

國立臺灣大學管理學院資訊管理學系

碩士論文

Department of Information Management

College of Management

National Taiwan University

Master Thesis



舞臺劇產業中銷售資訊對銷售速率的影響：

藉由台灣劇團資料之實證研究

Impact of Performance Ticket Sales Information on Sales
Speed: An Empirical Study Based on a Taiwanese Theater

孫珮瑜

Pei-Yu Sun

指導教授：孔令傑 博士

Adviser: Ling-Chieh Kung, Ph.D.

中華民國 106 年 7 月

July, 2017

國立臺灣大學碩士學位論文
口試委員會審定書



Impact of Performance Ticket Sales Information on
Sales Speed: An Empirical Study Based on a
Taiwanese Theater

本論文係孫珮瑜君（學號 R04725009）在國立臺灣大學資訊管理學系、所完成之碩士學位論文，於民國 106 年 6 月 28 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

孔令儀

張永寧


莊能鈞

柯貞如

所長：

李益坤

謝辭



碩士班這兩年能夠加入資訊經濟與決策最佳化實驗室中的的大家庭，並得到熱誠的孔令傑老師的指教。這段時間中，老師一直以亦師亦友的態度，不斷地給予我機會去審思自己並提點不足的地方，提供我許多進步、發展的空間。除了課堂中的專業知識，如作業研究、資訊經濟等，老師更鼓勵我們不停突破自己的舒適圈，挑戰自己的極限。在您的帶領之下，碩一的時候，我就在決策分析研討會上得到最佳論文獎並在 PACIS 國際會議上用英文向台下的學者們講述我們的研究。碩二時開始撰寫論文，在期間老師指引我發想題目、提點架構與寫作方式、並逐字審閱字句。專案部分也讓我參與神腦國際與 Coursera 的資料分析專案，讓我實作如何清理、分析、並呈現資料。另外，我很感激老師您選擇我來當服務學習三的助教。藉由帶領著大學生推廣程式教育，我深刻瞭解回饋社會不僅是影響他人更是帶給自己快樂並適時地替自己充電。此外也特別感謝我的口試委員，林真如老師、莊皓鈞老師與裘家寧老師，在口試時給予我許多建議與肯定。

當然，我要謝謝無時無刻都給予我支持並鼓勵我時刻充實、突破自我的家人。感謝冠宇學長、何禾學姊、偉宏學長與騏瑋學長，幫我指引出碩班研究之路的明燈。感謝同屆的實驗室夥伴們，在研究的路上互相扶持和成長。謝謝雪兒、韋志和宸安在神腦專案時的合作與互助。謝謝維哲和千瑜在碩一時一起征服 PACIS 研討會。謝謝韋志、柏宣和宸安的帶領，讓日子過得快樂又有深度。感謝學弟妹敬傑、佩蓉、子翔和鑑霖，每次的互動都讓我有學習。特別感謝一路走來心靈夥伴，維哲。謝謝你這些年的陪伴與鼓勵，有你我的世界多了許多色彩。未來我們也將一同互助學習，一同成長。

孫珮瑜 謹致

于臺大資訊管理研究所

民國一百零六年七月

中文摘要

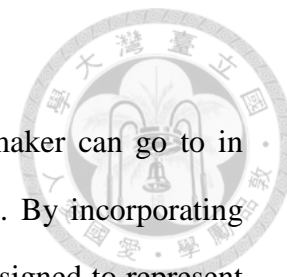


本研究資料透過銷售資訊的揭露，期望能幫助決策者最大化獲利。我們藉由迴歸模型分析一個台灣劇團的售票資料，研究消費者購票瞬間觀察到的各票種售票狀況如何影響消費者購買意願，並探討決策者持有此資訊後，能如何透過動態調整各票種的可售數量來最大化預期營收。我們發現票種的銷售量與銷售速度成正相關，但是和其他票種的累積銷售量呈負相關。這樣的現象尤其在票種之間的票價差距越大的時候越趨明顯。我們的研究也證實銷售速度會因為銷售期間的增加而降低。

透過需求預測模型，我們可以根據當前的銷售來計算並調整不同票種最適座位數量。我們也考量了動態決策的情境，使決策者能動態分配不同票種的座位數量，並最佳化獲利。在我們的數值實驗中，透過銷售資訊的揭露，本研究發現考慮銷售狀況對銷售速度影響的動態座位數量調整能使獲利提高 4.60%。

關鍵字：票券銷售、銷售資訊揭露、購買意願不確定性、動態座位分配、迴歸分析

Thesis Abstract



The data in this study highlight the lengths that a decision maker can go to in order to maximize its profit through sales information disclosure. By incorporating sales information of price bands into a single regression model, designed to represent the example of Taiwanese theater ticket theoretical framework formalizes how decision makers are able to affect customers' valuation. It is verified that sales of a price band has a positive relation to its sales speed while the cumulated sales of other price bands have negative effects. The relationships intensify when the price of the price bands are closer to one another. Our analyses further support that the sales speed decreases when the length of selling period escalates.

The estimated demand system allows for the calculation of capacity allocation for each price bands based on its current sales. A scenario in which decision makers can dynamically allocate ticket capacities of various price bands of the same performance to optimize profit is investigated. The profit difference between the model with the sales information disclosure is 4.6% higher than the suboptimal model.

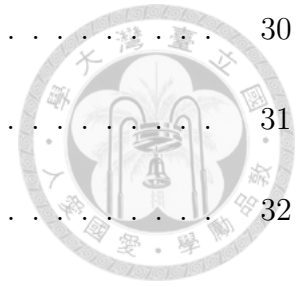
Keywords: Ticket Sales, Sales Information Disclosure, Valuation Uncertainty, Dynamic Seat Allocation, Regression Analysis



Contents

1	Introduction	3
1.1	Background and motivation	3
1.2	Research objectives	6
1.3	Research plan	7
2	Literature review	9
2.1	Perishable products	9
2.2	Sales information disclosure	11
2.3	Ticket selling	12
3	Problem Description and Formulation	17
4	Analysis	23
4.1	Data cleansing	23
4.2	An illustrative example of data processing	26
4.2.1	Testing Hypothesis 1 to 3	27

4.2.2	Testing Hypothesis 4	30
4.3	Exploratory data analysis	31
4.4	Regression analysis	32
4.4.1	Testing Hypothesis 1 to 3	33
4.4.2	Testing Hypothesis 4	39
5	Application: Dynamic Seat Allocation	43
5.1	Exploratory data analysis	44
5.2	Numerical experiment	46
5.3	Quantity allocation based on Lagrangian relaxation	50
6	Conclusions and Future Works	55
6.1	Conclusions	55
6.2	Future works	56
A	Figures of Exploratory Data Analysis	59
	Bibliography	63





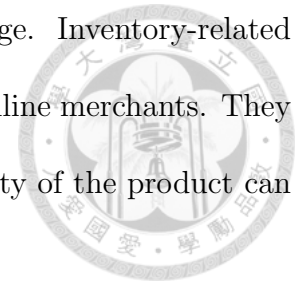
Chapter 1

Introduction

1.1 Background and motivation

Online environment allows merchants to play various inventory-related selling tactics by deciding whether to divulge certain information to customers. As a top online auction platform, eBay shows on the products' detailed page the amount sold, the number of views per hour, and the quantity left when the products have high selling speed. China's market has a more visible mechanism. Both Taobao and JD.com, two dominant online marketplaces in China, disclose inventory information directly on advertisements and on search result lists using a bar chart in bright red showing the percentage of products sold versus inventory left. The phenomenon is also found on hotel booking websites. Hotels.com reveals the information that hotels were booked several hours ago beside the hotel names on the search result list. As for Booking.com, it shows on the hotel's detailed page that whether the room is in high demand. Group buying websites like Groupon, GOMAJI, and LivingSocial especially utilize the inventory-disclosure mechanism. Peo-

ple can see the amount of products sold almost on every web page. Inventory-related information disclosure tactics seem to be widely used by existing online merchants. They seem to believe that disclosing related information on the popularity of the product can trigger higher sales.



Academic evidence of the above practice were provided in several streams of literature. Bikhchandani et al. (1992) define that an informational cascade occurs when an individual, having observed the actions of those ahead of him, finds it optimal to follow the behavior of the preceding individual without regard to his own information. Watts (2002) studies a network setting in which an agent is able to affect its neighbors only. It is found that limited numbers of initial shocks have the potential to trigger a global network's information cascade. Herding and information cascade might also occur among professional analysts (Alevy et al., 2007; Welch, 2000). Ho et al. (2014) state that sales velocity indicates the popularity of the product. Showing high sales velocity would have a positive effect on purchase likelihood of customers that it even reverses the participants' preferences for a product with a higher sales rank.

Both literature and practical applications have the information that sales outcome has effects on customers' purchase decisions; however, limited literature on performance ticket selling discuss this issue. They normally assume people know their valuation (Courty, 2002; Leslie, 2004; Rascher et al., 2007) and only price change affect customer utility (Drayer et al., 2012; McAfee and Te Velde, 2006; You, 1999). As for the business aspect, we can see from the websites of ticket selling like Ticketmaster and UDN ticket selling that websites disclose cumulated sales outcome. Even though some do not directly show any number related to inventory on the detailed product page until customers click into

the show and choose the specific date of the performance, the websites need to disclose cumulated sales outcome during the process of seat selection. Normally ticket websites show the information by disclosing a seat map with different color as background indicating various price bands and seats with red color are sold seats. Customers can directly understand the sales data and percentage of all price bands at a glance.

It is intriguing that people of both academic and business fields in performance ticket pricing have limited research whether people will be influenced by other people under valuation. Theater performance is a type of product with high valuation. The quality of performance might be affected by the location, seat, casts, actors' emotion, and the entire crew. In comparison to the uncertainty people might obtain from general online product purchase, performance ticket buying should have higher risks. Customers might be more influenced by other people's opinions; therefore, cumulated sales outcome might play a role in affecting the sales.

Performance ticket purchase has some other interesting characteristics that are different from most general products selling. First, performance tickets are perishable and have no salvage value. Second, there are multiple price bands in a show. Customers have different valuations for each price band based on their quality and customer preference. Lastly, there is no inventory replenishment in performance. People might buy the tickets of higher or lower price when their most desirable price is sold out.

1.2 Research objectives



In this study, we investigate the relationship between ticket sales outcome and customer valuation and how the relationship affects the profit of a ticket seller. We obtain a performance data set spanned from 2008 to 2012 of a local theater group in Taiwan. They had been actively performing and producing innovative and educational kids' drama for sixteen years. The ticket selling data set contains detailed information of ticket sales. From the data set, we obtain the information of the sales volumes of all price bands for each performance in each day.

Aforementioned literature shows that customers' valuation may be affected by sales outcome. Utilizing the data and regression model, how essential inventory information disclosure is to performance ticket selling can be estimated. In addition to the cumulated sales outcome data, we also add other related performance information like location and selling period into the model and examine the interaction effect between basic performance information and cumulated sales. The relationships of the target price band and other price bands are also investigated.

In addition to proving the relation between sales speed and cumulated sales data, an operational application is considered. In a performance, each price band has its own capacity and different price bands have various cumulated sales. We study a scenario in which allocated seats of a price band can be transferred to another price band while considering the potential influence on customers' willingness to buy. The capacity of the price band that is not selling so well can be transferred to one that is with better sales performance. With quantity control, we examine how the additional information about

the relationship between sales outcome and customers' valuation may contribute to a ticket seller's profit. .



1.3 Research plan

The remainder of the thesis is organized as follows. In Chapter 2, we review some related works about perishable products, inventory disclosure, and ticket pricing. In Chapter 3, a proposed regression model describes the relationship of customer utility between sales speed. Chapter 4 shows the analysis based on the proposed models. An application based on our research finding is examined in Chapter 5. Chapter 6 concludes.





Chapter 2

Literature review

Our research discusses the relationship of ticket cumulated sales to its sales speed; therefore, we focus on three streams of literature: perishable product, sales information disclosure, and ticket selling.

2.1 Perishable products

Many inventory models assume that stock items can be stored indefinitely to meet future demands. However, certain types of inventories undergo change in storage so that in time they may become partially or entirely unfit for consumption (Nahmias, 1982). Those assets are called perishable products. Some examples are fresh food, fashion or seasonal goods, and travel-related products like hotels or air tickets. Theater tickets, of course, is a kind of perishable good.

Retailers offer discounts on perishable products nearing their shelf-lives to encourage consumers to buy under the assumption that customers are less willing to purchase per-

ishable products when the expiration date is near. The assumption is verified by Tsiros and Heilman (2005). They conduct a survey of 300 customers of a grocery shop and found that the willingness-to-pay of the customers decreases throughout the course of the product's shelf life. In situations of discount offering for perishable goods, Sezen (2004) uses expected profit approach to identify the timing and quantum of discount for perishable commodities. Ramanathan (2006) extends the expected profit approach presented in Sezen (2004) by including decisions on the quantities of the perishable products to be stocked by the retailer. Zhao and Zheng (2000) conduct numerical studies to show the revenue impact of dynamic pricing on perishable assets. Their examples show that using optimal dynamic optimal policies achieves 2.4 – 7.3% revenue improvement over the optimal single price policy. Price changes become even more critical when the reservation price distribution shifts over time. In the case, the revenue increase can be as high as 100%. These results explain why yield management has become so essential to fashion retailing and travel service industries.

Weatherford and Bodily (1992) propose Perishable-Asset Revenue Management (PARM) and define it as optimal revenue management of perishable assets through price segmentation. Fourteen taxonomies are investigated such as discount price classes, group reservations, and overbooking. The fourteen taxonomies are mostly price-related; however, the price for theater are fixed in many circumstances. As Gallego and Van Ryzin (1994) state, in reality, fixed-priced policies appears to be more appealing for dynamic pricing needs constant adjustment and is undesirable for practical application. The option of dynamic pricing is thus eliminated in our situation. We focus on whether information disclosure of cumulated sales will affect the customers' willingness to pay for various price

bands and leads to modification of sales speed.



2.2 Sales information disclosure

Aforementioned, we observe several e-commerce platforms utilizing sales information disclosure method to urge customers' purchase desire. Academic evidence are also provided by several streams of literature. We know that people will be affected by others' decisions, especially when entering a new arena of products. Word of mouth (WOM) is a classic example. It has been proven that WOM influences people's short-term (Herr et al., 1991; Bone, 1995; Anderson, 1998) and long-term judgements (Bone, 1995). A WOM data, collected from the Yahoo! movies website, show that WOM information offers significant explanatory power for both aggregate and weekly box office revenue, especially in the early weeks after a movie is released (Liu, 2006).

The situation is the same when disclosing sales information. Bikhchandani et al. (1992) define information cascade as an action when a person follows the behavior of the preceding individual without considering his own information. The cascades can explain not only conformity but the rapid spread of new behaviors and the new behaviors might be positive or negative (Liang et al., 2014). The early demand realization attracts more customers when it is good; however, when the early demand realization is bad, it works in the opposite way to depress future demand. An experiment with 870 participants shows the same result (Tseng, 2016).

Chen et al. (2011) combine the effects of WOM and observational learning, which is the effect that customers are influenced by others' actual purchase decisions. An intriguing

ing finding is that, while negative WOM is more influential than positive WOM, positive observational learning information significantly increases sales but negative observational learning information has no effect. The finding conflicts with that of Liang et al. (2014). Chen et al. (2011) argue that consumers might think that other people might have different individual preferences; therefore, bad sales does not necessarily indicate low quality.

In sum, sales outcome reveals the popularity of the product and the power of information disclosure of accumulated sales outcome seems to influence customers' decision making. The effect can be strong. Ho et al. (2014) show that the positive effect of disclosing high sales velocity on purchase likelihood of customers can be so strong that it reverses the participants' preferences for a product with a higher sales rank.

2.3 Ticket selling

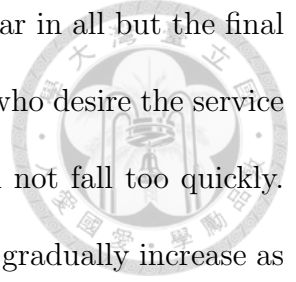
Ticket selling owns the prerequisite circumstances identified by Kimes (1989). There are six prerequisites: (1) The ability to segment markets – By separating consumers into different groups, managers can have different marketing strategies and varying prices across groups. (2) Perishable inventory – This is a common issue faced by several facets of the tourism and hospitality industries (hotels/motels, airlines, and car rental agencies). (3) Product sold in advance – This issue deals specifically with time and uncertainty of sales. (4) Low marginal sales costs – Servicing additional customers will not cost the firm a large amount of money. (5) High marginal production costs – Creating additional inventory is difficult. For example, it is practically impossible for a hotel to quickly add rooms or an airplane to quickly add seats. Therefore, as inventory runs low, managers may

have the opportunity to increase prices. (6) Fluctuating demand – The hotel industry, in particular, experiences frequent fluctuations according to season and day of the week. Tickets for airlines, hotels, sports, and performance are all included.



Many past studies examine the relationship between price policy and revenue potential. Using censored regression and elasticity analysis, Rascher et al. (2007) demonstrates that variable pricing would have yielded approximately \$590,000 additional ticket revenue for each major league team in 1996. Huntington (1993) indicates that theaters offering a range of ticket prices tend to return a higher box office revenue compared to those theaters that only offered tickets at a single price. Boyd and Boyd (1998) suggest that whenever secondary market sellers can resell tickets for a profit, it indicates that tickets are not priced optimally. Conversely, there are other events with high numbers of unsold seats, which indicates the tickets are priced too high (Howard and Crompton, 2004). Shapiro and Drayer (2012) find that it is essential to assess demand-based pricing to see if dynamic ticket pricing policy is appropriate to fit for every professional league because each team has a different market of consumers with varying perceptions of value and willingness to pay for tickets. They obtain the data of San Francisco Giants and find that ticket prices are significantly higher through dynamic ticket pricing compared with the fixed season ticket price. Even though these prices were still lower than secondary market prices for comparable tickets, the willingness to pay for the same tickets are better captured using dynamic pricing strategies.

Time's effect on ticket pricing is also investigated. People hold different opinions on how time variable affects pricing. Rosen and Rosenfield (1997) assume that those attending later performances are paid to wait, getting their tickets at lower prices. A



policy of declining prices over time allows the waiting market to clear in all but the final period. Different preference groups are served sequentially. People who desire the service the most are willing to buy early if they expect that the price will not fall too quickly. As for the case of sports tickets, the average dynamic-priced tickets gradually increase as the game drew closer (Shapiro and Drayer, 2012). Fans choosing to attend a game at the last minute are asked to pay a premium for that convenience. This phenomenon seems to target the customers Courty (2003) considers to be the typical price insensitive, last-minute buyer or what he calls the “busy professional”. Courty (2003) suggests that some people simply appreciate the convenience of having flexibility with their social calendar and are willing to pay a premium for that. He also categorizes another groups of people to be “diehard fans”. Those are the people who plan their schedule in advance and is not willing to pay anything if they cannot commit earlier. He suggests that the number of diehard fans is greater than busy professionals so it is optimal for ticket provider to focus on diehard fans and leave busy professionals to brokers.

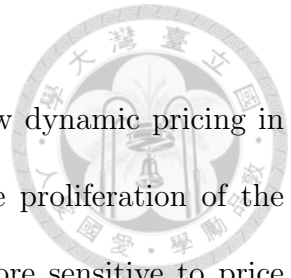
Even though many articles support that dynamic pricing helps companies gain profit, research indicates that consumers will respond negatively to paying different prices for the same product (Hall and Hitch, 1939; Kung et al., 2002). Furthermore, with the proliferation of Internet, it is easier for customers to search for the product of their choice within their desired price range (Kung et al., 2002). In some established industries, such as auto retailing and term insurance, consumers pay lower prices as a result of the Internet. Car buyers use the Internet to gather information and borrow the negotiating clout of an online buying service. The Internet has increased consumers’ sensitivity to price and changes in price (Kotler, 2009).

In the case of theater, we believe the effect of negative response of customers to the price change is more significant. In reality, theater tickets are mainly sold through the Internet. It is extremely easy for customers to check the price. In addition, the existing sales method of most performance producers is that ticket prices are set from the beginning. They maintain consistency of prices between seat locations. Theater suppliers can only change price slightly by offering discounts.

A more related work is by Rosen and Rosenfield (1997). In the article, price discrimination among ticket service classes is analyzed when aggregate demand is known and individual preferences are private information. Serving customers in cheap second-class seats limits the sellers' ability to extract surplus from expensive first-class seats because of some switch to the lower class. The article assumes that there are two different types of customers: high- and low-type. They want to know how high- and low-type tickets' quantity and price are decided, focusing on substitution effect of the tickets at the same time. A buyer's reserve prices for each type of seat are increasing functions of the intensity of the demand. The demand has two parameters α and β that depend on the specific service, seat quality, and prices of complements. The conclusion is that marginal buyers of first-class seats switch to second class rather than no purchase when price rises. The marginal second-class buyers are less likely to cease purchasing. Customers with the highest general tastes for the basic service never leave the market when prices of their most preferred class of seat increases. The seller can charge them more and deter substitution to the other class by increasing the interclass quality difference. The cross effects of quality on price in the socially optimum marginal benefit calculations turn out to be symmetrically negative in both equations because an increase in one service quality

reduces consumer surplus in the other quality.

Streams of literature related to ticket selling investigate on how dynamic pricing in regards to time improves ticket sellers' profit. However, with the proliferation of the Internet, customers can check prices with convenience and are more sensitive to price change. It is especially true when it comes to theater ticket selling. In this industry, price bands are predefined and will not alter throughout the entire sales period. Instead of focusing on pricing strategies, we turn to quantity control. Rosen and Rosenfield (1997) utilizes fixed-price strategy and assumes there are two segments of pricing with cross effects in the market; however, they do not investigate the relation of cumulated sales to customer utility and focus on the effect of price fluctuation. In this study, we investigate how cumulated sales outcome leverage the revenue of performance ticket retailers.





Chapter 3

Problem Description and Formulation

In this study, we investigate the relationship of cumulated sales outcome to customer utility and how the relationship affects the profit of a ticket seller. We obtain a performance data set spanned from 2008 to 2012 of a local theater group in Taiwan. They had been actively performing and producing innovative and educational kids' drama for sixteen years. The ticket selling data set contains detailed information of ticket sales of every price band. For example, ticket time, selling period, and ticket sold per day. The data set is used to verify our hypotheses.

We define a *show* to have one or multiple *performances*. Take Notre Dame de Paris for example. It's *performance period* is from February 24 to March 5 and there are thirteen performances during the period. Shows have great extends of uncertainty in terms of cast, story, and performance location. Neglecting one of the key factors of the show might lead to mismatch between expected quality and the actual quality of the performance. We

assume that there are fixed portions of knowledgeable customers who understand all the factors of theaters' various price bands; therefore, those who are not familiar with the performance quality will consider cumulated sales outcome as a useful information.

This paper examines how cumulated sales outcome affects customer valuation on different price bands. The cumulated sales outcome mentioned is translated into the percentage of the sold quantities divided by the total quantity. Let's use an example to explain the logic. There are two products in the market. Product A and B are sold for 12 and 20 products respectively in the same period. It seems that product B has a better sales; however, product A has 20 products in total and product B has 50. It is intuitive that people will consider product A as a more popular product. People not only look at the sales number but also the total inventory prepared because the risk of having stock out will also drive people to buy products. The data provided by the theater does not have the exact value of customer utility of various tickets. Therefore, we use tickets sold each day to be a proxy of customers' willingness to buy. If the tickets of the price band is popular and people are more prone to purchase, the quantity tickets sold today will increase and vice versa. We believe that higher sales outcome leads to higher customer utility (sales speed) and Hypothesis 1 is stated.

Hypothesis 1: A price band's cumulated sales positively affect sales speed of this band. More precisely, the higher a price band's cumulated sales, the higher the sales speed of this band.

A show usually contains multiple price bands. When a show started to sell tickets, people evaluate various price bands, compare the utility of each price band, and purchase the one with maximum utility. If one of the price bands has a better sales outcome,

people will have been more willing to purchase it and less willing to purchase the others. It is easy to justify that there are cross effects between various price bands; therefore, the cumulated sales outcome of other price bands are taken into consideration when examining a specific price band. Hypothesis 2 is made based on the assumption.

Hypothesis 2: Other price bands' cumulated sales negatively affect sales speed of this band. More precisely, the higher other price bands' cumulated sales, the lower the sales speed of this band.

The data set contains other information like the sales period, time, and location of the performances. We hypothesize that when the selling period become longer, the average sales speed will become slower. Hypothesis 3 is therefore formed.

Hypothesis 3: Sales period negatively affect sales speed of the performance. More precisely, the higher a performance's sales period, the slower the sales speed of the performance.

Hsieh (2015) says that the total sales of the performances in the afternoon and in Taichung are the highest. As a result, we consider time and region of the performances factors that might influence our hypotheses. We add them as control factors of our model. The data set spanned from 2008 to 2012; therefore, $Year_k$ is also added to be another control variable. Other control variables includes $Show_k$, $PerformanceId_k$, and $Week_{kt}$, where k stands for performance k and t is week t . $Week_{kt}$ stands for the t week since the start of ticket sales of performance k .

From the hypotheses, we formulate the equation

$$\frac{y_{ikt}}{s_{ikt} + q_{ikt}} = \beta_0 + \sum_{i \in N} \beta_i^S \left(\frac{s_{ikt}}{s_{ikt} + q_{ikt}} \right) + \beta_1^P Period_k + \beta_2^P Week_{kt} + \beta_1^T Morning_k + \beta_2^T Evening_k$$

$$+\beta_1^R Northern_k + \beta_2^R Southern_k + \beta_1^T Year_k + \beta_1^D Show_k + \beta_2^D PerformanceId_k + \epsilon,$$

where N is the set of price bands of a performance. For example, performance A has three price bands: 300, 500, and 600; therefore N is 1, 2, and 3. Three price bands ordered from the highest price band to the lowest. The price band 600's i is 1, 500 is 2, and 300 is 3. y_{ikt} is the tickets' sales speed in performance k during the week t for price band i . The notation β_j^S stands for sales related variable, while P is period, T is time, R is region, and D is identity (ID) related. Both *Time* and *Region* are factored variables. s_{ikt} is cumulated tickets sold of the price band i of performance k in week t . As for, q_{ikt} is tickets' quantity left of the price band i . We let *Afternoon* _{k} and *Central* _{k} as the factors for *Time* _{k} and *Region* _{k} for performance k . y_{ikt} and s_{ikt} are both divided by $s_{ikt} + q_{ikt}$, which is the capacity, to make performances of various capacity to have the same scale.

As we mentioned before, one of the unique characters of performance ticket selling is that a performance has different price bands. We examine the interaction between prices. We hypothesize that the relationships of a price band to another should be stronger when their prices are closer to each other. Hypothesis 4 is added.

Hypothesis 4: The effects of a price band to another are stronger when their prices are closer to each other.

Using the historical ticket selling data from a Taiwanese theater, we use regression model to verify how cumulated sales outcome affect customer utility and how various price band's sales outcome ($\frac{s_{ikt}}{s_{ikt}+q_{ikt}}$) affect the sales speed of other price bands ($\frac{y_{jkt}}{s_{jkt}+q_{jkt}}$, where $j \in N$ and $j \neq i$). There are performances with two price bands to six price bands. The number of variables will be different when the show have different numbers of price bands. In the first part of Chapter 4, we aggregate all other price bands to a new

variable; as a result, performances with different numbers of price bands can be examined in the same model. Hypothesis 1 to 3 are examined. In the second part, we focus on verifying Hypothesis 4 and examine the relationships of price bands by categorizing the price bands to those with closer and farer prices to a specific price band.







Chapter 4

Analysis

4.1 Data cleansing

The theater data set is from 2008 to 2012. In prior to making analysis, we adjust and clean the data set in order to fit the hypotheses we made.

First, there are some days with no historical sales. What we need in our research is sales speed so that we can understand customer preference. In order to measure how cumulated sales will affect the sales speed, we add the missing dates back to the historical sales and set the sales amount to be zero.

Second, we find that when the tickets of a price band are about to be sold out, sales distortion occurs. In Figure 4.1, the sales speed of the price bands with higher cumulated sales becomes slower in comparison to its previous sales speed. The sales speed of other price bands accelerates. Instead of flocking to purchase the price band's tickets, people turn to other price bands. One possible reason might be that when the tickets are about

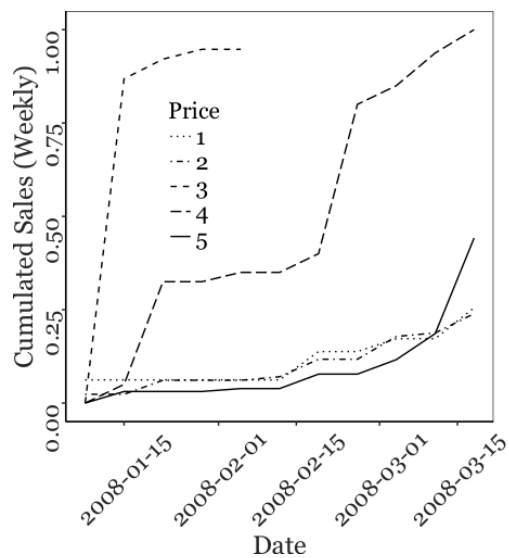
to sold out, the remaining tickets scatter over the entire segments; therefore, groups of people cannot purchase the seats linked together. People will either purchase tickets of other price bands or make no purchase. We assume that the distortion effect occurs when the tickets are sold 90% and ignore the data when any one of the price bands' cumulated sales exceed 90%. The most important reason to ignore data of those cumulated sales exceed 90% is that if a ticket is sold out, the sales speed will be zero afterwards; however, the cumulated sales will be 100%. The data without a doubt will lead to biased results.

Third, there are times when people ask for refunds. We consider refunds as negative sales. The reason is that the data set does not document every single transaction of the ticket sales. It shows the aggregate sales of one day. When refunds are higher than tickets sold, the sales data that day is documented as negative sales. When the sales are positive, the information might include refunds with less quantity than sales.

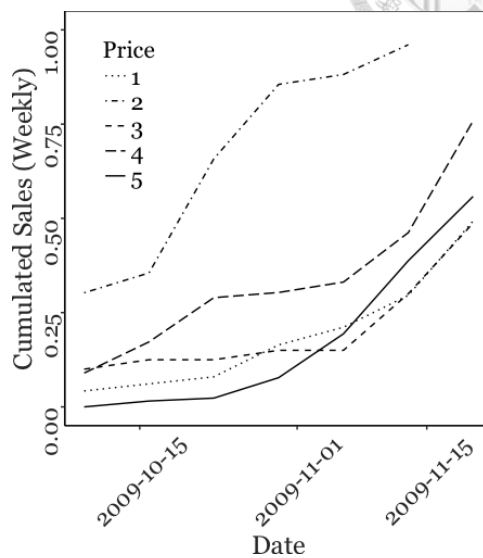
Fourth, we find that the time unit of the sales speed is important. Take Show 4 as an example, when we look at sales speed data (Number of tickets sold, $NumSold_{ikt}$) of Show 4, we find that there are many zeros in the data (Figure 4.2a). As a result, we aggregate sales data by week to decrease the quantity of zeros (Figure 4.2b).

Lastly, the interactions between the price bands of a performance is also investigated in the paper. We, therefore, select the performances who has sufficient numbers of price bands, which are the ones greater than 3 price bands. After eliminating the unfit performances, there are 30 shows and 324 performances in total out of the five-year data set.

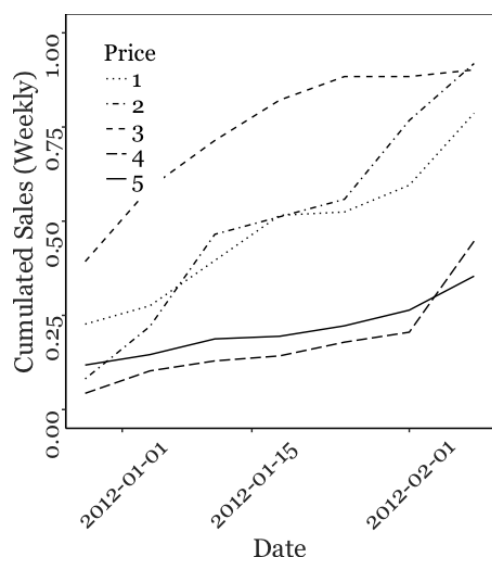
The variable mentioned in the hypotheses are the cumulated sales outcomes ($\frac{s_{ikt}}{s_{ikt}+q_{ikt}}$) of the investigated price band and other price bands, sales period, time (morning, af-



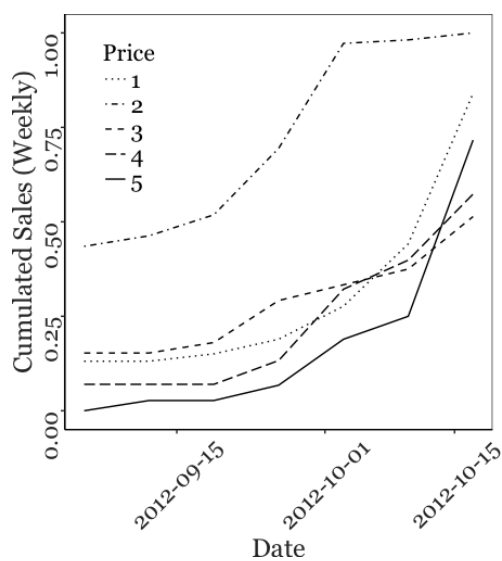
(a) Performance 26



(b) Performance 133



(c) Performance 277



(d) Performance 348

Figure 4.1: Weekly Cumulated Sales of Performances.

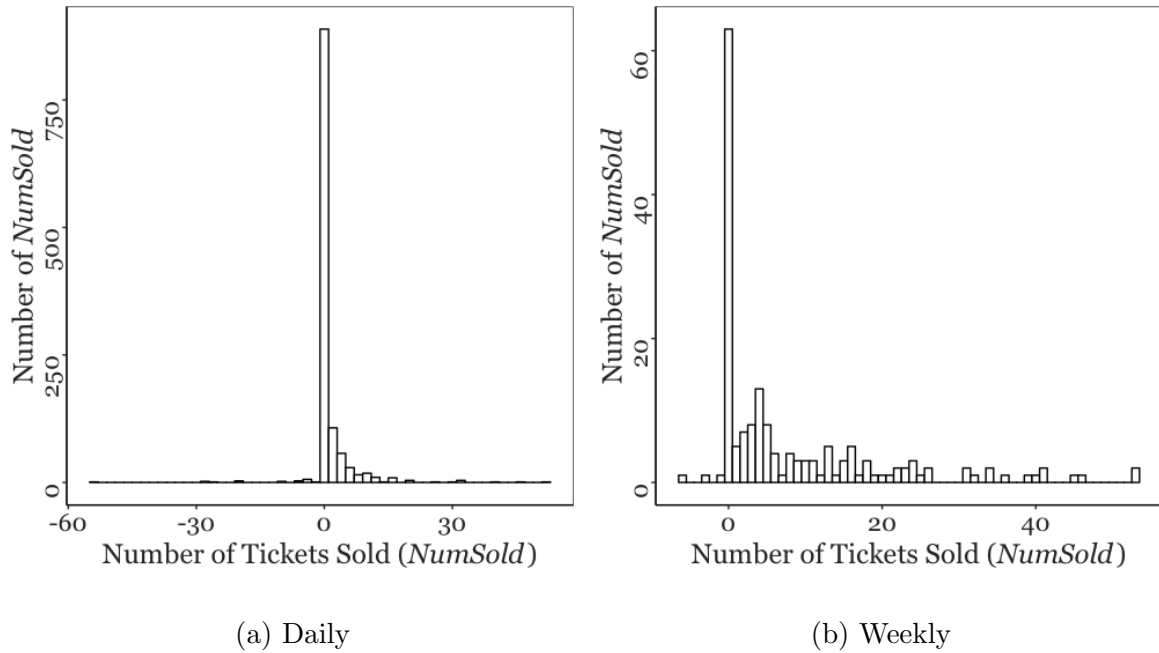


Figure 4.2: Aggregate Weekly Sales of Show 4.

ternoon, and evening), and region (Northern, Central, and Southern). In addition, we add $Year_k$ as another variable because the data set spans various year. We there let cumulated sales outcomes of the price band i ($\frac{s_{ikt}}{s_{ikt}+q_{ikt}}$) as $CumSelf_{ikt}$ and cumulated sales outcomes of other price bands ($j \in N$, and $j \neq i$) as $CumOther_{ikt}$ (Table 4.1).

4.2 An illustrative example of data processing

Here we use an example with three price bands to let readers understand our data processing. The section is separated in to two parts. The first part focuses on the relationship of the cumulated sales of a specific price band in a performance and other price bands. The second part focuses on the interaction between a specific price band and the price bands closer or farer to it. As a result, the regression models in the the separate sections are different and the data transformations are different.

Variable	Type	Possible Values
$CumSelf_{ikt}$	Numerical	0 – 0.9
$CumOther_{jkt}$	Numerical	0 – 0.9
$Period_k$	Numerical	13 – 359
$Week_{kt}$	Categorical	1 – 53
$Time_k$	Categorical	$Morning_k, Afternoon_k, Evening_k$
$Region_k$	Categorical	$Northern_k, Central_k, Southern_k$
$Year_k$	Numerical	2008 – 2012
$Show_k$	Categorical	1 – 30
$PerformanceId_k$	Categorical	1 – 324



Table 4.1: Type and Possible Values of Variables

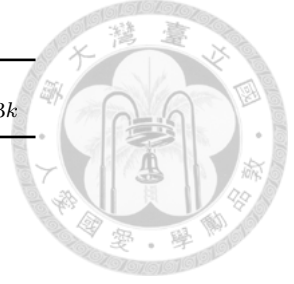
4.2.1 Testing Hypothesis 1 to 3

Table 4.2 lists an example data set for the first part. There are three price bands in the performance. P_{ik} stands for price band i of performance k . P_{1k} is the price band with the highest price and P_{2k} is the second highest price band. P_{3k} is the price band with the lowest price for performance k . $Week_{kt}$ denotes the number of week counted from the start of the sales.

From the data set of local Taiwanese theater, we can also obtain the capacity for each price band (Table 4.3).

We let $NumSold = \frac{SalesofP_{ik}}{Capacity_{ik}}$ (Table 4.4).

However, normally people can only see the cumulated sales of tickets rather than the



$Week_{kt}$	Sales of P_{1k}	Sales of P_{2k}	Sales of P_{3k}
1	0	3	5
2	4	2	1
3	4	3	0
4	2	1	5
5	2	0	0

Table 4.2: Sales of the Example Data Set 1

Capacity of P_{1k}	Capacity of P_{2k}	Capacity of P_{3k}
12	80	120

Table 4.3: Capacity of the Example Data Set 1

$Week_{kt}$	$NumSold$ for P_{1k}	$NumSold$ for P_{2k}	$NumSold$ for P_{3k}
1	$\frac{0}{12} = 0$	$\frac{3}{80} = 0.0375$	$\frac{5}{120} = 0.041$
2	$\frac{4}{12} = 0.33$	$\frac{2}{80} = 0.0250$	$\frac{1}{120} = 0.008$
3	$\frac{4}{12} = 0.33$	$\frac{3}{80} = 0.0375$	$\frac{0}{120} = 0$
4	$\frac{2}{12} = 0.17$	$\frac{1}{80} = 0.0125$	$\frac{5}{120} = 0.041$
5	$\frac{2}{1} = 0.17$	$\frac{0}{80} = 0$	$\frac{0}{120} = 0$

Table 4.4: $NumSold$ for Example Data Set 1

$Week_{kt}$	Cumulated Sales for P_{1k}	Cumulated Sales for P_{2k}	Cumulated Sales for P_{3k}
1	0	3	5
2	4	5	6
3	8	8	6
4	10	9	11
5	12	9	11

Table 4.5: Cumulated Sales of Example Data Set 1

$Week_{kt}$	$CumSelf_{1kt}$	$CumOther_{1kt}$
1	$\frac{0}{12} = 0$	$\frac{3+5}{80+120} = 0.04$
2	$\frac{4}{12} = 0.33$	$\frac{5+6}{80+120} = 0.055$
3	$\frac{8}{12} = 0.67$	$\frac{8+6}{80+120} = 0.07$
4	$\frac{10}{12} = 0.83$	$\frac{9+11}{80+120} = 0.1$
5	$\frac{12}{12} = 1$	$\frac{9+11}{80+120} = 0.1$

Table 4.6: $CumSelf_{1kt}$ and $CumOther_{1kt}$ of Example Data Set 1

sales speed. We transform sales data of P_{ik} into cumulated sales (Table 4.5).

Let we are examining the price band 1 now; therefore, $CumSelf_{1kt}$ is P_{1k} and $CumOther_{1kt}$ is the combination of P_{2k} and P_{3k} where its formulation is $(\frac{\sum_{j \neq i}(s_j)}{\sum_{j \neq i}(s_j + q_j)})$.

Aforementioned, the data that has $CumSelf > 0.9$ will be deleted. In Table 4.6, we find the $CumSelf$ of the last row exceed 0.9 and the data and the data afterwards are ignored.

$Week_{kt}$	Sales of P_{1k}	Sales of P_{2k}	Sales of P_{3k}	Sales of P_{4k}
1	0	3	5	0
2	1	2	2	0
3	5	6	0	1
4	0	0	3	2
5	4	0	0	1

Table 4.7: Sales of the Example Data Set 2

$Capacity$ of P_{1k}	$Capacity$ of P_{2k}	$Capacity$ of P_{3k}	$Capacity$ of P_{4k}
10	20	30	100

Table 4.8: Capacity of the Example Data Set 2

4.2.2 Testing Hypothesis 4

When it comes to the second part, we investigate if the effects of a price band to another are stronger when their prices are closer to each other. Price bands need to be separately examined. We categorize other price bands as two types: Near and Far. For P_{1k} , the price bands that are near is P_{2k} and the rest are far. For P_{2k} , the near price bands are P_{1k} and P_{3k} . The $CumSold$ for Near and Far are $CumNear_{ikt}$ and $CumFar_{ikt}$ separately. The calculation of the both variable are the same as that to $CumOther_{ikt}$

We use another example data set (Table 4.7) here and list $CumSelf_{2kt}$ and $CumNear_{2kt}$ and $CumFar_{2kt}$ are listed in Table 4.9.

Week	$CumSelf_{2kt}$	$CumNear_{2kt}$	$CumFar_{2kt}$
1	$\frac{3}{20} = 0.15$	$\frac{0+5}{20+30} = 0.10$	$\frac{0}{100} = 0$
2	$\frac{5}{20} = 0.25$	$\frac{1+7}{20+30} = 0.16$	$\frac{0}{100} = 0$
3	$\frac{11}{20} = 0.55$	$\frac{6+7}{20+30} = 0.26$	$\frac{1}{100} = 0.01$
4	$\frac{11}{20} = 0.55$	$\frac{6+10}{20+30} = 0.32$	$\frac{3}{100} = 0.03$
5	$\frac{11}{20} = 0.55$	$\frac{10+10}{20+30} = 0.40$	$\frac{4}{100} = 0.04$



Table 4.9: $CumSelf_{2kt}$, $CumNear_{2kt}$, and $CumFar_{2kt}$ of Example Data Set 2

4.3 Exploratory data analysis

Prior to getting into detail, we describe the data set to let readers understand the data more. The data set contains 36,283 rows of data. There are 30 shows and 324 performances in total out of the five-year data set. The average numbers of performance in a show is 10.8.

First, we test on $CumSelf_{ikt}$ and aggregate other price bands to $CumOther_{ikt}$ to put data of different numbers of price bands in one model. Variables used in this part is $CumSelf_{ikt}$, $CumOther_{ikt}$, $Period_k$, $Time_k$, $Region_k$, and $Year_k$.

Before looking at $CumSelf_{ikt} \times Capacity_{ik}$ and $CumOther_{ikt} \times Capacity_{ik}$, we investigate the numbers of tickets sold without dividing capacity, which we interpret it as $NumSold_{ikt}$. The results of both $NumSold_{ikt}$ and cumulated sales of both variables are listed in Table 4.10. The variation of all four variables are huge and from Figure A.1a to Figure A.1d, we can conclude that all distributions of the variables are exponential. For $Period_k$, the variation is also large and its distribution is right skewed (Figure A.1e).

Variable	Mean	Standard Deviation	Maximum	Minimum
$NumSold_{ikt} (y_{ikt})$	6.625	15.578	347	-158
$CumSelf_{ikt} \times Capacity_{ik}$	39.050	42.111	437	0
$CumOther_{ikt} \times Capacity_{ik}$	175.590	162.877	1390	0
$Period_k$	144.843	87.734	359	13

Table 4.10: Exploratory Analysis of the Variables of the First Part

As for the control variables, we look at how many performances are in different categories or values of $Time_k$, $Region_k$, and $Year_k$. The performances played in the morning are 83, in the afternoon are 122, and in the evening are 119 (Figure A.1f). There are 284 performances took place in Northern Taiwan, 28 in Central, and 12 in Southern (Figure A.1g). When it comes to $Year_k$, we find that number of performances played has an increasing trend and the quantities are 44, 44, 70, 72, and 94 (Figure A.1h).

4.4 Regression analysis

There are two parts in this section. Part A mainly discusses whether cumulated sales outcome affects weekly sales speed under various circumstances. Part B focuses on the interactions between price bands and whether price bands with closer prices have larger effects on the sales speed of the targeted price band.

The main difference between the two parts is the form of other price bands. In our data set the shows have either 4, 5, or 6 numbers of price bands in a performance. In part A, we mainly discuss the effect of cumulated sales to sales speed; therefore, instead of

separating the performances having different numbers of price bands, we aggregate all the other price bands to be $(\frac{\sum_{j \neq i}(s_j)}{\sum_{j \neq i}(s_j + q_j)})$ and name it as $CumOther_{ikt}$. After this aggregation, we can put data of various numbers of price bands in the same model. In the second part, we aggregate the prices that are close to the target price band to $CumNear_{ikt}$ and the ones farer to $CumFar_{ikt}$.

4.4.1 Testing Hypothesis 1 to 3

There are four models in part A. We first focus on whether $CumSelf_{ikt}$ is having a positive relation to $WeeklySales_{ikt}$ and if $CumOther_{ikt}$ has a negative relation. This is directly investigated by model A_1 . We add variables that belong to the basic performance information people can get when buying tickets into consideration in Model A_2 . $Period_k$, $Time_k$, $Region_k$, and $Year_k$ are added. The interaction effects of performance information to $CumSelf_{ikt}$ are further added in Model A_3 to see more detailed interactions. We further add $Show_k$, $PerformanceId_k$, and $Week_{kt}$ as control variables to Model A_3 as Model A_3^D to control the variation among performances. The effect of $CumSelf_{ikt}$ and $CumOther_{ikt}$ to $WeeklySales_{ikt}$ might not be linear. Model A_4 is then proposed to examine the relationship. We take $Afternoon_k$ and $Central_k$ as factors for $Time_k$ and $Region_k$. The reference table of variable names and math notations is in Table 4.11.

Model A_1 :

$$WeeklySales_{ikt} = \beta_0 + \beta_1^S CumSelf_{ikt} + \beta_2^S CumOther_{ikt} + \epsilon$$

Model A_2 :

$$WeeklySales_{ikt} = \beta_0 + \beta_1^S CumSelf_{ikt} + \beta_2^S CumOther_{ikt} + \beta_1^P Period_k$$

Variable Name	Math Notation
$WeeklySales_{ikt}$	$\left(\frac{y_{ikt}}{s_{ikt}+q_{ikt}}\right)$
$CumSelf_{ikt}$	$\left(\frac{s_{ikt}}{s_{ikt}+q_{ikt}}\right)$
$CumOther_{ikt}$	$\left(\frac{\sum_{j \neq i}(s_{jkt})}{\sum_{j \neq i}(s_{jkt}+q_{jkt})}\right)$



Table 4.11: Variable Names and Notations

$$\begin{aligned}
& +\beta_1^T Morning_k + \beta_2^T Evening_k + \beta_1^R Northern_k + \beta_2^R Southern_k \\
& +\beta_1^Y Year_k + \epsilon
\end{aligned}$$

Model A_3 :

$$\begin{aligned}
WeeklySales_{ikt} &= \beta_0 + \beta_1^S CumSelf_{ikt} + \beta_2^S CumOther_{ikt} + \beta_1^P Period_k \\
& +\beta_1^T Morning_k + \beta_2^T Evening_k + \beta_1^R Northern_k + \beta_2^R Southern_k + \beta_1^Y Year_k \\
& \beta_1^{PI} Period_k \times CumSelf_{ikt} + \beta_1^{TI} Morning_k \times CumSelf_{ikt} + \beta_2^{TI} Evening_k \times CumSelf_{ikt} \\
& +\beta_1^{RI} Northern_k \times CumSelf_{ikt} + \beta_2^{RI} Southern_k \times CumSelf_{ikt} + \epsilon
\end{aligned}$$

Model A_3^D :

$$\begin{aligned}
WeeklySales_{ikt} &= \beta_0 + \beta_1^S CumSelf_{ikt} + \beta_2^S CumOther_{ikt} + \beta_1^P Period_k + \beta_2^P Week_{kt} \\
& +\beta_1^T Morning_k + \beta_2^T Evening_k + \beta_1^R Northern_k + \beta_2^R Southern_k + \beta_1^Y Year_k \\
& +\beta_1^D Show_k + \beta_2^D PerformanceId_k + \beta_1^{PI} Period_k \times CumSelf_{ikt} \\
& +\beta_1^{TI} Morning_k \times CumSelf_{ikt} + \beta_2^{TI} Evening_k \times CumSelf_{ikt} \\
& +\beta_1^{RI} Northern_k \times CumSelf_{ikt} + \beta_2^{RI} Southern_k \times CumSelf_{ikt} + \epsilon
\end{aligned}$$

In Model A_3 and A_3^D , we added the interaction effects and are denoted by β_j^{XI} , where $X \in \{S, P, T, R\}$.

We can find that the three models all have a significant correlation between cumulated sales to sales speed and a negative correlation between other price band's cumulated sales to sales speed (Table 4.12). Hypothesis 1 and 2 are therefore proven. As for Hypothesis 3, the results of Model A_2 and A_3 both show that the length of sales period negatively affects sales speed and verified Hypothesis 3; however, when control variables $Show_k$, $PerformanceId_k$, and $Week_{kt}$ are added, $Period_k$ becomes positive. $Week_{kt}$ are mostly negative with the level of $Week_{k1}$, which indicates that the sales speed of Week 1 is normally the fastest.

It is also considered that if the effect of Hypothesis 1 and 2 are not only linearly correlated. We add two variables: $CumSelf_{ikt}^2$ and $CumOther_{ikt}^2$ to Model A_3 as Model A_4 .

Model A_4 :

$$\begin{aligned}
 WeeklySales_{ikt} = & \beta_0 + \beta_1^S CumSelf_{ikt} + \beta_2^S CumOther_{ikt} + \beta_1^P Period_k + \beta_2^P Week_{kt} \\
 & + \beta_1^T Morning_k + \beta_2^T Evening_k + \beta_1^R Northern_k + \beta_2^R Southern_k + \beta_1^Y Year_k \\
 & + \beta_1^D Show_k + \beta_2^D PerformanceId_k + \beta_1^{PI} Period_k \times CumSelf_{ikt} \\
 & + \beta_1^{TI} Morning_k \times CumSelf_{ikt} + \beta_2^{TI} Evening_k \times CumSelf_{ikt} \\
 & + \beta_1^{RI} Northern_k \times CumSelf_{ikt} + \beta_2^{RI} Southern_k \times CumSelf_{ikt} \\
 & + \beta_3^S CumSelf_{ikt}^2 + \beta_4^S CumOther_{ikt}^2 + \epsilon
 \end{aligned}$$

The previous variables maintain the same effects in Model A_4 (Table 4.12). If we fix all other variables and the equation that only consider the effect of $CumSelf_{ikt}$ and

Regression	Model A ₁	Model A ₂	Model A ₃	Model A ₃ ^D
(Intercept)	$1.11 \times 10^{-2***}$	9.84×10^{-1}	-4.16×10^{-1}	3.16×10^1
<i>CumSelf</i> _{ikt}	$1.49 \times 10^{-1***}$	$1.46 \times 10^{-1***}$	$3.02 \times 10^{-1***}$	$3.30 \times 10^{-1***}$
<i>CumOther</i> _{ikt}	$-5.66 \times 10^{-2***}$	$-7.41 \times 10^{-2***}$	$-6.56 \times 10^{-2***}$	$1.21 \times 10^{-2*}$
<i>Period</i> _k		$-1.85 \times 10^{-4***}$	$-7.39 \times 10^{-5***}$	$1.38 \times 10^{-4***}$
<i>Year</i> _k		-4.61×10^{-4}	2.19×10^{-4}	-1.64×10^{-2}
Time _k				
<i>Morning</i> _k		$-2.66 \times 10^{-3*}$	$4.54 \times 10^{-3*}$	-5.69×10^{-3}
<i>Evening</i> _k		$-4.14 \times 10^{-3***}$	$-5.76 \times 10^{-3***}$	-4.74×10^{-3}
Region _k				
<i>Northern</i> _k		-6.97×10^{-4}	$7.92 \times 10^{-3**}$	$-1.64 \times 10^{-1***}$
<i>Southern</i> _k		-2.66×10^{-3}	-1.01×10^{-3}	$-1.44 \times 10^{-1***}$
<i>CumSelf</i> _{ikt} × <i>Period</i> _k			$-5.08 \times 10^{-4***}$	$-6.41 \times 10^{-4***}$
<i>CumSelf</i> _{ikt} × <i>Time</i> _k				
<i>CumSelf</i> _{ikt} × <i>Morning</i> _k			$-2.37 \times 10^{-2***}$	$-2.91 \times 10^{-2***}$
<i>CumSelf</i> _{ikt} × <i>Evening</i> _k			6.45×10^{-3}	6.51×10^{-3}
<i>CumSelf</i> _{ikt} × <i>Region</i> _k				
<i>CumSelf</i> _{ikt} × <i>Northern</i> _k			$-5.58 \times 10^{-2***}$	$-3.90 \times 10^{-2***}$
<i>CumSelf</i> _{ikt} × <i>Southern</i> _k			7.81×10^{-3}	3.58×10^{-2}
Other Control Variables				
<i>Show</i> _k				
<i>PerformanceId</i> _k				
<i>Week</i> _{kt}				
R ²	12.50%	18.54%	20.83%	40.68%
Adjusted R ²	12.49%	18.51%	20.78%	39.60%

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

Table 4.12: The Results of Part A

$CumOther_{ikt}$ will be

$$WeeklySales_{ikt} = 0.469CumSelf_{ikt} - 0.157CumSelf_{ikt}^2 + 0.247CumOther_{ikt} - 0.369CumOther_{ikt}^2.$$



The effects of $CumSelf_{ikt}^2$ and $CumOther_{ikt}^2$ are both interactive effects of that of $CumSelf_{ikt}$ and $CumOther_{ikt}$.

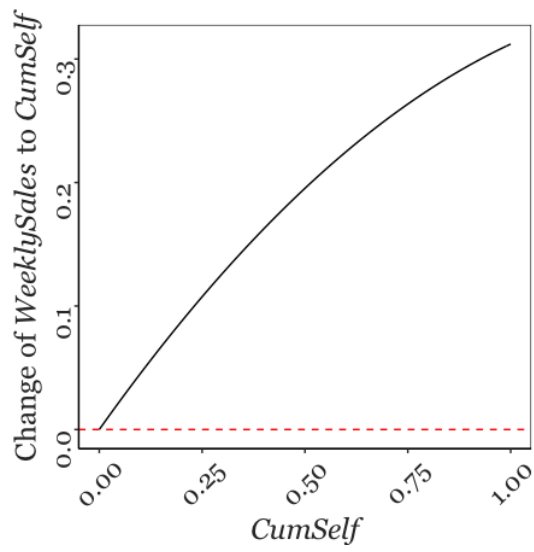
We can obtain the first derivatives of y of $CumSelf_{ikt}$ and $CumOther_{ikt}$.

$$\frac{\partial WeeklySales_{ikt}}{\partial CumSelf_{ikt}} = 0.469 - 0.314CumSelf_{ikt}.$$

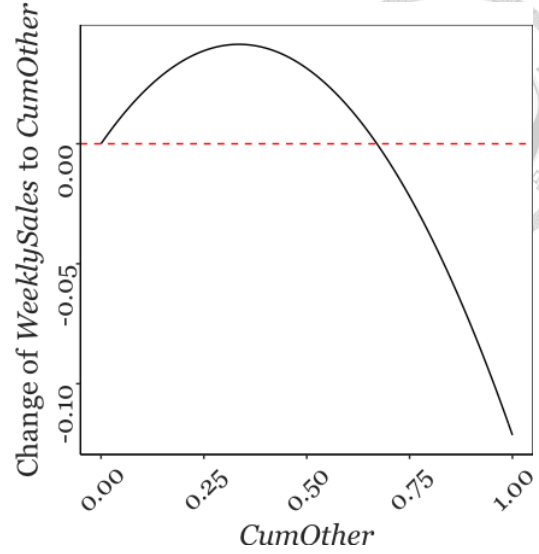
$$\frac{\partial WeeklySales_{ikt}}{\partial CumOther_{ikt}} = 0.247 - 0.738CumOther_{ikt}.$$

Both $CumSelf_{ikt}$ and $CumOther$ are concave functions. Negative $CumSelf_{ikt}^2$ explains that the effect of $CumSelf_{ikt}$ continues to increase positively with decreasing margin (Figure 4.3a). As for the relationship of $CumOther_{ikt}$ to $WeeklySales_{ikt}$ is that it first decreases positively and then when cumulated sales of other price bands is high, customer tend to have less incentive to purchase the price band's ticket (Figure 4.3b). The situation is due to the interaction of two effects. One is that when people find that other price bands are selling good, they consider the show worthy to watch and purchase their most preferable price band of the show. The second one is that when people find that other price bands are selling good, they turn to buy the price bands with higher sales. From Graph 4.3b, we can conclude that the first effect is stronger during the start of the sales period and the second effect is stronger in the end of the sales period.

We verify the correlation coefficients between variables (Figure 4.4). We find that no two variables are highly correlated.



(a) $CumSelf_{ikt}$



(b) $CumOther_{ikt}$

Figure 4.3: Change of $WeeklySales_{ikt}$.

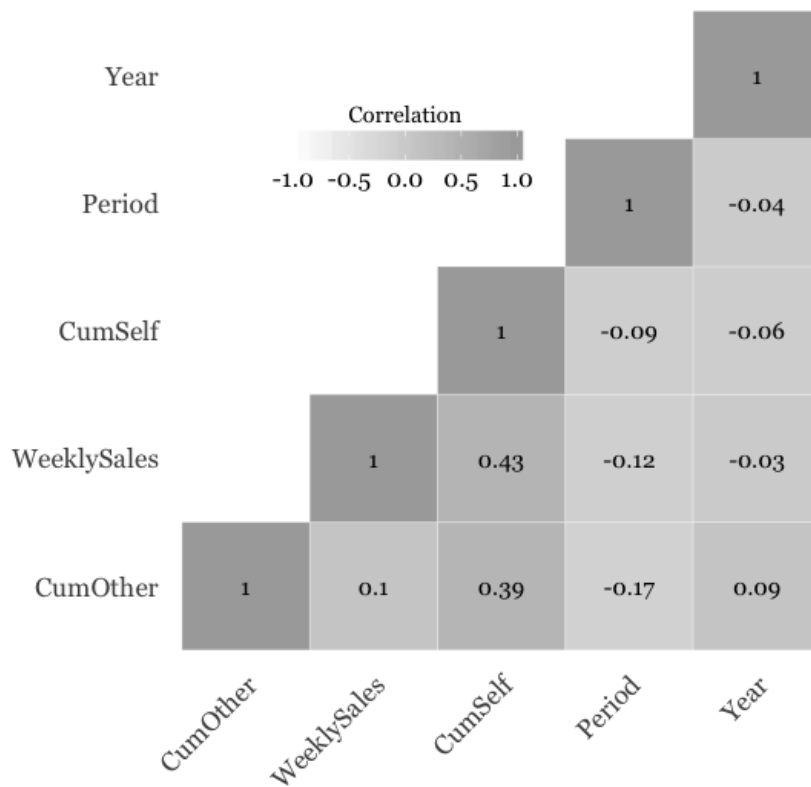


Figure 4.4: Correlation Coefficients of Variables

4.4.2 Testing Hypothesis 4

When it comes to part B, we investigate if the effects of a price band to another are stronger when their prices are closer to each other. Price bands need to separately examined. We categorize other price bands as two types: Near and Far. For P_{1k} , the price bands that are near is P_{2k} and the rest are far. For P_{2k} , the near price bands are P_{1k} and P_{3k} . The $CumSold_{ikt}$ for Near and Far are $CumNear_{ikt}$ and $CumFar_{ikt}$ separately.

Model B:

$$\begin{aligned}
 WeeklySales_{ikt} = & \beta_0 + \beta_1^S CumSelf_{ikt} + \beta_2^S CumNear_{ikt} + \beta_3^S CumFar_{ikt} \\
 & + \beta_1^P Period_k + \beta_2^P Week_{kt} + \beta_1^T Morning_k + \beta_2^T Evening_k + \beta_1^R Northern_k + \beta_2^R Southern_k \\
 & + \beta_1^Y Year_k + \beta_1^D Show_k + \beta_2^D PerformanceId_k + \beta_1^{PI} Period_k \times CumSelf_{ikt} \\
 & + \beta_1^{TI} Morning_k \times CumSelf_{ikt} + \beta_2^{TI} Evening_k \times CumSelf_{ikt} \\
 & + \beta_1^{RI} Northern_k \times CumSelf_{ikt} + \beta_2^{RI} Southern_k \times CumSelf_{ikt} + \epsilon
 \end{aligned}$$

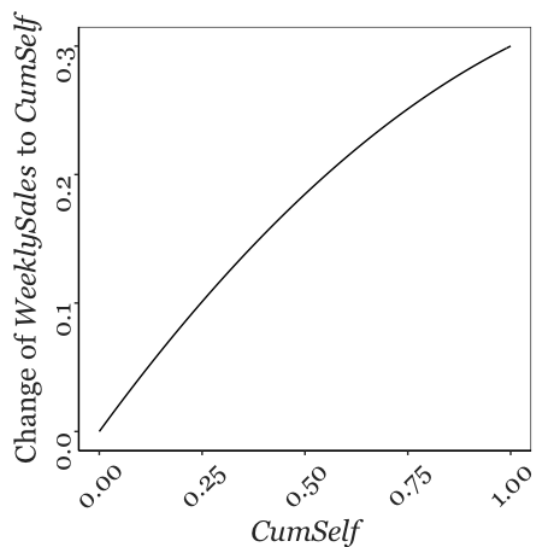
From Table 4.13 we find that the correlation between $CumSelf_{ikt}$, $CumNear_{ikt}$, and $CumFar_{ikt}$ to $WeeklySales_{ikt}$ has the same pattern as that of Model A_3^D and A_4 . From Figure 4.5, we can also find that Model B has the same pattern as A_4 . $CumSelf_{ikt}$, $CumNear_{ikt}$, and $CumFar_{ikt}$ are all concave. $CumNear_{ikt}$ and $CumFar_{ikt}$ both are with decreasing marginal increase first and then reach an U-turn and decrease with increasing margin. Interestingly, $CumFar_{ikt}$ has larger effects compare with $CumNear_{ikt}$. Hypothesis 4 is not verified and can be further investigated in the future.



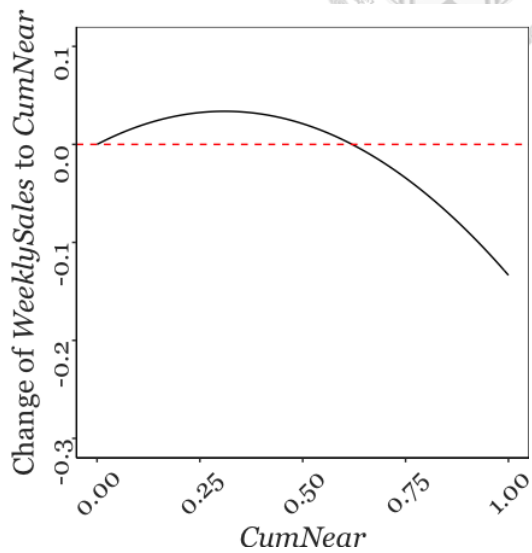
Regression	Model <i>B</i>
(Intercept)	-1.66×10^1
<i>CumSelf</i> _{<i>ikt</i>}	$4.39 \times 10^{-1***}$
<i>CumNear</i> _{<i>ikt</i>}	$2.18 \times 10^{-1***}$
<i>CumFar</i> _{<i>ikt</i>}	$5.31 \times 10^{-1***}$
<i>CumSelf</i> _{<i>ikt</i>} ²	$-1.39 \times 10^{-1***}$
<i>CumNear</i> _{<i>ikt</i>} ²	$-3.51 \times 10^{-1***}$
<i>CumFar</i> _{<i>ikt</i>} ²	$-8.36 \times 10^{-1***}$
<i>Period</i> _{<i>k</i>}	$2.49 \times 10^{-4***}$
<i>Year</i> _{<i>k</i>}	8.31×10^{-3}
<i>Time</i>_{<i>k</i>}	
<i>Morning</i> _{<i>k</i>}	-8.43×10^{-3}
<i>Evening</i> _{<i>k</i>}	-5.90×10^{-4}
<i>Region</i>_{<i>k</i>}	
<i>Northern</i> _{<i>k</i>}	$-7.86 \times 10^{-2**}$
<i>Southern</i> _{<i>k</i>}	$-6.13 \times 10^{-2*}$
<i>CumSelf</i>_{<i>ikt</i>} × <i>Period</i>_{<i>k</i>}	$-6.44 \times 10^{-4***}$
<i>CumSelf</i>_{<i>ikt</i>} × <i>Time</i>_{<i>k</i>}	
<i>CumSelf</i> _{<i>ikt</i>} × <i>Morning</i> _{<i>k</i>}	$-1.79 \times 10^{-2**}$
<i>CumSelf</i> _{<i>ikt</i>} × <i>Evening</i> _{<i>k</i>}	$-1.36 \times 10^{-2**}$
<i>CumSelf</i>_{<i>ikt</i>} × <i>Region</i>_{<i>k</i>}	
<i>CumSelf</i> _{<i>ikt</i>} × <i>Northern</i> _{<i>k</i>}	$-4.86 \times 10^{-2***}$
<i>CumSelf</i> _{<i>ikt</i>} × <i>Southern</i> _{<i>k</i>}	8.18×10^{-3}
Other Control Variables	
<i>Show</i> _{<i>k</i>}	
<i>PerformanceId</i> _{<i>k</i>}	
<i>Week</i> _{<i>kt</i>}	
<i>R</i> ²	48.12%
Adjusted <i>R</i> ²	47.16%

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

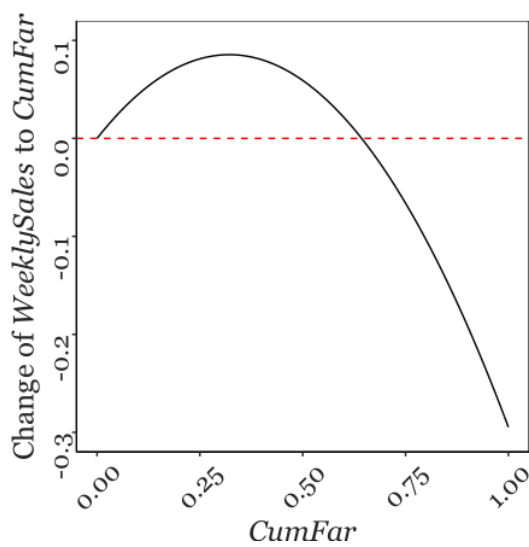
Table 4.13: The Results of Part B



(a) $CumSelf_{ikt}$



(b) $CumNear_{ikt}$



(c) $CumFar_{ikt}$

Figure 4.5: Change of $WeeklySales_{ikt}$ for Model B .





Chapter 5

Application: Dynamic Seat Allocation

After verifying our hypotheses, we come up with a scenario that can apply our findings. When a show with several price bands has been selling tickets for some time, people can verify if a specific price band is selling well or not with its cumulated sales outcome. If a price band is selling very well and is about to sold out and another price band is not selling as well and has a lot of tickets remain to be sold, we can do dynamic seat allocation. We consider a case which is possible to adjust a portion of ticket capacity of the price bands selling inferior to the price bands that is about to sell out. That is to say in the end the fewer tickets are likely to be unsold.

We use regression models to predict the future sales and adjust ticket capacity of various price bands based on the prediction. As we find that cumulated sales of price bands affect customer utility, we compared the profit difference of using the correct regression and the one neglecting cumulated sales.

A Performance' Number of Price Bands	Four	Five	Six
Number of Performances	54	158	112

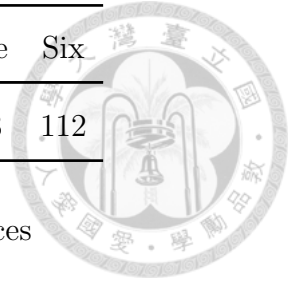


Table 5.1: Number of Price Bands of Performances

As previously mentioned, each performance has various numbers of price bands. We select some performances of five price bands as example for in our data set the performances with five price bands have the largest quantity (Table 5.1).

5.1 Exploratory data analysis

Having the same procedure as Chapter 4, we conduct exploratory data analysis in prior to implementing the application. The data with five price bands contains 20,125 rows of data and is examined. We examine the sales of performances by adding up $NumSold_{ikt}$ of the five price bands. Then, we separately analyze each price band's $NumSold_{ikt}$ and $WeeklySales_{ikt}$. Due to the resemblance of distribution, we listed out only the information of $NumSold_{ikt}$ in Table 5.2. The distributions are also all exponential and are very similar. Figure A.2a is taken as an example. As for the cumulated sales of the five price bands the variation is also huge and Figure A.2b is taken as an example. $Period_k$ is also right skewed for the performances with five price bands. From Table 5.2, we can identify that variations of the variables are all high.

$Time_k$ and $Region_k$ have resembling patterns to the first part (Figure A.2d and A.2e). The performances played in the morning is still the least (26). In the evening, there are 58 performances played during this time, which is slightly lower than the ones played in the



Variable	Mean	Standard Deviation	Maximum	Minimum
$NumSold_{kt}$ of Performance k	35.378	66.779	728	-164
$NumSold_{1kt}$ of P_{1k}	3.409	8.664	110	-33
$NumSold_{2kt}$ of P_{2k}	4.250	9.655	138	-8
$NumSold_{3kt}$ of P_{3k}	3.669	8.500	100	-26
$NumSold_{4kt}$ of P_{4k}	3.637	7.994	71	-6
$NumSold_{5kt}$ of P_{5k}	2.964	7.795	123	-27
$CumP_{1kt} \times Capacity_{1k}$	57.699	56.872	261	0
$CumP_{2kt} \times Capacity_{2k}$	75.800	68.662	331	0
$CumP_{3kt} \times Capacity_{3k}$	70.460	74.712	549	0
$CumP_{4kt} \times Capacity_{4k}$	66.463	72.878	509	0
$CumP_{5kt} \times Capacity_{5k}$	50.686	54.196	391	0
$Period_k$	161.224	92.433	359	33

Table 5.2: Exploratory Analysis of the Variables of the Second Part

afternoon (59). Northern Taiwan continues to be the location where most performances take place (111). Central Taiwan has 23 performances and Southern Taiwan only has 9. $Year_k$ of the performances of the five price bands is very different from that of the first part. Numbers of performances of five price bands progressively decrease from 42 in 2008, 34 in 2009, 13 in 2010 to 9 in 2011. In 2012, it suddenly escalates to 45 (Figure A.2f).

5.2 Numerical experiment

Hypotheses 1 and 2 show that cumulated sales of a price band is positive to sales speed and the cumulated sales of other price bands are negative to sales speed. We estimate the demand distributions of price bands with the findings and obtain the optimal ticket quantities for price bands. Some people might only use basic performance information to obtain the optimal ticket quantities and the estimated profit is inaccurate. We compared the profit with the mistaken quantities to the correct ones and the difference is 4.60% on average.

We let demand functions be

$$D_{ikt} = \mu_{ikt} + \epsilon_{ikt},$$

where i stands for the i th price band of the performances and are ranked from the highest price to the lowest. The price band with the highest price's demand function is

$$D_{1kt} = \mu_{1kt} + \epsilon_{1kt}.$$

μ_{ikt} is expected demand. This can be found out using regression model. As for ϵ_{ikt} , it is a value with 0 mean and is normally distributed. The standard deviation is estimated by using $WeeklySales_{ikt}$ of the price bands: we use the standard deviation of y_i as the

standard deviation of D_{ikt} .

There are two regression models with different variables. Model O_i is interpreted as the old model, where people only consider basic information on the performances, which includes $Period_k$, $Time$, $Region$, and $Year_k$. Model N_i considers not only basic information of the performances but also the sales data of the price bands and other price bands. We utilize Model B_i as Model N_i .

Model O_i :

$$\begin{aligned} WeeklySales_{ikt} = & \beta_0 + \beta_1^P Period_k + \beta_1^T Morning_k + \beta_2^T Evening_k \\ & + \beta_1^R Northern_k + \beta_2^R Southern_k + \beta_1^Y Year_k + \epsilon \end{aligned}$$

Model N_i :

$$\begin{aligned} WeeklySales_{ikt} = & \beta_0 + \beta_1^S CumSel_{ikt} + \beta_2^S CumNear_{ikt} + \beta_3^S CumFar_{ikt} \\ & + \beta_4^S CumSel_{ikt}^2 + \beta_5^S CumNear_{ikt}^2 + \beta_6^S CumFar_{ikt}^2 + \beta_1^P Period_k \\ & + \beta_1^T Morning_k + \beta_2^T Evening_k + \beta_1^R Northern_k + \beta_2^R Southern_k + \beta_1^Y Year_k \\ & + \beta_1^{PI} Period_k \times CumP_{ikt} + \beta_1^{TI} Morning_k \times CumP_{ikt} + \beta_2^{TI} Evening_k \times CumP_{ikt} \\ & + \beta_1^{RI} Northern_k \times CumP_{ikt} + \beta_2^{RI} Southern_k \times CumP_{ikt} + \epsilon \end{aligned}$$

In combination of ϵ_{ikt} and the μ_{ikt} we derived from Model O_i and Model N_i , we find the distributions of the price bands' demands. A price band's quantity of Model O_i is stated as q_{ikt}^O and that of Model N_i is q_{ikt}^N , where i in S is noted as the five price bands ordering from the highest price band to the lowest. We normalize the capacity of each price bands to be 1 and for five price bands the total capacity of a performance is 5.

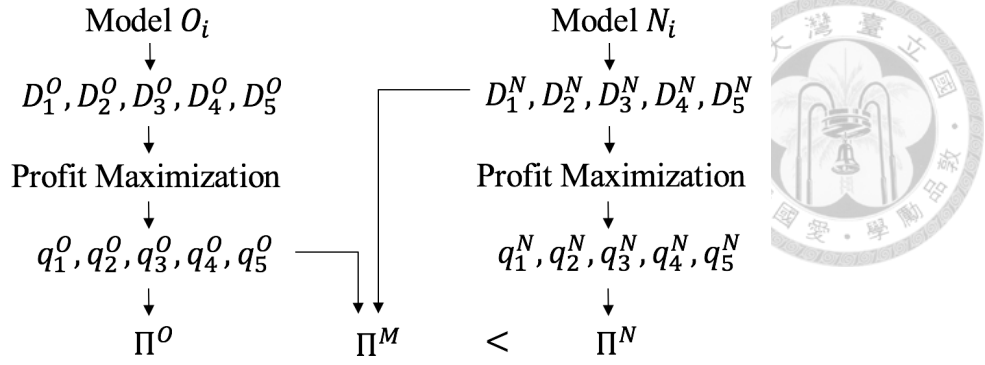


Figure 5.1: Process of Profit Prediction.

There are five known distributions and with different q_{ikt} . The expected profit can be derived with the profit function

$$\Pi_k = \max \left[\sum_{i \in N} r_i \mathbb{E}[\min\{q_{ikt}, D_{ikt}\}] \right] = \max \left[r_i \left\{ \int_0^{q_{ikt}} x f(x) dx + \int_{q_{ikt}}^{K_{0kt}} q_{ikt} f(x) dx \right\} \right].$$

$$s.t. \sum_{i \in N} q_{ikt} \leq K_{0kt},$$

where $K_{0kt} = K - \sum_{i \in N} s_{ikt}$.

In the situation of Model O_i , people only look at the information of a performance. They use the model to predict μ_{ikt}^O and estimated the demand distributions; therefore, they will find the allocated quantities (q_{ikt}^O) using the information and obtain the profit, Π_k^O . Model N_i adds in the neglected variables that influence consumers' preference and predict μ_{ikt}^N to estimate the correct demand distribution. Another set of quantities, q_{ikt}^N are derived and the estimated profit for Model N_i is Π_k^N . Π_k^O is biased for the distributions estimated in Model O_i is not as fit to the reality to Model N_i . We calculate the suboptimal profit Π_k^M using the correct demand distributions and the inaccurate q_{ikt}^O and compare with Π_k^N to see the influence of using the inappropriate model.

Table 5.3 lists all the μ_{ikt} of Model O_i and N_i . We can find that adding *CumSold*

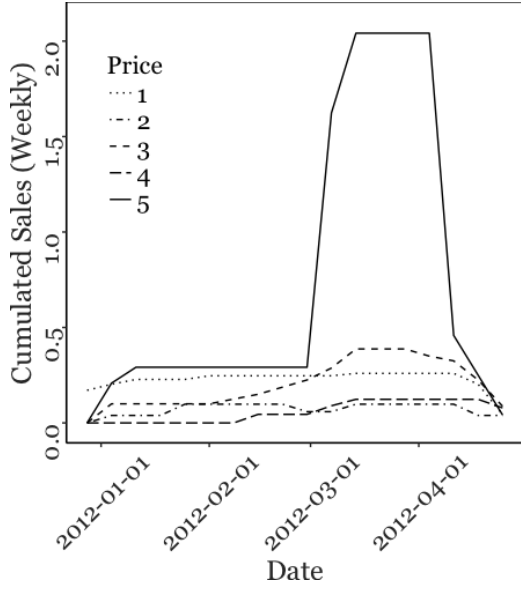
Show	μ_{1kt}^O	μ_{1kt}^N	μ_2^O	μ_{2kt}^N	μ_{3kt}^O	μ_{3kt}^N	μ_{4kt}^O	μ_{4kt}^N	μ_{5kt}^O	μ_{5kt}^N	$\sum_{i \in N} \mu_{ikt}^O$	$\sum_{i \in N} \mu_{ikt}^N$
4	0.020	0.082	0.021	0.012	0.011	-0.012	0.013	0.039	0.062	0.044	0.126	0.166
6	0.032	0.030	0.038	0.078	0.034	0.029	0.145	0.237	0.032	0.008	0.282	0.381
9	0.189	0.189	0.090	0.090	0.134	0.134	0	0	0.313	0.008	0.656	0.725
10	0.010	0.016	0.021	0.007	0.031	0.045	0.005	0.040	0.014	0.019	0.282	0.381
13	0.147	0.079	0.201	0.233	0.115	0.114	0.096	-0.020	0.071	0.100	0.630	0.506
23	-0.002	0	-0.010	-0.065	-0.011	-0.035	-0.002	-0.008	0.010	0.006	0.015	-0.102
24	0.002	-0.059	0	-0.039	-0.006	-0.108	-0.001	-0.020	-0.131	-0.048	-0.135	-0.273
26	0.016	0.016	0.136	0	0.046	0.073	0.001	0.004	0.041	0.032	0.240	0.125
27	0.076	0.204	0.109	0.153	0.096	0.285	0.079	0.260	0.394	0.604	0.754	1.506
28	0.022	0.088	0.015	0.067	0.016	0.067	0.011	0.039	0.017	0.100	0.081	0.361

Table 5.3: Predicted μ_{ikt} of New and Old Models

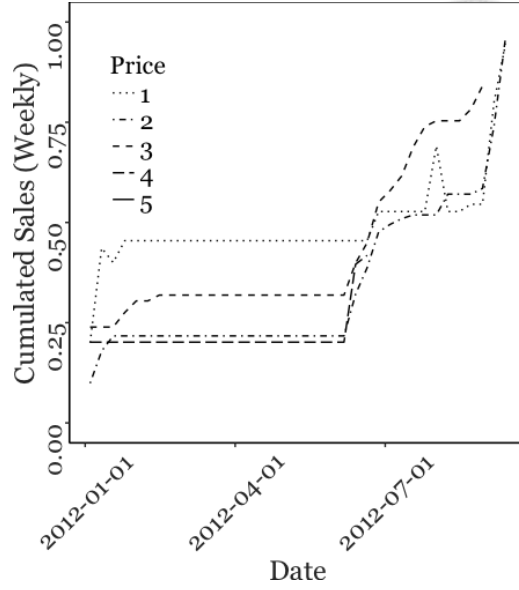
and $CumOther_{ikt}$ make the μ_{ikt}^N more affected by the sales. The result of show 27 is the most significant. The performance 289 is a performance of show 24. It has slower cumulated sales of price bands (Figure 5.2a) and results to $\mu_{ikt}^N < \mu_{ikt}^O$ for all i . Note that the cumulated sales of performance 289 is due to refund. As for the performance 337 (Figure 5.2b) in show 27, it has an extremely high sales and leads to $\mu_{ikt}^N > \mu_{ikt}^O$. Those price bands with worse sales end up with lower μ_{ikt}^N compared with μ_{ikt}^O and the price bands with higher sales results to $\mu_{ikt}^N > \mu_{ikt}^O$.

Here, we use 0.05 as the step to numerically search the best profit for all models. ¹Ten shows are chosen to be fitted and on average Π_k^M is 4.60% less than Π_k^N . Six out of ten

¹Theoretically, we may use Lagrangian relaxation and some knowledge about the newsvendor model to solve the optimization problem. Nevertheless, for our particular data set and the process of data analysis, this method will fail. In Section 5.3 we describe the method and explain why it fails.



(a) Performance 289 of Show 24



(b) Performance 337 of Show 27

Figure 5.2: Weekly Cumulated Sales of Performances.

shows, which is 60% of the total shows, have differences in profit. The average difference of the shows whose $\Pi_k^N > \Pi_k^M$ is 7.67%. Note that show 27 has a very significant difference due to the fact that the differences of μ_{ikt}^N and μ_{ikt}^O are huge (Table 5.3). The optimal quantities for Model O_i , Model N_i , Π_k^N , and Π_k^M are listed in Table 5.4.

5.3 Quantity allocation based on Lagrangian relaxation

In section 5.1, we utilize line search with 0.05 steps to find the optimal Π_k^N and Π_k^M . Line search takes much time to finish the entire process. We think of using Newsvendor Model, which is a more efficient way to approach the solution.

Show	q_{1kt}^{O*}	q_{2kt}^{O*}	q_{3kt}^{O*}	q_{4kt}^{O*}	q_{5kt}^{O*}	q_{1kt}^{N*}	q_{2kt}^{N*}	q_{3kt}^{N*}	q_{4kt}^{N*}	q_{5kt}^{N*}	Π_k^{N*}	Π_k^M
4	0.05	0.05	0.15	0.15	0.10	0.10	0.05	0.15	0.15	0.10	61.93	61.60
6	0.05	0.05	0.10	0.15	0.10	0.05	0.10	0.10	0.20	0.10	136.46	127.48
9	0.15	0.10	0.15	0.05	0.25	0.15	0.10	0.15	0.05	0.25	223.45	223.45
10	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	35.13	35.13
13	0.15	0.25	0.10	0.10	0.10	0.10	0.25	0.10	0.05	0.10	203.73	193.99
23	0.05	0.10	0.10	0.05	0.10	0.05	0.10	0.10	0.05	0.10	28.01	28.01
24	0.05	0.05	0.05	0.05	0.15	0.05	0.05	0.05	0.05	0.20	12.06	12.00
26	0.05	0.10	0.10	0.05	0.10	0.05	0.05	0.10	0.10	0.10	49.38	49.38
27	0.10	0.10	0.10	0.10	0.30	0.15	0.10	0.25	0.20	0.50	540.47	383.58
28	0.05	0.05	0.05	0.05	0.05	0.10	0.05	0.05	0.05	0.10	119.25	113.76

Table 5.4: Quantities of New and Old Models and the Corresponding Profits

Demand functions of five price bands are

$$D_{ikt} = \mu_{ikt} + \epsilon_i,$$

where μ is the result of expected demand. This can be found out using the regression model we obtain in the last section. As for ϵ , it is a value with 0 mean and normally distributed. The standard deviation can be found from the theater data.

The expected profit of the theater is stated as

$$\Pi = \max\left[\sum_{i \in N} r_i \mathbb{E}[\min\{q_{ikt}, D_{ikt}\}]\right].$$

$$s.t. \sum_{i \in N} q_{ikt} \leq K_{0kt}$$

, where $K_{0kt} = K - \sum_{i \in N} s_{ikt}$.

Using Lagrange relaxation, we can get

$$\Pi_k = \max \left[\sum_{i \in N} r_i \mathbb{E}[\min\{q_{ikt}, D_{ikt}\}] + \lambda \left(\sum_{i \in N} q_{ikt} - K_{0kt} \right) \right].$$



With newsvendor model, we can derive

$$\Pi_k = \max \left[\sum_{i \in N} r_i \left\{ \int_0^{q_{ikt}} x f(x) dx + \int_{q_{ikt}}^{K_{0kt}} q_{ikt} f(x) dx \right\} + \lambda \left(\sum_{i \in N} q_{ikt} - K_{0kt} \right) \right].$$

The equation can be transformed into

$$\Pi_k = \max \left[\sum_{i \in N} r_i \left\{ \int_0^{q_{ikt}} x f(x) dx + q_{ikt} [1 - F(q_{ikt})] \right\} + \lambda \left(\sum_{i \in N} q_{ikt} - K_{0kt} \right) \right].$$

Let there are only two price bands for simplification. With first derivation of q_{1kt} , q_{2kt} , and λ to Π , three equations can be found.

$$\begin{aligned} \frac{\partial \Pi_k}{\partial q_{1kt}^*} &= r_{1k} [1 - F_{1k}(q_{1kt}^*)] + \lambda \\ \frac{\partial \Pi_k}{\partial q_{2kt}^*} &= r_{2k} [1 - F_{2k}(q_{2kt}^*)] + \lambda \\ q_{1kt}^* + q_{2kt}^* &= K_{0kt} \end{aligned}$$

The relationship of

$$\frac{r_{1k}}{r_{2k}} \times \frac{1 - F_{1k}(q_{1kt}^*)}{1 - F_{2k}(q_{2kt}^*)} = 1$$

is therefore derived.

We let $\frac{1 - F_{2k}(q_{2kt}^*)}{1 - F_{1k}(q_{1kt}^*)}$ as

$$h(q_{1kt}^*) = \frac{1 - F_{2k}(q_{2kt}^*)}{1 - F_{1k}(q_{1kt}^*)} = \frac{r_{1k}}{r_{2k}} = 1.$$

Three situations might take place:

(1) if $h(0) > \frac{r_{1k}}{r_{2k}} \Rightarrow q_{1kt}^* = 0$,

(2) if $h(K_{0kt}) < \frac{r_{1k}}{r_{2k}} \Rightarrow q_{1kt}^* = K_{0kt}$, and

(3) if $h(0) \leq \frac{r_{1k}}{r_{2k}} \leq h(K_{0kt}) \Rightarrow 0 \leq q_{1kt}^* \leq K_{0kt}$.

When we combine Lagrange relaxation with newsvendor model, it becomes a much more efficient way to calculate the optimal quantities for price bands. However, it can be used only when the sum of all quantities are the remaining capacity of the performance. Lagrange relaxation will result to $\sum_{i \in N} q_{ikt}^* = K_{0kt}$, which is not necessarily the case in our data for we delete the data that $CumSelf_{ikt} > 0.9$ and $CumOther_{ikt} > 0.9$ considering distortion effect when the tickets are about to be sold out. Due to the constraint, we only use line search model in this paper.







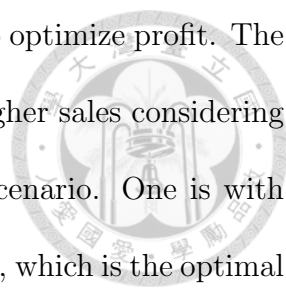
Chapter 6

Conclusions and Future Works

6.1 Conclusions

Theater performance is a type of product with high valuation in comparison to general products. Customers might be more influenced by precedented opinions. Both literature and practical applications conclude that sales information has effects on customers' purchase decisions; however, limited literature on performance ticket selling discuss this issue. We approach the study with a data set of a local theater group of Taiwan using regression. The hypotheses are supported that cumulated sales of a price band has a positive relation to its sales speed and the cumulated sales of other price bands have negative effects. A further hypothesis, that the price bands with closer price have larger effects on one another is also verified. We also obtain the conclusion that length of selling period poses a negative effect to sales speed.

A dynamic seat allocation application is proposed, in which decision makers can switch

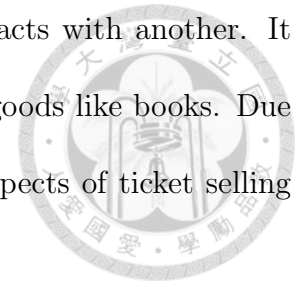


ticket capacities of different price bands of the same performances to optimize profit. The price band with fewer sales can switch capacity to the ones with higher sales considering the effects on customer valuation. There are two models in the scenario. One is with basic performance information and another adds in sales information, which is the optimal model in the case. Ten performances with five price bands are selected to test the profit difference obtained by the two models. With quantity control, we find that the model with additional sales information gains 4.60% of profit on average compared with the suboptimal model and 60% of the performances have profit difference.

6.2 Future works

A number of research problems are uncovered due to data limitation which presents future opportunities for future research in this area. The data we obtained is limited with only price and date of the performances. There are still a lot of data that can be tracked as many performance tickets are sold by online systems. Customers need to click many steps before making a purchase. When those customer behaviors are documented, people can discuss topics related to purchase process and link the information to ticket sales speed and cumulated sales. Another information we do not have is seat numbers. When seat number data is combined with sales data, we can map our analysis on to the scope of actual theater in a much more realistic scope. A price band is consist of plenty of seats with varied views. Most consumers think of the seat location and the cost-performance ratio in corresponds first when making a purchase. The result of the correlation might vary among theaters. Ticket selling has many interesting characteristics. As mentioned

before, it has various price bands in a performance and one interacts with another. It has many uncertainties in comparison to that of normal physical goods like books. Due to the growing convenience of data acquisition, there are many aspects of ticket selling that can be discussed in the future.







Appendix A

Figures of Exploratory Data

Analysis

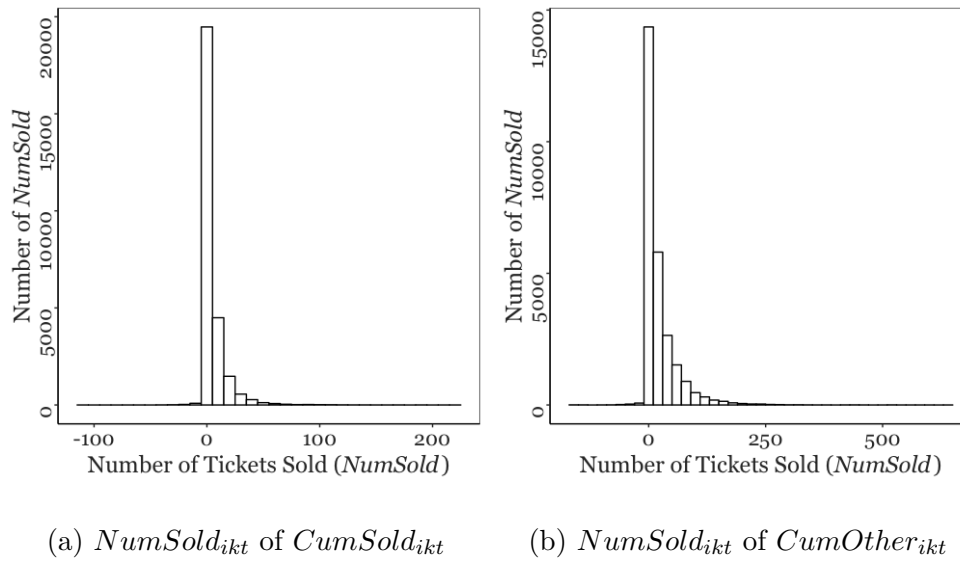
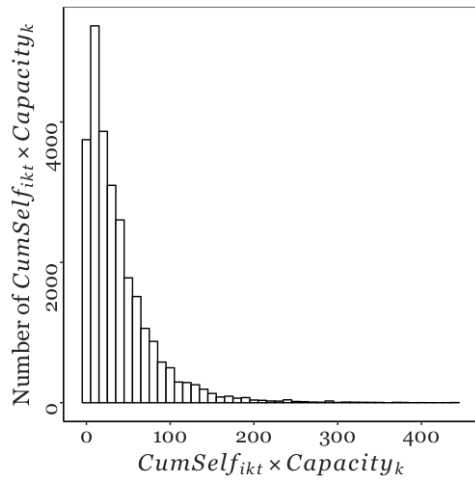
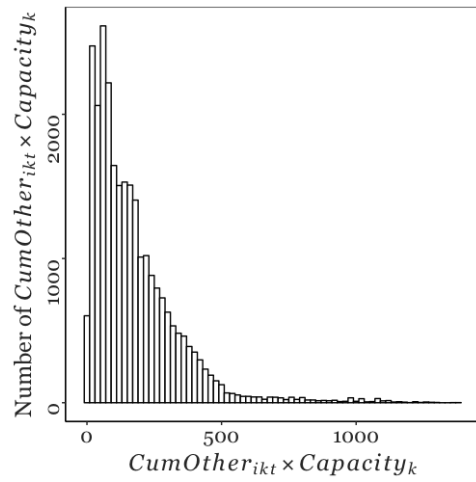


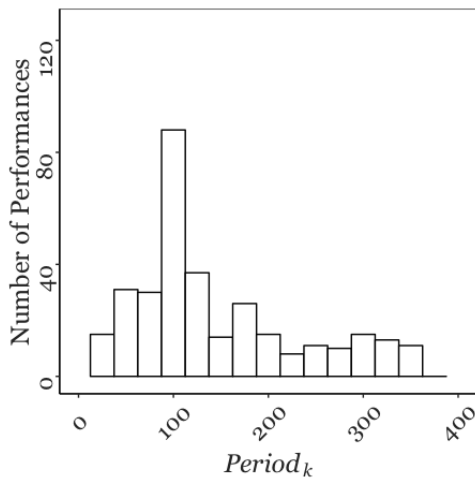
Figure A.1: The First Part of Experiment Analysis.



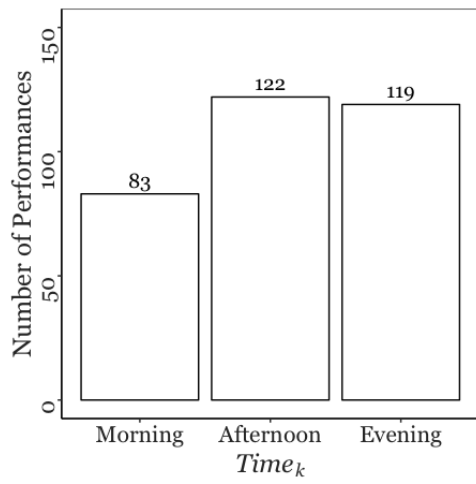
(c) Cumulated Sales of $CumSelf_{ikt}$



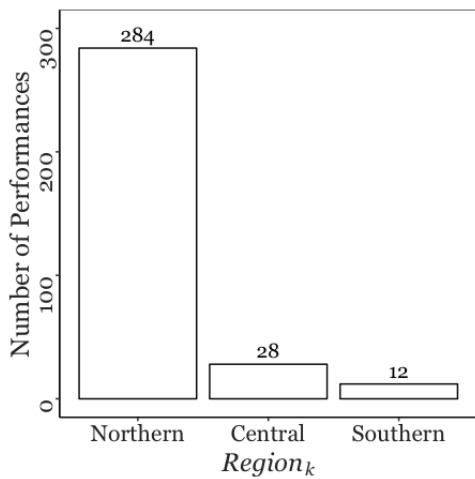
(d) Cumulated Sales of $CumOther_{ikt}$



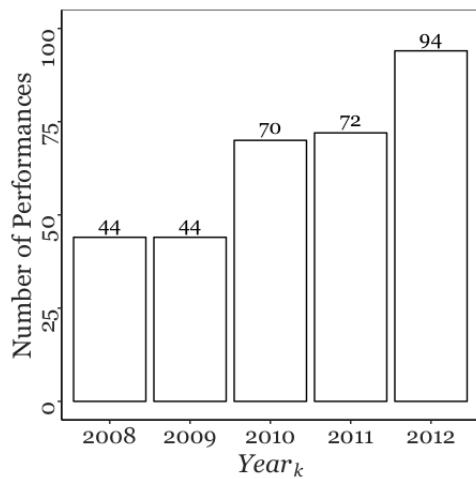
(e) $Period_k$



(f) $Time_k$

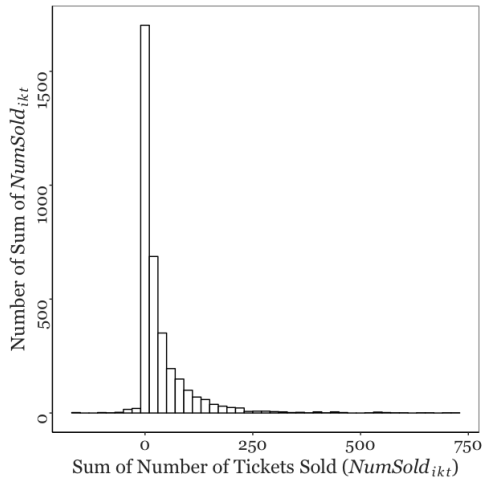


(g) $Region_k$

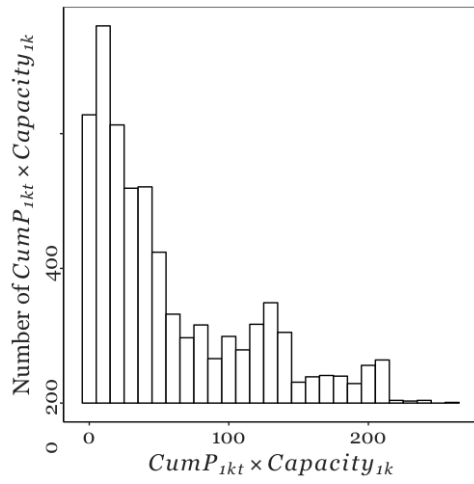


(h) $Year_k$

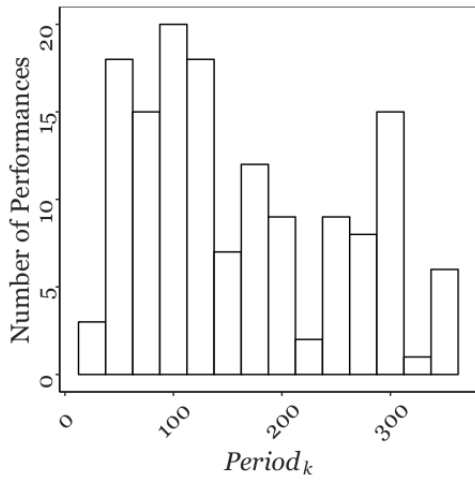
Figure A.1: The First Part of Experiment Analysis (cont.).



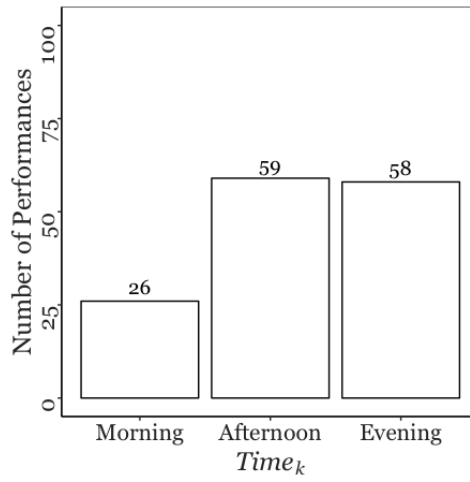
(a) $NumSold_{ikt}$ of Performances



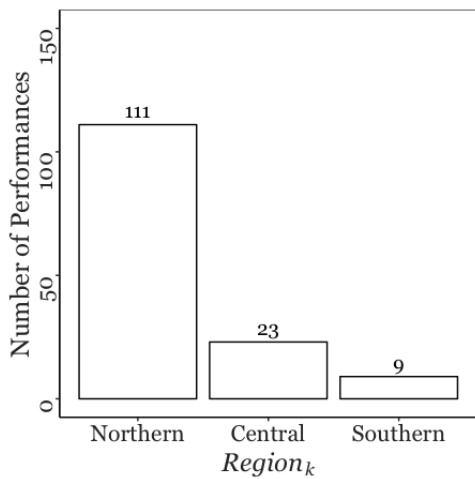
(b) Cumulated Sales of P_{1kt}



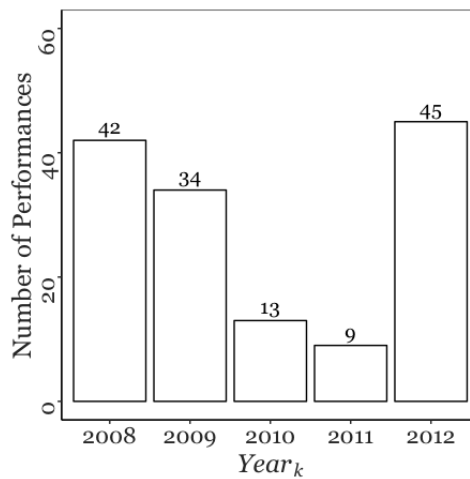
(c) $Period_k$



(d) $Time_k$



(e) $Region_k$



(f) $Year_k$


Figure A.2: The Second Part of Experiment Analysis





Bibliography

- Alevy, J. E., M. S. Haigh, J. A. List. 2007. Information cascades: Evidence from a field experiment with financial market professionals. *The Journal of Finance* **62**(1) 151–180.
- Anderson, E. W. 1998. Customer satisfaction and word of mouth. *Journal of service research* **1**(1) 5–17.
- Bikhchandani, S., D. Hirshleifer, I. Welch. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy* 992–1026.
- Bone, P. F. 1995. Word-of-mouth effects on short-term and long-term product judgments. *Journal of business research* **32**(3) 213–223.
- Boyd, D. W., L. A. Boyd. 1998. The home field advantage: Implications for the pricing of tickets to professional team sporting events. *Journal of Economics and Finance* **22**(2-3) 169–179.
- Chen, Y., Q. Wang, J. Xie. 2011. Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of marketing research* **48**(2) 238–254.
- Courty, P. 2002. Ticket pricing under demand uncertainty .

- 
- Courty, P. 2003. Some economics of ticket resale. *The Journal of Economic Perspectives* **17**(2) 85–97.
- Drayer, J., D. A. Rascher, C. D. McEvoy. 2012. An examination of underlying consumer demand and sport pricing using secondary market data. *Sport Management Review* **15**(4) 448–460.
- Gallego, G., G. Van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management science* **40**(8) 999–1020.
- Hall, R. L., C. J. Hitch. 1939. Price theory and business behaviour. *Oxford economic papers* (2) 12–45.
- Herr, P. M., F. R. Kardes, J. Kim. 1991. Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnostics perspective. *Journal of consumer research* **17**(4) 454–462.
- Ho, E., T. Kowatsch, A. Ilic. 2014. The sales velocity effect on retailing. *Journal of Interactive Marketing* **28**(4) 237–256.
- Howard, D. R., J. L. Crompton. 2004. Tactics used by sports organizations in the united states to increase ticket sales. *Managing Leisure* **9**(2) 87–95.
- Hsieh, C. Y. 2015. The impact of pricing on tickets sales in the theater industry.
- Huntington, P. A. 1993. Ticket pricing policy and box office revenue. *Journal of Cultural Economics* **17**(1) 71–87.
- Kimes, S. E. 1989. The basics of yield management .

Kotler, P. 2009. *Marketing management: A south Asian perspective*. Pearson Education India.



Kung, M., K. B. Monroe, J. L. Cox. 2002. Pricing on the internet. *Journal of Product & Brand Management* **11**(5) 274–288.

Leslie, P. 2004. Price discrimination in Broadway theater. *RAND Journal of Economics* 520–541.

Liang, X., L. Ma, L. Xie, H. Yan. 2014. The informational aspect of the group-buying mechanism. *European Journal of Operational Research* **234**(1) 331–340.

Liu, Y. 2006. Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing* **70**(3) 74–89.

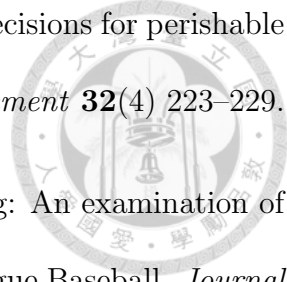
McAfee, R. P., V. Te Velde. 2006. Dynamic pricing in the airline industry. *forthcoming in Handbook on Economics and Information Systems, Ed: TJ Hendershott, Elsevier* .

Nahmias, S. 1982. Perishable inventory theory: A review. *Operations research* **30**(4) 680–708.

Ramanathan, R. 2006. Stocking and discounting decisions for perishable commodities using expected profit approach. *International Journal of Retail & Distribution Management* **34**(2) 172–184.

Rascher, D. A., C. D. McEvoy, M. Nagel, M. T. Brown. 2007. Variable ticket pricing in major league baseball. *Journal of Sport Management, Forthcoming* .

Rosen, S., A. M. Rosenfield. 1997. Ticket pricing 1. *The Journal of Law and Economics* **40**(2) 351–376.

- 
- Sezen, B. 2004. Expected profit approach used in discount pricing decisions for perishable products. *International Journal of Retail & Distribution Management* **32**(4) 223–229.
- Shapiro, S. L., J. Drayer. 2012. A new age of demand-based pricing: An examination of dynamic ticket pricing and secondary market prices in Major League Baseball. *Journal of Sport Management* **26** 532–546.
- Tseng, Y. C. 2016. The influence of herd behavior on sales volume and inventory in online shopping. *Department of Internet Business, Chung Yuan Christian University* 1–122.
- Tsiros, M., C. M. Heilman. 2005. The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories. *Journal of marketing* **69**(2) 114–129.
- Watts, D. J. 2002. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences* **99**(9) 5766–5771.
- Weatherford, L. R., S. E. Bodily. 1992. A Taxonomy and Research Overview of Perishable-Asset Revenue Management: Yield Management, Overbooking, and Pricing. *Operations Research* **40**(5) 831–844.
- Welch, I. 2000. Herding among security analysts. *Journal of Financial economics* **58**(3) 369–396.
- You, P. S. 1999. Dynamic pricing in airline seat management for flights with multiple flight legs. *Transportation Science* **33**(2) 192–206.
- Zhao, W., Y. S. Zheng. 2000. Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management science* **46**(3) 375–388.