5.3, we omit it here. ∎

Appendix: Bernstein-von Mises Theorem ([45])

Let $(X, \mathscr{A})$ be a measurable space and $P_\theta | \mathscr{A}$, $\theta \in \Theta$, a family of probability measures, where Θ is an open subset of $\mathbb{R}^k$. Let θ be a random variable with prior distribution $\lambda | \mathscr{B}^k \cap \Theta$. Assume that λ has a finite density ρ with respect to the Lebesgue measure, which is positive on Θ and zero on Θ^c. Let $R_{n,\boldsymbol{x}}$ be the posterior distribution of θ for the sample size n given $\boldsymbol{x} \in X^n$, which is defined as

$$R_{n,\boldsymbol{x}}(B) = \frac{\int_B \Pi_{i=1}^n p(x_i, \sigma) \rho(\sigma) \, d\sigma}{\int \Pi_{i=1}^n p(x_i, \sigma) \rho(\sigma) \, d\sigma}, \quad B \in \mathscr{B}^k,$$

where $p(\cdot, \theta)$ is a density of $P_\theta | \mathscr{A}$ with respect to a dominating measure. Let $Q_{n,\boldsymbol{x}}$ be the normal distribution centered at the maximum likelihood estimator $\theta_n(\boldsymbol{x})$ with covariance matrix $\Gamma_n(\boldsymbol{x})^{-1}$, where

$$\Gamma_n(\boldsymbol{x}) = \left(\sum_{\nu=1}^n \frac{\partial^2}{\partial \theta_i \partial \theta_j} f(x_\nu, \theta) \Big|_{\theta=\theta_n(\boldsymbol{x})} \right)_{i,j=1,\cdots,k},$$

which is positive definite and $f(x_\nu, \theta) = -\log p(x_\nu, \theta)$.

Define the variational distance between the measures $R_{n,\boldsymbol{x}}$ and $Q_{n,\boldsymbol{x}}$ as

$$d(R_{n,\boldsymbol{x}}, Q_{n,\boldsymbol{x}}) = \sup \left\{ |R_{n,\boldsymbol{x}}(B) - Q_{n,\boldsymbol{x}}(B)| : \ B \in \mathscr{B}^k \right\}.$$

If the regularity conditions which depend on an integer $s \geq 2$ in [45] Section 4 hold (also see below), then for every compact subset K of Θ there exists a constant $c_K > 0$ with

$$\sup_{\theta \in K} P_\theta^n \left\{ \boldsymbol{x} \in X^n : d(R_{n,\boldsymbol{x}}, Q_{n,\boldsymbol{x}}) > c_K n^{-1/2} \right\} = \mathcal{O}\left(n^{-s/2} \right).$$

[153]

Regularity Conditions ([45])

In the following s denotes an integer with $s \geq 2$. Let $f'(x,\theta) = (\partial/\partial\theta) f(x,\theta)$ and $f''(x,\theta) = (\partial^2/\partial\theta^2) f(x,\theta)$.

1. $\theta \to P_\theta$ is continuous on Θ with respect to the supremum-metric on $\{P_\theta : \theta \in \Theta\}$.

2. For each $x \in X$, $\theta \to f(x,\theta)$ is continuous on $\bar{\Theta}$.

3. For every $\theta \in \Theta$, there exists an open neighborhood U_θ of θ such that

$$\sup \{\mathbb{E}_\theta |f(\cdot,\tau)|^s : \sigma, \tau \in U_\theta\} < \infty.$$

4. For every $(\theta, \tau) \in \Theta \times \bar{\Theta}$, $\theta \neq \tau$, there exists neighborhood $U_{\theta,\tau}$ of θ and $V_{\theta,\tau}$ of τ such that for all neighborhood V of τ with $V \subset V_{\theta,\tau}$,

$$\sup \left\{ \mathbb{E}_\sigma |\inf_{\delta \in V} f(\cdot,\delta)|^s : \sigma \in U_{\theta,\sigma} \right\} < \infty.$$

5. For each $x \in X$, $\theta \to f(x,\theta)$ is twice differentiable in Θ.

6. For every $\theta \in \Theta$, there exists an open neighborhood U_θ of θ such that

   (a) $\inf \{\lambda_0(\tau) : \tau \in U_\theta\} > 0$, where $\lambda_0(\tau)$ is the smallest eigenvalue of $\mathbb{E}_\tau f''(\cdot,\tau)$.

   (b) $\sup \{\mathbb{E}_\tau \|f''(\cdot,\tau)\|^s : \tau \in U_\theta\} < \infty$.

7. For every $\theta \in \Theta$, there exists an open neighborhood U_θ of θ and a measurable function $k_\theta : X \to \bar{\mathbb{R}}$ such that

   (a) for every $\tau \in \Theta$ there exists an open neighborhood V_τ of τ with $\sup \{\mathbb{E}_\sigma k_\theta^s : \sigma \in V_\tau\} < \infty$.

   (b) $\|f''(x,\tau) - f''(x,\sigma)\| \leq \|\tau - \sigma\| k_\theta(x)$ for all $\tau, \sigma \in U_\theta$, $x \in X$.

8. The probability measure $\lambda|\mathscr{B}^k$ has a finite Lebesgue-density ρ, which is positive on Θ and zero on Θ^c.

[154]

9. For every $\theta \in \Theta$ there exists an open neighborhood U_θ of θ and a constant $c_\theta > 0$ such that

$$|\log \rho\,(\sigma) - \log \rho\,(\tau)| \leq \|\sigma - \tau\|\, c_\theta, \quad \text{for all } \sigma, \tau \in U_\theta.$$

Statement of Local Limit Theorem ([104] Chapter VII)

Let $\{X_n\}$ be a sequence of independent random variables with a common distribution and density $p\,(x)$, such that

$$\mathbb{E}|X_1|^3 < \infty\ ,\ \mathbb{E}X_1 = 0\ ,\ \mathbb{E}X_1^2 = 1\ ,\ \sup_x p\,(x) \leq C.$$

Let $p_n\,(x)$ be the density of the random variable $\frac{1}{\sqrt{n}} \sum_{j=1}^n X_j$. Then

$$\sup_x \left| p_n\,(x) - \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \right| \leq \frac{A\beta_3}{\sqrt{n}} \max\left(1, C^3\right),$$

and for all x

$$\left| p_n\,(x) - \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \right| \leq \frac{A\beta_3^{2m-1} \max\left(1, C^{2m+1}\right)}{\sqrt{n}\,(1 + |x|^m)},$$

for $m = 2$ and $m = 3$, where $\beta_3 = \mathbb{E}|X_1|^3$ and A is an absolute constant.

[155]

B. Appendix for Chapter 2

Appendix A: Cournot Competing Retailers

To characterize the game, we use backward induction and in this section, we first look at the third and forth-stage equilibrium of the supplier and retailers. Because production cost c is realized and publicly known the beginning of stage three, the game is exactly the same as in [50] and we keep our notations consistent with this paper. $\sum_{j \in K} Y_j$ is a sufficient statistic for estimating θ and the equilibrium wholesale price w is a monotone function of $\sum_{j \in K} Y_j$ which implies that the retailers can have unbiased inference of market condition via realized w. Given (Y_i, w) observed by retailer i and the other retailers' order quantity q_l, $l \neq i$, her expected profit with respect of order quantity q_i is

$$
\mathbb{E}[\pi_i | Y_i, w] = \left(a + \mathbb{E}[\theta | Y_i, w] - q_i - \sum_{l \neq i} \mathbb{E}[q_l | Y_i, w] - w \right) q_i
$$
$$
= \left(a + \frac{1}{k+s} \sum_{j \in K} Y_j - q_i - \sum_{l \neq i} \mathbb{E}[q_l | Y_i, w] - w \right) q_i.
$$

And the unique Cournot-Bayesian equilibrium strategies are that for $i \in K$,

$$
q_i^*(Y_i, w) = \frac{1}{n+1} \left(a - w + A_1^k \sum_{j \in K} Y_j \right), \tag{B.1}
$$

where $A_1^k = \frac{1}{k+s}$ and for $l \in N \setminus K$,

$$
q_l^*(Y_l, w) = \frac{1}{n+1} \left(a - w + B_1^k \sum_{j \in K} Y_j + B_2^k Y_l \right), \tag{B.2}
$$

[156]

where $B_1^k = \frac{k+2s}{(k+s)(n+k+1+2s)}$ and $B_2^k = \frac{n+1}{n+k+1+2s}$. And correspondingly, the retailer's equilibrium profit is $\mathbb{E}[\pi_i^* | Y_i, w] = [q_i^*(Y_i, w)]^2$, $i \in N$. A direct property of $q_l^*(Y_l, w)$ is that

$$\mathbb{E}[q_l^*(Y_l, w) | Y_j, j \in K] = q_i^*(Y_i, w),$$

where $i \in K$ and $l \in N \setminus K$. Therefore, given the shared signals Y_j, $j \in K$, the supplier expects to have the same expected order quantity from all the retailers. [50] has shown that the supplier should not do price discrimination upon the retailers' information-sharing decisions. Whence given the shared information, the expected total order quantity is

$$\mathbb{E}[D_S | (Y_j), j \in K] = \mathbb{E}\left[\sum_{i \in N} q_i^*(Y_i, w) | (Y_j), j \in K\right] = \frac{n}{n+1}\left(a - w + A_1^k \sum_{j \in K} Y_j\right), \quad \text{(B.3)}$$

and the expected profit for realized cost c is

$$\mathbb{E}[\pi_S | c, (Y_j), j \in K] = (w - c)\mathbb{E}[D_S | (Y_j), j \in K],$$

thus the optimal wholesale price is

$$w^*(c, (Y_j), j \in K) = \frac{1}{2}\left(a + c + A_1^k \sum_{j \in K} Y_j\right) \quad \text{(B.4)}$$

The corresponding supplier's equilibrium expected profit is

$$\mathbb{E}[\pi_S^* | c, (Y_j), j \in K] = \frac{n}{4(n+1)}\left(a + A_1^k \sum_{j \in K} Y_j - c\right)^2, \quad \text{(B.5)}$$

[157]

Then the expected profit difference between hedging and not hedging is

$$\mathbb{E}[\pi_S^*|\bar{c}, (Y_j), j \in K] - \mathbb{E}[\mathbb{E}_\xi[\pi_S^*|c(\xi), (Y_j), j \in K]]$$

$$= \frac{n}{4(n+1)} \left(2 \left(a + A_1^k \sum_{j \in K} Y_j \right) (\mathbb{E}_\xi[c] - \bar{c}) - \left(\mathbb{E}_\xi[c^2] - \bar{c}^2 \right) \right).$$

Similar to (2.1), the supplier chooses to hedge if and only if

$$\sum_{j \in K} Y_j \geq (k+s) \left(\frac{\mathbb{E}_\xi[c^2] - \bar{c}^2}{2 \left(\mathbb{E}_\xi[c] - \bar{c} \right)} - a \right) = (k+s)(t - a) \triangleq t_Y(k). \tag{B.6}$$

Specially, when there is no retailer sharing information, the left hand side of the above criteria is substituted by zero, which means that the supplier chooses to hedge if $a \geq t$ and not to hedge if $a < t$. Therefore from (B.5), the supplier's expected equilibrium profit given that K $(K \geq 1)$ retailers share information is

$$\Pi_S(k) = \mathbb{E}_\theta \left[\mathbb{E}_{\sum_{j \in K} Y_j | \theta} \left[\mathbb{E}_\xi \left[\pi_S^* | c, (Y_j), j \in K \right] \Big| \theta \right] \right]$$

$$= \frac{n}{4(n+1)} \mathbb{E}_\theta \left[\mathbb{E}_\xi \left[\int_{-\infty}^{t_Y(k)} \left(a + A_1^k Y - c \right)^2 dF_k(Y) + \int_{t_Y(k)}^{\infty} \left(a + A_1^k Y - \bar{c} \right)^2 dF_k(Y) \right] \right]$$

$$= \frac{n}{4(n+1)} \mathbb{E}_\theta \left[(a - \bar{c})^2 + 2A_1^k k\theta(a - \bar{c}) + (A_1^k)^2(k^2\theta^2 + k\sigma^2) \right.$$

$$- 2(a + A_1^k t_Y(k)) \left(\mathbb{E}_\xi[c] - \bar{c} \right) F_k(t_Y(k))$$

$$\left. + 2 \int_{-\infty}^{t_Y(k)} A_1^k (\mathbb{E}_\xi[c] - \bar{c}) F_k(Y) dY + (\mathbb{E}_\xi[c^2] - \bar{c}^2) F_k(t_Y(k)) \right]$$

$$= \frac{n}{4(n+1)} \left((a - \bar{c})^2 + (A_1^k)^2(k^2\sigma_0^2 + k\sigma^2) + 2A_1^k(\mathbb{E}_\xi[c] - \bar{c}) \cdot \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(Y) dY \right), \tag{B.7}$$

[158]

where $F_k(\cdot)$ is the cdf of distribution $N(k\theta, k\sigma^2)$ and $\widetilde{F}_k(\cdot)$ is the cdf of $N(0, k^2\sigma_0^2 + k\sigma^2)$.

And when there is no information sharing,

$$
\Pi_S(0) = \begin{cases} \dfrac{n}{4(n+1)}(a-\bar{c})^2, & \text{if } a \geq t, \\[2ex] \dfrac{n}{4(n+1)}\left(a^2 - 2\mathbb{E}_\xi[c]a + \mathbb{E}_\xi[c^2]\right), & \text{if } a < t. \end{cases} \tag{B.8}
$$

The retailers' equilibrium profits given realized cost c and shared signals $(Y_j)_{j \in K}$, $k \geq 1$ are

$$
\mathbb{E}_{(Y_m)_{m \in N \setminus K}}[\pi_i^* | c, (Y_j), j \in K] = \frac{1}{4(n+1)^2}\left(a + A_1^k \sum_{j \in K} Y_j - c\right)^2, \quad i \in K, \tag{B.9}
$$

and

$$
\mathbb{E}_{(Y_m)_{m \in N \setminus K}}[\pi_l^* | c, Y_l, (Y_j), j \in K]
$$

$$
= \frac{1}{4(n+1)^2}\left(a + (2B_1^k - A_1^k)\sum_{j \in K} Y_j + 2B_2^k Y_l - c\right)^2, \quad l \in N \setminus K. \tag{B.10}
$$

Given (B.6), (B.9) and (B.10), we can calculate retailers' expected profits given k retailers share information. Let $\sum_{j \in K} Y_j = Y(k)$, then $Y(k)|\theta \sim N(k\theta, k\sigma^2)$. For the retailer who shares information, i.e., $i \in K$, the expected profit is:

$$\Pi_R^S(k) = \mathbb{E}_\theta \left[\mathbb{E}_\xi \left[\mathbb{E}_{\sum_{j \in K} Y_j | \theta} \left[\mathbb{E}_{(Y_m)_{m \in N \setminus K}} [\pi_i^* | c, (Y_j), j \in K] \right] \middle| \theta \right] \right]$$

$$= \frac{1}{4(n+1)^2} \mathbb{E}_\theta \left[\mathbb{E}_\xi \left[\int_{-\infty}^{t_Y(k)} \left((A_1^k Y)^2 + 2A_1^k (a-c) Y + (a-c)^2 \right) dF_k(Y) \right. \right.$$

$$\left. \left. + \int_{t_Y(k)}^{\infty} \left((A_1^k Y)^2 + 2A_1^k (a-\bar{c}) Y + (a-\bar{c})^2 \right) dF_k(Y) \middle| \theta \right] \right]$$

$$= \frac{1}{4(n+1)^2} \mathbb{E}_\theta \left[\int_{-\infty}^{\infty} \left((A_1^k Y)^2 + 2A_1^k (a-\bar{c}) Y + (a-\bar{c})^2 \right) dF_k(Y) \right.$$

$$\left. - \int_{-\infty}^{t_Y(k)} \left(2A_1^k (\mathbb{E}_\xi[c] - \bar{c}) Y + 2a(\mathbb{E}_\xi[c] - \bar{c}) - (\mathbb{E}_\xi[c^2] - \bar{c}^2) \right) dF_k(Y) \right]$$

$$= \frac{1}{4(n+1)^2} \left((a-\bar{c})^2 + (A_1^k)^2(k^2\sigma_0^2 + k\sigma^2) + 2A_1^k(\mathbb{E}_\xi[c] - \bar{c}) \cdot \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(Y) dY \right),$$

$$(B.11)$$

where $\widetilde{F}_k(\cdot)$ is the cdf of $N(0, k^2\sigma_0^2 + k\sigma^2)$ and the above calculation involves two helpful facts:

$$\mathbb{E}_\theta [F_k(t)] = \widetilde{F}_k(y) \ , \ \mathbb{E}_\theta \left[\int_{-\infty}^{t} y dF_k(y) \right] = t \widetilde{F}_k(t) - \int_{-\infty}^{t} \widetilde{F}_k(y) \, dy.$$

For the retailer who does not share information, i.e., $l \in N \setminus K$,

$$\Pi_R^N(k) = \mathbb{E}_\theta \left[\mathbb{E}_\xi \left[\mathbb{E}_{Y_l, \sum_{j \in K} Y_j | \theta} \left[\mathbb{E}_{(Y_m)_{m \in N \setminus K}} [\pi_l^* | c, (Y_j), j \in K] \right] \middle| \theta \right] \right]$$

$$= \frac{1}{4(n+1)^2} \mathbb{E}_\theta \left[\mathbb{E}_\xi \left[\mathbb{E}_{Y_l | \theta} \left[\int_{-\infty}^{t_Y(k)} \left(a + (2B_1^k - A_1^k) Y + 2B_2^k Y_l - c \right)^2 dF_k(Y) \right. \right. \right.$$

$$\left. \left. \left. + \int_{t_Y(k)}^{\infty} \left(a + (2B_1^k - A_1^k) Y + 2B_2^k Y_l - \bar{c} \right)^2 dF_k(Y) \right] \right] \right]$$

$$\overset{(B.11)}{=} \frac{1}{4(n+1)^2} \left((a - \bar{c})^2 + (2B_1^k - A_1^k)^2 (k^2 \sigma_0^2 + k\sigma^2) \right.$$

$$- \mathbb{E}_\theta \left[\int_{-\infty}^{t_Y(k)} \left(2 \left(2B_1^k - A_1^k \right) \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) Y + 2a \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) - \left(\mathbb{E}_\xi \left[c^2 \right] - \bar{c}^2 \right) \right) dF_k \left(Y \right) \right]$$

$$+ \mathbb{E}_\theta \left[\mathbb{E}_\xi \left[\mathbb{E}_{Y_l | \theta} \left[4 \left(B_2^k \right)^2 Y_l^2 + 4 \left(2B_1^k - A_1^k \right) B_2^k \cdot \int_{-\infty}^{\infty} Y \cdot Y_l \ dF_k \left(Y \right) \right. \right. \right.$$

$$\left. \left. \left. \left. + 4B_2^k \cdot \int_{-\infty}^{t_Y(k)} \left(a - c \right) Y_l \ dF_k(Y) + 4B_2^k \cdot \int_{t_Y(k)}^{\infty} \left(a - \bar{c} \right) Y_l \ dF_k(Y) \right] \right] \right] \right)$$

$$= \frac{1}{4(n+1)^2} \left((a - \bar{c})^2 + (2B_1^k - A_1^k)^2 (k^2 \sigma_0^2 + k\sigma^2) \right.$$

$$- 2 \left(2B_1^k - A_1^k \right) \left(\mathbb{E}_\xi[c] - \bar{c} \right) \left(t_Y \left(k \right) \widetilde{F}_k \left(t_Y \left(k \right) \right) - \int_{-\infty}^{t_Y(k)} \widetilde{F}_k \left(Y \right) dY \right)$$

$$- 2 \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) \left(a - t \right) \widetilde{F}_k \left(t_Y \left(k \right) \right)$$

$$+ 4 \left(B_2^k \right)^2 \left(\sigma_0^2 + \sigma^2 \right) + 4 \left(2B_1^k - A_1^k \right) B_2^k \cdot k\sigma_0^2 - 4A_1^k B_2^k \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) t_Y \left(k \right) \widetilde{F}_k \left(t_Y \left(k \right) \right)$$

$$\left. + 4A_1^k B_2^k \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) \cdot \int_{-\infty}^{t_Y(k)} \widetilde{F}_k \left(Y \right) dY \right)$$

$$= \frac{1}{4(n+1)^2} \left((a - \bar{c})^2 + (2B_1^k - A_1^k)^2 (k^2 \sigma_0^2 + k\sigma^2) + 4 \left(2B_1^k - A_1^k \right) B_2^k \cdot k\sigma_0^2 \right.$$

$$+ 4 \left(B_2^k \right)^2 \left(\sigma_0^2 + \sigma^2 \right) + 2A_1^k (\mathbb{E}_\xi[c] - \bar{c}) \times$$

$$\left. \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(Y) dY - 2 \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) \left(\left(2B_1^k - A_1^k \right) - A_1^k + 2A_1^k B_2^k \right) t_Y \left(k \right) \widetilde{F}_k \left(t_Y \left(k \right) \right) \right)$$

$$= \frac{1}{4(n+1)^2} \left((a - \bar{c})^2 + (2B_1^k - A_1^k)^2 (k^2 \sigma_0^2 + k\sigma^2) + 4 \left(2B_1^k - A_1^k \right) B_2^k \cdot k\sigma_0^2 \right.$$

$$\left. + 4 \left(B_2^k \right)^2 \left(\sigma_0^2 + \sigma^2 \right) + 2A_1^k (\mathbb{E}_\xi[c] - \bar{c}) \cdot \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(Y) dY \right), \qquad \text{for } k \geq 1, \tag{B.12}$$

where the last equality is because $B_1^k - A_1^k + A_1^k B_2^k = \frac{k+2s}{(k+s)(n+k+1+2s)} - \frac{1}{k+s} + \frac{n+1}{(k+s)(n+k+1+2s)} = 0$.

Specially, for $k = 0$ we have

$$\Pi_R^N(0) = \begin{cases} \frac{1}{4(n+1)} \left((a - \bar{c})^2 + 4\left(B_2^0\right)^2 \left(\sigma_0^2 + \sigma^2\right) \right), & \text{if } a \geq t, \\ \frac{1}{4(n+1)} \left((a^2 - 2\mathbb{E}_\xi[c]a + \mathbb{E}_\xi[c^2]) + 4\left(B_2^0\right)^2 \left(\sigma_0^2 + \sigma^2\right) \right), & \text{if } a < t. \end{cases} \tag{B.13}$$

Then we investigate $\Pi_R^S(k) - \Pi_R^N(k-1)$ to determine the information sharing equilibrium.

$$\Pi_R^S(k) - \Pi_R^N(k-1)$$

$$= \frac{1}{4(n+1)^2} \Bigg[\left((A_1^k)^2(k^2\sigma_0^2 + k\sigma^2) - (2B_1^{k-1} - A_1^{k-1})^2\left((k-1)^2\sigma_0^2 + (k-1)\sigma^2\right) \right.$$

$$\left. - 4(2B_1^{k-1} - A_1^{k-1})B_2^{k-1}(k-1)\sigma_0^2 - 4(B_2^{k-1})^2(\sigma_0^2 + \sigma^2) \right)$$

$$+ 2(\mathbb{E}_\xi[c] - \bar{c})\left(A_1^k \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(Y)dY - A_1^{k-1} \int_{-\infty}^{t_Y(k-1)} \widetilde{F}_{k-1}(Y)dY \right) \Bigg]$$

$$= \frac{1}{4(n+1)^2} \Bigg[- \frac{\sigma_0^4\sigma^2\left(\left((2n+1)k - n\right)\sigma_0^2 + 2n\sigma^2\right)\left(\left((2n+3)k + n\right)\sigma_0^2 + 2(n+2)\sigma^2\right)}{\left(k\sigma_0^2 + \sigma^2\right)\left((k-1)\sigma_0^2 + \sigma^2\right)\left((n+k)\sigma_0^2 + 2\sigma^2\right)^2}$$

$$+ 2(\mathbb{E}_\xi[c] - \bar{c})\left(\int_{-\infty}^{t-a} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right) \Bigg]$$

$$= \frac{1}{4(n+1)^2} \Bigg[- \frac{\sigma_0^4\sigma^2\left(\left((2n+1)k - n\right)\sigma_0^2 + 2n\sigma^2\right)\left(\left((2n+3)k + n\right)\sigma_0^2 + 2(n+2)\sigma^2\right)}{\left(k\sigma_0^2 + \sigma^2\right)\left((k-1)\sigma_0^2 + \sigma^2\right)\left((n+k)\sigma_0^2 + 2\sigma^2\right)^2}$$

$$+ 2(\mathbb{E}_\xi[c] - \bar{c})\left(\int_{-\infty}^{-|t-a|} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right) \Bigg], \quad k > 1,$$

[162]

where $\sigma_k^2 = \frac{k\sigma_0^2}{k+\sigma^2/\sigma_0^2}$ and $\Phi(\cdot)$ is the cdf of standard normal distribution. And

$$\Pi_R^S(1) - \Pi_R^N(0)$$

$$= \begin{cases} \frac{1}{4(n+1)^2}\left(-(4(B_2^0)^2 - (A_1^1)^2)(\sigma_0^2 + \sigma^2) + 2(\mathbb{E}_\xi[c] - \bar{c})\int_{-\infty}^{t-a}\Phi\left(\frac{x}{\sigma_1}\right)dx\right), & \text{if } a \geq t, \\ \frac{1}{4(n+1)^2}\left(-(4(B_2^0)^2 - (A_1^1)^2)(\sigma_0^2 + \sigma^2) + 2(\mathbb{E}_\xi[c] - \bar{c})\int_{t-a}^{\infty}\overline{\Phi}\left(\frac{x}{\sigma_1}\right)dx\right), & \text{if } a < t. \end{cases}$$

$$= \frac{1}{4(n+1)^2}\left(-\left(4(B_2^0)^2 - (A_1^1)^2\right)(\sigma_0^2 + \sigma^2) + 2(\mathbb{E}_\xi[c] - \bar{c})\int_{-\infty}^{-|t-a|}\Phi\left(\frac{x}{\sigma_1}\right)dx\right).$$

We notice that we still have a unified form for the profit difference of $k > 1$ and $k = 1$, that is,

$$\Pi_R^S(k) - \Pi_R^N(k-1)$$

$$= \frac{1}{4(n+1)^2}\left[-\frac{\sigma_0^4\sigma^2\left(\left((2n+1)k - n\right)\sigma_0^2 + 2n\sigma^2\right)\left(\left((2n+3)k + n\right)\sigma_0^2 + 2(n+2)\sigma^2\right)}{\left(k\sigma_0^2 + \sigma^2\right)\left((k-1)\sigma_0^2 + \sigma^2\right)\left((n+k)\sigma_0^2 + 2\sigma^2\right)^2}\right.$$

$$\left. + 2(\mathbb{E}_\xi[c] - \bar{c})\left(\int_{-\infty}^{-|t-a|}\Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right)dx\right)\right]. \tag{B.14}$$

Similarly, we can also write down the marginal utilities of the other retailers and the supplier respectively when one additional retailer shares information. For the retailers who have already shared their information:

$$\Pi_R^S(k) - \Pi_R^S(k-1) = \frac{1}{4(n+1)^2}\left[\frac{\sigma_0^4\sigma^2}{(k\sigma_0^2 + \sigma^2)((k-1)\sigma_0^2 + \sigma^2)}\right.$$

$$\left. + 2(\mathbb{E}_\xi[c] - \bar{c})\left(\int_{-\infty}^{-|t-a|}\Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right)dx\right)\right].$$

[163]

And for the retailers who do not share information,

$$\Pi_R^N(k) - \Pi_R^N(k-1) = \frac{1}{4(n+1)^2} \left[\frac{\sigma_0^4 \sigma^2}{(k\sigma_0^2 + \sigma^2)((k-1)\sigma_0^2 + \sigma^2)} \right.$$

$$- \frac{1}{((k-1)\sigma_0^2 + \sigma^2)((n+k)\sigma_0^2 + 2\sigma^2)^2(k\sigma_0^2 + \sigma^2)((n+k+1)\sigma_0^2 + 2\sigma^2)^2} \times$$

$$\left(4(n+1)^2 \sigma_0^6 \sigma^2 \left(\left(2k^2(k+1) + 2k(k+1)n + n^2 \right) \sigma_0^6 \right. \right.$$

$$+ 2(k(4k+3) + 2(k+1)n)\sigma_0^4 \sigma^2 + (5 + 10k + 2n)\sigma_0^2 \sigma^4 + 4\sigma^6 \big)$$

$$\left. \left. + 2(\mathbb{E}_\xi[c] - \bar{c}) \left(\int_{-\infty}^{-|t-a|} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right) \right) \right].$$

For the supplier:

$$\Pi_S(k) - \Pi_S(k-1) = \frac{n}{4(n+1)} \left[\frac{\sigma_0^4 \sigma^2}{(k\sigma_0^2 + \sigma^2)((k-1)\sigma_0^2 + \sigma^2)} \right.$$

$$\left. + 2(\mathbb{E}_\xi[c] - \bar{c}) \left(\int_{-\infty}^{-|t-a|} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right) \right].$$

Simple calculation gives the following relation, for any $k \in \{1, \cdots, n\}$,

$$\Pi_R^S(k) - \Pi_R^N(k-1) < \Pi_R^N(k) - \Pi_R^N(k-1) < \Pi_R^S(k) - \Pi_R^S(k-1)$$

$$= \frac{1}{n(n+1)} (\Pi_S(k) - \Pi_S(k-1)).$$

[164]

Appendix B: Duopoly Information-Sharing Game

Scenario (S, S). This is the case when both retailers share the market information to the supplier. Following the analysis in section 2.3.1, given the realization of production cost c, the equilibrium wholesale price is $w_m^* = (\tilde{a}_m + c)/2$ and the supplier's profit in market m is $\pi_{Sm}(w_m^*; \tilde{a}_m, c) = (\tilde{a}_m - c)^2/8$, $m \in \{1, 2\}$, where $\tilde{a}_m = a_m + \theta_m$. The retailer m's corresponding order quantity is $q_m^* = (\tilde{a}_m - w_m^*)/2 = (\tilde{a}_m - c)/4$ and profit is $\pi_{Rm}(q_m^*; \tilde{a}_m) = (\tilde{a}_m - c)^2/16$. Therefore the supplier's expected total profit when he does not hedge is

$$\pi_S^{nh*}(\tilde{a}_1, \tilde{a}_2) = \mathbb{E}_\xi \left[\frac{(\tilde{a}_1 - c)^2 + (\tilde{a}_2 - c)^2}{8} \right] = \frac{1}{8} \left(\tilde{a}_1^2 + \tilde{a}_2^2 - 2(\tilde{a}_1 + \tilde{a}_2) \mathbb{E}_\xi[c] + 2\mathbb{E}_\xi[c^2] \right),$$

on the other hand, if the supplier hedges, the expected total profit is

$$\pi_S^{h*}(\tilde{a}_1, \tilde{a}_2) = \frac{1}{8} \left(\tilde{a}_1^2 + \tilde{a}_2^2 - 2(\tilde{a}_1 + \tilde{a}_2) \bar{c} + 2\bar{c}^2 \right).$$

thus we get

$$\pi_S^{h*}(\tilde{a}_1, \tilde{a}_2) - \pi_S^{nh*}(\tilde{a}_1, \tilde{a}_2) = \frac{1}{4} \left((\mathbb{E}_\xi[c] - \bar{c})(\tilde{a}_1 + \tilde{a}_2) - (\mathbb{E}_\xi[c^2] - \bar{c}^2) \right),$$

whence the supplier chooses to hedge if and only if

$$\frac{\tilde{a}_1 + \tilde{a}_2}{2} \geq \frac{\mathbb{E}_\xi[c^2] - \bar{c}^2}{2(\mathbb{E}_\xi[c] - \bar{c})} \triangleq t, \tag{B.15}$$

which is quite similar to (2.1) but just substitute the single market size to the average size of the two markets (and it is straightforward to see that when there are more than two markets, the left hand side of the above criteria will be extended to the average size of all the markets). Discussions fall into two cases: (A) Low average market size with $2(t - \delta) < a_1 + a_2 < 2t$ and (B) High average market size with $2t < a_1 + a_2 < 2(t + \delta)$. In the first case, the

[165]

supplier will only hedge when both of the markets turn out to be good (that is, $\widetilde{a}_m = a_m + \delta$, $m = 1, 2$) and in the second case, the supplier will hedge unless both of the markets are bad. The other two cases where $a_1 + a_2 < 2(t - \delta)$ (average market size is too low) and $a_1 + a_2 > 2(t + \delta)$ (average market size is too high) are off our interest, because in those two cases, the supplier's hedging decision would be independent of realized market conditions.

Case (A). When we incorporate hedging decision, the supplier's expected total profit is

$$
\mathbb{E}\left[\pi_S^{nh*}\left(\widetilde{a}_1, \widetilde{a}_2\right) \mathbb{1}\{\bar{a} < t\}\right] + \mathbb{E}\left[\pi_S^{h*}\left(\widetilde{a}_1, \widetilde{a}_2\right) \mathbb{1}\{\bar{a} \geq t\}\right]
$$

$$
= \mathbb{E}\left[\frac{1}{8}\left(\widetilde{a}_1^2 + \widetilde{a}_2^2\right)\right] - \mathbb{E}\left[\left(\frac{1}{2}\bar{a}\mathbb{E}_\xi[c] - \frac{1}{4}\mathbb{E}_\xi[c^2]\right)\mathbb{1}\{\bar{a} < t\}\right] - \mathbb{E}\left[\left(\frac{1}{2}\bar{a}\bar{c} - \frac{1}{4}\bar{c}^2\right)\mathbb{1}\{\bar{a} \geq t\}\right]
$$

$$
= \frac{1}{8}\Big[a_1^2 + a_2^2 + 2\delta^2 - \left(2\left(1 - \alpha\right)\mathbb{E}_\xi[c] + 2\alpha\bar{c}\right)\left(a_1 + a_2\right) + 4\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)\delta
$$

$$
+ 2(1 - \alpha)\mathbb{E}_\xi[c^2] + 2\alpha\bar{c}^2\Big]. \tag{B.16}
$$

For the retailer m, if the supplier does not hedge, given the realization of the market condition, the expected profit is

$$
\pi_{Rm}^{nh*}\left(\widetilde{a}_1, \widetilde{a}_2\right) = \frac{1}{16}\left(\widetilde{a}_m^2 - 2\widetilde{a}_m\mathbb{E}_\xi[c] + \mathbb{E}_\xi\left[c^2\right]\right),
$$

and if the supplier hedges, we have

$$
\pi_{Rm}^{h*}\left(\widetilde{a}_1, \widetilde{a}_2\right) = \frac{1}{16}\left(\widetilde{a}_m^2 - 2\widetilde{a}_m\bar{c} + \bar{c}^2\right).
$$

When the realized first market condition is, say $\widetilde{a}_m$, then the probability that the supplier will not hedge is

$$
P\left(\widetilde{a}_{-m} < 2t - \widetilde{a}_m|\widetilde{a}_m\right) = \begin{cases} 1 - 2\alpha, & \widetilde{a}_m = a_m + \delta, \\ \\ 1, & \widetilde{a}_m = a_m - \delta. \end{cases}
$$

[166]

Therefore, the total expected profit of retailer m is

$$
\begin{aligned}
\pi_{Rm}^* &= \mathbb{E}_{\widetilde{a}_m}\left[P\left(\widetilde{a}_{-m} < 2t - \widetilde{a}_m\right)\left(\frac{1}{16}\left(\widetilde{a}_m^2 - 2\widetilde{a}_m\mathbb{E}_\xi\left[c\right] + \mathbb{E}_\xi\left[c^2\right]\right)\right) \right. \\
&\quad \left. + \left(1 - P\left(\widetilde{a}_{-m} < 2t - \widetilde{a}_m\right)\right)\left(\frac{1}{16}\left(\widetilde{a}_m^2 - 2\widetilde{a}_m\bar{c} + \bar{c}^2\right)\right) \right] \\
&= \frac{1}{2}\left[(1 - 2\alpha)\left(\frac{1}{16}\left((a_m + \delta)^2 - 2(a_m + \delta)\mathbb{E}_\xi\left[c\right] + \mathbb{E}_\xi\left[c^2\right]\right)\right) \right. \\
&\quad \left. + 2\alpha\left(\frac{1}{16}\left((a_m + \delta)^2 - 2(a_m + \delta)\bar{c} + \bar{c}^2\right)\right) \right] \\
&\quad + \frac{1}{2}\left(\frac{1}{16}\left((a_m - \delta)^2 - 2(a_m - \delta)\mathbb{E}_\xi\left[c\right] + \mathbb{E}_\xi\left[c^2\right]\right)\right) \\
&= \frac{1}{16}\left(a_m^2 + \delta^2 - \left(2(1 - \alpha)\mathbb{E}_\xi[c] + 2\alpha\bar{c}\right)a_m + 2\alpha\delta\left(\mathbb{E}_\xi[c] - \bar{c}\right) + (1 - \alpha)\mathbb{E}_\xi[c^2] + \alpha\bar{c}^2\right).
\end{aligned}
$$

(B.17)

Case (B)**.** For the other case, the supplier's expected profit is

$$
\begin{aligned}
&\mathbb{E}\left[\pi_S^{nh*}\left(\widetilde{a}_1, \widetilde{a}_2\right)\mathbb{1}\{\bar{a} < t\}\right] + \mathbb{E}\left[\pi_S^{h*}\left(\widetilde{a}_1, \widetilde{a}_2\right)\mathbb{1}\{\bar{a} \geq t\}\right] \\
=&\mathbb{E}\left[\frac{1}{8}\left(\widetilde{a}_1^2 + \widetilde{a}_2^2\right)\right] - \mathbb{E}\left[\left(\frac{1}{2}\bar{a}\mathbb{E}_\xi[c] - \frac{1}{4}\mathbb{E}_\xi[c^2]\right)\mathbb{1}\{\bar{a} < t\}\right] - \mathbb{E}_{\bar{a}}\left[\left(\frac{1}{2}\bar{a}\bar{c} - \frac{1}{4}\bar{c}^2\right)\mathbb{1}\{\bar{a} \geq t\}\right] \\
=&\frac{1}{8}\left[a_1^2 + a_2^2 + \delta^2 + \delta^2 - \left(2\alpha\mathbb{E}_\xi[c] + 2\left(1 - \alpha\right)\bar{c}\right)\left(a_1 + a_2\right)\right. \\
&\quad \left. + 4\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)\delta + 2\alpha\mathbb{E}_\xi[c^2] + 2\left(1 - \alpha\right)\bar{c}^2\right].
\end{aligned}
$$

(B.18)

When the realized first market condition is, say $\widetilde{a}_m$, then the probability that the supplier will not hedge is

$$
P\left(\widetilde{a}_{-m} < 2t - \widetilde{a}_m | \widetilde{a}_m\right) = \begin{cases} 0, & \widetilde{a}_m = a_m + \delta, \\ 2\alpha, & \widetilde{a}_m = a_m - \delta. \end{cases}
$$

Therefore, the total expected profit of retailer m is

$$
\begin{aligned}
\pi_{Rm}^* = \mathbb{E}_{\tilde{a}_m} & \left[P\left(\tilde{a}_{-m} < 2t - \tilde{a}_m\right) \left(\frac{1}{16} \left(\tilde{a}_m^2 - 2\tilde{a}_m \mathbb{E}_\xi\left[c\right] + \mathbb{E}_\xi\left[c^2\right]\right) \right) \right. \\
& \left. + \left(1 - P\left(\tilde{a}_{-m} < 2t - \tilde{a}_m\right)\right) \left(\frac{1}{16} \left(\tilde{a}_m^2 - 2\tilde{a}_m \bar{c} + \bar{c}^2\right) \right) \right] \\
= \frac{1}{2} & \left[2\alpha \left(\frac{1}{16} \left((a_m - \delta)^2 - 2(a_m - \delta)\mathbb{E}_\xi\left[c\right] + \mathbb{E}_\xi\left[c^2\right]\right) \right) \right. \\
& \left. + (1 - 2\alpha) \left(\frac{1}{16} \left((a_m - \delta)^2 - 2(a_m - \delta)\bar{c} + \bar{c}^2\right) \right) \right] \\
& + \frac{1}{2} \left(\frac{1}{16} \left((a_m + \delta)^2 - 2(a_m + \delta)\bar{c} + \bar{c}^2\right) \right) \\
= \frac{1}{16} & \left(a_m^2 + \delta^2 - \left(2\alpha\mathbb{E}_\xi[c] + 2(1 - \alpha)\bar{c}\right) a_m + 2\alpha\delta\left(\mathbb{E}_\xi[c] - \bar{c}\right) + \alpha\mathbb{E}_\xi[c^2] + (1 - \alpha)\bar{c}^2\right).
\end{aligned}
$$

(B.19)

Scenario (S, N). This is the case when retailer 1 shares the market information to the supplier while retailer 2 does not. Following the analysis in section 2.3.1, given the realization of production cost c, and reported market condition $\tilde{a}_1$, the equilibrium wholesale price for market 1 is $w_1^* = (\tilde{a}_1 + c)/2$ and the supplier's profit in market 1 is $\pi_{S1}(w_1^*; \tilde{a}_1, c) = (\tilde{a}_1 - c)^2/8$. And correspondingly, retailer 1's expected profit is $\pi_{R1}(q_1^*; \tilde{a}_1, c) = (\tilde{a}_1 - c)^2/16$. We know that no matter whether the retailer shares the information or not, for given wholesale price w, retailer m's optimal order quantity is $(\tilde{a}_m - w)/2$. When $\tilde{a}_1 = a_1 + \delta$, then the supplier's expected profit in market 2 is

$$
\mathbb{E}_{\tilde{a}_2}\left[\frac{1}{2}(w - c)(\tilde{a}_2 - w) \right] = 2\alpha \left(\frac{1}{2}(w - c)(a_2 + \delta - w) \right) \\
+ (1 - 2\alpha)\left(\frac{1}{2}(w - c)(a_2 - \delta - w) \right).
$$

Therefore the optimal wholesale price is $w_2^* = (a_2 + (4\alpha - 1)\delta + c)/2$ and supplier's expected profit in market 2 is $\pi_{S2}(w_2^*; a_1 + \delta, c) = \frac{1}{8}(a_2 + (4\alpha - 1)\delta - c)^2$. The retailer 2's profit

is $\pi_{R2}\left(q_2^*; a_1 + \delta, c\right) = \frac{1}{16}\left((a_2 + (4\alpha - 1)\delta - c)^2 + 32\alpha(1 - 2\alpha)\delta^2\right)$. On the other hand, if $\tilde{a}_1 = a_1 - \delta$, we can similarly get $w_2^* = \left(a_2 - (4\alpha - 1)\delta + c\right)/2$ and supplier's expected profit in market 2 is $\pi_{S2}\left(w_2^*; a_1 - \delta, c\right) = \frac{1}{8}\left(a_2 - (4\alpha - 1)\delta - c\right)^2$, while retailer 2's profit is $\pi_{R2}\left(q_2^*; a_1 - \delta, c\right) = \frac{1}{16}\left((a_2 - (4\alpha - 1)\delta - c)^2 + 32\alpha(1 - 2\alpha)\delta^2\right)$. Therefore the supplier's expected total profit when he does not hedge is

$$\pi_S^{nh*}\left(\tilde{a}_1\right) = \begin{cases} \mathbb{E}_\xi\left[\dfrac{(a_1 + \delta - c)^2}{8} + \dfrac{(a_2 + (4\alpha - 1)\delta - c)^2}{8}\right], & \text{if } \tilde{a}_1 = a_1 + \delta, \\[2mm] \mathbb{E}_\xi\left[\dfrac{(a_1 - \delta - c)^2}{8} + \dfrac{(a_2 - (4\alpha - 1)\delta - c)^2}{8}\right], & \text{if } \tilde{a}_1 = a_1 - \delta. \end{cases}$$

$$= \begin{cases} \dfrac{1}{8}\left((a_1 + \delta)^2 + (a_2 + (4\alpha - 1)\delta)^2\right) - \dfrac{1}{4}\left(a_1 + a_2 + \delta + (4\alpha - 1)\delta\right)\mathbb{E}_\xi[c] \\[2mm] \quad + \dfrac{1}{4}\mathbb{E}_\xi[c^2], \text{ if } \tilde{a}_1 = a_1 + \delta, \\[2mm] \dfrac{1}{8}\left((a_1 - \delta)^2 + (a_2 - (4\alpha - 1)\delta)^2\right) - \dfrac{1}{4}\left(a_1 + a_2 - \delta - (4\alpha - 1)\delta\right)\mathbb{E}_\xi[c] \\[2mm] \quad + \dfrac{1}{4}\mathbb{E}_\xi[c^2], \text{ if } \tilde{a}_1 = a_1 - \delta, \end{cases}$$

on the other hand, if the supplier hedges, the expected total profit is

$$\pi_S^{h*}\left(\tilde{a}_1\right) = \begin{cases} \dfrac{1}{8}\left((a_1 + \delta)^2 + (a_2 + (4\alpha - 1)\delta)^2\right) - \dfrac{1}{4}\left(a_1 + a_2 + \delta + (4\alpha - 1)\delta\right)\bar{c} \\[2mm] \quad + \dfrac{1}{4}\bar{c}^2, \text{ if } \tilde{a}_1 = a_1 + \delta, \\[2mm] \dfrac{1}{8}\left((a_1 - \delta)^2 + (a_2 - (4\alpha - 1)\delta)^2\right) - \dfrac{1}{4}\left(a_1 + a_2 - \delta - (4\alpha - 1)\delta\right)\bar{c} \\[2mm] \quad + \dfrac{1}{4}\bar{c}^2, \text{ if } \tilde{a}_1 = a_1 - \delta. \end{cases}$$

thus we get

$$\pi_S^{h*}\left(\tilde{a}_1\right) - \pi_S^{nh*}\left(\tilde{a}_1\right)$$

$$= \begin{cases} \dfrac{1}{4}\left((\mathbb{E}_\xi[c] - \bar{c})\left(a_1 + a_2 + \delta + (4\alpha - 1)\delta\right) - \left(\mathbb{E}_\xi[c^2] - \bar{c}^2\right)\right), & \text{if } \tilde{a}_1 = a_1 + \delta, \\[2mm] \dfrac{1}{4}\left((\mathbb{E}_\xi[c] - \bar{c})\left(a_1 + a_2 - \delta - (4\alpha - 1)\delta\right) - \left(\mathbb{E}_\xi[c^2] - \bar{c}^2\right)\right), & \text{if } \tilde{a}_1 = a_1 - \delta. \end{cases}$$

[169]

whence the supplier chooses to hedge if and only if

$$
\begin{cases}
\dfrac{(a_1 + a_2 + 4\alpha\delta)}{2} \geq \dfrac{\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2}{2\left(\mathbb{E}_\xi\left[c\right] - \bar{c}\right)} = t, & \text{if } \tilde{a}_1 = a_1 + \delta, \\[3mm]
\dfrac{(a_1 + a_2 - 4\alpha\delta)}{2} \geq \dfrac{\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2}{2\left(\mathbb{E}_\xi\left[c\right] - \bar{c}\right)} = t, & \text{if } \tilde{a}_1 = a_1 - \delta.
\end{cases}
\tag{B.20}
$$

Therefore we can do further case discussions:

Case $(A1)$. $2\left(t - 2\alpha\delta\right) < a_1 + a_2 < 2t$. In this case, the supplier would hedge when observes $\tilde{a}_1 = a_1 + \delta$ and the supplier's expected profit is:

$$
\pi_S^{(S,N)*} = \frac{1}{8}\Big(a_1^2 + a_2^2 - (a_1 + a_2)\left(\mathbb{E}_\xi[c] + \bar{c}\right) + \mathbb{E}_\xi[c^2] + \bar{c}^2
$$
$$
+ \left(\delta + (4\alpha - 1)\delta\right)\left(\mathbb{E}_\xi[c] - \bar{c}\right) + \delta^2 + (4\alpha - 1)^2\delta^2\Big). \tag{B.21}
$$

The expected profits for the two retailers are

$$
\pi_{R1}^{(S,N)*} = \frac{1}{2}\frac{(a_1 + \delta - \bar{c})^2}{16} + \frac{1}{2}\mathbb{E}_\xi\left[\frac{(a_1 - \delta - c)^2}{16}\right], \tag{B.22}
$$

$$
\pi_{R2}^{(S,N)*} = \frac{1}{2}\pi_{R2}\left(q_2^*; a_1 + \delta, \bar{c}\right) + \frac{1}{2}\mathbb{E}_\xi\left[\pi_{R2}\left(q_2^*; a_1 - \delta, c\right)\right]
$$
$$
= \frac{1}{2}\left(\frac{1}{16}\left((a_2 + (4\alpha - 1)\delta - \bar{c})^2\right)\right) + \frac{1}{2}\mathbb{E}_\xi\left[\frac{1}{16}\left((a_2 - (4\alpha - 1)\delta - c)^2\right)\right] + 2\alpha(1 - 2\alpha)\delta^2.
$$
$$
\tag{B.23}
$$

Case $(A2)$. $2\left(t - \delta\right) < a_1 + a_2 < 2\left(t - 2\alpha\delta\right)$. In this case, with only one retailer shares information, the supplier would never hedge, whence

$$
\pi_{R1}^{(S,N)*} = \frac{1}{2}\mathbb{E}_\xi\left[\frac{(a_1 + \delta - c)^2}{16} + \frac{(a_1 - \delta - c)^2}{16}\right], \tag{B.24}
$$

$$
\pi_{R2}^{(S,N)*} = \frac{1}{2}\mathbb{E}_\xi\left[\pi_{R2}\left(q_2^*; a_1 + \delta, c\right) + \pi_{R2}\left(q_2^*; a_1 - \delta, c\right)\right]
$$
$$
= \frac{1}{2}\mathbb{E}_\xi\left[\frac{1}{16}\left((a_2 + (4\alpha - 1)\delta - c)^2\right) + \frac{1}{16}\left((a_2 - (4\alpha - 1)\delta - c)^2\right)\right] + 2\alpha(1 - 2\alpha)\delta^2.
$$
$$
\tag{B.25}
$$

[170]

According to the classic literature like [50], sharing information would be a pure loss for the retailer. Therefore, it is clear that in Case $(A2)$, (S, N) (or (N, S)) cannot be equilibrium, (N, N) is an equilibrium and it is unclear that whether (S, S) could be an equilibrium.

Case $(B1)$. $2t < a_1 + a_2 < 2(t + 2\alpha\delta)$. In this case, the supplier's hedging strategy is the same as in Case $(A1)$. Therefore the supplier and the retailers' expected profits also keep the same as in Case $(A1)$.

Case $(B2)$. $2(t + 2\alpha\delta) < a_1 + a_2 < 2(t + \delta)$. In this case, with only one retailer shares information, the supplier would always hedge, whence

$$\pi_{R1}^{(S,N)*} = \frac{1}{2}\left(\frac{(a_1 + \delta - \bar{c})^2}{16} + \frac{(a_1 - \delta - \bar{c})^2}{16}\right), \tag{B.26}$$

$$\begin{aligned}\pi_{R2}^{(S,N)*} &= \frac{1}{2}\left(\pi_{R2}\left(q_2^*; a_1 + \delta, \bar{c}\right) + \pi_{R2}\left(q_2^*; a_1 - \delta, \bar{c}\right)\right)\\ &= \frac{1}{2}\left(\frac{1}{16}\left((a_2 + (4\alpha - 1)\delta - \bar{c})^2\right) + \frac{1}{16}\left((a_2 - (4\alpha - 1)\delta - \bar{c})^2\right)\right) + 2\alpha(1 - 2\alpha)\delta^2.\end{aligned}$$
$$\tag{B.27}$$

For the same reasons as Case $(A2)$, (S, N) (or (N, S)) cannot be equilibrium, (N, N) is an equilibrium and it is unclear that whether (S, S) could be an equilibrium.

Scenario (N, S). This scenario is a symmetric analog to **Scenario** (S, N) above.

Scenario (N, N). When both of the retailers do not share information, for any realized production cost c, the supplier's expected total profit has form

$$\pi_S^{(N,N)}(w_1, w_2) = \mathbb{E}_{\tilde{a}_1, \tilde{a}_2}\left[\frac{1}{2}(w_1 - c)(\tilde{a}_1 - w_1) + \frac{1}{2}(w_2 - c)(\tilde{a}_2 - w_2)\right],$$

which leads to optimal wholesale prices $w_1^* = (a_1 + c)/2$ and $w_2^* = (a_2 + c)/2$. Thus the supplier's corresponding optimal expected profit when he does not hedge and hedges respectively are

$$\pi_S^{nh*} = \frac{1}{8} \left(a_1^2 + a_2^2 - 2 \left(a_1 + a_2 \right) \mathbb{E}_\xi \left[c \right] + 2 \mathbb{E}_\xi \left[c^2 \right] \right),$$
$$\pi_S^{h*} = \frac{1}{8} \left(a_1^2 + a_2^2 - 2 \left(a_1 + a_2 \right) \bar{c} + 2 \bar{c}^2 \right).$$

Whence the supplier prefers hedging when $\pi_S^{h*} \geq \pi_S^{nh*}$ which is equivalent to

$$\frac{a_1 + a_2}{2} \geq \frac{\mathbb{E}_\xi[c^2] - \bar{c}^2}{2 \left(\mathbb{E}_\xi[c] - \bar{c} \right)} = t. \tag{B.28}$$

The corresponding optimal profit of retailer m has the form

$$\pi_{Rm}^{(N,N)*} = \begin{cases} \mathbb{E}_\xi \left[\frac{((a_m - c)^2 + 4\delta^2)}{16} \right], & \text{if } \frac{a_1 + a_2}{2} < t, \\ \frac{((a_m - \bar{c})^2 + 4\delta^2)}{16}, & \text{if } \frac{a_1 + a_2}{2} \geq t. \end{cases} \tag{B.29}$$

Equilibrium Analysis

For **Case** $(A1)$, $2 \left(t - 2\alpha\delta \right) < a_1 + a_2 < 2t$, from (B.22) and (B.29), we have

$$\pi_{R1}^{(S,N)*} - \pi_{R1}^{(N,N)*} = \frac{1}{32} \left(2 \left(\mathbb{E}_\xi[c] - \bar{c} \right) a_1 - \left(\mathbb{E}_\xi \left[c^2 \right] - \bar{c}^2 \right) + 2\delta \left(\mathbb{E}_\xi[c] - \bar{c} - 3\delta \right) \right). \tag{B.30}$$

From (B.17) and (B.23), we have

$$\pi_{R2}^{(S,S)*} - \pi_{R2}^{(S,N)*} = \frac{1}{32} \left(1 - 2\alpha \right) \left(\left(\mathbb{E}_\xi[c^2] - \bar{c}^2 \right) - 2 \left(\mathbb{E}_\xi[c] - \bar{c} \right) a_2 + 2\delta \left(\mathbb{E}_\xi[c] - \bar{c} - 24\alpha\delta \right) \right), \tag{B.31}$$

[172]

To break the tie, we assume that if a retailer is indifferent between sharing and not sharing information, the retailer would keep the information private. If (N, N) is an equilibrium, then both (B.30) and its analogue for retailer 2 should be non-positive, which leads to

$$a_1, a_2 \leq t - \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right). \tag{B.32}$$

If (S, S) is an equilibrium, then both (B.42) and its analogue for retailer 1 should be positive, which leads to

$$a_1, a_2 < t + \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right). \tag{B.33}$$

We notice that the parameter region where (S, S) is an equilibrium overlaps the region where (N, N) is an equilibrium, therefore we may conduct Pareto refinement in the overlapped region. From (B.17) and (B.29), we have

$$\pi_{Rm}^{(S,S)*} - \pi_{Rm}^{(N,N)*} = \frac{1}{16} \left(\alpha \left(2 \left(\mathbb{E}_\xi[c] - \bar{c}\right) a_m - \left(\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2\right)\right)\right.$$
$$\left. +2\alpha\delta \left(\mathbb{E}_\xi[c] - \bar{c}\right) - 3\delta^2\right) , \ m = 1, 2. \tag{B.34}$$

Therefore, (S, S) is ruled out after the Pareto refinement if and only if $\pi_{Rm}^{(S,S)*} \leq \pi_{Rm}^{(N,N)*}$, $m = 1, 2$, which is equivalent to

$$a_1, a_2 \leq t - \delta \left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)}\right). \tag{B.35}$$

We know that $\alpha \in \left(0, \frac{1}{2}\right)$, whence condition (B.32) is stronger than (B.35), i.e., once (N, N) is an equilibrium, it Pareto dominates the outcome (S, S).

If (S, N) is an equilibrium, then (B.30) should be positive while (B.31) should be non-positive, which leads to

$$a_1 > t - \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right) \text{ and } a_2 \geq t + \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right). \tag{B.36}$$

Symmetrically, if (N, S) is an equilibrium, then

$$a_1 \geq t + \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right) \text{ and } a_2 > t - \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right). \tag{B.37}$$

Parameter regions for equilibrium (S, N) and (N, S) also overlap. We have that

$$\pi_{R1}^{(S,N)*} - \pi_{R1}^{(N,S)*} = \frac{1}{8} (1 - 2\alpha) \delta \left(\mathbb{E}_\xi[c] - \bar{c} - 12\alpha\delta \right), \tag{B.38}$$

$$\pi_{R2}^{(S,N)*} - \pi_{R2}^{(N,S)*} = -\frac{1}{8} (1 - 2\alpha) \delta \left(\mathbb{E}_\xi[c] - \bar{c} - 12\alpha\delta \right), \tag{B.39}$$

which implies that neither (S, N) nor (N, S) can Pareto dominate the other, therefore in the overlapped parameter regions of equilibrium (S, N) and (N, S), the two equilibriums are both possible. We notice that the above analysis indicates that the parameter regions:

$$\begin{aligned} a_1 &\geq \max \left\{ t - \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right), t + \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right) \right\}, \\ a_2 &\leq \min \left\{ t - \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right), t + \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right) \right\}, \end{aligned} \tag{B.40}$$

and

$$\begin{aligned} a_1 &\leq \min \left\{ t - \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right), t + \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right) \right\}, \\ a_2 &\geq \max \left\{ t - \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right), t + \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right) \right\}, \end{aligned} \tag{B.41}$$

[174]

admit no pure Nash equilibrium. We use (B.40) as an example to explain the intuition ((B.41) has an analogous explanation). In the case of (B.40), the bigger player, retailer 1 tends to deviate her information-sharing policy from the smaller player, retailer 2, while retailer 2 tends to follow retailer 1's policy, which leads to no pure Nash equilibrium. Specifically, without any information sharing, the supplier would not hedge which is not in favor of retailer 1's interest, therefore, if retailer 2 holds information, retailer 1 would like to build up an information-sharing channel in order to incentivize supplier to hedge when retailer 1's market turns out to be high type. However, if retailer 2 shares information, the retailer 1's information sharing benefit of changing supplier's hedging decision would be offset such that the cost of giving supplier the information advantage in setting wholesale price outweighs any benefits, which discourages retailer 1's information sharing. On the other hand, for the smaller player, retailer 2, without information sharing, supplier's no-hedging decision aligns with retailer 2's interest, therefore when retailer 1 holds information, there is no incentive for retailer 2 to set up the information-sharing channel. However, if retailer 1 shares information, retailer 2 is hurt due to supplier's twisted hedging decision. Retailer 2 has a smaller market than retailer 1, therefore the benefit of offsetting retailer 1's influence on supplier's hedging decision by sharing information can still outweigh the cost of more severe double marginalization if the supplier is more informed of retailer 2's market. Slightly different but analogous phenomenon also takes place in **Case** ($B1$) below.

For **Case** ($A2$), $2(t - \delta) < a_1 + a_2 < 2(t - 2\alpha\delta)$, as we have indicated before, (N, N) is for sure an equilibrium, (S, N) or (N, S) cannot be equilibrium, now we check whether (S, S) could be an equilibrium. From (B.17) and (B.25), we know that

$$
\pi_{R2}^{(S,S)*} - \pi_{R2}^{(S,N)*} = \frac{1}{16}\alpha \left(2 \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) a_2 - \left(\mathbb{E}_\xi \left[c^2 \right] - \bar{c}^2 \right) + 2\delta \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) \right.
$$
$$
\left. - 24 \left(1 - 2\alpha \right) \delta^2 \right), \quad \text{(B.42)}
$$

Symmetrically, we can get the profit difference for retailer 1. Therefore, (S, S) is equilibrium if and only if

$$a_1, a_2 > t - \delta \left(1 - \frac{12 \left(1 - 2\alpha \right) \delta}{\mathbb{E}_\xi[c] - \bar{c}} \right). \tag{B.43}$$

Because retailers' profit functions in equilibrium outcome (S, S) and (N, N) have the same function forms as in Case $(A1)$, condition (B.35) still holds here.

For **Case** $(B1)$, $2t < a_1 + a_2 < 2 \left(t + 2\alpha\delta \right)$, from (B.22) and (B.29), we have

$$\pi_{R1}^{(S,N)*} - \pi_{R1}^{(N,N)*} = \frac{1}{32} \left(\left(\mathbb{E}_\xi \left[c^2 \right] - \bar{c}^2 \right) - 2 \left(\mathbb{E}_\xi[c] - \bar{c} \right) a_1 + 2\delta \left(\mathbb{E}_\xi[c] - \bar{c} - 3\delta \right) \right), \tag{B.44}$$

and from (B.19) and (B.23), we have

$$\pi_{R2}^{(S,S)*} - \pi_{R2}^{(S,N)*} = \frac{1}{32} \left(1 - 2\alpha \right) \left(2 \left(\mathbb{E}_\xi[c] - \bar{c} \right) a_2 - \left(\mathbb{E}_\xi[c^2] - \bar{c}^2 \right) + 2\delta \left(\mathbb{E}_\xi[c] - \bar{c} - 24\alpha\delta \right) \right). \tag{B.45}$$

By exchanging a_1 and a_2, we can get the analogue results for scenario (N, S).

If (N, N) is an equilibrium, then both (B.44) and its analogue for retailer 2 should be non-positive, which leads to

$$a_1, a_2 \geq t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right). \tag{B.46}$$

If (S, S) is an equilibrium, then both (B.45) and its analogue for retailer 1 should be positive, which leads to

$$a_1, a_2 > t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right). \tag{B.47}$$

[176]

Similar as in Case $(A1)$, from (B.19) and (B.29), we have

$$
\pi_{Rm}^{(S,S)*} - \pi_{Rm}^{(N,N)*} = \frac{1}{16} \left(\alpha \left(\left(\mathbb{E}_\xi \left[c^2 \right] - \bar{c}^2 \right) - 2 \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) a_m \right) \right.
$$
$$
\left. + 2\alpha\delta \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right) - 3\delta^2 \right) \ , \ m = 1, 2. \quad \text{(B.48)}
$$

Therefore, (S, S) is ruled out after the Pareto refinement if and only if $\pi_{Rm}^{(S,S)*} \leq \pi_{Rm}^{(N,N)*}$, $m = 1, 2$, which is equivalent to

$$
a_1, a_2 \geq t + \delta \left(1 - \frac{3\delta}{2\alpha \left(\mathbb{E}_\xi \left[c \right] - \bar{c} \right)} \right). \quad \text{(B.49)}
$$

We know that $\alpha \in \left(0, \frac{1}{2} \right)$, whence condition (B.46) is stronger than (B.49), i.e., once (N, N) is an equilibrium, it Pareto dominates the outcome (S, S).

If (S, N) is an equilibrium, then (B.44) should be positive while (B.45) should be non-positive, which leads to

$$
a_1 < t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right) \text{ and } a_2 \leq t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right). \quad \text{(B.50)}
$$

Symmetrically, if (N, S) is an equilibrium, then

$$
a_1 \leq t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right) \text{ and } a_2 < t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right). \quad \text{(B.51)}
$$

Similar as Case $(A1)$, neither of (S, N) nor (N, S) would be ruled out after Pareto refinement. Analogous to Case $(A1)$, in the following parameter regions, pure Nash equilibrium does not exist:

$$
a_1 \geq \max \left\{ t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right), t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right) \right\},
$$
$$
a_2 \leq \min \left\{ t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right), t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right) \right\},
$$
$$
\quad \text{(B.52)}
$$

[177]

and

$$a_1 \leq \min\left\{t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right), t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)\right\},$$
$$a_2 \geq \max\left\{t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right), t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)\right\}. \tag{B.53}$$

For **Case** $(B2)$, $2(t + 2\alpha\delta) < a_1 + a_2 < 2(t + \delta)$, as we have indicated before, (N, N) is for sure an equilibrium, (S, N) or (N, S) cannot be equilibrium, now we check whether (S, S) could be an equilibrium. From (B.19) and (B.27), we know that

$$\pi_{R2}^{(S,S)*} - \pi_{R2}^{(S,N)*} = \frac{1}{16}\alpha\left(\left(\mathbb{E}_\xi[c^2] - \bar{c}^2\right) - 2\left(\mathbb{E}_\xi[c] - \bar{c}\right)a_2 + 2\delta\left(\mathbb{E}_\xi[c] - \bar{c}\right)\right.$$
$$\left. -24(1 - 2\alpha)\delta^2\right), \quad \text{(B.54)}$$

Symmetrically, we can get the profit difference for retailer 1. Therefore, (S, S) is equilibrium if and only if

$$a_1, a_2 < t + \delta\left(1 - \frac{12(1 - 2\alpha)\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right). \tag{B.55}$$

Because retailers' profit functions in equilibrium outcome (S, S) and (N, N) have the same function forms as in Case $(B1)$, condition (B.49) still holds here.

Supply Chain Efficiency

Case $(A1)$. $2(t - 2\alpha\delta) < a_1 + a_2 < 2t$. Under scenario (S, N), the total welfare is

$$\pi_{SC}^{(S,N)*} = \pi_S^{(S,N)*} + \pi_{R1}^{(S,N)*} + \pi_{R2}^{(S,N)*}$$
$$= \frac{1}{16}\left(3\left(a_1^2 + a_2^2 - \left(\mathbb{E}_\xi[c] + \bar{c}\right)(a_1 + a_2) + \mathbb{E}_\xi[c^2] + \bar{c}^2\right) + 12\alpha\delta\left(\mathbb{E}_\xi[c] - \bar{c}\right)\right.$$
$$\left. +2\left(3 + 4\alpha - 8\alpha^2\right)\delta^2\right).$$

Under scenario (S, S), the total welfare is

$$\pi_{SC}^{(S,S)*} = \pi_S^{(S,S)*} + \pi_{R1}^{(S,S)*} + \pi_{R2}^{(S,S)*}$$
$$= \frac{3}{16} \left(a_1^2 + a_2^2 - 2 \left((1 - \alpha) \mathbb{E}_\xi [c] + \alpha \bar{c} \right) (a_1 + a_2) \right.$$
$$\left. + 2 \left((1 - \alpha) \mathbb{E}_\xi [c] + \alpha \bar{c} + 2\alpha\delta \left(\mathbb{E}_\xi [c] - \bar{c} \right) + \delta^2 \right) \right).$$

Under scenario (N, N), the total welfare is

$$\pi_{SC}^{(N,N)*} = \pi_S^{(N,N)*} + \pi_{R1}^{(N,N)*} + \pi_{R2}^{(N,N)*} = \frac{3}{16} \left(a_1^2 + a_2^2 - 2 \left(a_1 + a_2 \right) \mathbb{E}_\xi [c] + 2\mathbb{E}_\xi \left[c^2 \right] \right) + \frac{\delta^2}{2}.$$

Therefore, (S, N) increases total welfare (i.e., $\pi_{SC}^{(S,N)*} \geq \pi_{SC}^{(N,N)*}$) if and only if

$$a_1 + a_2 \geq 2t - 4\alpha\delta + \frac{2\delta^2 \left(1 - 4\alpha + 8\alpha^2 \right)}{3 \left(\mathbb{E}_\xi [c] - \bar{c} \right)}.$$

And (S, S) increases total welfare (i.e., $\pi_{SC}^{(S,S)*} \geq \pi_{SC}^{(N,N)*}$) if and only if

$$a_1 + a_2 \geq 2t - 2\delta + \frac{\delta^2}{3\alpha \left(\mathbb{E}_\xi [c] - \bar{c} \right)}.$$

Combining the equilibrium analysis in Section 2.5.1, we summarize the conditions under which (S, N), (N, S) and (S, S) emerge as equilibrium (after Pareto refinement) and decreases the system's efficiency:

Case $(A2)$. $2 \left(t - \delta \right) < a_1 + a_2 < 2 \left(t - 2\alpha\delta \right)$. As we have indicated before, (S, N) and (N, S) would never emerge as an equilibrium under this scenario, while the profit functions for both sharing outcome (S, S) and (N, N) are the same as in Case $(A1)$, therefore all the results in Case $(A1)$ also hold here.

[179]

Case $(B1)$**.** $2t < a_1 + a_2 < 2(t + 2\alpha\delta)$. Under scenario (S, S), the total welfare is

$$
\begin{aligned}
\pi_{SC}^{(S,S)*} &= \pi_S^{(S,S)*} + \pi_{R1}^{(S,S)*} + \pi_{R2}^{(S,S)*} \\
&= \frac{3}{16} \left(a_1^2 + a_2^2 - 2 \left(\alpha \mathbb{E}_\xi [c] + (1 - \alpha) \bar{c} \right) (a_1 + a_2) \right. \\
&\quad \left. + 2 \left(\alpha \mathbb{E}_\xi [c] + (1 - \alpha) \bar{c} + 2\alpha\delta \left(\mathbb{E}_\xi [c] - \bar{c} \right) + \delta^2 \right) \right).
\end{aligned}
$$

Under scenario (S, N), the total welfare is

$$
\begin{aligned}
\pi_{SC}^{(S,N)*} &= \pi_S^{(S,N)*} + \pi_{R1}^{(S,N)*} + \pi_{R2}^{(S,N)*} \\
&= \frac{1}{16} \left(3 \left(a_1^2 + a_2^2 - \left(\mathbb{E}_\xi [c] + \bar{c} \right) (a_1 + a_2) + \mathbb{E}_\xi [c^2] + \bar{c}^2 \right) \right. \\
&\quad \left. + 12\alpha\delta \left(\mathbb{E}_\xi [c] - \bar{c} \right) + 2 \left(3 + 4\alpha - 8\alpha^2 \right) \delta^2 \right).
\end{aligned}
$$

Under scenario (N, N), the total welfare is

$$
\pi_{SC}^{(N,N)*} = \pi_S^{(N,N)*} + \pi_{R1}^{(N,N)*} + \pi_{R2}^{(N,N)*} = \frac{3}{16} \left(a_1^2 + a_2^2 - 2 (a_1 + a_2) \bar{c} + 2\bar{c}^2 \right) + \frac{\delta^2}{2}.
$$

Therefore, (S, S) increases total welfare (i.e., $\pi_{SC}^{(S,S)*} \geq \pi_{SC}^{(N,N)*}$) if and only if

$$
a_1 + a_2 \leq 2t + 2\delta - \frac{\delta^2}{3\alpha \left(\mathbb{E}_\xi [c] - \bar{c} \right)}.
$$

And (S, N) increases total welfare (i.e., $\pi_{SC}^{(S,N)*} \geq \pi_{SC}^{(N,N)*}$) if and only if

$$
a_1 + a_2 \leq 2t + 4\alpha\delta - \frac{2\delta^2 \left(1 - 4\alpha + 8\alpha^2 \right)}{3 \left(\mathbb{E}_\xi [c] - \bar{c} \right)}.
$$

Case $(B2)$**.** $2(t + 2\alpha\delta) < a_1 + a_2 < 2(t + \delta)$. Due to the similar arguments as in Case $(A2)$, only (S, S) and (N, N) could be potential equilibriums and all the results hold the same as in Case $(B1)$.

[180]

Appendix C: Duopoly Stackelberg Information-Sharing Game

We consider the following scenario where retailer 1 is the Stackelberg leader in the information-sharing game.

For **Case** $(A1)$, $2\left(t - 2\alpha\delta\right) \leq a_1 + a_2 < 2t$, given that retailer 1 shares information, retailer 2 would share information if and only if $a_2 < t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$. On the other hand, given that retailer 1 does not share information, retailer 2 would share information if and only if $a_2 > t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$. We know that $t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right) \geq t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$ is equivalent to $\mathbb{E}_\xi[c] - \bar{c} \geq \left(12\alpha + \frac{3}{2}\right)\delta$.

Therefore, when $\mathbb{E}_\xi[c] - \bar{c} \geq \left(12\alpha + \frac{3}{2}\right)\delta$, if $a_2 \leq t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will follow retailer 1's information-sharing strategy; if $t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right) < a_2 < t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will always share information; if $a_2 \geq t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will always conduct the opposite sharing strategy against retailer 1. Anticipating retailer 2's strategies, we then conduct discussions on retailer 1's policies. When $a_2 < t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare the equilibrium profits of sharing outcomes (S, S) and (N, N). We have

$$\pi_{R1}^{(S,S)*} - \pi_{R1}^{(N,N)*} = \frac{1}{16}\left(2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)a_1 - \alpha\left(\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2\right) + 2\alpha\delta\left(\mathbb{E}_\xi[c] - \bar{c}\right) - 3\delta^2\right),$$

whence retailer 1 would choose to share information and end up in the equilibrium outcome (S, S) when $a_1 > t - \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)}\right)$, while retailer 1 would choose the alternative outcome (N, N) when $a_1 \leq t - \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)}\right)$. When $t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right) \leq a_2 < t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare equilibrium outcomes of (S, S) and (N, S). Based on the calculations before, we know that retailer 1 would share information when $a_1 < t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$ and hold information when $a_1 \geq t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$. Lastly, when $a_2 \geq t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare equilibrium outcomes of (S, N) and (N, S). We have

$$\pi_{R1}^{(S,N)*} - \pi_{R1}^{(N,S)*} = \frac{1}{8}\left(1 - 2\alpha\right)\delta\left(\mathbb{E}_\xi[c] - \bar{c} - 12\alpha\delta\right),$$

[181]

therefore when $\mathbb{E}_\xi[c] - \bar{c} > 12\alpha\delta$, retailer 1 would share information and end up in equilibrium (S, N); otherwise when $\mathbb{E}_\xi[c] - \bar{c} \leq 12\alpha\delta$, retailer 1 would hold information and end up in equilibrium (N, S). Because we now assume that $\mathbb{E}_\xi[c] - \bar{c} \geq \left(12\alpha + \frac{3}{2}\right)\delta$, retailer 1 would share information and retailer 2 would hold information.

When $\mathbb{E}_\xi[c] - \bar{c} < \left(12\alpha + \frac{3}{2}\right)\delta$, if $a_2 < t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will follow retailer 1's information-sharing strategy; if $t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right) \leq a_2 \leq t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will always hold information; if $a_2 > t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will always conduct the opposite sharing strategy against retailer 1. Anticipating retailer 2's strategies, we then conduct discussions on retailer 1's policies. When $a_2 < t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare the equilibrium profits of sharing outcomes (S, S) and (N, N). Same as before, we have

$$\pi_{R1}^{(S,S)*} - \pi_{R1}^{(N,N)*} = \frac{1}{16}\left(2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)a_1 - \alpha\left(\mathbb{E}_\xi[c^2] - \bar{c}^2\right) + 2\alpha\delta\left(\mathbb{E}_\xi[c] - \bar{c}\right) - 3\delta^2\right),$$

whence retailer 1 would choose to share information and end up in the equilibrium outcome (S, S) when $a_1 > t - \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)}\right)$, while retailer 1 would choose the alternative outcome (N, N) when $a_1 \leq t - \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)}\right)$. When $t + \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right) \leq a_2 \leq t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare equilibrium outcomes of (S, N) and (N, N). Based on the calculations before, we know that retailer 1 would share information when $a_1 > t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$ and hold information when $a_1 \leq t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$. Lastly, when $a_2 > t - \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare equilibrium outcomes of (S, N) and (N, S). We have

$$\pi_{R1}^{(S,N)*} - \pi_{R1}^{(N,S)*} = \frac{1}{8}\left(1 - 2\alpha\right)\delta\left(\mathbb{E}_\xi[c] - \bar{c} - 12\alpha\delta\right),$$

therefore when $\mathbb{E}_\xi[c] - \bar{c} > 12\alpha\delta$, retailer 1 would share information and end up in equilibrium (S, N); otherwise when $\mathbb{E}_\xi[c] - \bar{c} \le 12\alpha\delta$, retailer 1 would hold information and end up in equilibrium (N, S).

For **Case** $(A2)$, $2(t - \delta) < a_1 + a_2 < 2(t - 2\alpha\delta)$, given that retailer 1 shares information, according to (B.42) and (B.43), retailer 2 would share information if and only if $a_2 > t - \delta\left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$. On the other hand, if $a_2 \le t - \delta\left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$, i.e., retailer 2 chooses to hold information, there would be no incentive for retailer 1 to share the information in the first place, because retailer 1's information alone has no influence on supplier's hedging decision while it does provide the supplier better knowledge to exploit more profit from the downstream. In the other scenario, given that retailer 1 holds information, there is no incentive for retailer 2 to share information, because retailer 2 has no influence on supplier's hedging decision in this case, i.e., the only possible outcome would be (N, N). In summary, retailer 1 would choose to share information if and only if retailer 2 prefers outcome (S, S) over (S, N) and retailer 1 prefers outcome (S, S) over (N, N). The equilibrium profits for outcomes (S, S) and (N, N) keep the same as in Case $(A1)$, therefore we conclude that retailer 1 would choose to share information and end up in outcome (S, S) if and only if $a_1 > t - \delta\left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c]-\bar{c})}\right)$ and $a_2 > t - \delta\left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$. Otherwise, the equilibrium outcome is (N, N).

For **Case** $(B1)$, $2t \le a_1 + a_2 < 2(t + 2\alpha\delta)$, given that retailer 1 shares information, retailer 2 would share information if and only if $a_2 > t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$. On the other hand, given that retailer 1 does not share information, retailer 2 would share information if and only if $a_2 < t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$. From the above $t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) > t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$ is equivalent to $\mathbb{E}_\xi[c] - \bar{c} < \left(12\alpha + \frac{3}{2}\right)\delta$.

Therefore, when $\mathbb{E}_\xi[c] - \bar{c} < \left(12\alpha + \frac{3}{2}\right)\delta$, if $a_2 < t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$, retailer 2 will conduct the opposite strategy against retailer 1; if $t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right) \le a_2 \le t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$, retailer 2 will always hold information; if $a_2 > t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c]-\bar{c}}\right)$, retailer 2 will always follow retailer

[183]

1's strategy. Anticipating retailer 2's strategies, we then conduct discussions on retailer 1's policies. When $a_2 < t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare the equilibrium profits of sharing outcomes (S, N) and (N, S). We have

$$\pi_{R1}^{(S,N)*} - \pi_{R1}^{(N,S)*} = \frac{1}{8}(1 - 2\alpha)\delta\left(\mathbb{E}_\xi[c] - \bar{c} - 12\alpha\delta\right),$$

therefore when $\mathbb{E}_\xi[c] - \bar{c} > 12\alpha\delta$, retailer 1 would share information and end up in equilibrium (S, N); otherwise when $\mathbb{E}_\xi[c] - \bar{c} \leq 12\alpha\delta$, retailer 1 would hold information and end up in equilibrium (N, S). When $t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right) \leq a_2 \leq t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare equilibrium outcomes of (S, N) and (N, N). Based on the calculations before, we know that retailer 1 would share information when $a_1 < t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$ and hold information when $a_1 \geq t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$. Lastly, when $a_2 > t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 1 needs to compare equilibrium outcomes of (S, S) and (N, N). We have

$$\pi_{R1}^{(S,S)*} - \pi_{R1}^{(N,N)*} = \frac{1}{16}\left(\alpha\left(\mathbb{E}_\xi[c^2] - \bar{c}^2\right) - 2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)a_1 + 2\alpha\delta\left(\mathbb{E}_\xi[c] - \bar{c}\right) - 3\delta^2\right),$$

therefore when $a_1 \geq t + \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)}\right)$, retailer 1 would hold information and end up in the equilibrium (N, N); when $a_1 < t + \delta\left(1 - \frac{3\delta}{2\alpha\left(\mathbb{E}_\xi[c] - \bar{c}\right)}\right)$, retailer 1 would share information and end up in the equilibrium (S, S).

When $\mathbb{E}_\xi[c] - \bar{c} \geq \left(12\alpha + \frac{3}{2}\right)\delta$, if $a_2 \leq t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will conduct the opposite sharing strategy against retailer 1; if $t - \delta\left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right) < a_2 < t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will always share information; if $a_2 \geq t + \delta\left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}}\right)$, retailer 2 will always follow retailer 1's strategy. Anticipating retailer 2's strategies, we then conduct discussions

[184]

on retailer 1's policies. When $a_2 \le t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right)$, retailer 1 needs to compare equilibrium outcomes of (S, N) and (N, S). We have

$$\pi_{R1}^{(S,N)*} - \pi_{R1}^{(N,S)*} = \frac{1}{8} (1 - 2\alpha) \delta (\mathbb{E}_\xi[c] - \bar{c} - 12\alpha\delta),$$

therefore when $\mathbb{E}_\xi[c] - \bar{c} > 12\alpha\delta$, retailer 1 would share information and end up in equilibrium (S, N); otherwise when $\mathbb{E}_\xi[c] - \bar{c} \le 12\alpha\delta$, retailer 1 would hold information and end up in equilibrium (N, S). Because we assume that $\mathbb{E}_\xi[c] - \bar{c} > \left(12\alpha + \frac{3}{2} \right) \delta$, retailer 1 would share information while retailer 2 would hold information. When $t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right) < a_2 < t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right)$, retailer 1 needs to compare equilibrium outcomes of (S, S) and (N, S). Based on the calculations before, we know that retailer 1 would share information when $a_1 > t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right)$ and hold information when $a_1 \le t - \delta \left(1 - \frac{24\alpha\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right)$. Lastly, when $a_2 \ge t + \delta \left(1 - \frac{3\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right)$, retailer 1 needs to compare the equilibrium profits of sharing outcomes (S, S) and (N, N). Same as before, we have

$$\pi_{R1}^{(S,S)*} - \pi_{R1}^{(N,N)*} = \frac{1}{16} \left(\alpha \left(\mathbb{E}_\xi[c^2] - \bar{c}^2 \right) - 2\alpha \left(\mathbb{E}_\xi[c] - \bar{c} \right) a_1 + 2\alpha\delta \left(\mathbb{E}_\xi[c] - \bar{c} \right) - 3\delta^2 \right),$$

whence retailer 1 would choose to hold information and end up in the equilibrium outcome (N, N) when $a_1 \ge t + \delta \left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c] - \bar{c})} \right)$, while retailer 1 would choose the alternative outcome (S, S) when $a_1 < t + \delta \left(1 - \frac{3\delta}{2\alpha(\mathbb{E}_\xi[c] - \bar{c})} \right)$.

For **Case** $(B2)$, $2(t + 2\alpha\delta) \le a_1 + a_2 < 2(t + \delta)$, the analysis is analogous to Case $(A2)$. Given that retailer 1 shares information, according to (B.54) and (B.55), retailer 2 would share information if and only if $a_2 < t + \delta \left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right)$. On the other hand, if $a_2 \ge t - \delta \left(1 - \frac{12(1-2\alpha)\delta}{\mathbb{E}_\xi[c] - \bar{c}} \right)$, i.e., retailer 2 chooses to hold information, there would be no incentive for retailer 1 to share the information in the first place. In the other scenario, given that retailer 1 holds information, there is no incentive for retailer 2 to share information

and the only possible outcome would be (N, N). In summary, retailer 1 would choose to share information if and only if retailer 2 prefers outcome (S, S) over (S, N) and retailer 1 prefers outcome (S, S) over (N, N). The equilibrium profits for outcomes (S, S) and (N, N) keep the same as in Case $(B1)$, therefore we conclude that retailer 1 would choose to share information and end up in outcome (S, S) if and only if $a_1 < t + \delta \left(1 - \frac{3\delta}{2\alpha \left(\mathbb{E}_\xi [c] - \bar{c} \right)} \right)$ and $a_2 < t + \delta \left(1 - \frac{12(1 - 2\alpha)\delta}{\mathbb{E}_\xi [c] - \bar{c}} \right)$. Otherwise, the equilibrium outcome is (N, N).

Appendix D: Proofs

Proof Proof for Proposition 2.3.1: First we consider the case when $\mathbb{E}_a[a] \geq t$. From (2.2) and (2.4), we have

$$
\pi_R^{s*} - \pi_R^{ns*} = \frac{1}{16b}\Bigg[-3\mathbb{E}_a\left[a^2\right] + 3\mathbb{E}_a\left[a\right]^2 + 2\mathbb{E}_a\left[a\right]\bar{c} - \bar{c}^2
$$
$$
- 2\left(\mathbb{E}_\xi\left[c\right]\int_{\underline{a}}^t a \, dF(a) + \bar{c}\int_t^{\bar{a}} a \, dF(a)\right) + \mathbb{E}_\xi\left[c^2\right] F(t) + \bar{c}^2 \bar{F}(t)\Bigg]. \quad \text{(B.56)}
$$

We notice that $\mathbb{E}_a\left[a^2\right] - \mathbb{E}_a\left[a\right]^2 = Var(a)$, then

$$
(B.56) = \frac{1}{16b}\Bigg[-3Var(a) + 2\mathbb{E}_a\left[a\right]\bar{c} - \bar{c}^2 - 2\left(\mathbb{E}_\xi\left[c\right]\int_{\underline{a}}^t a \, dF(a) + \bar{c}\int_t^{\bar{a}} a \, dF(a)\right)
$$
$$
+ \mathbb{E}_\xi\left[c^2\right] F(t) + \bar{c}^2 \bar{F}(t)\Bigg]
$$
$$
= \frac{1}{16b}\Bigg[-3Var(a) + 2\mathbb{E}_a\left[a\right]\bar{c} - 2\left((\mathbb{E}_\xi\left[c\right] - \bar{c})\int_{\underline{a}}^t a \, dF(a) + \bar{c}\int_{\underline{a}}^{\bar{a}} a \, dF(a)\right)
$$
$$
+ \left(\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2\right) F(t)\Bigg]
$$
$$
= \frac{1}{16b}\Bigg[-3Var(a) - 2\left(\mathbb{E}_\xi\left[c\right] - \bar{c}\right)\int_{\underline{a}}^t a \, dF(a) + \left(\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2\right) F(t)\Bigg]. \quad \text{(B.57)}
$$

Here we use the formulas $\int_{\underline{a}}^t a \, dF(a) = tF(t) - \int_{\underline{a}}^t F(a) \, da$ and $t = \frac{\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2}{2\left(\mathbb{E}_\xi\left[c\right] - \bar{c}\right)}$, then

$$
(B.57) = \frac{1}{16b}\Bigg[-3Var(a) - 2\left(\mathbb{E}_\xi\left[c\right] - \bar{c}\right)\left(tF(t) - \int_{\underline{a}}^t F(a) \, da\right) + \left(\mathbb{E}_\xi\left[c^2\right] - \bar{c}^2\right) F(t)\Bigg]
$$
$$
= \frac{1}{16b}\Bigg[2\left(\mathbb{E}_\xi\left[c\right] - \bar{c}\right)\int_{\underline{a}}^t F(a) \, da - 3Var(a)\Bigg].
$$

Therefore, voluntary information sharing is possible if and only if under the two conditions mentioned above,

$$
2\left(\mathbb{E}_\xi\left[c\right] - \bar{c}\right)\int_{\underline{a}}^t F(a) \, da > 3Var(a). \quad \text{(B.58)}
$$

[187]

Then we consider the other case when $\mathbb{E}_a[a] < t$. From (2.2) and (2.5), we have

$$\pi_R^{s*} - \pi_R^{ns*} = \frac{1}{16b} \Bigg[-3\mathbb{E}_a[a^2] + 3\mathbb{E}_a[a]^2 + 2\mathbb{E}_a[a]\mathbb{E}_\xi[c] - \mathbb{E}_\xi[c^2]$$
$$- 2\left(\mathbb{E}_\xi[c] \int_{\underline{a}}^{t} a\, dF(a) + \bar{c}\int_{t}^{\bar{a}} a\, dF(a)\right) + \mathbb{E}_\xi[c^2]F(t) + \bar{c}^2\bar{F}(t)\Bigg]. \quad \text{(B.59)}$$

Following the similar process, we can get the sufficient and necessary condition that leads to voluntary information sharing when $\mathbb{E}_a[a] < t$:

$$2(\mathbb{E}_\xi[c] - \bar{c})\int_{t}^{\bar{a}} \bar{F}(a)\, da > 3Var(a), \quad \text{(B.60)}$$

where $\bar{F}(\cdot) = 1 - F(\cdot)$. ∎

Proof Proof for Corollary 2.3.1: Let $\mathbb{E}[a] = t$, then condition (B.58) (similarly for (B.60)) directly leads to the following simpler necessary condition for voluntary information sharing:

$$\mathbb{E}_\xi[c] - \bar{c} > \frac{3Var(a)}{2\int_{\underline{a}}^{\mathbb{E}[a]} F(a)\, da}. \quad \text{(B.61)}$$

Let σ_0 denote the standard deviation of random variable a, then for example if a follows a truncated normal distribution, the right-hand side of (B.61) is less than $3\sqrt{2\pi}\sigma_0/2$; if a follows a uniform distribution, the right-hand side of (B.61) is $2\sqrt{3}\sigma_0$. Because (B.61) is a necessary condition to (B.58), if (B.61) does not hold, one can direct conclude that there would not be information sharing. Furthermore, when a follows a symmetric distribution, we can actually get a distribution-free criteria based upon (B.61) by calculating $\inf_{F(\cdot)}\left\{3Var(a)/2\int_{\underline{a}}^{\mathbb{E}[a]} F(a)da\right\}$. Let $\tilde{a} = a - \mathbb{E}[a]$ which follows distribution $\widetilde{F}(\cdot)$, then we have

$$\int_{\underline{a}}^{\mathbb{E}[a]} F(a)da = \int_{\underline{a}-\mathbb{E}[a]}^{0} \widetilde{F}(a)da = -\int_{\underline{a}-\mathbb{E}[a]}^{0} a\widetilde{f}(a)da = \frac{1}{2}\mathbb{E}[|\tilde{a}|] \leq \frac{1}{2}\sqrt{Var(\tilde{a})} = \frac{1}{2}\sigma_0,$$

[188]

where the inequality comes from Hölder's inequality. Therefore, $3Var(a)/2 \int_{\underline{a}}^{\mathbb{E}[a]} F(a)da \geq 3\sigma_0$, that is, if $\mathbb{E}_\xi[c] - \bar{c} < 3\sigma_0$, there would not be voluntary information sharing. $\blacksquare$

Proof Proof for Corollary 2.4.1: It is straightforward to see that the existence of the information-sharing channel cannot make the supplier worse, because the supplier can simply neglect any shared information and act as if there is no information-sharing channel. If no-information-sharing is not an equilibrium, i.e., $\Pi_R^S(1) - \Pi_R^N(0) > 0$, there must exist an integer $k^* \in \{1, \cdots, n-1\}$ such that $\Pi_R^S(l) > \Pi_R^N(l-1)$ for all $l \leq k^*$ and $\Pi_R^S(k^*+1) \leq \Pi_R^N(k^*)$, then k^* out of n retailers sharing information is a Nash equilibrium. We claim that it Pareto improves the no-information-sharing outcome.

From (B.11) and (B.12), for any $k \in \{1, \cdots, n\}$, we have

$$\Pi_R^N(k) - \Pi_R^S(k) = \frac{4\sigma_0^4 \sigma^2 (n+1)^2 ((k+1)\sigma_0^2 + \sigma^2)}{(k\sigma_0^2 + \sigma^2)((n+k+1)\sigma_0^2 + 2\sigma^2)^2} > 0. \tag{B.62}$$

Therefore for any $l \leq k^*$, $\Pi_R^S(l) \geq \Pi_R^N(l-1) > \Pi_R^S(l-1)$, which leads to $\Pi_R^S(k^*) > \Pi_R^S(1) \geq \Pi_R^N(0)$. On the other hand, $\Pi_R^N(k^*) > \Pi_R^S(k^*) > \Pi_R^N(0)$, so we conclude that the information sharing equilibrium where k^* out of n retailers share information Pareto dominates the no-information-sharing outcome. $\blacksquare$

Proof Proof for Lemma 2.4.2: Here we only prove $\Delta_I^{SN}(k) < \Delta_I^{NN}(k)$, the rest of the relations are straightforward. Direct calculation gives

$$\Delta_I^{NN}(k) - \Delta_I^{SN}(k) = \frac{4(n+1)^2 \sigma^2 \sigma_0^4 (\sigma^2 + (k+1)\sigma_0^2)}{(\sigma^2 + k\sigma_0^2)(2\sigma^2 + (n+k+1)\sigma_0^2)^2} > 0.$$

The proof is completed. $\blacksquare$

Proof   Proof for Proposition 2.4.2: From (B.1), (B.2) and (B.4), we know that for realized production cost c and shared information $\{Y_j\}_{j \in K}$,

$$Q(k) = \frac{1}{n+1}\left(\frac{n(a-c)}{2} + \left(kA_1^k + (n-k)B_1^k - \frac{n}{2}A_1^k\right) \cdot \sum_{j \in K} Y_j + B_2^k \cdot \sum_{i \in N\setminus K} Y_i\right).$$

On the other hand, we know that

$$\sum_{i \in N\setminus K} Y_i \,\bigg|\, \{Y_j\}_{j \in K} \sim N\left((n-k)A_1^k \cdot \sum_{j \in K} Y_j \,,\, (n-k)\sigma^2 \cdot \frac{A_1^k}{A_1^n}\right),$$

$$\sum_{j \in K} Y_j \sim N\left(0, k^2\sigma_0^2 + k\sigma^2\right),$$

then we can get

$$
\begin{aligned}
CS(k) &= \frac{1}{2}\mathbb{E}\left[Q^2\right] = \frac{n^2}{8(n+1)^2} \cdot \mathbb{E}\left[\left(a - c + A_1^k \cdot \sum_{j \in K} Y_j\right)^2\right] + \frac{(n-k)\sigma^2}{2(n+1)^2} \cdot \frac{A_1^k}{A_1^n} \cdot \left(B_2^k\right)^2 \\
&= \frac{n^2}{8(n+1)^2} \cdot \left((a - \bar{c})^2 + 2\left(\mathbb{E}_\xi[c] - \bar{c}\right)A_1^k \cdot \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(x)\,dx + A_1^k \cdot k\sigma_0^2\right) \\
&\quad + \frac{(n-k)\sigma^2}{2(n+1)^2} \cdot \frac{A_1^k}{A_1^n} \cdot \left(B_2^k\right)^2 \\
&= \frac{n^2}{8(n+1)^2} \cdot \left((a - \bar{c})^2 + 2\left(\mathbb{E}_\xi[c] - \bar{c}\right)A_1^k \cdot \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(x)\,dx + A_1^k \cdot k\sigma_0^2\right. \\
&\quad \left. + \frac{4(n-k)\sigma^2}{n^2} \cdot \frac{A_1^k}{A_1^n} \cdot \left(B_2^k\right)^2\right),
\end{aligned}
$$

[190]

therefore

$$CS(k) - CS(k-1)$$

$$= \frac{n^2}{8(n+1)^2} \cdot \left(2\left(\mathbb{E}_\xi[c] - \bar{c}\right) \left(A_1^k \int_{-\infty}^{t_Y(k)} \widetilde{F}_k(x)\, dx - A_1^{k-1} \int_{-\infty}^{t_Y(k-1)} \widetilde{F}_{k-1}(x)\, dx \right) \right.$$

$$+ \left(A_1^k \cdot k\sigma_0^2 + \frac{4(n-k)\sigma^2}{n^2} \cdot \frac{A_1^k}{A_1^n} \cdot \left(B_2^k\right)^2 \right)$$

$$\left. - \left(A_1^{k-1} \cdot (k-1)\sigma_0^2 + \frac{4(n-k+1)\sigma^2}{n^2} \cdot \frac{A_1^{k-1}}{A_1^n} \cdot \left(B_2^{k-1}\right)^2 \right) \right)$$

$$\equiv \frac{n^2}{8(n+1)^2} \cdot \left(\Delta_I^{CS}(k) + 2(\mathbb{E}_\xi[c] - \bar{c}) \left(\int_{-\infty}^{t-a} \Phi\left(\frac{x}{\sigma_k}\right) - \Phi\left(\frac{x}{\sigma_{k-1}}\right) dx \right) \right), \quad k > 1,$$

and when $k = 1$,

$$CS(1) - CS(0)$$

$$= \frac{n^2}{8(n+1)^2} \cdot \left(A_1^1 \cdot \sigma_0^2 + \frac{4(n-1)\sigma^2}{n^2} \cdot \frac{A_1^1}{A_1^n} \cdot \left(B_2^1\right)^2 - 4\sigma^2 \cdot \frac{A_1^0}{A_1^n} \cdot \left(B_2^0\right)^2 \right.$$

$$\left. + 2\left(\mathbb{E}_\xi[c] - \bar{c}\right) \int_{-\infty}^{|t-a|} \Phi\left(\frac{x}{\sigma_1}\right) dx \right)$$

$$\equiv \frac{n^2}{8(n+1)^2} \cdot \left(\Delta_I^{CS}(1) + 2\left(\mathbb{E}_\xi[c] - \bar{c}\right) \int_{-\infty}^{|t-a|} \Phi\left(\frac{x}{\sigma_1}\right) dx \right).$$

We can write the above in a more concise way as

$$CS(k) - CS(k-1) = \frac{n^2}{8(n+1)^2} \left(\Delta_I^{CS}(k) + \Delta_H(k) \right).$$

[191]

We notice that

$$\Delta_I^{CS}(k) - \Delta_I^{SN}(k)$$

$$= 4(n+1)^2 \sigma^2 \left(\frac{\sigma_0^4 \left((n(n-1)+k-1)\sigma^2 + (k-1)n(n+1)\sigma_0^2 \right)}{\left((k-1)\sigma_0^2 + \sigma^2 \right) \left(n(n+k)\sigma_0^2 + 2n\sigma^2 \right)^2} \right.$$

$$\left. + \frac{(n-k)\left(n\sigma_0^2 + \sigma^2 \right)\sigma_0^4}{n^2 \left(k\sigma_0^2 + \sigma^2 \right) \left((n+k+1)\sigma_0^2 + 2\sigma^2 \right)^2} \right) > 0,$$

therefore the condition for $CS(k) \geq CS(k-1)$ is weaker than $\Pi_R^S(k) \geq \Pi_R^N(k-1)$. On the other hand, according to the proof of Corollary 2.4.1, when no-information-sharing is not a Nash equilibrium, there must exists a $k^* \in \{1, \cdots, n\}$ such that $\Pi_R^S(l) \geq \Pi_R^N(l-1)$ holds for all $l \leq k^*$. Therefore for such k^*, we have $CS(l) > CS(l-1)$ hold for all $l \leq k^*$, which directly leads to $CS(k^*) > CS(0)$. ∎

C. Appendix for Chapter 3

C.1 A Constant-Factor Approximation ($K = T$)

In this section, for any $\epsilon > 0$, we present a simple $\frac{1-\epsilon}{2}$-approximation scheme for DISPLAY-OPT-T that runs in polynomial time. The approach begins by considering each stage independently of all others, and deriving near-optimal assortments for these T single-stage sub-problems. We then show how to string together these T assortments to produce a feasible sequence for our original problem, which achieves the desired performance guarantee.

The stage-t assortment problem. For each stage $t \in [T]$, we define the stage-t assortment problem as

$$R_t^{\text{myopic}} = \max_{\substack{S_t \subseteq [n]: \\ w(S) \geq W_t}} R(S_t), \tag{C.1}$$

where $R(S) = \sum_{i \in S} \frac{\rho_i}{1+w(S)}$ denotes the expected revenue earned from offering assortment S. In Appendix C.1.1, we show that even this single-stage variant of our problem is NP-Hard via a relatively straightforward reduction from the 2-partition problem, which is well-known to be NP-Hard ([105]). Nonetheless, a fully polynomial-time approximation scheme (FPTAS) can easily be achieved, whose exact nature is formally stated in the following lemma.

Lemma C.1.1 *For any stage $t \in [T]$ and $\epsilon > 0$, there is an algorithm that returns an assortment $\tilde{S}_t \subseteq [n]$ with total weight $w(\tilde{S}_t) \geq W_t$ and expected revenue $R(\tilde{S}_t) \geq (1 - \epsilon) \cdot R_t^{\text{myopic}}$, whose running time is $O(\frac{n}{\epsilon}^{O(1)})$.*

The above result can be established with basic adaptations of the dynamic programming ideas presented in [93], who show how to solve problem (C.1) with a flipped weight constraint of $w(S) \leq W_t$. Hence for brevity, we omit its proof.

Stitching the stage-t assortments together. Consider the T assortments $\tilde{S}_1, \ldots, \tilde{S}_T$ derived from applying Lemma C.1.1 for each stage $t \in [T]$. We recursively build the sequence of assortments $\hat{\mathcal{S}} = (\hat{S}_1, \ldots, \hat{S}_T)$ as follows. We set $\hat{S}_1 = \tilde{S}_1$, and then for stage $t \geq 2$, we set

$$
\hat{S}_t = \begin{cases} \hat{S}_{t-1}, & \text{if } w(\hat{S}_{t-1}) \geq W_t \\ \hat{S}_{t-1} \cup \tilde{S}_t, & \text{otherwise.} \end{cases}
$$

By construction, we clearly have that $\hat{\mathcal{S}} \in \mathcal{F}$, and the following lemma shows that this sequence of assortments achieves the desired guarantee with respect to its expected revenue.

Lemma C.1.2 $\mathcal{R}(\hat{\mathcal{S}}) \geq \frac{1-\epsilon}{2} \cdot \mathcal{R}(\mathcal{S}^*)$.

C.1.1 NP-Hardness of the single-stage problem

Our proof utilizes a reduction from the the 2-partition problem, where the input is a sequence of m integers $S = \{a_1, \ldots, a_m\}$ whose sum is $\sum_{i \in [m]} a_i = 2L$ for some $L \in \mathbb{Z}_+$. The goal is to find a subset $S \subset [m]$ such that $\sum_{i \in S} a_i = L$.

Given an arbitrary instance of 2-partition, we create the following instance of problem (C.1). We create a product for each of the m integers, whose associated weight is $w_i = a_i$ and whose revenue is $r_i = 0$. Additionally, we create a "special product", given index s, whose revenue and weight is $r_s = w_s = 1$. Finally, we let $W_t = L$. In this case, problem (C.1) reduces to

$$
\max_{\substack{S \subseteq [m]: \\ w(S) \geq L}} \sum_{i \in S} \frac{1}{2 + w(S)}, \tag{C.2}
$$

[194]

since the special product must be included in the optimal assortment. Letting S^* denote the optimal solution to the above problem, it is straightforward to see that a partition exists if and only if $\frac{1}{1+w(S^*)} = \frac{1}{1+L}$, in which case S^* reveals the partition.

C.1.2 Proof of Lemma C.1.2

To begin, we show the following two intermediate claims, and note that we can assume without loss of generality that $w(\tilde{S}_{t-1}) \leq w(\tilde{S}_t)$ and that $R(\tilde{S}_{t-1}) \geq R(\tilde{S}_t)$ for any $t \in [T-1]$.

Claim C.1.3 *For any stage $t \in [T]$, we have $w(\hat{S}_t) \leq 2w(\tilde{S}_t)$.*

Proof We prove the result by induction over the t. The base of $t = 1$ holds trivially since we set $\hat{S}_1 = \tilde{S}_1$, and so we proceed to the general case of $t \geq 2$. If $\hat{S}_t = \hat{S}_{t-1}$, we have that

$$w(\hat{S}_t) = w(\hat{S}_{t-1}) \leq 2w(\tilde{S}_{t-1}) \leq 2w(\tilde{S}_t),$$

where the first inequality uses the induction hypothesis and the second uses the fact that $w(\tilde{S}_{t-1}) \leq w(\tilde{S}_t)$. On other hand, if $\hat{S}_t = \hat{S}_{t-1} \cup \tilde{S}_t$, we must have that $w(\hat{S}_{t-1}) < W_t$, and so

$$w(\hat{S}_t) \leq w(\hat{S}_{t-1}) + w(\tilde{S}_t) < W_t + w(\tilde{S}_t) \leq 2w(\tilde{S}_t),$$

where the last inequality uses the feasibility of $\tilde{S}_t$ for stage t.

∎

Claim C.1.4 *For any stage $t \in [T]$, we have $R(\hat{S}_t) \geq \frac{1}{2} \cdot R(\tilde{S}_t)$.*

[195]

Proof We prove the result by induction over the t. The base of $t = 1$ holds trivially since we set $\hat{S}_1 = \tilde{S}_1$, and so we proceed to the general case of $t \geq 2$. If $\hat{S}_t = \hat{S}_{t-1} \cup \tilde{S}_t$, we have that

$$
\begin{aligned}
R(\hat{S}_t) &= \frac{1}{1 + w(\hat{S}_t)} \cdot \sum_{i \in \hat{S}_t} \rho_i \\
&\geq \frac{1}{1 + 2w(\tilde{S}_t)} \cdot \sum_{i \in \hat{S}_t} \rho_i \\
&\geq \frac{1}{1 + 2w(\tilde{S}_t)} \cdot \sum_{i \in \tilde{S}_t} \rho_i \\
&\geq \frac{1}{2} \cdot R(\tilde{S}_t).
\end{aligned}
$$

The first inequality uses Claim C.1.3, and the second uses the fact that $\hat{S}_t \supseteq \tilde{S}_t$. On the other hand, if $\hat{S}_t = \hat{S}_{t-1}$, we have that

$$
R(\hat{S}_t) = R(\hat{S}_{t-1}) \geq \frac{1}{2} \cdot R(\tilde{S}_{t-1}) \geq \frac{1}{2} \cdot R(\tilde{S}_t),
$$

where the first inequality uses the induction hypothesis. ∎

So, from Claim C.1.4, we get that

$$
\mathcal{R}(\hat{S}) = \sum_{t \in [T]} \lambda_t \cdot R(\hat{S}_t) \geq \frac{1}{2} \cdot \sum_{t \in [T]} \lambda_t \cdot R(\tilde{S}_t) \geq \frac{1 - \epsilon}{2} \cdot \mathcal{R}(\mathcal{S}^*),
$$

as desired.

C.2 Proofs from Section 3.2

C.2.1 Problem DISPLAY-OPT-K is strongly NP-Hard.

Our proof utilizes a reduction from the the 3-partition problem, where the input is a sequence of $3m$ integers $S = \{a_1, \ldots, a_{3m}\}$ whose sum is $\sum_{i \in S} a_i = mL$ for some $L \in \mathbb{Z}_+$. The goal is to find a partition of S into m sets $A_1, \ldots A_m$ such that $\sum_{i \in A_1} a_i = \ldots = \sum_{i \in A_m} a_i = L$.

Given an arbitrary instance of 3-partition, we create the following instance of our assortment problem. We create m stages and set $\lambda_t = 2 + tL$ for each $t \in [m]$. We create a product for each of the $3m$ integers, whose associated weight is $w_i = a_i$ and revenue is $r_i = 0$. Additionally, we create a "special product", given index s, whose revenue and weight is $r_s = w_s = 1$. Finally, for each $t \in [m]$, we let $W_t = 1 + tL$. The optimal sequence of assortments is once again denoted as $\mathcal{S}^* = (S_1^*, \ldots, S_m^*)$.

Claim C.2.1 *We have $\mathcal{R}(\mathcal{S}^*) = m$ if and only if a valid 3-partition exists.*

Proof First, note that $\mathcal{S}^*$ clearly always adds the special product to the first stage, since it is the only product with a non-zero revenue. As a result, we will implicitly assume that any feasible sequence of assortments adds the special products to stage one, and let $S_t \subseteq [3m]$ denote the assortment decisions across stages $t \in [m]$. The expected revenue of any feasible sequence of assortments is therefore

$$\mathcal{R}(\mathcal{S}) = (2 + L) \cdot \frac{1}{2 + w(S_1)} + (2 + 2L) \cdot \frac{1}{2 + w(S_2)} + \ldots + (2 + mL) \cdot \frac{1}{2 + w(S_m)} \quad \text{(C.3)}$$

Let $A_1, \ldots, A_m$ correspond to a valid 3-partition, and consider the sequence of assortments $\mathcal{S}'$, where for stage $t \in [m]$, we set $S_t' = \biguplus_{\tau \leq t} A_\tau$. We clearly have that $S_t' \subset S_{t-1}'$ for any stage $t \in [m-1]$, and also, since $w(S_t') = tL$, we know that the stage-t weight constraint is

[197]

met for each stage $t \in [m]$. Finally, given the expected revenue expression in (C.3), we see that $\mathcal{R}(\mathcal{S}') = m$.

On the other hand, assume that the sequence of assortments $\mathcal{S}'$ satisfies $\mathcal{R}(\mathcal{S}') = m$. Now, since feasibility of $\mathcal{S}'$ requires that $w(S_t) \geq tL$, we have that $\mathcal{R}(\mathcal{S}') = m$ if and only if $w(S_t') = tL$ for each $t \in [m]$. This directly implies that $A_1 = S_1$ and $A_t = S_t \setminus S_{t-1}$ for $t \geq 2$ is a valid 3-partition. ∎

Claim C.2.2 *If a valid 3-partition does not exist, then $\mathcal{R}(\mathcal{S}^*) \leq m - \frac{1}{3 + (m-1) \cdot L}$.*

Proof If a valid 3-partition does not exist, then is some stage t, the optimal sequence $\mathcal{S}^*$ must satisfy $w(S_t^*) > tL$. In this case, it is straightforward to see that the best expected revenue is achieved by a sequence of assortment for which only the stage $m - 1$ weight constraint is not satisfied with equality, and instead satisfies $w(S_{m-1}^*) = (m-1) \cdot L + 1$. In this case, we have

$$
\begin{aligned}
\mathcal{R}\left(\mathcal{S}^*\right) &= m - 1 + (2 + (m-1) \cdot L) \cdot \frac{1}{3 + (m-1) \cdot L} \\
&= m - \frac{1}{3 + (m-1) \cdot L},
\end{aligned}
$$

∎

as desired.

Summary. Combining Claims C.2.1 and C.2.2, we see that an FPTAS for problem DISPLAY-OPT-K, run with $\epsilon < \frac{1}{m \cdot (3 + (m-1) \cdot L)}$, would yield a sequence of assortments $\mathcal{S}$ with expected revenue $\mathcal{R}(\mathcal{S}) \geq m \cdot (1 - \epsilon) > m - \frac{1}{3 + (m-1) \cdot L}$ if a valid partition exists. Consequently, this FPTAS would return a valid 3-partition if one exists, in a pseudo-polynomial running time. Since the 3-partition problem is strongly NP-Hard, this implies that problem DISPLAY-OPT-K cannot admit an FPTAS unless $P = NP$.

C.2.2   Proof of Lemma 3.2.2

Case 1: $\tilde{S}$ returned by Step 2 (line 9). In this case, we get that $\hat{w}(\tilde{S}) \geq W_t$. To show the upper bound on the total weight, we let $\ell \in \tilde{S}$ denote the final product added to $\tilde{S}$ before the reaching the return statement, and let $q(\ell) \in [\mathcal{Q}]_0$ be the class weight class to which product l belongs. In this case, we get that

$$
\begin{aligned}
W_t \;>\;& \hat{w}(\tilde{S} \setminus \{\ell\}) \\
\geq\;& \hat{w}\left(\biguplus_{q \in Q_A(\tilde{S})} \tilde{S}_q \right) \\
\geq\;& \frac{1}{\epsilon} \cdot w_\ell,
\end{aligned}
$$

where the last inequality follows since $\tilde{S}$ must contain at least $\frac{1}{\epsilon}$ active classes indexed higher than $q(l)$. As such, we get that

$$
\frac{1}{1+\epsilon} \cdot w(\tilde{S}) \leq \hat{w}(\tilde{S}) \leq W_t + w_\ell \leq (1+\epsilon) \cdot W_t,
$$

and so the condition of the lemma statement is satisfied.

Case 2: $\tilde{S}$ returned by Step 2 (line 13). Moreover, this return statement is reached only if $|Q_A(\tilde{S})| \leq \frac{1}{\epsilon}$, and so the condition of the lemma statement holds.

C.2.3   Proof of Lemma 3.2.3

We show will show that each of the three candidate sequences of assortments satisfies the lemma statement, if it is indeed chosen to be $\hat{\mathcal{S}}$.

Candidate 1 feasibility. If $\hat{\mathcal{S}} = \mathcal{S}^{(1)}$, then we know that $w(S_t^{(1)}) \geq \hat{w}(S_t^{(1)}) \geq W_t$ for each $t \in [T]$, and hence the weight constraint in each stage is satisfied. Moreover, since $k_{t+1,q}^* \geq k_{t,q}^*$ for any stage $t \in [T-1]$ and class $q \in [\mathcal{Q}]_0$, we are also guaranteed that $S_1^{(1)} \subseteq S_2^{(1)} \subseteq \ldots \subseteq S_T^{(1)}$. Hence $\mathcal{S}^{(1)} \in \mathcal{F}$.

Candidate 1 revenue. We have that

$$
\begin{aligned}
\mathcal{R}(\mathcal{S}^{(1)}) &\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^{(1)})} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in \mathcal{C}_q[k_{t,q}^*]} \rho_i \\
&\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{(1 + 5\epsilon)} \cdot \frac{1}{1 + w(S_t^*)} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in \mathcal{C}_q[k_{t,q}^*]} \rho_i \\
&\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{(1 + 5\epsilon)} \cdot \frac{1}{1 + w(S_t^*)} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in S_{t,q}^*} \rho_i \\
&\geq (1 - 5\epsilon) \cdot \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^*)} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in S_{t,q}^*} \rho_i \\
&= (1 - 5\epsilon) \cdot \mathcal{R}(\mathcal{S}^*).
\end{aligned}
$$

The first inequality follows since $\mathcal{S}^{(1)}$ adds at least $k_{t,q}^*$ items from each class. The second inequality follows because

$$
\begin{aligned}
w(S_t^{(1)}) &\leq \sum_{\substack{q \in [\mathcal{Q}]_0: \\ k_{t,q}^* \leq \frac{1}{\epsilon}}} w_{\min} \cdot (1+\epsilon)^{q+1} \cdot k_{t,q}^* + \sum_{\substack{q \in [\mathcal{Q}]_0: \\ k_{t,q}^* > \frac{1}{\epsilon}}} w_{\min} \cdot (1+\epsilon)^{q+1} \cdot \left((1+\epsilon) \cdot k_{t,q}^* + \underbrace{1}_{\epsilon k_{t,q}^* > 1} \right) \\
&\leq (1+\epsilon) \cdot \left(\sum_{\substack{q \in [\mathcal{Q}]_0: \\ k_{t,q}^* \leq \frac{1}{\epsilon}}} w_{\min} \cdot (1+\epsilon)^q \cdot k_{t,q}^* + \sum_{\substack{q \in [\mathcal{Q}]_0: \\ k_{t,q}^* > \frac{1}{\epsilon}}} w_{\min} \cdot (1+\epsilon)^q \cdot \left((1+2\epsilon) \cdot k_{t,q}^* \right) \right) \\
&\leq (1 + 5\epsilon) \cdot w(\mathcal{S}_t^*).
\end{aligned}
$$

for any stage $t \in [T]$. The second inequality uses the fact that $\sum_{i \in C_q[k^*_{t,q}]} \rho_i \geq \sum_{i \in S^*_{t,q}} \rho_i$ for any stage $t \in [T]$ and class $q \in [\mathcal{Q}]_0$, since $C_q[k^*_{t,q}]$ is precisely the $k^*_{t,q}$ items from C_q with the largest ρ-value.

Candidate 2 feasibility. If $\mathcal{S}' = S^{(2)}$, then it is again trivial that the weight constraint in each stage is satisfied. Moreover, it is easy to see that $S^{(2)}_1 \subseteq S^{(2)}_2 \subseteq \ldots \subseteq S^{(2)}_T$, based on how Candidate 2 is defined and the fact that fill events only add products to each class. Hence $\mathcal{S}^{(2)} \in \mathcal{F}$.

Candidate 2 revenue. To establish the desired guarantee for $\mathcal{R}(\mathcal{S}^{(2)})$ requires first proving the following two intermediate claims, which together allow us to relate the total weight of $S^{(2)}_t$ to that of S^*_t in each stage $t \in [T]$. For this purpose, let $T_{\text{fill}} \subseteq [T]$ give the stages in which fill events were invoked in creating $\mathcal{S}^{(2)}$. Furthermore, for stage $t \in T_{\text{fill}}$, let $w^{\text{fill}}_t = w(S^{(2)}_t) - w(S^{(1)}_t)$ denote the total weight added on top of $S^{(1)}$ by fill events in stages $1, \ldots, t$.

Claim C.2.3 *For any stage $t \in T_{\text{fill}}$, we have $w^{\text{fill}}_t \leq 4\epsilon W_t$.*

Proof Consider arbitrary stage $t \in T_{\text{fill}}$, and note that that $w(S^{(1)}_t) \geq \sum_{q \in [\mathcal{Q}]_0} w_{\min} \cdot (1 + \epsilon)^q \cdot k^*_{t,q}$. Additionally, we know that $W_t \leq w(S^*_t) \leq \sum_{q \in [\mathcal{Q}]_0} w_{\min} \cdot (1 + \epsilon)^{q+1} \cdot k^*_{t,q}$ by the feasibility of S^*_t. Putting these two strings of inequalities together yields that $w(S^{(1)}_t) \geq (1 - \epsilon) \cdot W_t$. Moreover, by Lemma 3.2.2 we have that $w(S^{(2)}_t) \leq (1+3\epsilon) \cdot W_t$, since the fill event must have returned an assortment that abides by the stage t weight constraint. So, putting everything together gives that

$$w^{\text{fill}}_t = w(S^{(2)}_t) - w(S^{(1)}_t) \leq (1 + 3\epsilon) \cdot W_t - (1 - \epsilon) \cdot W_t = 4\epsilon W_t,$$

as desired. ∎

The second claim builds off of the first, relating the total weights in each stage under $\mathcal{S}^{(2)}$ and $\mathcal{S}^*$.

Claim C.2.4 *For any stage $t \in [T]$, we have $w(S_t^{(2)}) \leq (1 + 9\epsilon) \cdot w(S_t^*)$.*

Proof If $S_t^{(2)} = \text{FILL}(S_t^{\text{temp}}, t)$, then by Lemma 3.2.2, we must have that $w(S_t^{(2)}) \leq (1 + 3\epsilon) \cdot W_t \leq (1 + 3\epsilon) \cdot w(S_t^*)$. Otherwise, we have that $S_t^{(2)} = S_t^{\text{temp}} = S_t^{(1)} \cup S_{t-1}^{(2)}$. For any stage $t \in [T]$, let $\tau_{\text{last}} = \max\{\tau \in T_{\text{fill}} : \tau < t\}$ denote the latest stage among stages $1, \ldots, t$ in which a fill event was invoked in building Candidate 2. We have

$$
\begin{aligned}
w(S_t^{(2)}) &= w(S_t^{(1)} \cup S_{t-1}^{(2)}) \\
&= w(S_t^{(1)}) + w(S_{\tau_{\text{last}}}^{(2)} \setminus S_t^{(1)}) \\
&\leq w(S_t^{(1)}) + w(S_{\tau_{\text{last}}}^{(2)}) - w(S_{\tau_{\text{last}}}^{(1)}) \\
&= w(S_t^{(1)}) + w_{\tau_{\text{last}}}^{\text{fill}} \\
&\leq (1 + 5\epsilon) \cdot w(S_t^*) + 4\epsilon W_t \\
&\leq (1 + 9\epsilon) \cdot w(S_t^*).
\end{aligned}
$$

The first inequality follows because $S_t^{(1)} \supseteq S_{\tau_{\text{last}}}^{(1)}$ since $\tau_{\text{last}} < t$, while the second inequality uses Claim C.2.3 along with the fact that $w(S_t^{(1)}) \leq (1 + 5\epsilon) \cdot w(S_t^*)$ as was established in the revenue proof for Candidate 1. ∎

[202]

With these intermediate claims in-hand, we now consider the expected revenue earned under $\mathcal{S}^{(2)}$:

$$
\begin{aligned}
\mathcal{R}(\mathcal{S}^{(2)}) &= \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^{(2)})} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in \mathcal{C}_q[k_{t,q}^{(2)}]} \rho_i \\
&\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^{(2)})} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in \mathcal{C}_q[k_{t,q}^*]} \rho_i \\
&\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + 9\epsilon} \cdot \frac{1}{1 + w(S_t^*)} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in \mathcal{C}_q[k_{t,q}^*]} \rho_i \\
&\geq (1 - 9\epsilon) \cdot \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^*)} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in S_{t,q}^*} \rho_i \\
&= (1 - 9\epsilon) \cdot \mathcal{R}(\mathcal{S}^*) .
\end{aligned}
$$

The first inequality uses the fact that $k_{t,q}^{(2)} \geq k_{t,q}^*$ for any stage $t \in [T]$ and class $q \in [\mathcal{Q}]_0$, since fill events only add products in descending ρ-order. The second inequality uses Claim C.2.4.

Candidate 3 feasibility. In this case, we show that the weight constraint in each stage is satisfied by first considering stages $t \in T^<$, and then stages $t \notin T^<$.

- Stages $t \in T^<$: Recall that for stages $t \in T^<$, a fill event was executed and stopped at line 13; returning an assortment whose rounded weight does not satisfy the stage-t

weight constraint, and that has fewer than $\frac{1}{\epsilon}$ active classes, all of which are infrequent. With this in mind, we have

$$
\begin{aligned}
w(S_t^{(3)}) \;=\;& w\left(\biguplus_{q\notin Q_A^<} S_{t,q}^{(2)}\right) + w\left(\biguplus_{q\in Q_A^<} \mathcal{C}_q^*\left(k_{q,t}^{(2)}\right)\right) \\[2ex]
\;=\;& w\left(\biguplus_{q\notin Q_A^<} S_{t,q}^{(2)}\right) + w\left(\biguplus_{\substack{q\in Q_A^<:\\ k_{t,q}^{(2)}\le\frac{1}{\epsilon}}} S_q^*\left(k_{q,t}^{(2)}\right)\right) + w\left(\biguplus_{\substack{q\in Q_A^<:\\ k_{t,q}^{(2)}>\frac{1}{\epsilon}}} \mathcal{C}_q^*\left(k_{q,t}^{(2)}\right)\right) \\[2ex]
\;\ge\;& w\left(\biguplus_{q\notin Q_A^<} S_{t,q}^{(2)}\right) + w\left(\biguplus_{\substack{q\in Q_A^<:\\ k_{t,q}^{(2)}\le\frac{1}{\epsilon}}} S_{t,q}^*\right) + w\left(\biguplus_{\substack{q\in Q_A^<:\\ k_{t,q}^{(2)}>\frac{1}{\epsilon}}} \mathcal{C}_q^*\left(k_{q,t}^{(1)}\right)\right) \\[2ex]
\;\ge\;& w\left(\biguplus_{q\notin Q_A^<} S_{t,q}^{(2)}\right) + w\left(\biguplus_{\substack{q\in Q_A^<:\\ k_{t,q}^{(2)}\le\frac{1}{\epsilon}}} S_{t,q}^*\right) + \sum_{\substack{q\in Q_A^<:\\ k_{t,q}^{(2)}>\frac{1}{\epsilon}}} w_{\min}\cdot(1+\epsilon)^{q+1}\cdot k_{t,q}^* \\[2ex]
\;\ge\;& w\left(\biguplus_{q\notin Q_A^<} S_{t,q}^{(2)}\right) + w\left(\biguplus_{q\in Q_A^<} S_{t,q}^*\right) \\[2ex]
\;=\;& w\left(\biguplus_{\substack{q\notin Q_A^<:\\ q\in Q_A(S_t^*)}} \mathcal{C}_q\right) + w\left(\biguplus_{q\in Q_A^<} S_{t,q}^*\right) \\[1ex]
\;\ge\;& w(S_t^*) \\[1ex]
\;\ge\;& W_t.
\end{aligned}
$$

The first inequality follows since we generally have that $k_{t,q}^{(2)} \ge k_{t,q}^{(1)}$ (where $k_{t,q}^{(1)} = |S_{t,q}^{(1)}|$), and because for all classes $q\in Q_A^<$ we must have that $k_{t,q}^{(2)} \ge k_{t,q}^*$ if $k_{t,q}^{(2)} \le \frac{1}{\epsilon}$, and hence by definition of $\frac{1}{\epsilon}$-capped class-q assortments we have that $S_q^*(k_{q,t}^{(2)}) \supseteq S_{t,q}^*$. The second inequality follows by definition of $k_{t,q}^{(1)}$, and the third inequality follows by definition of

[204]

$k_{t,q}^*$. To see the final equality, note that any non-empty class $q \notin Q_A^<$ must be exhausted under $\mathcal{S}^{(2)}$ by construction of a fill event.

- Stages $t \notin T^<$: For stages $t \in T^<$, we must have that $\hat{w}(S_t^{(2)}) \geq W_t$ by construction. Moreover, letting $k_{t,q}^{(3)} = |S_{t,q}^{(3)}|$, it is easy to see that every stage-class pair, we have that $k_{t,q}^{(2)} = k_{t,q}^{(3)}$, i.e. Candidates 2 and 3 add the same number of products from each class in each stage. Consequently, we get that $\hat{w}(S_t^{(2)}) = \hat{w}(S_t^{(3)}) \geq W_t$.

Finally, given that Candidates 2 and 3 add the same number of products from each class in each stage, it is easy to see that $S_1^{(3)} \subseteq S_2^{(3)} \subseteq \ldots \subseteq S_T^{(3)}$, and so $\mathcal{S}^{(3)} \in \mathcal{F}$.

Candidate 3 revenue. We have

$$
\begin{aligned}
\mathcal{R}(\mathcal{S}^{(3)}) &= \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^{(3)})} \cdot \left(\sum_{q \in Q_A^<} \sum_{i \in \mathcal{C}_q^*(k_{t,q}^{(2)})} \rho_i + \sum_{q \in Q_A^<} \sum_{i \in \mathcal{C}_q[k_{t,q}^{(2)}]} \rho_i \right) \\
&\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^{(3)})} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in S_{t,q}^*} \rho_i \\
&\geq (1 - \epsilon) \cdot \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^{(2)})} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in S_{t,q}^*} \rho_i \\
&\geq (1 - 10\epsilon) \cdot \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(S_t^*)} \cdot \sum_{q \in [\mathcal{Q}]_0} \sum_{i \in S_{t,q}^*} \rho_i \\
&= (1 - 10\epsilon) \cdot \mathcal{R}(\mathcal{S}^*).
\end{aligned}
$$

The first inequality the fact that $k_{t,q}^{(2)} \geq k_{t,q}^*$ combined with the fact that

$$
\sum_{i \in \mathcal{C}_q^*(k')} \rho_i \geq \sum_{i \in S_{t,q}^*[k]} \rho_i
$$

for any $k, k' \in [|\mathcal{C}_q|]_0$ such that $k \leq k'$. The second inequality uses the fact that Candidates 2 and 3 add the same number of products from each class in each stage, and so $w(S_t^{(3)}) \leq (1 + \epsilon) \cdot w(S_t^{(2)})$ for any stage $t \in [T]$. The final inequality uses Claim C.2.4.

[205]

C.2.4 Proof of Lemma 3.2.4

Recall that $T^< = \{t \in [T] : w(S_t^{(2)}) < W_t\}$ and that for any $t \in T^<$, the assortment $S_t^{(2)}$ must be the result of a fill event that concluded with fewer than $\frac{1}{\epsilon}$ active stages (since feasibility with respect to the weight constraint was never achieved during the fill event). As such, letting $\tau_{\text{before}} = \max\{\tau \leq t : \tau \in T^< \cup \{0\}\}$ if and $\tau_{\text{after}} = \min\{\tau > t : \tau \in T^< \cup \{T+1\}\}$ (we define $S_0^2 = S_{T+1}^{(2)} = \emptyset$), we have that

$$
|\{q \in Q_A^< : k_{t,q}^{(2)} > 0\}| = |Q_A(S_t^{(2)}) \cap Q_A^<|
$$

$$
= \left| Q_A(S_t^{(2)}) \cap \left(\biguplus_{\substack{\tau \leq t: \\ \tau \in T^<}} Q_A(S_\tau^{(2)}) \right) \right| + \left| Q_A(S_t^{(2)}) \cap \left(\biguplus_{\substack{\tau > t: \\ \tau \in T^<}} Q_A(S_\tau^{(2)}) \right) \right|
$$

$$
= \left| Q_A(S_t^{(2)}) \cap Q_A(S_{\tau_{\text{before}}}^{(2)}) \right| + \left| Q_A(S_t^{(2)}) \cap Q_A(S_{\tau_{\text{after}}}^{(2)}) \right|
$$

$$
\leq \frac{2}{\epsilon}.
$$

The second equality follows because $Q_A^< = \biguplus_{t \in T^<} Q_A(S_t^{(2)})$, while the third holds since we know that any class $q \notin Q_A(S_{\tau_{\text{before}}}^{(2)})$ that is active in stage $t < \tau_1$ must be exhausted in stage τ_1 by definition of a fill event. Along these same lines, any class $q \notin Q_A(S_{\tau_{\text{after}}}^{(2)})$ that is active in stage $t > \tau_{\text{after}}$ must be empty in stage τ_{after}. The lone inequality follows since we have $\max\{|Q_A(S_{\tau_{\text{before}}}^{(2)})|, |Q_A(S_{\tau_{\text{after}}}^{(2)})|\} \leq \frac{1}{\epsilon}$, as noted at the onset of this proof.

C.2.5 Proof of Lemma 3.2.5

Monoticity. To begin, note that we must have $K \leq K^+$, since $S \subseteq S^+$. We consider two cases:

- For any class $q < [q_{\max}(K^+) - L + 1]^+$, we have that $k_q^{+\uparrow} = |\mathcal{C}_q| \geq k_q^\uparrow$. The lone inequality follows because $K \leq K^+$ implies that $q_{\max}(K) \leq q_{\max}(K^+)$.

- The remaining cases follow directly from Lemma 2 of [98], and hence we omit their proof for brevity.

Total weight. We have that

$$
\begin{aligned}
w(S^\uparrow) &\leq w(S) + L \cdot \text{Power}_2\left[\frac{\epsilon}{L} \cdot w(S)\right] + \sum_{q < q_{\max}(K)-L+1} w(|\mathcal{C}_q|) \\
&\leq w(S) + 2\epsilon \cdot w(S) + \sum_{q < q_{\max}(K)-L+1} w(|\mathcal{C}_q|) \\
&\leq w(S) + 2\epsilon \cdot w(S) + n \cdot w_{\min} \cdot (1+\epsilon)^{q_{\max}(K)-L} \\
&= w(S) + 2\epsilon \cdot w(S) + \epsilon \cdot w_{\min} \cdot (1+\epsilon)^{q_{\max}(K)} \\
&\leq (1+3\epsilon) \cdot w(S),
\end{aligned}
$$

where the last inequality follows since S must include at least one product from class $q_{\max}(K)$.

C.2.6 Proof of Lemma 3.2.6

We consider the total number of distinct three parameter triples $(Q_{\text{CAP}}, \{A_q\}_{q \in Q_{\text{CAP}}}, K^\uparrow)$ that result from rounding the parameter triples $(Q_{\text{CAP}}, \{A_q\}_{q \in Q_{\text{CAP}}}, K)$ for all assortment $S \in \mathcal{U}$. Combining these three counts yields the desired result.

(i) For Q_{CAP}, the total number of possibilities is

$$
O(\mathcal{Q}^{\frac{2}{\epsilon}}) = O\left(\left(\frac{1}{\epsilon}\log\left(\frac{w_{\max}}{w_{\min}}\right)\right)^{O(\frac{1}{\epsilon})}\right)
$$

(ii) For each Q_{CAP}, there are a total of

$$
O\left(\left(n^{\frac{1}{\epsilon}}\right)^{|Q_{\text{CAP}}|}\right) = O\left(n^{O(\frac{1}{\epsilon^2})}\right)
$$

[207]

options for the collection of such assortments $\{A_q\}_{q \in Q_{\mathrm{CAP}}}$.

(iii) For each $(Q_{\mathrm{CAP}}, \{A_q\}_{q \in Q_{\mathrm{CAP}}})$, we must enumerate over all $\mathcal{K}^{\uparrow} = \{K^{\uparrow} : K \in \mathcal{K}\}$. Replicating the arguments of Lemma 4 of [98], it is possible to show that

$$|\mathcal{K}^{\uparrow}| = O\left(\log\left(\frac{nw_{\max}}{w_{\min}}\right) \cdot 2^{O\left(\frac{L}{\epsilon}\right)}\right) = O\left(\log\left(\frac{w_{\max}}{w_{\min}}\right) \cdot n^{O\left(\frac{1}{\epsilon^2}\right)}\right)$$

C.2.7 Proof of Lemma 3.2.7

For each stage $t \in [T]$, we know that $\hat{S}_t$ is represented by the parameter triplet

$$\left(\hat{Q}_{\mathrm{CAP}}, \left\{S_q^*\left(\min\{\frac{1}{\epsilon}, \hat{k}_{t,q}\}\right)\right\}_{q \in \hat{Q}_{\mathrm{CAP}}}, \hat{K}_t\right),$$

where $\hat{Q}_{\mathrm{CAP}} = \{q \in Q_A^{\lessgtr} : k_{t,q}^{(2)} > 0\}$ ($= \emptyset$ unless $\hat{S} = S^{(3)}$) and $\hat{k}_{t,q} = |\hat{S}_{t,q}|$. Consider the sequence of assortment $\hat{S}^{\uparrow} = (\hat{S}_1^{\uparrow}, \ldots, \hat{S}_T^{\uparrow})$, where $\hat{S}_t^{\uparrow}$ is the assortment corresponding the parameter triplet

$$\left(\hat{Q}_{\mathrm{CAP}}, \left\{S_q^*\left(\min\{\frac{1}{\epsilon}, \hat{k}_{t,q}\}\right)\right\}_{q \in \hat{Q}_{\mathrm{CAP}}}, \hat{K}_t^{\uparrow}\right),$$

i.e. we have simply up-rounded the utilization vectors that arise from the assortment $\hat{S}_t$. The following claim concerning $\hat{S}^{\uparrow}$ is enough to establish the lemma; the first two conditions establish feasibility for the dynamic program given in (3.3), and the third condition establishes the revenue guarantee.

Claim C.2.5 *The sequence of assortments $\hat{S}^{\uparrow}$ satisfies the following three conditions:*

(i) For any stage $t \in [T]$, we have that $\hat{S}_t^{\uparrow} \in \mathcal{U}_{\mathrm{small}}$

(ii) $\hat{S}^{\uparrow} \in \mathcal{F}$

(iii) $\mathcal{R}(\hat{S}^{\uparrow}) \geq (1 - 3\epsilon) \cdot \mathcal{R}(\hat{S})$

Proof We prove each of the three conditions as follows:

[208]

- Condition (i): By Lemma 3.2.4, we have that $|\hat{Q}_{\text{CAP}}| \leq \frac{2}{\epsilon}$. Furthermore, we also clearly have that $|S_q^*\left(\min\{\frac{1}{\epsilon}, \hat{k}_{t,q}\}\right)| \leq \frac{1}{\epsilon}$ for any $q \in \hat{Q}_{\text{CAP}}$.

- Condition (ii): Since the up-rounding scheme only added products to each class, we have that $w(\hat{S}_t^{\uparrow}) \geq w(\hat{S}_t) \geq W_t$ for each stage $t \in [T]$, and hence the weight constraints are satisfied. Furthermore, by the monotonicity property of Lemma 3.2.5, we have that $\hat{S}_t^{\uparrow} \subseteq \hat{S}_{t+1}^{\uparrow}$ for any $t \in [T-1]$.

- Condition (iii): We have that

$$
\begin{aligned}
\mathcal{R}(\hat{\mathcal{S}}^{\uparrow}) &= \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(\hat{S}_t^{\uparrow})} \cdot \sum_{i \in \hat{S}_t^{\uparrow}} \rho_i \\
&\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w(\hat{S}_t^{\uparrow})} \cdot \sum_{i \in \hat{S}_t} \rho_i \\
&\geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + (1 + 3\epsilon) \cdot w(\hat{S}_t)} \cdot \sum_{i \in \hat{S}_t} \rho_i \\
&= (1 - 3\epsilon) \cdot \mathcal{R}(\hat{\mathcal{S}}),
\end{aligned}
$$

where the first inequality follows since the up-rounding ensures that $\hat{S}_t \subseteq \hat{S}_t^{\uparrow}$, and the second inequality follows by the total weight property of Lemma 3.2.5.

∎

C.3 Proofs from Section 3.3

C.3.1 Proof of Theorem 3.3.1:

Let $\mathcal{S}^* = (S_1^*, \cdots, S_T^*)$ be the optimal solution to DISPLAY-OPT-1 with feasible space

$$
\mathcal{F}_1 = \left\{ (S_1, \cdots, S_T) : S_t \subseteq [n], \sum_{t \in [T]} \lambda_t \cdot cs\left(w\left(S_t\right)\right) \geq \alpha_1, S_1 \subseteq \cdots \subseteq S_T \right\},
$$

[209]

and define $W_t^* = w(S_t^*)$. Then it is straightforward to see that $\mathcal{S}^*$ is the optimal solution to the DISPLAY-OPT-T problem with feasible space

$$\mathcal{F}_T^* = \left\{ (S_1, \cdots, S_T) : S_t \subseteq [n], w(S_t) \geq W_t^*, S_1 \subseteq \cdots \subseteq S_T \right\}.$$

Here notice that any sequence $\mathcal{S} \in \mathcal{F}_T^*$ must satisfy $w(S_t) \geq w(S_t^*) = W_t^*$, $\forall t \in [T]$, or equivalently $cs(w(S_t)) \geq cs(w(S_t^*))$ due to the monotonicity assumption on $cs(\cdot)$, which directly implies that $\mathcal{S} \in \mathcal{F}_1$. Therefore, $\mathcal{S}$ cannot return a higher revenue than $\mathcal{S}^*$, which will otherwise contradict the optimality of $\mathcal{S}^*$ to DISPLAY-OPT-1.

Then according to the construction of $\mathcal{U}$ and $\mathcal{U}_{\text{small}}$ in Section 3.2.3, there must exist sequences of assortments $\hat{\mathcal{S}} = \left(\hat{S}_1, \cdots, \hat{S}_T \right) \in \bigtimes_{t=1}^T \mathcal{U}$ and $\widetilde{\mathcal{S}} = \left(\widetilde{S}_1, \cdots, \widetilde{S}_T \right) \in \bigtimes_{t=1}^T \mathcal{U}_{\text{small}}$ such that $\hat{S}_1 \subseteq \hat{S}_2 \subseteq \cdots \subseteq \hat{S}_T$, $\widetilde{S}_1 \subseteq \widetilde{S}_2 \subseteq \cdots \subseteq \widetilde{S}_T$, and

$$\mathcal{R}\left(\widetilde{\mathcal{S}} \right) \geq (1 - 3\epsilon) \cdot \mathcal{R}\left(\hat{\mathcal{S}} \right) \geq (1 - 3\epsilon)(1 - 10\epsilon) \cdot \mathcal{R}(\mathcal{S}^*) > (1 - 13\epsilon) \cdot \mathcal{R}(\mathcal{S}^*),$$

where the first inequality comes from Lemma 3.2.7, and the second inequality comes from Lemma 3.2.3. At the same time,

$$w\left(\widetilde{S}_t \right) \geq w\left(\hat{S}_t \right) \geq W_t^* , \ t \in [T],$$

where the first inequality is due to the *Weight added* property in Lemma 3.2.5, and the second inequality comes from the construction process of $\hat{\mathcal{S}}$ in Section 3.2.2. This directly implies that $cs\left(\widetilde{\mathcal{S}} \right) \geq cs(\mathcal{S}^*) \geq \alpha_1$.

Consider computing $\mathcal{V}_{\mathcal{U}_{\text{small}}}\left(1, \lfloor \mathcal{R}(\mathcal{S}^*) / (1 + \delta)^T \rfloor_{1+\delta}, \emptyset \right)$. We know that any nontrivial feasible solution $\mathcal{S}$ to the Bellman's equation (3.4) will result in a revenue such that $\mathcal{R}(\mathcal{S}) \geq \mathcal{R}(\mathcal{S}^*) / (1 + \delta)^{T+1}$. On the other hand, $\widetilde{\mathcal{S}}$ is clearly a feasible solution to (3.4) when $\mathcal{R}(\mathcal{S}^*) / (1 + \delta)^{T+1} < (1 - 13\epsilon) \cdot \mathcal{R}(\mathcal{S}^*)$. That is, when $(1 + \delta)^{T+1} > 1 / (1 - 13\epsilon)$,

[210]

(3.4) is guaranteed to admit a nontrivial solution. Suppose the output of (3.4) is $\widetilde{\mathcal{S}}^*$, then according to the formulation of (3.4), $cs\left(\widetilde{\mathcal{S}}^*\right) \geq cs\left(\widetilde{\mathcal{S}}\right) \geq \alpha_1$. At the same time $\mathcal{R}\left(\widetilde{\mathcal{S}}^*\right) \geq \mathcal{R}\left(\mathcal{S}^*\right)/(1+\delta)^{T+1}$.

We know that $\mathcal{R}\left(\mathcal{S}^*\right)$ takes value from set $\{0\} \cup [r_{\min}w_{\min}/\left(1+w_{\min}\right), r_{\max}]$, therefore $\lfloor \mathcal{R}\left(\mathcal{S}^*\right)/(1+\delta)^T \rfloor_{1+\delta}$ takes value from a $(1+\delta)-$power $\mathsf{Dom}^\delta = \{0\}\cup\left\{(1+\delta)^k : \underline{k} \leq k \leq \overline{k}\right\}$ where $(1+\delta)^{\underline{k}} = \left\lfloor r_{\min}w_{\min}/\left((1+\delta)^T \cdot (1+w_{\min})\right)\right\rfloor_{1+\delta}$ and $(1+\delta)^{\overline{k}} = \left\lfloor r_{\max}/(1+\delta)^T \right\rfloor_{1+\delta}$. The cardinality of Dom^δ is $O\left(\log_{1+\delta}\left(r_{\max}/r_{\min}w_{\min}\right)\right)$. Therefore $\widetilde{\mathcal{S}}^*$ can be derived by solving $\mathcal{V}_{\mathcal{U}_{\text{small}}}(1, R, \emptyset)$ with at most $O\left(\log_{1+\delta}\left(r_{\max}/r_{\min}w_{\min}\right)\right)$ number of enumerations on R, which can be ended in a running time of $O\left(\left(|\mathcal{U}_{\text{small}}| \cdot \log_{1+\delta}\left(r_{max}/r_{min}w_{\min}\right)\right)^2 \cdot T\right) =$ $O\left(|\mathcal{I}|^{O(1)} \cdot \frac{1}{\delta^2} \cdot n^{O\left(\frac{1}{\epsilon^2}\right)}\right)$. Let $\delta = 14\epsilon/(T+1)$, then we can compute a nontrivial $\widetilde{\mathcal{S}}^*$ in a running time of $O\left(|\mathcal{I}|^{O(1)} \cdot n^{O\left(\frac{1}{\epsilon^2}\right)}\right)$ with

$$\mathcal{R}\left(\widetilde{\mathcal{S}}^*\right) \geq \mathcal{R}\left(\mathcal{S}^*\right)/(1+\delta)^{T+1} = \mathcal{R}\left(\mathcal{S}^*\right)/(1+14\epsilon/(T+1))^{T+1} \geq (1-14\epsilon) \cdot \mathcal{R}\left(\mathcal{S}^*\right),$$

and $cs\left(\widetilde{\mathcal{S}}^*\right) \geq \alpha_1$.

C.4 Proofs from Section 3.4

C.4.1 Proof of Theorem 3.4.1:

Let $\mathcal{S}^*$ denote the true optimal solution to DISPLAY-OPT-K. Then following the similar argument as in the proof of Theorem 3.3.1, there exists a proxy assortment $\widetilde{\mathcal{S}} = \left(\widetilde{S}_1, \cdots, \widetilde{S}_T\right) \in \bigtimes_{t\in[T]} \mathcal{U}_{\text{small}}$ such that $\mathcal{R}\left(\widetilde{\mathcal{S}}\right) > (1-13\epsilon) \cdot \mathcal{R}\left(\mathcal{S}^*\right)$ and $w\left(\widetilde{S}_t\right) \geq W_t^*$, $t \in [T]$ which means that $\widetilde{\mathcal{S}}$ is a feasible solution to (3.5) and (3.6). Therefore, (3.5) and (3.6) are guaranteed to generate nontrivial solutions. Again, similar to the proof of Theorem 3.3.1, when $\delta = 14\epsilon/(T+1)$, then we can compute a nontrivial $\widetilde{\mathcal{S}}^*$ in a running time of $O\left(|\mathcal{I}|^{O(1)} \cdot n^{O\left(\frac{1}{\epsilon^2}\right)}\right)$ with $\mathcal{R}\left(\widetilde{\mathcal{S}}^*\right) \geq (1-14\epsilon) \cdot \mathcal{R}\left(\mathcal{S}^*\right)$ and $\widetilde{\mathcal{S}}^* \in \mathcal{F}_K$.

[211]

C.5 Proofs from Section 3.5

C.5.1 Proof of Lemma 3.5.1

Clearly, we have $\hat{S}_1 \subseteq \cdots \subseteq \hat{S}_T$, then we show the efficacy of $\hat{\mathcal{S}}$ in both feasibility and revenue.

$\hat{\mathcal{S}}$ **feasibility.** For any stage t,

$$
\left| \hat{S}_t \setminus \hat{S}_{t-1} \right| = \left| \biguplus_{q \in [\mathcal{Q}]_0} \left(\mathcal{C}_q \left[\hat{k}_{t,q} \right] \setminus \mathcal{C}_q \left[\hat{k}_{t-1,q} \right] \right) \right|
$$

$$
= \sum_{q \in [\mathcal{Q}]_0, k_{t,q}^* > \frac{1}{\epsilon}} \left(\lceil (1-\epsilon) \cdot k_{t,q}^* \rceil - \lceil (1-\epsilon) \cdot k_{t-1,q}^* \rceil \right) + \sum_{q \in [\mathcal{Q}]_0, k_{t,q}^* \leq \frac{1}{\epsilon}} \left(k_{t,q}^* - k_{t-1,q}^* \right)
$$

$$
\leq \sum_{q \in [\mathcal{Q}]_0} k_{t,q}^* - k_{t-1,q}^* = \left| S_t^* \setminus S_{t-1}^* \right| \leq C,
$$

therefore, the capacity constraints are satisfied. On the other hand,

$$
w\left(\hat{S}_t \right) = w \left(\biguplus_{q \in [\mathcal{Q}]_0, k_{t,q}^* \leq \frac{1}{\epsilon}} \mathcal{C}_q \left(k_{t,q}^* \right) \bigcup \biguplus_{q \in [\mathcal{Q}]_0, k_{t,q}^* > \frac{1}{\epsilon}} \mathcal{C}_q \left(\lceil (1-\epsilon) \cdot k_{t,q}^* \rceil \right) \right)
$$

$$
\leq w \left(\biguplus_{q \in [\mathcal{Q}]_0} \mathcal{C}_q \left(k_{t,q}^* \right) \right) \leq (1+\epsilon) \cdot w\left(S_t^* \right)
$$

$$
w\left(\hat{S}_t \right) = w \left(\biguplus_{q \in [\mathcal{Q}]_0, k_{t,q}^* \leq \frac{1}{\epsilon}} \mathcal{C}_q \left(k_{t,q}^* \right) \bigcup \biguplus_{q \in [\mathcal{Q}]_0, k_{t,q}^* > \frac{1}{\epsilon}} \mathcal{C}_q \left(\lceil (1-\epsilon) \cdot k_{t,q}^* \rceil \right) \right)
$$

$$
\geq w \left(\biguplus_{q \in [\mathcal{Q}]_0} \mathcal{C}_q \left(\lceil (1-\epsilon) \cdot k_{t,q}^* \rceil \right) \right) \geq (1-\epsilon)^2 \cdot w \left(\biguplus_{q \in [\mathcal{Q}]_0} S_{t,q}^* \right) > (1-2\epsilon) \cdot w\left(S_t^* \right).
$$

[212]

Therefore, for $k \in [K]$,

$$\sum_{t \in [t_k : t_{k+1})} \lambda_t \cdot cs\left(w\left(\hat{S}_t\right)\right) > \sum_{t \in [t_k : t_{k+1})} \lambda_t \cdot cs\left((1 - 2\epsilon) \cdot w\left(S_t^*\right)\right)$$

$$> (1 - 2\epsilon) \cdot \sum_{t \in [t_k : t_{k+1})} \lambda_t \cdot cs\left(w\left(S_t^*\right)\right) \geq (1 - 2\epsilon) \cdot \alpha_k,$$

where the second inequality is due to Assumption 3.1.2. So the customer satisfaction constraint for each customer group is violated by a factor of $(1 - 2\epsilon)$.

$\hat{\mathcal{S}}$ **revenue.**

$$\mathcal{R}\left(\hat{\mathcal{S}}\right) = \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + w\left(\hat{S}_t\right)} \cdot \sum_{i \in \hat{S}_t} \rho_i \geq \sum_{t \in [T]} \lambda_t \cdot \frac{1}{1 + (1 + \epsilon) \cdot w\left(S_t^*\right)} \cdot (1 - \epsilon) \cdot \sum_{i \in S_t^*} \rho_i$$

$$> \frac{1 - \epsilon}{1 + \epsilon} \cdot \mathcal{R}\left(\mathcal{S}^*\right) > (1 - 2\epsilon) \cdot \mathcal{R}\left(\mathcal{S}^*\right),$$

where in the first inequality, $w\left(\hat{S}_t\right) < (1 + \epsilon) \cdot w\left(S_t^*\right)$ comes from the result that we just derived above, and $\sum_{i \in \hat{S}_t} \rho_i \geq (1 - \epsilon) \cdot \sum_{i \in S_t^*} \rho_i$ comes from the fact that in each class q, $\hat{S}_{t,q}$ is picked following the ρ =order and $|\hat{S}_{t,q}| \geq (1 - \epsilon) \cdot |S_{t,q}^*|$.

C.5.2 Proof of Lemma 3.5.2

$\hat{\mathcal{S}}^{\text{small}}$ **feasibility.** It is easy to see that $\hat{S}_1^{\text{small}} \subseteq \cdots \subseteq \hat{S}_T^{\text{small}}$. We then argue that the cardinality constraints are also satisfied. For any stage t and class q, we claim $k_{tq}^{\downarrow} \leq k_{tq}$. According to the definitions of $k_{tq}^{\downarrow}$ and k_{tq}, the claim is only not straightforward for class

$q \geq [q_{\max}(K_t) - L + 1]^+$ such that $k_{tq} > \frac{1}{\epsilon}$. However, since $w\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)[k_{tq}]\right) = w\left(\hat{S}_{t,q} \setminus \hat{S}_{t-1,q}\right) \geq w_q^{\downarrow}$, we have $k_{tq}^{\downarrow} \leq k_{tq}$. Therefore

$$
\left| \hat{S}_t^{\text{small}} \setminus \hat{S}_{t-1}^{\text{small}} \right| = \left| \biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \left(\hat{S}_{t,q}^{\text{small}} \setminus \hat{S}_{t-1,q}^{\text{small}} \right) \right| + \left| \biguplus_{q=1}^{[q_{\max}(K_t)-L]^+} \left(\hat{S}_{t,q}^{\text{small}} \setminus \hat{S}_{t-1,q}^{\text{small}} \right) \right|
$$

$$
= \sum_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} k_{tq}^{\downarrow} + \sum_{\tau=1}^{t} \left| \biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\hat{S}_{t,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}} \right) \setminus \left(\hat{S}_{t-1,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}} \right) \right|
$$

$$
= \sum_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} k_{tq}^{\downarrow} + \sum_{\tau=1}^{t} l_{t\tau} - \sum_{\tau=1}^{t-1} l_{(t-1)\tau}
$$

$$
\leq \sum_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} k_{tq} + \sum_{\tau=1}^{t} \left| \biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\hat{S}_{t,q} \setminus \hat{S}_{\tau-1,q} \right) \right|
$$

$$
- \sum_{\tau=1}^{t-1} \left| \biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\hat{S}_{t-1,q} \setminus \hat{S}_{\tau-1,q} \right) \right|
$$

$$
= \left| \biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \left(\hat{S}_{t,q} \setminus \hat{S}_{t-1,q} \right) \right|
$$

$$
+ \sum_{\tau=1}^{t} \left| \biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\hat{S}_{t,q} \setminus \hat{S}_{\tau-1,q} \right) \setminus \left(\hat{S}_{t-1,q} \setminus \hat{S}_{\tau-1,q} \right) \right|
$$

$$
= \left| \hat{S}_t \setminus \hat{S}_{t-1} \right| \leq C.
$$

Therefore, $\hat{S}^{\text{small}}$ satisfies the cardinality constraints. Then we check the customer satisfaction constraints.

$$
w\left(\hat{S}_t^{\text{small}}\right) \geq w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \hat{S}_{t,q}^{\text{small}}\right) = w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t}\left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right)
$$

$$
= w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t}\left(\mathcal{C}_q \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\left[k_{\tau,q}^{\downarrow}\right]\right)
$$

$$
\geq \frac{1}{1+\epsilon} \cdot w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t}\left(\mathcal{C}_q \setminus \hat{S}_{\tau-1,q}\right)\left[k_{\tau,q}^{\downarrow}\right]\right)
$$

$$
\geq \frac{1}{1+\epsilon} \cdot \left(w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t}\left(\hat{S}_{\tau,q} \setminus \hat{S}_{\tau-1,q}\right)\right)\right.
$$

$$
\left. - \sum_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \sum_{\tau=1}^{t} \text{Power}_2\left[\frac{\epsilon}{L} \cdot w(\hat{S}_\tau \setminus \hat{S}_{\tau-1})\right]\right)
$$

$$
> (1-\epsilon) \cdot \left(w\left(\hat{S}_t\right) - 2\epsilon \cdot w\left(\hat{S}_t\right)\right) > (1-3\epsilon) \cdot w\left(\hat{S}_t\right),
$$

where the first inequality is due to the fact that within each weight class q, the real weights won't differentiate from each other by a factor of $(1+\epsilon)$, and the second inequality comes from the definition of $k_{\tau,q}^{\downarrow}$. Therefore,

$$
\sum_{t\in[t_k:t_{k+1})} \lambda_t \cdot cs\left(w\left(\hat{S}_t^{\text{small}}\right)\right) > \sum_{t\in[t_k:t_{k+1})} \lambda_t \cdot cs\left((1-3\epsilon) \cdot w\left(\hat{S}_t\right)\right)
$$

$$
> (1-3\epsilon) \cdot \sum_{t\in[t_k:t_{k+1})} \lambda_t \cdot cs\left(w\left(\hat{S}_t\right)\right) \geq (1-5\epsilon) \cdot \alpha_k,
$$

[215]

where the second inequality comes from Assumption 3.1.2, and the last inequality comes from Lemma 3.5.1. We can also give an upper bound on $w\left(\hat{S}_t^{\text{small}}\right)$.

$$
\begin{aligned}
w\left(\hat{S}_t^{\text{small}}\right) &\leq w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \hat{S}_{t,q}^{\text{small}}\right) + w\left(\biguplus_{q=0}^{[q_{\max}(K_t)-L]^+} \mathcal{C}_q\right) \\
&\leq (1+\epsilon)\cdot w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \hat{S}_{t,q}^{\text{small}}\right) \leq (1+\epsilon)\cdot w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \hat{S}_{t,q}\right) \\
&\leq (1+\epsilon)\cdot w\left(\hat{S}_t\right),
\end{aligned}
$$

where the second inequality comes from the definition of L, and the third inequality is due to the fact that for $q \geq [q_{\max}(K_t) - L + 1]^+$, $\hat{S}_{t,q}^{\text{small}} \subseteq \hat{S}_{t,q}$. Then according to Lemma 3.5.1, we hvae $w\left(\hat{S}_t^{\text{small}}\right) \leq (1+2\epsilon)\cdot w(S_t^*)$.

$\hat{\mathcal{S}}^{\text{small}}$ **revenue.** For $q > [q_{\max}(K_t) - L + 1]^+$, according to the definition of $k_{tq}^{\downarrow}$, when $k_{tq} \leq \frac{1}{\epsilon}$, $k_{tq}^{\downarrow} = k_{tq}$. What is more complicated is when $k_{tq} > \frac{1}{\epsilon}$, we have

$$
0 < w\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)[k_{tq}]\right) - w\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)\left[k_{tq}^{\downarrow}\right]\right) < \text{Power}_2\left[\frac{\epsilon}{L}\cdot w(\hat{S}_t \setminus \hat{S}_{t-1})\right] \quad \text{(C.4)}
$$

$$
0 < \rho\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)[k_{tq}]\right) - \rho\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)\left[k_{tq}^{\downarrow}\right]\right) < \text{Power}_2\left[\frac{\epsilon}{L}\cdot \rho(\hat{S}_t \setminus \hat{S}_{t-1})\right], \quad \text{(C.5)}
$$

where $\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)[k_{tq}] = \hat{S}_{t,q} \setminus \hat{S}_{t-1,q}$.

Therefore

$$
\begin{aligned}
w\left(\hat{S}_t^{\text{small}}\right) &= w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \hat{S}_{t,q}^{\text{small}}\right) + w\left(\biguplus_{q=1}^{[q_{\max}(K_t)-L]^+} \hat{S}_{t,q}^{\text{small}}\right) \\
&\leq w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t} \left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right) + n \cdot w_{\min} \cdot (1+\epsilon)^{q_{\max}(K_t)-L} \\
&\leq \left(1 + n \cdot (1+\epsilon)^{-L}\right) \cdot w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t} \left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right) \\
&= (1+\epsilon) \cdot w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t} \left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right) \\
&\leq (1+\epsilon) \cdot \left(w\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t} \left(\hat{S}_{\tau,q} \setminus \hat{S}_{\tau-1,q}\right)\right)\right. \\
&\quad \left. + \sum_{\tau=1}^{t} \sum_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \text{Power}_2\left[\frac{\epsilon}{L} \cdot w(\hat{S}_\tau \setminus \hat{S}_{\tau-1})\right]\right) \\
&\leq (1+\epsilon) \cdot \left(w\left(\hat{S}_t\right) + \sum_{\tau=1}^{t} 2\epsilon \cdot w\left(\hat{S}_\tau \setminus \hat{S}_{\tau-1}\right)\right) \\
&= (1+\epsilon)(1+2\epsilon) \cdot w\left(\hat{S}_t\right) < (1+4\epsilon) \cdot w\left(\hat{S}_t\right).
\end{aligned}
$$

And on the other hand, since for $q > [q_{\max}(K_t) - L + 1]^+$, $k_{\tau q}^{\downarrow} \leq k_{\tau q}$, $\tau \in [t]$, we have $\hat{S}_{t-1,q}^{\text{small}} \subseteq \hat{S}_{t-1,q}$. Notice that $w\left(\hat{S}_{t,q}^{\text{small}} \setminus \hat{S}_{t-1,q}^{\text{small}}\right) = w\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}^{\text{small}}\right)\left[k_{tq}^{\downarrow}\right]\right)$, then $\rho\left(\hat{S}_{t,q}^{\text{small}} \setminus \hat{S}_{t-1,q}^{\text{small}}\right) \geq \rho\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)\left[k_{tq}^{\downarrow}\right]\right)$, where the inequality is due to the fact that for each q, $\mathcal{C}_q$ is ranked by ρ−order. Combining with (C.5), we have

$$
\begin{aligned}
\rho\left(\hat{S}_{t,q}^{\text{small}} \setminus \hat{S}_{t-1,q}^{\text{small}}\right) &\geq \rho\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)\left[k_{tq}^{\downarrow}\right]\right) \\
&> \rho\left(\left(\mathcal{C}_q \setminus \hat{S}_{t-1,q}\right)[k_{tq}]\right) - \text{Power}_2\left[\frac{\epsilon}{L} \cdot \rho(\hat{S}_t \setminus \hat{S}_{t-1})\right] \\
&= \rho\left(\hat{S}_{t,q} \setminus \hat{S}_{t-1,q}\right) - \text{Power}_2\left[\frac{\epsilon}{L} \cdot \rho(\hat{S}_t \setminus \hat{S}_{t-1})\right]. \quad\quad \text{(C.6)}
\end{aligned}
$$

[217]

The above inequalities imply the following,

$$
\rho\left(\hat{S}_t^{\text{small}}\right) = \rho\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t} \left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right) + \rho\left(\biguplus_{\tau=1}^{t} \biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \hat{S}_{t,q}^{\text{small}}\right)
$$

$$
= \rho\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t} \left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right)
$$

$$
+ \sum_{\tau=1}^{t} \rho\left(\biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\hat{S}_{t,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right) \cup \hat{S}_{\tau-1,q}^{\text{small}}\right)
$$

$$
= \rho\left(\biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \biguplus_{\tau=1}^{t} \left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right)
$$

$$
+ \sum_{\tau=1}^{t} \left\{ \rho\left(\left(\biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\mathcal{C}_q \setminus \mathcal{C}_q\left[k_{\tau-1,q}^\downarrow\right]\right)\right)[l_{t\tau}]\right)\right.
$$

$$
\left. + \rho\left(\biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \biguplus_{i=1}^{\tau-1} \left(\hat{S}_{i,q}^{\text{small}} \setminus \hat{S}_{i-1,q}^{\text{small}}\right)\right)\right\}
$$

$$
\geq \rho\left(\biguplus_{\tau=1}^{t} \biguplus_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \left(\hat{S}_{\tau,q}^{\text{small}} \setminus \hat{S}_{\tau-1,q}^{\text{small}}\right)\right)
$$

$$
+ \sum_{\tau=1}^{t} \left\{ \rho\left(\left(\biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\mathcal{C}_q \setminus \mathcal{C}_q\left[k_{\tau-1,q}\right]\right)\right)[l_{t\tau}]\right)\right.
$$

$$
\left. + \rho\left(\biguplus_{i=1}^{\tau-1} \biguplus_{q=[q_{\max}(K_{\tau-1})-L+1]^+}^{[q_{\max}(K_\tau)-L]^+} \left(\hat{S}_{i,q}^{\text{small}} \setminus \hat{S}_{i-1,q}^{\text{small}}\right)\right)\right\} \tag{C.7}
$$

[218]

where the first inequality is due to the fact that $k^{\downarrow}_{\tau-1,q} \leq k^{\downarrow}_{\tau-1,q}$ and each class $\mathcal{C}_q$ is $\rho-$ordered.

Then

$$
\begin{aligned}
(\text{C.7}) \geq & \sum_{\tau=1}^{t} \left(\rho \left(\underset{q=[q_{\max}(K_t)-L+1]^+}{\overset{q_{\max}(K_t)}{\biguplus}} \left(\hat{S}_{\tau,q} \setminus \hat{S}_{\tau-1,q} \right) \right) \right. \\
& \left. - \sum_{q=[q_{\max}(K_t)-L+1]^+}^{q_{\max}(K_t)} \text{Power}_2 \left[\frac{\epsilon}{L} \cdot \rho(\hat{S}_\tau \setminus \hat{S}_{\tau-1}) \right] \right) \\
& + \sum_{\tau=1}^{t} \left\{ \rho \left(\left(\underset{q=[q_{\max}(K_{\tau-1})-L+1]^+}{\overset{[q_{\max}(K_\tau)-L]^+}{\biguplus}} \left(\mathcal{C}_q \setminus \mathcal{C}_q \left[k_{\tau-1,q} \right] \right) \right) \left[l_{t\tau} \right] \right) \right. \\
& + \sum_{i=1}^{\tau-1} \left(\rho \left(\underset{q=[q_{\max}(K_{\tau-1})-L+1]^+}{\overset{[q_{\max}(K_\tau)-L]^+}{\biguplus}} \left(\hat{S}_{i,q} \setminus \hat{S}_{i-1,q} \right) \right) \right. \\
& \left. \left. - \sum_{q=[q_{\max}(K_\tau)-L+1]^+}^{q_{\max}(K_\tau)} \text{Power}_2 \left[\frac{\epsilon}{L} \cdot \rho(\hat{S}_i \setminus \hat{S}_{i-1}) \right] \right) \right\} \\
\geq & \rho \left(\underset{q=[q_{\max}(K_t)-L+1]^+}{\overset{q_{\max}(K_t)}{\biguplus}} \hat{S}_{t,q} \right) - 22\epsilon \cdot \rho \left(\hat{S}_t \right) \\
& + \sum_{\tau=1}^{t} \left\{ \rho \left(\underset{q=[q_{\max}(K_{\tau-1})-L+1]^+}{\overset{[q_{\max}(K_\tau)-L]^+}{\biguplus}} \left(\hat{S}_{t,q} \setminus \hat{S}_{\tau-1,q} \right) \right) \right. \\
& \left. + \rho \left(\underset{q=[q_{\max}(K_{\tau-1})-L+1]^+}{\overset{[q_{\max}(K_\tau)-L]^+}{\biguplus}} \hat{S}_{\tau-1,q} \right) - 2\epsilon \cdot \rho \left(\hat{S}_{\tau-1} \right) \right\} \\
= & \rho \left(\underset{q=[q_{\max}(K_t)-L+1]^+}{\overset{q_{\max}(K_t)}{\biguplus}} \hat{S}_{t,q} \right) - 2\epsilon \cdot \rho \left(\hat{S}_t \right) + \rho \left(\underset{q=1}{\overset{[q_{\max}(K_\tau)-L]^+}{\biguplus}} \hat{S}_{t,q} \right) - 2\epsilon \cdot \sum_{\tau=1}^{t-1} \rho \left(\hat{S}_\tau \right) \\
\geq & \left(1 - 2T \cdot \epsilon \right) \cdot \rho \left(\hat{S}_t \right).
\end{aligned}
$$

[219]

Therefore, we have

$$
\begin{aligned}
\mathcal{R}\left(\hat{\mathcal{S}}^{\text{small}}\right) = \sum_{t \in [T]} \lambda_t \cdot \frac{\rho\left(\hat{S}_t^{\text{small}}\right)}{1 + w\left(\hat{S}_t^{\text{small}}\right)} \cdot &\geq \sum_{t \in [T]} \lambda_t \cdot \frac{(1 - 2T \cdot \epsilon) \cdot \rho\left(\hat{S}_t\right)}{1 + (1 + 4\epsilon) \cdot w\left(\hat{S}_t\right)} \\
> \frac{1 - 2T \cdot \epsilon}{1 + 4\epsilon} \cdot \mathcal{R}\left(\hat{\mathcal{S}}\right) &> \frac{1 - 2T \cdot \epsilon}{1 + 4\epsilon} \cdot (1 - 2\epsilon) \cdot \mathcal{R}\left(\mathcal{S}^*\right) \\
> (1 - 2\,(T + 3) \cdot \epsilon) \cdot &\,\mathcal{R}\left(\mathcal{S}^*\right).
\end{aligned}
$$

C.5.3 Proof of Theorem 3.5.3:

The proof follows closely as Theorem 3.4.1.

[220]