Università
Ca'Foscari
Venezia

Scuola Dottorale di Ateneo
Graduate School

Dottorato di ricerca
in Management
Ciclo XXIX
Anno di discussione 2016/2017

# *Three Essays on Revenue Management in Nonprofit Performing Arts Organizations*

SETTORE SCIENTIFICO DISCIPLINARE DI AFFERENZA: SECS-S/06

**Tesi di Dottorato di Baldin Andrea, matricola 815183**

Coordinatore del Dottorato

**Prof.ssa Anna Comacchio**

Supervisore del Dottorando

**Prof. Andrea Ellero**

Co-tutore del Dottorando

**Prof.ssa Daniela Favaretto**

# Three Essays on Revenue Management in Nonprofit Performing Arts Organizations

Andrea Baldin

A thesis submitted to the Graduate School of Ca' Foscari University of Venice, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2016

## Acknowledgements

First and foremost, I would like to express my gratitude to my two advisors: Andrea Ellero and Daniela Favaretto. They have provided competent guidance and continuous support for all my initiatives.

A special thanks to Trine Bille for her great support during my visiting period at Copenhagen Business School and even after I left her Department. Collaborating with her was really inspiring.

I would like to thank the members of the Programming, Strategic Analytics and Sales Department of the Royal Danish Theatre for their interest in this dissertation and for the data provided to me.

I gratefully acknowledge Elisabetta Lazzaro (Utrecht University) and Marilena Vecco (Erasmus University Rotterdam) for their valuable comments to the manuscript and their constructive suggestions.

Finally, I am grateful for financial support by the Department of Management at Ca' Foscari University of Venice.

# Contents

# Introduction

The purpose of this dissertation is to demonstrate the effectiveness of Revenue Management (RM) techniques in nonprofit performing arts organizations, although such organizations are characterized by multiple and conflicting objectives other than revenue.

RM can be defined as a set of optimization strategies that match supply and demand by acting on the prices and the availability capacity (Smith *et al.*, 1992). The aim of RM is to maximize the revenue by allocating the right capacity to the right customers at the right price. Since its inception in the 1980's, RM practices have quickly developed in the transportation industry, and a considerable amount of research has been accomplished in the past three decades looking at the issues detectable in this industry. However, there are other sectors, including the performing arts organizations, that offer a considerable number of RM-type problems that have not yet been fully addressed. Indeed, theaters as providers of services are endowed with the features that complies with Kimes' (1989) preconditions for a successfull application of RM: the capacity constraint, that makes the marginal costs of providing the service to one more customer much smaller than the average cost, unless capacity is full; the perishability of the product offered (i.e, once the performance starts, an unfilled seat is worthless and the revenue is lost forever); high fixed costs and low variable costs (once a performance has been staged, the cost of an additional performance is relatively small).

A decidedly small amount of research have been conducted concerning RM in nonprofit performing arts organization. One reason that may explain this lack in attention is that RM is not a pervasive practice in most of these organizations. Moreover, it can be argued that RM practices are in conflict with the nonprofit nature of public-subsidized theatres (Lariviere, 2012). This contradiction can be resolved by considering that nonprofit organizations engage also in for-profit activities, in order to generate excess revenue to subsidize activities more involved with their mission (De Vericourt and Sousa Lobo, 2009). Following this perspective, it can be claimed that RM is a tool for the achievement of a nonprofit performing arts firm's objectives. These are, according to literature (Hansmann, 1981; Luksetich and Lande, 1995): the maximization of the budget to administer; the maximization of the quality of the services provided and the maximization of the audience. For example, an opera house can invest the revenue generated by a popular performance (which is expected to be a high-demand event), in the production of other opera activities that are artistically important but less lucrative, satisfying the quality goal. Moreover, the price discrimination practice, which is an example of an RM technique, enables people who are supposed to be less able to pay, to attend a theatrical production (audience maximizer goal).

Thus, contextualizing the theory and practice of RM in the field of nonprofit performing arts offers an interesting and unique laboratory for the research, but faces also specific challenges and requires a special attention. Apart from the multidimensional nature of their objectives, it is necessary to consider other distinguishing characteristics of this sector. Cultural institutions take the form of "hybrid organizations" (Glynn, 2000) in which two identity elements - the normative artistry and the utilitarian economics - coexist and may come into conflict with each other. In this framework, managers of cultural organizations deal with five polarities, outlined by Lampel *et al.* (2000), that shape organizational practices[1]. Although the dualism between economic and artistic/cultural imperatives is present in non profit performing arts organizations, we believe that RM practices can work in such kind of organizations as an aid to theatre managers in managing demand and taking decision on variables (quantity and price) that must be quantified and that can take different values according to the preference over the different goals to achieve. Indeed, the same RM models applied in commercial sector can be adapted in non profit sector in order to accomodate an utility function that falls along some continuum between conflicting objectives.
Another peculiar characteristic of these organizations, compared to traditional RM industries, regards the nature of the product offered. Cultural goods are experiential goods which value falls to a large extent outside the boundaries of purely economic value. Their value is more related to abstract, subjective and experience-related aspects of the product. Therefore, it is difficult to understand why people choose to consume what they do (Caves, 2000). Similarly, for customer it is difficult to assess the quality and the value of a cultural product before committing to consume it. This leads to the demand uncertainty faced by performing arts organizations (the so-called "nobody knows" property). Given the difficulty of defining the notion of quality, in this work we consider the quality as given, at least when we consider it as an objective of the theatre. A different discourse should be made when the quality is incorporated in the empirical analysis of consumer demand: as Throsby (1990) notes, it is extremely hard to take into account the quality in its dimension derived from a theory of aestheticism. However, some components of quality can be observed in advance and used in the empirical analysis of demand: genre of the production, popularity, period of creation and so on. Moreover, as we do in the

[1]This polarities are: artistic values versus mass entertainment, product differentiation versus market innovation, demand analysis versus market construction, vertical integration versus flexible specialization and individual inspiration versus creative system

first paper, a measure of quality, that is supposed to influence the demand, can be incorporated through professional reviews (e.g. in the newspaper) and audience evaluations (word of mouth mechanism).

The dissertation is composed of three papers, each of which examines a RM-type problem topic that a nonprofit theatre may face. In conformity with the RM logic, in these papers we consider the earned-revenue derived from the theatre activities, without considering the other sources of revenue. This clarification allows to generalize the results to other countries, even if the case studies (for the first and third paper) are Danish, as the main difference between Europe and American non profit performing arts organization derives by a different source of non-earned income (direct subsidies by government for the former; private donations encouraged by tax deduction for the latter; see Brooks, 2006).

Research on RM blends element of several disciplines such as marketing, operations management, operational research, microeconomics, behavioural economics and industrial organization. Despite its multidisciplinary nature, the mainstream RM literature is characterized by the usage of quantitative analytical methods. This approach is justified by the problem-solving nature of RM as a discipline that focuses on how the demand-management decisions (price, quantity and structural decisions) are scientifically made, through a "technologically sophisitcated, detailed, and intensely operational approach" (Van Ryzin and Talluri, 2005). As such, RM leverages tools from statistics, econometrics and operational research literature in order to model demand, estimate and forecast market response, and find solutions to complex decision problems[2]. In compliance with the dominant practice, this dissertation makes use of quantitative methods adopting tools from both microeconometrics (first paper) and operational research (second paper), also integrating both of them (third paper).

The first paper investigates customers choice behaviour with respect to the purchase of a theater ticket. In particular, it analyzes the extent to which the different attributes that are source of price discrimination affect the choice of the ticket, and how this effect differs among the theatregoers, assuming heterogeneity in preferences. To this purpose, two modelling approaches; multinomial logit (with socio-demographic characteristics) and latent class are applied to a dataset for the period 2010-2013 from the sale system of the Royal Danish National Theatre. Final results

[2]This does not exclude the possibility to adopt a qualitative approach to RM issues: for example Mitev (2009), using the Actor-Network Theory, analyzes the causes of the failure of a computerised reservation system (RM technology) in the National French rail company. In this perspective, the author is interested in how the introduction of a RM system is translated and interpreted by the organization and its end-users.

suggest that customers characteristics, in terms of age and frequency of theater attendance, characterize different patterns of behavior in the choice of theater ticket. Moreover, with the results of these models it is possible to estimate the willingness to pay of each choice attribute and how it differs among customer categories, providing so guidance to theatre managers in setting prices.

The second paper analyzes a structural-based RM problem. Assuming a theatre proposes both highbrow and lowbrow events, this paper tackles the issue of identifying the most efficient subset of the events scheduled by a theatre to offer as a subscription. The problem is formulated following the choice-based network RM perspective (Liu and Van Ryzin, 2008), relying on its definition of efficient offer set as the one that provides the most favourable trade-off between expected revenue and capacity consumption. Based on this approach, an integrated model that makes use of the super-efficiency data envelopment analysis (DEA) and a probabilistic approach is formulated. Indeed DEA seems to be a suitable tool in our context, given the multi-objective nature of the non profit organizations; whereas the probabilistic approach models the purchase decision on the basis of two random variables: available time and reservation price per perfomance. The results of different simulations are presented considering a range of values for the theatre capacity and the discount rate of the bundle. Moreover, an econometric analysis is carried out to obtain some insights into what determines the efficiency level of a subscription.

The third paper develops a bi-objective optimization model that simultaneously considers pricing and seat allocation, assuming that the theater wants to optimize both attendance and revenue. The proposed model integrates the demand forecast with a customer choice model, accounting for the difference in price sensitivity and seating area preference. Finally the model is validated with booking data from the Royal Danish Theatre during the period 2010-2016. Results obtained confirms the existence of a trade-off among the two theater objectives, each of which correspond to different pricing and allocation policies.

Overall, the three papers should demonstrate the potentiality of RM techniques for assisting theater managers in their decision-making process for what concerns the demand-management decisions problems:price, quantity and structural decisions. We think that this dissertation has added value to literature by intersecting two area of research that have interacted little with each other. Indeed, whereas RM literature has devoted greater attention to issues detectable in for profit industries, the culture economics literature has not give a great deal of attention to the potentiality of RM techniques and its effects to both the demand side and the achievement of the objectives of cultural organizations. The three papers presented here aim to demonstrate that this interaction is possible, by analyzing three issues that, at best of our knowledge, have never been considered in the literature of non profit performing arts organizations. However, further research is needed in this direction, especially in light of the fact that nowdays there is a tendency to reduce the public funds allocated to cultural organizations, forcing the latter to increase their self-earned income.[3]

# References


- Brooks, A. C. (2006). Nonprofit firms in the performing arts. *Handbook of the Economics of Art and Culture*, 1, 473-506.

- Caves, R. (2000). *Economics of the creative industries*. Cambridge, Mass: Harvard.

- De Vricourt, F., Lobo, M.S. (2009). Resource and revenue management in nonprofit operations. *Operations research*, 57, 1114-1128.

- Glynn, M. A. (2000). When cymbals become symbols: Conflict over organizational identity within a symphony orchestra. *Organization science*, 11, 285-298.

- Hansmann, H. (1981). Nonprofit enterprise in the performing arts. *The Bell Journal of Economics*, 341-361.

- Kimes, S. E. (1989). The basics of yield management. *The Cornell Hotel and Restaurant Administration Quarterly*, 30, 14-19.

- Lampel, J., Lant, T., Shamsie, J. (2000). Balancing act: Learning from organizing practices in cultural industries. *Organization science*, 11, 263-269.

- Lariviere, M. (2012). *Should a non-profit use revenue management techniques?*. Retrieved from https://operationsroom.wordpress.com/2012/10/25/should-a-non-profit-use-revenue-management-techniques/

- Liu, Q., Van Ryzin, G. J (2008). On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management*, 10, 288-310.

- Luksetich, W. A., Lange, M. D. (1995). A simultaneous model of nonprofit symphony orchestra behavior. *Journal of Cultural Economics*, 19, 49-68.


[3]The three papers are preliminary versions. Further comments received are going to be included in view of future submission to journals.

- Mitev, N. (2009). In and out of actor-network theory: a necessary but insufficient journey. *Information Technology & People*, 22, 9-25.

- Smith, B.C., Leimkuhler, J.F., Darrow, R.M. (1992). Yield management at American airlines. *Interfaces*, 22, 8-31.

- Throsby, C. D. (1990). Perception of Quality in Demand for the Theatre. *Journal of cultural economics*, 14, 65-82.

- Van Ryzin, G. J., Talluri, K. T. (2005). An introduction to revenue management. *Tutorials in operations research*, 142-195.

# Modelling preference heterogeneity for theater tickets: a discrete choice modelling approach on Royal Danish Theater booking data

Andrea Baldin*[1] and Trine Bille†[2]

[1] *Department of Management, Ca' Foscari University of Venice, Cannaregio 873, 30121 Venice, Italy*

[2] *Department of Management, Politics and Philosophy, Copenhagen Business School, Porcelænshaven 18B, 3.floor, 2000 Frederiksberg, Denmark*

**Abstract**

This paper analyzes the behavioral choice for theater tickets using a a rich data set for 2010-2013 from the sale system of the Royal Danish National Theatre. A consumer who decides to attend a theater production faces multiple sources of price variation that depends on: socio-economic characteristics, quality of the seat, day of the performance and timing of purchase. Except for the first case, factors of price differentiation involves a choice by the consumer among different ticket alternatives. Two modelling approaches are proposed in order to model ticket purchases: multinomial logit (MNL) with socio-demographic characteristics, and latent class. These models allow us explicitly to take into account consumers' preference heterogeneity with respect to the attributes associated with each ticket alternative In addition, the willingness-to-pay (WTP) of choice attributes is estimated. Understanding theater-goers' choice behavior and WTP for the quality of seat and the day of performance is important to policy makers and theater managers in adopting

*E-mail: andrea.baldin@unive.it

†E-mail: tbi.mpp@cbs.dk

Authors' contributions: Andrea Baldin contributes to the paper by conceiving and defining the design of the study, doing the literature review, analyzing and interpreting the data, estimating the model and most of the writing. Trine Bille contributes to the paper by collecting the data, writing the part related to the Royal Danish Theatre and its pricing policy, and revising the paper critically.

different pricing and marketing strategies. Final results suggest that customers' characteristics in terms of age, frequency of theater attendance and period of ticket purchase characterize different patterns of behavior in the choice of theater ticket.

**Keywords** Theater demand - Discrete choice models - Price discrimination - Willingness to pay

# 1 Introduction

During recent years, revenue management (RM) and price discrimination techniques have played an increasing role in the performing arts sector. Evidence shows how theaters can charge different prices for the same production. This practice is driven, on the one hand, by the social duty consisting of allowing consumer segments who are supposed to be less able to pay, to attend a theatrical production; and on the other hand, by the possibility of extracting part of the consumer's surplus. An example of the first situation is exhibited in discount tickets offered to certain social categories (e.g. students, youth, senior citizens...). In the latter case, the theater provide consumers with incentives to discriminate among themselves by offering a schedule of different prices according to the quality of the seat. In this way, each consumers will choose the seat location in the venue according to reservation price and preference. Another form of price discrimination is made explicit through a variation in prices, both in the full and in the discount ticket price, according to the day of the performance: for example, a ticket for a Saturday night performance is usually more expensive than a ticket for a weekday performance. This kind of differentiation refers to the peak-load pricing issue that takes into account the capacity constraint of the theater, increasing the price when the demand is high and decreasing it when deman is low.

The pricing strategies described above are perfectly coherent with the different objectives that are pursued by a non-profit performing arts organization, as described by Hansmann (1981): in enabling people with lower willingness to pay (WTP) to attend a performance, the theater satisfies the objective of maximizing the attendance, while the appropriation of consumers' surplus fulfills the budget goal, generating excess revenue to subsidize activities that can be less lucrative but artistically important (quality goal). After all, as Rosen and Rosenfield (1997) pointed out, price discrimination is observed in activities, such as the performing arts, where the marginal costs of providing the service to one more customer is smaller than the

average cost: the additional cost to fill one more seat in a theater is in fact quite small.
As attendees can choose among different ticket alternatives, it is crucial to understand their behavior in order to support pricing strategies. Using a unique sale system data set from the Royal Danish National Theatre during the period 2010/2011 to 2012/2013, we aim to analyze which attributes affect the choice of theater ticket. Indeed, the Royal Danish Theatre provides a good example of discriminatory pricing. Taking advantage of this rich data set, this study adopts revealed preference (RP) design approach (i.e choice based on actual market behavior) as opposed to a stated preference (SP) approach (i.e choice based on hypothetical scenarios). Both approaches are founded in the theory of consumer demand postulated by Lancaster (1966) and present advantages as well as disadvantages: in order to encompass enough variation in the level of attributes, RP requires a large amount of data, whereas SP is more flexible in its data requirements, providing new non-existing alternatives in the hypothetical scenarios presented to the respondents. However, the main drawback of SP is the risk of response bias under experimental conditions which, according to Carrier (2008), seem to be high for pricing applications. Given the details and wideness of our data set, we adopt an RP perspective in this study.
From a methodological point of view, we compare two different approaches to discrete choice analysis: MNL with socio-demographic specification, and latent class models (LCM). Whereas the MNL model includes interaction terms with socio-demographic terms in order to account for heterogeneous preference, the LCM approach allows the parameters of the utility function to vary across agents according to a probabilistic discrete distribution. As Green and Hensher (2003) pointed out, LCM is supported by strong statistical foundations and has a clear interpretation as it identifies different clusters of customers, each of which is characterized by a specific value of the parameters. Therefore LCM is appealing from both a marketing and a policy perspective as it distinguishes, along behavioral variables, distinct classes of customers characterized by different price sensibilities and WTP. Moreover, LCM overcomes the independence of irrilevant alternatives (IIA) restriction of MNL, according to which the odds of choosing one alternative over another alternative is not altered by the addition of a new alternative.
The assumption of heterogeneity seems to be realistic in the theater sector: empirical studies on demand for the performing arts have shown ambiguous values of price elasticity, in some cases even a positive elasticity, configuring the theatrical experience as a Veblen good (Laamanen, 2013). Indeed, many of these studies use aggregated data and the average price (revenue divided by attendance) to estimate

price elasticity. Studies that have accounted for different sources of price variation (as our data set allow us to do) produce different estimates of levels of price elasticity. Hence, the literature confirms how is heterogeneity among customers in the price sensitivity.

This paper aims to investigate this preference heterogeneity by analyzing the choice of ticket theater. Unlike previous research on theater demand, we consider the wide range of prices available to customers. For this aim, we adopt a discrete choice modeling approach and estimate the different WTP for the choice attributes. This approach is widely used in the transportation industries (airline and railway in particular); to the best of our knowledge, Willis and Snowball (2009) and Grisolia and Willis (2011a, 2011b, 2012) are the only researchers to have applied discrete choice modeling in the performing arts sector. They have investigated preference for the different attributes of theatrical production (as venue, repertory classification, word of mouth, type of play, author and review). In addition to the different attributes of theatrical production, we consider also the attributes that are sources of price differentiation: seat category, attributes of the different performances for the same production (day, premiere or not), consumer category. This is the main contribution of this study: providing a new segmentation of the theater demand. This may have important implications for policy makers and theatre managers, as the identification of market segments with different WTP for a theatrical attribute is relevant to pricing and marketing strategies.

The structure of the paper can be outlined as follows: Section 2 reviews the literature on demand for the performing arts and price discrimination in the theater sector; Section 3 offers a description of the Royal Danish Theatre and its price discrimination policy; Section 4 describes the models implemented; Section 5 describes the data set and the variables used; and Section 6 shows the final result. Finally, Section 7 provides some conclusions and implications of our research.

# 2 Literature review

This study follows two main streams of literature. The first relates to the determinants of demand for the performing arts. Many studies have aimed to identify elasticity with respect to price and/or income. This topic has been widely analyzed over a long period, starting from Gapinski (1984): we refer to Seaman (2006) for a comprehensive review of the literature on performing arts demand. In addition to price, other variables have been included as determinants of performing arts attendance, for example the price of substitutes (e.g. Colbert *et al*, 1998; Zieba

2009), quality indicators (e.g. Throsby, 1990; Urrutiaguer, 2002), type of play (e.g. Abbé-Decarroux, 1994; Corning and Levi, 2002) and socio-economic variables such as education level and availability of time (e.g. Werck and Heyndels, 2007; Swanson *et al*, 2008).

The papers most related to ours are those that infer consumer heterogeneity through attributes that underlie price discrimination. This implies the adoption of disaggregated data for price measure and demand. One of the classic segmentations is based on whether the consumer is a subscriber or not. Felton (1994) analyzed the demand of 25 large US orchestras and estimates two different regressions: the first considers only subscribers, whereas the second also includes the single ticket holders. The author obtained a lower price elasticity for the subscribers (-0.24) than the total attendance price elasticity (-0.85). Colbert *et al.* (1998), through a survey conducted among the audience of seven Canadian theaters, identified two segments of consumers in both the subscribers and non-subscribers groups according to their sensitivity to price: (a) those who show a high price elasticity are rich in time and poor in money; and (b) the reverse.

Abbé-Decarroux (1994) estimated demand for a Geneva theater company, distinguishing two kind of tickets: full-price tickets and reduced-price tickets, the latter for students, seniors and unemployed. As expected, a higher price elasticity (-2.45) was found for the latter consumer group, whereas there was price inelasticity for the former, full-price ticket group, for whom the price coefficient was not statistically significant. Schimmelpfennig's (1997) paper employed a non-parametric linear regression analysis to demand for the Royal Ballet Summer Season, a special event organized by the Royal Opera House, Covent Garden. The main characteristic of this paper is that it focused on the individual seat categories. Surprisingly, for both the productions examined, the Orchestra Stalls showed a higher price elasticity than the cheapest seat category (denoted as Rear Amphitheater), which is supposed to serve low-income consumers.

Corning and Levy (2002), instead of estimating different equations for subscribers and single-ticket holders, decided to model the effect of number of subscribers and price of subscription on the demand for single tickets, including them as explanatory variables in the single full-priced tickets equation. An interesting result was that subscription sales had a weak effect on the demand for single tickets, hence the two different segments had little overlap. A remarkable characteristic of their work was the inclusion of variables related to the time of performance (e.g matinee, evening, preview),which are shown to be highly correlated with scheduled price, and to seasonality effects (monthly dummy variables): the final results indicated a significant

positive effect for evening and weekend performances.
Laamanen [2013] used eight years of sales system data of the Finnish National Opera to estimate demand for opera for both premiere season and reprises. For demand for the former the price elasticity is fairly small (-0.69), whereas the demand for reprises is highly elastic (-3.99). What distinguishes this paper from the previous one is not only the estimation method based on censored quantile regression, which allows accounting for the capacity constraint, but also the disaggregation of ticket sales by area of seating and price category. In this way, the researcher was able to avoid bias estimation of price elasticity, which results when the average price ticket and aggregated data are used in the demand estimation.
From a methodological perspective, discrete choice models have already been used in the cultural economics domain: in particular LCM were employed to explain heterogeneity in culture consumption (Chan and Goldthorpe, 2005) and cinema attendance (Fernandez-Blanco *et al.*, 2009). In the theater sector, discrete choice modeling has been used in order to assess preference for theatrical attributes (as venue, repertory classification, word of mouth, type of play, author and review) and to estimate the WTP for each attribute. In particular, Willis and Snowball (2009) and Grisolia and Willis (2011a, 2011b) used an SP discrete choice experiment using MNL and mixed logit models; while Grisolia and Willis (2012) employ an LCM that allows audience segmentation according to preferences for attributes of theatrical productions (e.g. repertory classification, type of play, author, review). Their results suggest a heterogeneous effect of such attributes on consumer choice.
The second stream of literature relates to the application of RM and price discrimination techniques in performing arts organizations. As RM is an area of research that has wide application in in highly commercial industries (e.g. the airline and hotel industries) very little empirical research has been done in the cultural sector. Most RM research in this area has focused on the price discrimination practice implemented by theaters. Huntington (1993) used a variant of the hedonic price model to describe price differentiation by seat. This model implies that, if there are observable differences between seats, a price discrimination policy can be adopted. Moreover, the author showed that the price discrimination policy leads to a greater profit than the unique price policy. From a theoretical perspective, Rosen and Rosenfield (1997) described a model of price discrimination focusing on the issue on how the theater should sort and price seats in categories, in order to maximize revenue. In the model proposed, the theaters have two qualities of seat (high and low) and the seller knows the intensity of the aggregate demand for each category and its distribution. Leslie's (2004) paper is considered one of the most important

pieces of research on pricing strategies in the performing arts field. The author analyzed the price discrimination policy for the Broadway show *Seven Guitars*, estimating a structural econometric model based on individual consumer behaviour. Tereyagoglu *et al.* (2012) employed a competing hazard framework to model the ticket sales, where the customers race against each other for the same ticket. The aim of their work was to analyze how pricing and discount actions over time affect the timing of customers' purchasing, as well as the propensity of different categories of customers (subscribers and occasional buyers) to purchase a ticket.

This review of the literature highlights the need to use disaggregated data over price category and performance, in order to analyze consumer behavior towards the price discrimination policy. Given the structure of our data set, a discrete choice model that accounts for heterogeneous preference in an RP setting seems to be the most suitable approach.

# 3 The price discrimination policy of the Royal Danish Theater

The Royal Danish Theater was founded in 1748 and is the Danish national theater. It has three main stages in Copenhagen. The Old Stage from 1874, a new Royal Opera House from 2005 and a new Royal Playhouse from 2008. The Opera House and the Playhouse have a main stage and smaller stages for experimental productions. The Royal Danish Theater is one of the few theaters in the world offering opera, ballet and theater performances as well as classical concerts. Before the two new houses were built, the Old Stage offered opera, ballet and theatre performances. Now the Old Stage is the house where ballet is performed.

The price discrimination policy by seat tier has been refined in recent years. In 2010 the Opera House and the Old Stage offered 5 different price zones, whereas now the price variation involves 8 different seat categories. Figure 1 shows the price discrimination by seat categoria related to the Opera House in 2016. A different policy is adopted concerning the New Playhouse where discrimination by quality of seat (up to the maximum of 5 price zones in the theater) is applied to only a few productions. [1]

Besides by price zones, each ticket sold is characterized by the price type, which is connected to the characteristics of the buyers that influence the price charged. In this study we have excluded some price types, such as the categories for which the

[1] Clearly, our sample of productions includes only those for which price discrimination by seat quality is applied.

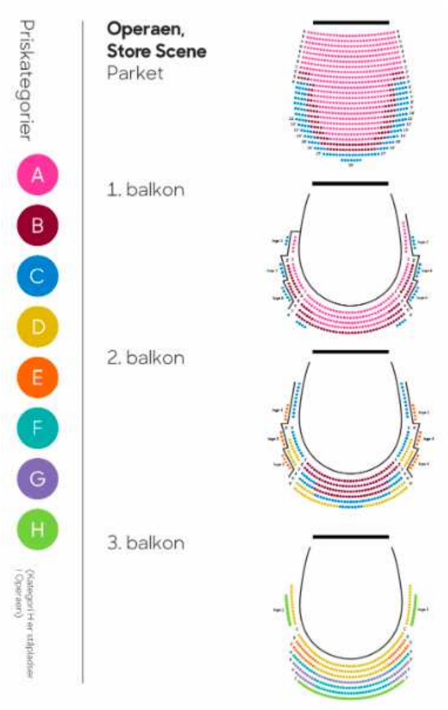


Figure 1: Price discrimination of the Opera House seats

ticket is free (press, sponsor, guests, attendant for disable, employees) and group sales. Moreover, we have excluded performances with a flat price, where the price is fixed regardless of seat choice[2], rush tickets (discounted by 50%) and those tickets that are discounted as the result of an advertising campaign. The logic behind this selection lies in the fact that these types of tickets either do not show a trade-off between price and seat tier, or do not give the customers the opportunity of a complete choice of seat category and/or day of performance. Table 1 shows the price types considered in our model.

| | **Price type** | **Price type group** | **Price type category** | **Discount (%)** |
|---|---|---|---|---|
| 1 | Standard | Standard | Standard | 0 |
| 2 | Youth[a]/Student | Social awareness | Discount | 50 |
| 3 | Senior citizen[b] | Social awareness | Discount | 50 |
| 4 | Theater card (Loyalty card) | Loyalty | Discount | 10 |
| 5 | Theater discount | Loyalty | Discount | 20 |
| 6 | Subscription Choose your own - youth | Subscription Choose your own | Subscription | 60 |
| 7 | Subscription Fixed - youth / student | Subscription Fixed | Subscription | 65 |
| 8 | Subscription Choose your own | Subscription Choose your own | Subscription | 10 |
| 9 | Subscription Fixed | Subscription fixed | Subscription | 15 |

[a] Under 25 years [b] Only for retirees

Table 1: Price type used by Royal Danish Theatre

Apart from the standard ticket, price types can be roughly divided into two cate-

[2]E.g. such performances as open dress rehearsals and previews before the opening night

gories: discount and subscription. Discounts can be applied to young people, students and senior citizens for social awareness purposes and also to those who sign up for a loyalty program. In the latter case, customers buy a loyalty card, which entitles them to some benefits, including a discount on the ticket price of the theater performance.
The Royal Danish Theater applies two kinds of subscriptions: a fixed subscription, in which the bundle of productions included is predetermined by the theater, and "choose your own" subscription, which allows the customer to choose the productions they want to see. In the latter case, subscribers commit to purchasing a pre-set quantity of tickets and, during the season, they freely choose the content of their bundle. In general terms, subscribers benefit from a discount with respect to the standard ticket price: this is an example of second-degree price discrimination, according to which the unit price varies depending on the quantity demanded.
Figure 2 shows how the sales of tickets are distributed among the different price types.

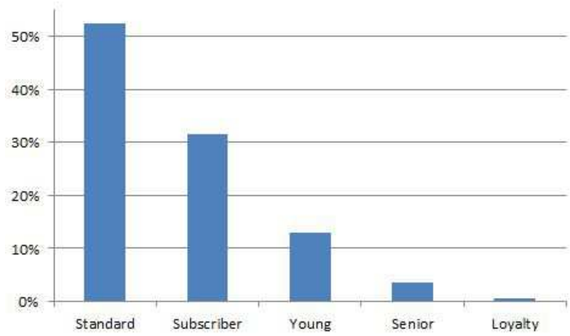


Figure 2: Percentage of ticket sold by price type

The low percentage of *senior* tickets is surprising. Indeed, many senior customers are subscribers, thus it is not convenient for the theater to offer a discount for senior customers for all of the productions. Senior customers are entitled to a discount of 50% only for some productions decided by the theater management. Given this, in our model these senior customers are representative of retiree customers who attend the theater only occasionally.
After deciding to attend a production, each consumer decides the day and the seat quality. Each day/seat combination has a different price that can be discounted according to Table 1.

# 4 Methodology

We consider a situation in which the consumer, after deciding which production to attend, evaluates a finite number of ticket alternatives, each of which differs by the quality of seat and day of performance (premiere, saturday evening and so on)[3]. Such combinations of seat and day of performance constitute the choice set $C$. According to the random utility theory, the utility of alternative $j$ received by the consumer $i$ is given by:

$$U_{ij} = V_{ij} + \epsilon_{ij} \tag{1}$$

The utility is partitioned in two components: the deterministic (or systematic) utility $V_{ij}$ that is observed by the analyst, and a residual term $\epsilon_{ij}$, which includes unobserved effects. It is assumed that the deterministic part is a linear function of the observed attributes of each alternative, so that the utility function of alternative $j$ can be written as:

$$U_{ij} = \beta' X_{ij} + \epsilon_{ij} \tag{2}$$

where $X_{ij}$ is a vector of values representing attributes of the alternative $j$ and $\beta'$ is a vector of the corresponding parameters to be estimated.
Hence, the probability that the individual $i$ chooses the alternative $j$ is given by:

$$P_{ij} = P(V_{ij}+\epsilon_{ij} \geq V_{ik}+\epsilon_{ik} \quad \forall k \neq j) = P((V_{ij}-V_{ik})+\epsilon_{ij} \geq \epsilon_{ik}) \quad k \neq j, \forall k \in C \tag{3}$$

We impose that the errors are independent and identically random variables distributed (i.i.d.) according to a Gumbel distribution. As the difference between two Gumbel variables is a logit random variable, the expression (3) takes the following form (McFadden, 1974):

$$P_{ij} = \frac{exp(\beta' X_j)}{\sum_{k \in C} exp(\beta' X_k)} \tag{4}$$

The coefficients of (4) are estimated by maximizing the likelihood function. The contribution to the likelihood for the individual $i$ is given by:

$$P_i = \prod_{j \in C} P_{ij}^{y_{ij}} \tag{5}$$

[3]Some studies in the transportation industry would suggest the ticket decision is a decision made at the lower nest, while the mode decision is made at the upper nest (Whelan *et al.*, 2008). Similarly, we could consider the decision on which production to attend as an upper nest decision. This kind of decision is the approach adopted by Grisolia and Willis (2011a, 2011b, 2012). However, this study is based on confirmed booking data, so we assume that the production decision has already been made.

where $y_{ij}$ is a dummy variable equal to 1 if individual $i$ makes choice $j$, 0 otherwise. Taking the log of both sides we obtain:

$$lnP_i = \sum_{j \in C} y_{ij} lnP_{ij} \tag{6}$$

which leads to the overall log-likelihood function for the sample:

$$lnL(\beta) = \sum_{i=1}^{n} \left( \sum_{j \in C} y_{ij} lnP_{ij} \right) \tag{7}$$

In the conditional logit and MNL models, parameters $\beta$ are assumed to be fixed among the population: this implies the same preference structure among customers, as the marginal utility of the attributes is the same in the population of theater-goers. This assumption seems unrealistic in the performing arts sector. Heterogeneity, can be efficiently addressed including in the model both socio-demographic characteristics and choice situation variables. Indeed, as discrete choice models work on the difference in utility, these variables that do not vary with the ticket alternative can enter the model in two ways: first, by interacting them with attributes of the alternative; and second, by including them in $J-1$ alternatives. In this way, these variables are able to affect the difference in utility.

In our application, we include as consumer characteristic variables the information derived by the type of ticket sold in terms of the discount that identifies consumer types (student, seniors, subscriber and so on), the period in which the ticket has been sold and whether the customer is inhabitant in Denmark or not. Moreover, the characteristics of the production are also used as variables to accomodate heterogenetiy among the population. Such variables indeed do not vary across ticket alternatives and are supposed to reflect consumer characteristics. Different productions, in terms of genre, newness, and whether highbrow or lowbrow, attract different consumers in terms of social class (see Sintas and Alvarez, 2005) and consequently are likely to affect the marginal utility of ticket attributes.

In our data set we have customers - as the subscribers - who repeat the choice more than once. Given the assumption of i.i.d. of the error component, it is not possible in MNL to account for correlation within individual preferences. However, we attempt to overcome this restriction by including variables related to the production: indeed, even if these variables are choice invariant, they can vary across the repeated chocies made by the same individuals. In such a way we consider the choices made by the same customers as choices made by different individuals that differ from each other by the values of the production variables.

Another way to incorporate preference heterogeneity is by using LCM. The logic underlying LCM is that the population can be sorted into $S$ classes, such that individuals within the same class have homogeneous preference. Therefore, each parameter $\beta$ takes $s$ different values with corresponding probabilities. The probability that alternative $j$ will be chosen by a randomly selected individual $i$ is given by:

$$P_{ij} = \sum_{s=1}^{S} P_{ij|s} \cdot M_i(s) \tag{8}$$

where $M_i(s)$ is the probability that the individual $i$ belongs to class $s$. In other terms, (5) is a sort of weighted average of different MNL models (as many as the number of classes), with the weights represented by the size of each class in the population. The analyst doesn't know to which class an individual belongs, however the likelihood of the individuals belonging to a class can be inferred through a probabilistic assignment process called membership function, which includes individual-specific variables. An MNL specification is a convenient form for the class membership model. Hence, the probability of individual $i$ to belong to the latent class $s$ is given by:

$$P_{is} = \frac{exp(\eta_s' Z_i)}{\sum_{s=1}^{S} exp(\eta_s' Z_i)} \quad \forall s \neq S; \quad \eta_S = 0 \tag{9}$$

where $Z_i$ is a vector of the values of the individual-specific variable for the individual $i$ while $\eta_s$ is the corresponding parameter for class $s$ to be estimated. Notice that for one latent class (the last one, $S$) the parameters are normalized to 0 to secure identification of the model (Greene and Hensher, 2003).

Including the membership function in (8) we obtain the probability of choosing alternative $j$ by individual $i$:

$$P_{ij} = \sum_{S=1}^{s} \left[ \frac{exp(\eta_s' Z_i)}{\sum_{s=1}^{S} exp(\eta_s' Z_i)} \right] \cdot \left[ \frac{exp(\beta_s' X_j)}{exp(\sum_{k \in C} \beta_s' X_k)} \right] \tag{10}$$

One feature of the LCM is noteworthy: with the presence of the membership function, the probability of selecting one alternative over another contains arguments that include the systematic utilities of the other alternatives available. Hence, unlike MNL, the IIA assumption can be relaxed (Boxall and Adamowicz, 2002). The parameters of the LCM are estimated maximizing the overall log-likelihood function for the sample:

$$lnL(\beta, \eta) = \sum_{i=1}^{N} ln \left[ \sum_{s=1}^{S} M_{is} \prod_{j \in C} P_{i|s}^{y_{ij}} \right] \tag{11}$$

In estimating (11), the number of classes $S$ is taken as given. Its determination is usually done through statistical criteria, such as Bayesian information criterion ($BIC$) and Akaike information criterion ($AIC$), which are used as a guide to determine the number of classes (e.g. Kamakura and Russel, 1989; Roeder *et al.*, 1999; Wedel and Kamakura, 2000]. These tests are calculated as follows:

$$AIC = -2LL + 2K$$

$$BIC = -2LL + Ln(N)K$$

where $LL$ is the value of the log-likelihood function, $K$ the number of parameters and $N$ the sample size. This tests are calculated for models with different numbers of classes. The final number of classes selected is the one for which the value of the test is the smallest.

# 5 Data set and variables

Our database consists of the ticket sales by the Royal Danish Theatre during the period 2010/2011 to 2012/2013. A total of 250 170 bookings records are included in the dataset, which involved 23 productions and 377 performances[4]. For each ticket reservation we have the following information, which allows us to identify the choice made and customer's characteristics: buyer's name and address, time and date of the purchase, price paid, price zone and price type.

The independent variables that enter in the model as choice attributes are:

- *Price* (in DKK[5])
- *Seat category*: a dummy variable for each seat category, ranked from 1 (the cheapest) to 5 (the most expensive)
- *Wkend*: takes value 1 when the performance is either a weekend matinee or is run on Friday/Saturday evening or in a public holiday day
- *Wkday*: takes value 1 when the performance is run during weekdays.

*Seat1* and *Wkday* are used as baselines in order to guarantee identification of the model.

The *Price* and *Seat category* variables aim to capture the trade-off behavior between cheap seats with low visibility and/or acoustics, and more expensive high-quality

[4] For the complete list of productions see the Appendix

[5] $1DKK \approx 0.13$€

seats. As the number of seat categories changed through the period under examination, we aggregated productions with more than five price zones into five seat categories.[6]

Table 2 provides an example of how the eight seat price categories of the production *Tannhäuser* have been aggregated. The baseline is the production *Boris Godunov*. Given a seat category and a customer category, *Price* can change also over time

| **Seat category** | **Price** | **Seat category** | **Price** | **New Seat category**[7] | **New price** |
|---|---|---|---|---|---|
| 1 | 115 | 1 | 125 | 1 (1,2) | 160 |
| 2 | 375 | 2 | 195 | 2 (3,4) | 345 |
| 3 | 565 | 3 | 295 | 3 (5) | 525 |
| 4 | 715 | 4 | 395 | 4 (6,7) | 720 |
| 5 | 895 | 5 | 525 | 5 (8) | 895 |
| | | 6 | 645 | | |
| | | 7 | 795 | | |
| | | 8 | 895 | | |

Table 2: Aggregation of seat categories

adapting it to the expected demand. As a matter of fact, if the demand for the first performances of a production is low, then the price for the following performances will be reduced. Conversely, price for the following performances will be raised if the expected demand increases. However, this fact does not affect the results o the model because for each customer we have information on the price charged in that moment for each seat category. Hence, our dataset includes the real choice set available in the moment in which the customer decides to buy a ticket.

The other two alternative variables reflect the choice of the day of performance. As Corning and Levi (2002) have shown, these variables affect performance-level demand. We have choosen only two variables to characterize the day of the performance: we have excluded a dummy indicating whether or not the performance is the opening performance, because no price discrimination is applied for such performances. Moreover, the weekend variable includes both the Friday/Saturday evening performance and the Sunday matinee. Indeed, from the data set we can observe that Sunday matinees constitute a small fraction of all the performances and they are not available for all productions. Moreover, we note that, essentially, Sunday matinee and Friday/Saturday night prices are homogeneous across productions.

[6]The rule of thumb followed is to consider as a baseline a production of the same genre in which the theatre was divided in 5 price zones: each new zone is associated with the baseline that has the smallest difference in price of a standard ticket. The price of the new seat categories is calculated as the average of the price of the original categories that were been aggregated to assemble the new categories.

In addition, in our model we have also included choice-invariant variables. These are related to customer characteristics, which are inferable by the ticket type purchased and the characteristics of the production. Concerning the first set we have:

- *Young*: takes value 1 when the customer is a student or a young person
- *Seniors*: takes value 1 when the customer is a senior citizen
- *Loyalty*: takes value 1 when the customer has bought a loyalty card
- *Subscriber*: takes value 1 when the customer is a subscriber
- *Foreign*: takes value 1 when the customer does not live in Denmark
- *Period*: a dummy variable for each period before the performance in which the ticket has been sold.

Essentialy, the variables used as customers' characteristics (except for *Foreign*) are those on the basis of which the theatre manager can implement a price discrimination. In this sense, the variables chosen are intended for the purpose of guiding theatre managers. A note concerning the last variable: we considered for each observation how many days before the performance the ticket was sold. Then we evaluated the distribution of these days among the observations and identified four quartiles, each representing a period in the sale horizon. Tickets were sold in one of four quartiles or periods before the performance: (1) 233 or more days beforehand; (2) from 64 to 232 days beforehand; (3) from 20 to 63 days beforehand; and (4) up to 19 days beforehand. These dummy variables were used in the MNL model, whereas the LCM model used continuous variables denoting how many days before the performance the ticket was sold.
We used the following variables for the production attributes, taken from Bille *et al.* (2015):

- *Opera*, *ballet*, *play*: dummy variables that capture the genre of the production and the customer's taste
- *Newness*: two dummy variables that measure the degree of newness/innovation in the performance
- *New DKT*: it takes value 1 when the production is run for the first time at the Royal Danish Theater
- *Review*: three dummy variable, for *bad*, *average* and *good* newspaper reviews of the performance

- *Audience evaluation*: three dummy variables, for *bad*, *average* and *good* audience evaluation of the performance

All these variables are included in the MNL model, whereas in the LCM we have included only the variables related to the genre of the production.
Some remarks about the production attribute variables: in Bille (2015), data for the audiences' evaluation of the productions were collected every season. Every season, a questionnaire was sent to the audiences of 5 operas, 5 plays and 5 ballets. For each production approximately 110 questionnaires were sent out, summing up to about 1650 questionnaires each season. During all the seasons the response rate has been around 52% (ranging from 49% to 60%). The quality of the performance was measured on a scale from 1 to 5 (where 1 is low quality and 5 high quality). Data for the professional reviewers' evaluations of the productions were collected every season as well. Similarly, reviews of the Royal Theater productions in all the major Danish newspaper (9 newspapers) were collected. Two independent researchers read all the reviews and rated the quality of the productions based on the reviewers' opinion. In this way the quality was indexed on a scale from 1 to 5, and in the case of inconsistent evaluations the two researchers agreed on the final index.
Based on these measures, we identified three categories for the audience's and the reviewer's evaluation variables: bad, average and good.
The degree of newness in the productions was assessed by an expert in theater science. A Mozart opera can be performed in a very traditional way or in an experimental or groundbreaking way, as can a brand new production. Newness refers to the impression of something *new* regards the direction, the manuscript, the actors, the stage design, the costumes, the music and so on. This variable takes two levels: traditional or innovative. Table 3 sumarizes the variables used in our models.

As already described, the combination of seating area and day of performance define the customer's choice set. One of the main difficulties in the model set-up is identification of the choice set of each booking. Indeed, the seat categories available for an individual depend on the choices made by individuals who have already bought a ticket. Because no performance had totally sold out, we do not have information on whether, at some stage of the sale period, a single region of the theater had sold out or, on the contrary, tickets for that zone were available at the time of the performance. However, we can observe that, in general, tickets for all the seat categories were still being sold in the last few days before the performance. Hence, we assume that for each individual the choice set includes all the seat categories. This assumption seems realistic as the theater management has confirmed how, in most cases, there are available seats for all the price zones just before the beginning

| Level | Variable | Description | Type |
|---|---|---|---|
| Alternatives | *Price* | Price in DKK | Continuous |
| | *Seat* | Seat category (5 level) | Dummy |
| | *Wkend* | Friday/Saturday evening, Sunday mattinee | Dummy |
| | *Wkday* | Weekdays | Dummy |
| Customer | *Young* | Under 25 years /Student | Dummy |
| | *Senior* | Retirees | Dummy |
| | *Subscribers* | Subscribers | Dummy |
| | *Loyalty* | Customers with a loyalty card | Dummy |
| | *Foreign* | Equal 1 if customer does not live in Denmark | Dummy |
| | *Period* | 4 Periods of purchasing (only MNL) | Dummy |
| | *Days* | No. of days before the performance the ticket has been sold (only LCM) | Continuous |
| Production | *Genre* | Opera, Ballet and Play | Dummy |
| | *Newness* | Degree of newness/innovation, 2 levels (only MNL) | Dummy |
| | *New DKT* | First time in Denmark (only MNL) | Dummy |
| | *Review* | Newspaper review, 3 level (only MNL) | Dummy |
| | *Evaluation* | Audience evaluation, 3 level (only MNL) | Dummy |

Table 3: Variables used in MNL and LC models

of the performance. The only exception was for the *senior* category, for which in some cases, the theater management chose not to make all the price zones available. Clearly, in estimating the model we take into account the situations in which this category has a reduced choice set.
The identification of the choice set by day of performance is easier: for each production we considered the last Friday/Saturday evening and weekday performance. Assuming this is the chronological order, all the bookings made after the last ticket sold of the last Friday/Saturday evening performance will have a reduced choice set as it will not include the weekday performance.
Table 4 illustrates how the choice set generation process works for the production *Così fan tutte*. In the context of this example, all bookings made after (a) face 5 alternatives instead of 10 (assuming that all the price zones of the theater are available).

| Date and time of performance | Dummy variable = 1 | Date last ticket sold |
|---|---|---|
| 11-10-2011 19:30 | Wkday | - |
| 14-10-2011 19:30 | Wkend | - |
| 16-10-2011 15:00 | Wkend | - |
| 25-10-2011 19:30 | Wkday | - |
| 27-10-2011 19:30 | Wkday | - |
| 30-10-2011 15:00 | Wkend | - |
| 02-11-2011 19:30 | Wkday | - |
| 06-11-2011 15:00 | Wkend | |
| 10-11-2011 19:30 | Wkday | - |
| 19-11-2011 19:30 | Wkend | 17-11-2011 10:38 (a) |
| 21-11-2011 19:30 | Wkday | - |

Table 4: Choice set generation process

# 6 Model estimation results

## 6.1 MNL model

The MNL model is estimated with Biogeme (Bierlaire, 2003)[8] and the results are shown in Table 5. The MNL logit model is linear in the parameters specification, including the characteristics of the alternatives and their interaction terms, in order to accomodate taste variations due to customers' and performances' characteristics. Models with different interaction terms are estimated and compared using the non-nested hypothesis test developed by Horowitz (1982).
Table 5 displays the significant coefficient[9] of the MNL specification that has shown a better-fitting model. The variables *play*, *seat*1, *period*1, *wkday*, *review bad* and *evaluation bad* are used as base variables to allow for identification of the model. In the final specification we allow the *price* sensitivity to take a different value according to the production characteristics and the period in which customer bought the ticket; whereas the marginal utility of the *seat* and *wkend* variables interact with the $foreign$ variable and the different customers' categories.

The *price* coefficient is negative, as expected. However, the heterogenetiy of the price sensibility for the theatrical experience is revealed through the coefficients of the interaction terms. In particular, the interaction with the period of purchasing reveals a pattern: the price coefficient increases as we consider bookings made long before the day of the performance, reaching a positive value in the first period of the time horizon (in the first period a coefficient of -0.00203+0.00216 = 0.00013). For this portion of consumers, the theatrical experience is configured as a Veblen good. Typically, the earlier ticket buyers are subscribers (Drake *et al.*, 2008; Tereyagoglu *et al.*, 2012) who, as empirical evidence has shown, are less responsive to ticket price changes (Felton, 1994). As Corning and Levy (2002) noticed, single-ticket purchasers have a higher opportunity cost of time compared to subscribers, so they prefer to preserve themselves for "flop": this can be done by buying the ticket at a later stage, after a period in which crucial information for the purchasing decision has been acquired. Moreover, this finding is consistent with the study by Swanson *et al.* (2008), in which the authors show that there is an association between the motivation to attend a theatrical performance and how far in advance ticket purchase decision is made: the stronger is the motivation, the further in advance the attendance to the performance is planned. Furthermore, as Drake *et al.* (2008)

[8] Biogeme is a free software specifically designed for discrete choice models. It can be downloaded from http://biogeme.epfl.ch/home.html)

[9] All variables are significant except the interaction terms between Price and Ballet and between Wkday and Loyalty

| | Coefficient | *t-stat* |
|---|---|---|
| **Price** | **-0.00203** | -17.07 |
| Price-Period1 | 0.00216 | 38.90 |
| Price-Period2 | 0.00139 | 36.37 |
| Price-Period3 | 0.000360 | 10.22 |
| Price-Aud. Evaluation average | $-0.000142^{**}$ | -1.93 |
| Price-Aud. Evaluation good | 0.000858 | 16.81 |
| Price-New DKT | 0.000918 | 21.64 |
| Price-Newness1 | -0.00106 | -16.96 |
| Price-Newness2 | -0.00128 | -21.70 |
| Price-Opera | -0.000574 | -14.39 |
| Price-Review average | 0.000377 | 8.85 |
| Price-Review good | 0.000517 | 9.64 |
| **Seat 2** | **0.648** | 36.90 |
| Seat 2 - Foreign | $0.0913^{*}$ | 2.26 |
| Seat 2 - Loyalty | 0.530 | 2.98 |
| Seat 2 - Senior | 0.543 | 3.67 |
| Seat 2 - Subscriber | 0.898 | 26.92 |
| Seat 2 - Young | -0.222 | -9.33 |
| **Seat 3** | **1.35** | 53.13 |
| Seat 3 - Foreign | $-0.0782^{*}$ | -2.04 |
| Seat 3 - Loyalty | $0.306^{**}$ | 1.80 |
| Seat 3 - Senior | 1.84 | 14.16 |
| Seat 3 - Subscriber | 1.03 | 31.38 |
| Seat 3 - Young | -0.622 | -24.94 |
| **Seat 4** | **1.80** | 52.78 |
| Seat 4 - Foreign | 0.151 | 4.17 |
| Seat 4 - Loyalty | 0.534 | 3.27 |
| Seat 4 - Senior | 1.52 | 11.66 |
| Seat 4 - Subscriber | 1.28 | 38.12 |
| Seat 4 - Young | -1.29 | -46.38 |
| **Seat 5** | **1.94** | 43.94 |
| Seat 5 - Foreign | 0.454 | 12.71 |
| Seat 5 - Loyalty | 0.448 | 2.74 |
| Seat 5 - Senior | 1.49 | 11.36 |
| Seat 5 - Subscriber | 1.40 | 39.08 |
| Seat 5 - Young | -1.48 | -46.38 |
| **Wkend** | **0.214** | 36.74 |
| Wkend - Foreign | 0.357 | 19.41 |
| Wkend - Senior | -0.751 | -30.91 |
| Wkend - Subscriber | -0.192 | -21.14 |
| Wkend - Young | -0.143 | -11.55 |
| No. of observations | | 250 170 |
| $\rho^2$ | | 0.083 |
| Adjusted $\rho^2$ | | 0.083 |
| Null log-likelihood | | - 573738.544 |
| Final log-likelihood | | -526059.818 |

$^{**}p = .10$ $^{*}p = .05$

For all the others variables $p = .001$

Table 5: Estimation of MNL model

claimed, there is a direct relationship between the demand rate and the inventory level: the seats that are already sold at a given price are more valuable than the ones that remain, as typically the latter are further away from the stage. Also, within the same seat tiers there are seats that guarantee a better viewing of the performance[10]. Figure 3 depicts the total sale of subscription and standard tickets in relation to the time before the performance. We consider the beginning of the time horizon to be 62 weeks before the show. For subscribers, the sale pattern reaches different peaks until around 30 weeks before the performance, after which it decreases monotonically. In contrast, for standard ticket buyers, from the beginning of the sale period the pattern increases monotonically increasing and reaches a peack one week before the performance.

Figure 4 shows the cumulative distribution over time for the sales of each seat

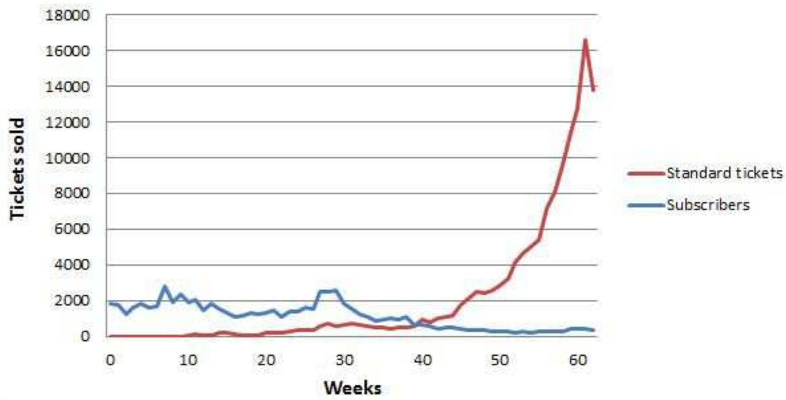


Figure 3: Total sales of subscription and standard tickets over time

category, considering the 60 weeks before the date of the performance. From Figure 4 it seems that the most valuable seats are sold in the beginning of the sale period, while, as we approach the date of the performance, there is an increase in lower quality seat sales. This pattern has been also found by Tereyagoglu *et al.* (2012) and can be explained by the finding that customers with higher valuation of the performance tend to buy the ticket in earlier periods, whereas customers with lower valuation tend to buy the ticket in later periods. According to Figure 4, when we consider 50% of total sales, that figure is achieved for the fifth seat category within 16 weeks before the performance, within 12 weeks for the fourth category, within 8 weeks for the third category, within 6 weeks for the second category and within only 4 weeks for the cheapest seat category. Concerning the interactions with the production characteristic, the price coefficient for *opera* is slightly smaller than for

[10]An exception occurs when the customer intentionally delays the ticket purchase when it is expected that the theater uses a discount policy for tickets sold very close to the performance.

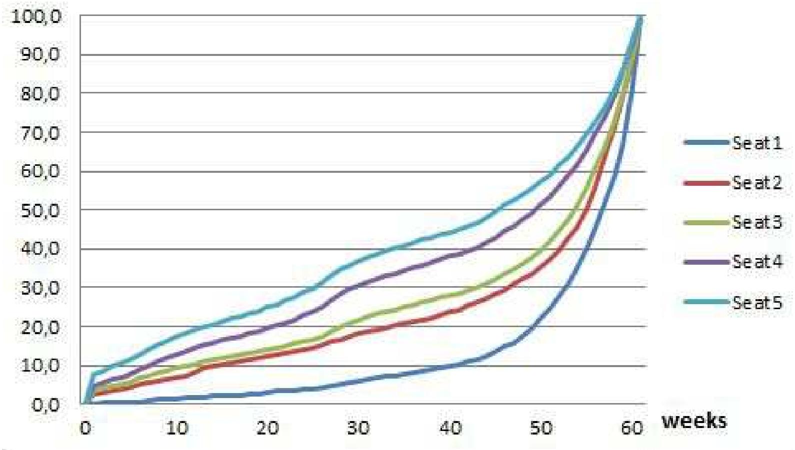


Figure 4: Cumulative distribution over time of each seat category

the *play* genre, while the interaction with *ballet* is not significant. Moreover, the price coefficient decreases as the degree of newness/innovation increases, showing that the audience prefer traditional and less risky productions than the more experimental. The quality of the production as reported by *reviews* has a positive impact on the customer's utility as well as, to a greater extent, those productions that are performed for the first time at the Royal Theatre. Surprisingly, the effect of audience evaluation does not monotonically increase: the *average evaluation* coefficient is negative where the *bad evaluation* is the baseline, although in terms of significance, the interaction with *average evaluation* produces the lowest absolute value of the $t - test$ (-1.93)
With regard to the seat quality, the coefficients reflect an expected pattern for the standard ticket buyers (which coefficient is the one without interaction terms), senior, subscribers and customers affiliated with a loyalty program: an increase of the quality of the seat leads to a greater utility. In particular, among these categories, for *senior* and *subscribers* this pattern is more evident, followed by *loyalty* and *standard*. Also, foreigner customers (87% of whom are standard ticket buyers), show a similar tendency, with a larger coefficient than Danish standard ticket buyers.
An explanation for the highest marginal utility of the *Senior* and *Subscribers* categories can be interpreted in the light of the well known theory of rational addiction developed by Stigler and Becker (1997): the consumption of cultural goods (a theater production in our case) increases the consumers' future capacity to appreciate it, through the "learning by doing" process. Hence, previous exposure to the cultural goods to leads to a growth in consumption and therefore to an increasing WTP. In this sense *subscribers* and *seniors* are the type of customers who have accumu-

lated consumption capital through their past consumption: the former because a subscription implies high frequency of theater attendance, the latter because of the age component. These customers, more than others, pay attention to seats that provide a better quality of theatrical experience from both the acoustic and visual perspectives.
The *young* category has the lowest value of marginal utility and does not increase monotonically with respect to seat quality, the largest value corresponding with the third seat category. Therefore, it seems likely that this category would not consider buying expensive seats and would pay little attention to the seat quality.
Figure 5 shows more graphically for each category the relationship between the utility function and the level of the seat attribute. As Figure 5 seems to suggest, except

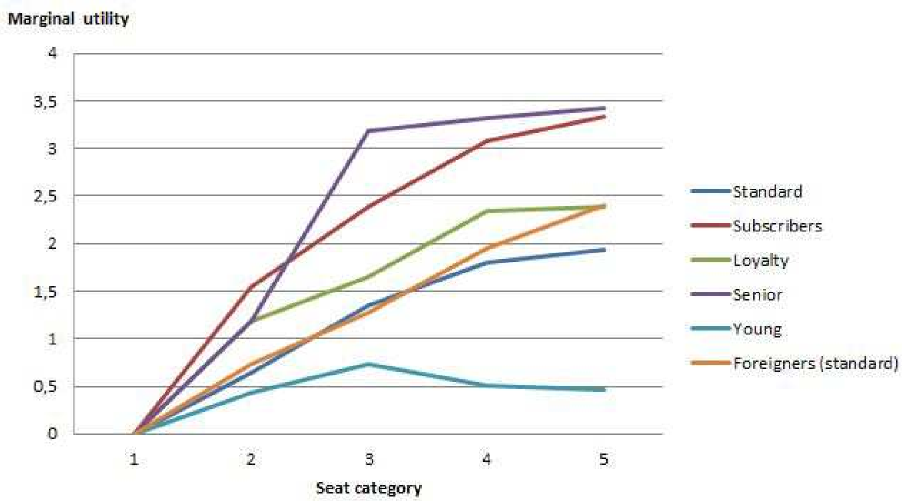


Figure 5: Quality of seat in relation to the utility function

for young customers, the relationship between quality of seat and customers' utility is approximately increasing and concave, meaning that as the level of seat quality increases, the difference in utility gets smaller and smaller.
Finally, we notice that, with the exception of *senior*, weekend performances are preferred over weekday performances, particularly by the *foreigners* standard ticket buyers, followed by Danish standard ticket buyers, *young* and *subscribers*. This result is probably due to to a greater flow of tourists in the city of Copenhagen during the weekend. The negative value for *senior* (0.213-0.510 = 0.297) can be explained by considering that this category is rich in time and therefore prefers the cheaper weekday alternatives. However, compared to the seat attributes, the day of the performance has a lower impact on explaining the choice of ticket.

## 6.2 Latent class estimation

In the LCM we aim to identify distinct groups of theatergoers according to their behavior with respect to the type of ticket purchased. We initially assess the number of classes in the LCM by BIC and AIC. These statistics indicate whether the complexity of the model, that is the number of parameters to be estimated, can be compensated by an improvement in the value of the log-likelihood. Table 6 summarizes the statistics for models with one, two, three and four classes. The results

| No. of classes | Log-likelihood value | AIC | BIC |
|---|---|---|---|
| 1 | -540535,3380 | 1081082,6759 | 1081145,2553 |
| 2 | -530590,2024 | 1061224,4048 | 1061453,8625 |
| 3 | -524502,4273 | 1049080,8545 | 1049477,1906 |
| 4 | -521576,2680 | 1043260,5359 | 1043823,7503 |

Table 6: Criteria for determining the optimal number of classes

show that as the number of classes increases, the model fits the data better. By further increasing the number of classes, we obtain the optimal model with seven classes. However, as the number of segments increases to more than four classes, we obtain some small segment sizes that make the parameter estimated unstable. For this reason, and also for an easier interpretation of the model, we adopt the four-class solution.

The explanatory variables of the choice model are *price*, *seat* and *wkend*, with $seat1$ and *wkday* set to 0 as base variables. We include the membership function in order to assign individuals to classes according to their characteristics and the choice situation. The variables employed for the membership function include the dummy variables related to the customer ticket's category and genre of the production. Moreover, we include a variable indicating how many days before the performance the ticket was bought: compared to the MNL model, where this is used as a categorical variable in 4 levels to be interacted with price, we use *days* as a continuous variable in the membership function that contributes to the class assignment of individuals.

Table 7 reports the results derived from the LCM, which was estimated using the software Latent Gold Choice (Vermunt and Magidson, 2005). Given that the magnitude of the coefficients of the choice model cannot be compared between different classes due to scale parameter [Carrier, 2008; Hetrakul and Cirillo, 2013], the different behavior of the classes is compared by their WTP for the choice attribute. As for the membership function, the coefficients indicate how much the variables account for the belonging to that particular class: the variables are interpreted in relation to Class 4 and normalized to zero for identification of the model.

In Table 7 we report for each parameter the result of the Wald test. Largely em-

| **Parameter** | **Class 1** | **Class 2** | **Class 3** | **Class 4** | Wald test | p-value |
|---|---|---|---|---|---|---|
| Price | -0.0039 | -0.0001 | -0.0003 | -0.1091 | 1224.54 | 0.000 |
| | *(-30.32)* | *(-0.29)* | *(-2.13)* | *(-11.45)* | | |
| Seat2 | 10.7610 | 0.0614 | 0.1507 | 7.4508 | 55.6901 | 0.000 |
| | *(1.38)* | *(0.69)* | *(7.01)* | *(1.18)* | | |
| Seat3 | 11.8896 | 1.8220 | 0.1364 | 28.2045 | 522.1306 | 0.000 |
| | *(1.52)* | *(18.84)* | *(4.28)* | *(11.74)* | | |
| Seat4 | 12.8816 | 2.8707 | -0.1115 | 27.6766 | 1055.6425 | 0.000 |
| | *(1.65)* | *(29.70)* | *(-2.36)* | *(10.72)* | | |
| Seat5 | 14.0908 | 2.2991 | -2.5306 | 34.1273 | 478.3154 | 0.000 |
| | *(1.82)* | *(18.25)* | *(-4.43)* | *(10.42)* | | |
| Wkend | 0.3495 | -0.0941 | 0.1055 | -0.4873 | 1691.2595 | 0.000 |
| | *(.30.32)* | *(-0.29)* | *(-2.13)* | *(-11.45)* | | |
| **Membership function** | | | | | | |
| Standard | 3.5465 | 10.2095 | 2.6567 | | 97.5406 | 0.000 |
| | *(6.75)* | *(1.98)* | *(4.83)* | | | |
| Subscribers | 2.0865 | 9.6714 | 0.7470 | | 154.7725 | 0.000 |
| | *(3.99)* | *(1.88)* | *(1.37)* | | | |
| Young | 0,4906 | -0.4604 | 1.0377 | | 36.1000 | 0.000 |
| | *(0.94)* | *(-0.85)* | *(1.91)* | | | |
| Senior | -20.0132 | -1.2175 | -18.6624 | | 18.4787 | 0.000 |
| | *(-3.31)* | *(-18.66)* | *(-3.12)* | | | |
| Loyalty | -3.4969 | 4.1490 | -4.7088 | | 32.0894 | 0.000 |
| | *(-0.39)* | *(0.41)* | *(-0.53)* | | | |
| Opera | 1.7714 | 5.0029 | 2.8948 | | 503.5944 | 0.000 |
| | *(6.75)* | *(1.98)* | *(4.83)* | | | |
| Ballet | 5.4112 | 7.0468 | 6.3155 | | 326.7014 | 0.000 |
| | *(1.08)* | *(1.41)* | *(1.26)* | | | |
| Play | -9.7429 | -4.8584 | -8.3682 | | 523.7187 | 0.000 |
| | *(-1.76)* | *(-0.88)* | *(-1.51)* | | | |
| Days | 0.0059 | 0.0025 | -0.0042 | | 977.2111 | 0.000 |
| | *(11.12)* | *(4.82)* | *(-7.77)* | | | |
| No. of observations | 250170 | | | | | |
| Adjusted $\rho^2$ | | 0.091 | | | | |

Table 7: Estimation of LCM

ployed in LCM, the Wald test is a test for the equality of effects between classes, indicating whether a variable is equal across classes and therefore, is class independent. In our model the null hypothesis is rejected for all the predictors and covariates, indicating that all the variables chosen are useful in discriminating individuals in classes. Classes are numbered in order of size.

Class 1 accounts for 48.4% of the market and exhibits an expected pattern: the price coefficient is negative and individuals in this class increase their utility as the quality of seat increases. Moreover, this class prefers weekend performances to weekday

ones. This class shows a high WTP for a theater ticket, but not the highest among the classes.
Instead, Class 2 is characterized by the largest WTP, as the price coefficient is negative but very close to zero. This class prefers weekday performances and the most expensive seats; however, it exhibits the greatest marginal utility for the fourth seat category. This class contributes 24.4% of the market.
Class 3 and Class 4 both exhibit low WTP compared to Class 1 and Class 2. However, they differ from each other significantly in some aspects: Class 3 is slightly smaller than Class 2, accounting for 24.1% of the market. The individuals of this class prefer weekend performance and the cheapest seats: the coefficient for the fourth and fifth seat categories is even negative, as if customers of this class would not consider buying expensive seats.
Class 4 is clearly the smallest, accounting for only 3.1% of the total market. Compared to individuals in Class 2, individuals in Class 4 prefer weekday performances and exhibit a stronger preference for the most expensive seats, even if its WTP for seat tiers is the lowest among all the classes.
In terms of customers' characteristics, the coefficients of the membership function show how Class 3 and Class 4 are strongly characterized by the age component: Class 3 can be considered representative of young customers as, compared with other categories, *young* has it largest coefficient in this class. This result confirms the low willingness of young customers to pay, which was found in the MNL model results. Class 4 comprises mainly by seniors customers: its coefficient is negative for all the other classes. However, considering that Class 4 is very small, we can deduce that a significant share of seniors are included in Class 2, given that in Class 1 and Class 3 its coefficients are decisively negative.
*Standard* and *subscribers* consistently show a positive coefficient, suggesting that these categories are distributed across the Class 1, Class 2 and Class 3; whereas the coefficient for the customers engaged in a *loyalty* program has a *zvalue* close to zero in all three classes that prevent us from making considerations.
The assignment of individuals to classes based on the maximum posterior probability can help us in understanding the class composition. In fact, once the parameters of the model are estimated, they can be used to calculate the conditional individual's probability of membership in each class by means of Bayes' theorem:

$$P(s \mid j, \hat{\eta}) = \frac{\hat{P}(j \mid s, \hat{\eta}) \cdot \hat{P}(s \mid \hat{\eta})}{\sum_{s=1}^{S} \hat{P}(j \mid s) \cdot M_{is}} \tag{12}$$

Equation (12) give us the probability that the individual belongs to class $s$ conditional on the choice made and his/her characteristics (which parameters are estimated). On the numerator we have the estimated choice probability for the choice made, given the class $s$, multiplied by the prior estimated class probability. On the denominator we find the probability to choose the alternative $j$ expressed, in the spirit of latent class, as a sum of MNL moderated by the size of each class. Indeed, the denominator is equal to expression (8) and (10).

Each individual is assigned to the latent class $s$, which provides the maximum value of (12). Based on this procedure, we can see how the categories are distributed across classes, as Figure 6 shows. Although in the MNL model the customers' be-

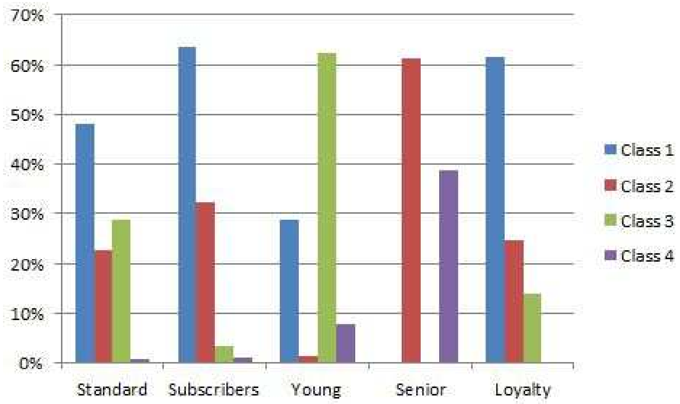


Figure 6: Distribution of customers across classes

havior is distinguished according to price category, in the LCM we can also observe some forms of heterogeneity within category, even if some patterns resulting from MNL are supported. The fact that almost all young customers (62.3%) are classified in Class 3 confirms their low WTP, the low utility gained by high quality seats and a preference for the weekend performances. However, a non-negligible share of young customers (28.7%) are found in Class 1. Probably, given the high value of WTP for Class 1, such customers are young subscribers. We should also consider the fact that in some cases the youth subscription makes it feasible for families to subscribe and include their children. In this case, the choice of these young customers depends on the one made by the family components that are subscribers.

*Subscribers* and *Loyalty* groups, which represent frequent theater attenders, are concentrated mainly in the Class 1 and Class 2 (in particular Class 1), confirming that these categories are characterized by high WTP and a preference for the most expensive seats.

Almost half (47.9%) of *Standard* ticket buyers, representative of infrequent theater attendance, are classified in Class 1; but significant shares are found also in Class 2 (22.6%) and Class 3 (28.7%). Hence, there is a sort of heterogeneity within this category, even if the majority are included in the two classes with the highest WTP. *Senior* customers are clearly split into Class 2 and Class 4, which are antithetical to each other from the WTP perspective, but similar in their preference for week-day performances. The latter aspect confirms the results of the MNL model. The majority (61.2%) of *Senior* customers fall in Class 2, confirming that this category has the greatest WTP. However, 38.7% of senior customers are found in the Class 4, which has the lowest willingness to pay. Both classes exhibit a preference for seats with high quality, but whereas in Class 2 we find that the fourth seat category has the highest coefficient, in Class 4 both the third and the fifth seat categories are preferred to the fourth. This different pattern can explain the MNL result in which, for this category, the marginal utility of the seat attribute does not increase monotonically with respect to its quality.

Concerning production genre, we notice a positive value of the coefficient for *opera* and *ballet* and a negative value of the coefficient for *play*, which suggests, by implication, that individuals of Class 4 are *play* attendants.
In Figure 7, the classification procedure is made on the basis of the production genre. In this way, we can verify whether or not customers' behavior is homogeneous across different types of theatrical productions. From Figure 7, we can observe

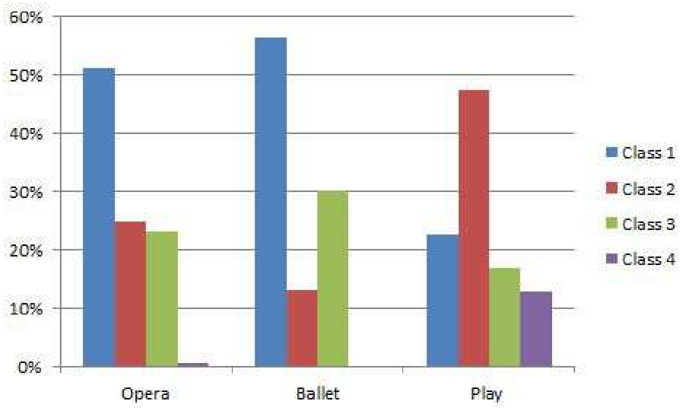


Figure 7: Distribution of latent classes across type of production

that people who attend *Opera* performances can be clustered into the four latent classes with about the same proportion resulting from the LCM. Hence, about half (51.2%) of the opera's customers belong to Class 1 and the other half are shared more or less equally by the Class 2 (24.8%) and Class 3 (23.2%). A high amount of *Ballet* customers belong to Class 1 (56.4%), but the remarkable aspect is that Class 3 individuals (30.2%) are more than twice as many as those in Class 2 (13.2%). *Plays* present a particular pattern: indeed, almost all of the individuals in Class

4 attend plays, that accounting for 13% of the total attendance of this production genre. However, their presence is counterbalanced by a large number (47.3%) in the class with the highest WTP.
Moreover, in plays we find, compared to other genres, a higher heterogeneity in terms of WTP, given the significant presence of Class 2 and Class 4 individuals. From this point of view, *Ballet* seems to be the most homogeneous genre as its share of Class 2 ticket buyers is low, whereas the presence of Class 3 individuals is substantial.

Finally, looking at the coefficients for the variable *day*, the negative coefficient in Class 3 suggests that the members of this class prefer to buy theater tickets in a period close to the date of the performance. The opposite holds for Class 1 and Class 2, for which the coefficients are positive. In Figure 8 we classify individuals by the purchase period, measured as days before the performance. What appears

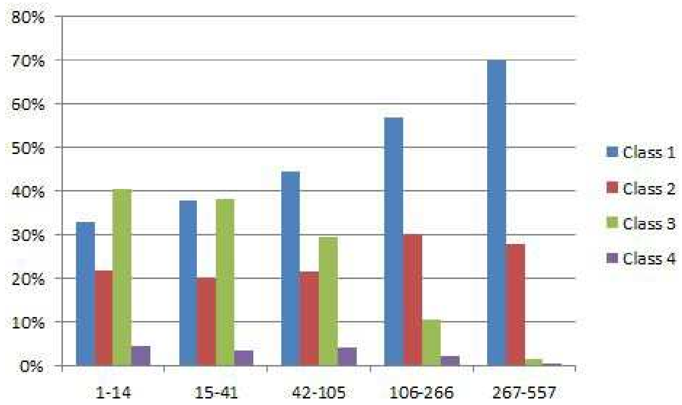


Figure 8: Distribution of latent class by purchase period

evident from Figure 8 is the opposite trend by classes with different WTP. At the beginning of the sale period, a large share of the tickets (97.8%) are sold to customers of Class 1 and Class 2, confirming the positive relationship between WTP and early ticket purchase. As we approach the day of the performance, the share of these two classes (in particular Class 1 as suggested by the magnitude of the *days* coefficient) decreases: within 14 days of the performance, 32.7% and 22% of the tickets sold are bought by customers that are classified in Class 1 and Class 2, respectively. Conversely, classes characterized by low WTP tend to purchase theater tickets in the latest stage of the sale period. This finding is more evident for Class 3. Indeed, the trend of Class 4 is quite stable from the middle of the sale period to the last days before the performance. Instead, the rate of individuals of Class 3 who buy a ticket increases as we approach to the day of performance: in the last

14 days, the relative majority (40.6 %) of customers who buy a ticket belong to this class.
In summary, the analysis suggests a typology of four classes.
Class 1 accounts for 48.4% of the market. This segment embraces theatergoers who have a high WTP for a theater ticket and gain a greater utility as the quality of the seat increases. They are early buyers and prefer weekend performances. This segment forms the majority of ticket buyers for opera and ballet productions, but not for plays. Individuals of this class are mainly standard ticket buyers, subscribers and customers enrolled in a loyalty program.
Class 2 accounts for 24.4% of the market and represents customers with the highest WTP. Like Class 1, this segment prefers the most expensive seats and tends to buy tickets in the early stage of the sale period. However, they prefer weekday performances and represents the majority of ticket buyers for plays. A big share of senior customers belong to Class 2 which also includes standard ticket buyers, subscribers and customers enrolled in a loyalty program.
Class 3 accounts for 24.1% of the market. It represents mainly young customers and standard ticket buyers with low WTP. Members of this segment prefer the cheapest seats and do not consider buying expensive seats. They prefer weekend performances and tend to buy the ticket in the latest stages of the sale period. This class is represented in all performances genres, but particularly ballet productions.
Class 4, with 3.1% of the market, is very small. It has the lowest WTP, and attends mainly plays. Like Class 2, it prefers the most expensive seats and weekday performances. Members of this class are used to buying tickets starting from the middle of the sale period, and are almost entirely senior customers.
Table 8 summarizes the main characteristics of the four classes identifed by the LCM. It is interesting to compare the classes obtained in this model with the one resulting

| | **1** | **2** | **3** | **4** |
|---|---|---|---|---|
| **Share** | 48.4 % | 24.4 % | 24.1 % | 3.1 % |
| **WTP** | High | Highest | Low | Lowest |
| **Seat** | Expensive | Expensive | Cheap | Expensive |
| **Day** | Weekend | Weekday | Weekend | Weekday |
| **Genre** | Opera and Ballet | Play | Opera, Ballet and Play | Play |
| **Purchase period** | Early buyer | Early buyer | Late buyer | Mid-late buyer |
| **Composition** | Standard,Subscribers,Loyalty | Senior and Subscribers | Young, Standard | Senior |

Table 8: Summary of latent classes

from Grisolia and Willis' (2012) model. Even though those authors considered a different set of theatre production attributes (price, review, word of mouth, repertory classification, author, review), by using the membership function we can find some similarities. Grisolia and Willis (2012) identified three classes of theatergoers: their

Class 3 denoted as the intellectual class, and formed of mature and high-frequency attendees with the largest WTP, seems to confirm the characteristics of our Class 2. Indeed, in our case senior (mature) people and subscribers (high-frequency attendees) are also characterized by the highest WTP. Grisolia and Willis' (2012) Class 2 can be associated with our Class 3: both classes are composed of young people, occasional attendee [11] who exhibit a low WTP. A different discourse can be made for Grisolia and Willis' (2012) Class 1 which, in their model, comprises affluent people who attend theater occasionally. In this case, we find no correspondence with our Class 1 because we do not have information about customers' income and, moreover, our Class 1 is comprises both subscribers and standard ticket buyers.

## 6.3 Statistical test to compare models

We can compare our models in terms of goodness of fit. In general terms, for both models the likelihood ratio test[12] indicates that these models are better than the null model, in which all parameters are set to zero. As can be seen from their log-likelihood values and $\rho$ squared, LCM performs better than MNL. As these two models are non-nested, we use the Horowitz test t0 compare model fits of MNL and LCM. The null hypothesis of the test is that the model with the lower adjusted rho-squared is preferred. The decision rule for which the null hypothesis is rejected is given by:

$$\phi[-(-2(\rho_H^2 - \rho_L^2) \cdot LL(0) + (K_H - K_L))^{1/2}] < \alpha$$

where $\phi$ is the standard normal cumulative distribution function, $\rho_H^2$ and $\rho_L^2$ are respectively the larger and the smaller values of adjusted rho-squared; $K_H$ and $K_L$ are the number of parameters in the model with the larger and smaller values of $\rho$ squared; and $\alpha$ is the significance level.

The null hypothesis is rejected, supporting the argument for which the LCM model fits the data better.

## 6.4 Willingness-to-pay measures

Measuring the WTP for the change in level of attributes is very important in order to adopt an appropriate pricing strategy. In the MNL framework, the WTP of an

[11] From Figure 6 we can see that a significant share of standard ticket buyers are included in Class 3, while it is not the same for subscribers and loyalty.

[12] This test is given by: $LR = -2 * (LL(\hat{\beta}) - LL(0))$ where $LL(\hat{\beta}))$ is the log-likelihood at the estimated parameters while $LL(0)$ is the log likelihood for the null model. LR is always positive, and distributed chi-squared with degrees of freedom equal to the number of parameters.

attribute $k$ is given by the ratio of the coefficient of the attribute ($\beta_k$) and the price coefficient $\beta_p$:

$$WTP_k = \frac{\beta_k}{\beta_p} \tag{13}$$

However, (13) provides a point estimate, while it is known that the parameters in (13) have a confidence interval and that they distribute asymptotically normal. A solution proposed in the literature is to calculate the confidence interval of the ratio using the Delta method, which allows accurate determination of the standard error of the ratio of two estimators (Daly *et al.*, 2012).
In particular, the standard error of the ratio between two estimated parameters can be measured by the following (Bliemer and Rose, 2013):

$$SE(\beta_k/\beta_p) = \sqrt{\frac{1}{\beta_p^2} \cdot \left[ SE(\beta_k)^2 - \frac{2\beta_k}{\beta_p} \cdot COV(\beta_k, \beta_p) + \left(\frac{\beta_k}{\beta_p}\right)^2 \cdot SE(\beta_p)^2 \right]} \tag{14}$$

Table 9 reports for each customer category the WTP (in Danish krone) and its confidence interval obtained with the MNL model, for switching from the first seat category to a higher quality seat and for switching from a weekday to a weekend performance. Apart from the standard ticket buyers, for the other customer categories it is taken into account that the coefficient attribute is obtained as the sum of the seat category coefficient and the coefficient of the interaction term. Hence, in calculating (13), the standard error of the sum of the two estimated parameters is considered. Given that price coefficients vary according to the production's characteristics, for simplicity we consider in Table 9 a ballet performance with bad review and evaluation; and that the customer buys the ticket in the last booking period. Clearly we don't report those attributes for which either the WTP is either negative (e.g. in the case of weekend performances for seniors), or the attribute coefficient is not significant (e.g weekend performances for loyalty customers).

Using the LC model, the WTP is obtained in a similar manner to MNL given that within each class the parameters are logit.
Table 10 shows the WTP for each latent class. We do not report the WTP for attributes that has a negative coefficient. From the LCM model, WTP in Class 1 and Class 2 is not statistically significant (with the exception of weekend performances for Class 1). For these attributes, we report the point estimate and not the confidence interval.

In general, the WTP values seem large. There can be various reason for this. Firstly, it might be that the customer pay little attention to price when they select the ticket as the result of high inelasticity of theater demand (Zieba, 2009; Grisolia

| Category | Attribute | WTP (DKK) | Standard error | *t*-ratio | 95% confidence interval | |
|---|---|---|---|---|---|---|
| Standard | Seat 2 | 319 | 15.22 | 20.96 | 289.17 | 348.83 |
| | Seat 3 | 665 | 30.40 | 21.87 | 605.42 | 724.58 |
| | Seat 4 | 887 | 39.33 | 22.55 | 809.91 | 964.09 |
| | Seat 5 | 956 | 39.44 | 24.24 | 878.70 | 1033.30 |
| | Wkend | 105 | 6.53 | 16.08 | 92.20 | 117.80 |
| Subscribers | Seat 2 | 762 | 42.82 | 17.79 | 678.07 | 845.93 |
| | Seat 3 | 1172 | 62.26 | 18.82 | 1049.97 | 1294.03 |
| | Seat 4 | 1517 | 78.81 | 19.25 | 1362.53 | 1671.47 |
| | Seat 5 | 1645 | 82.65 | 19.90 | 1483.01 | 1806.99 |
| | Wkend | 11 | 3.55 | 3.10 | 4.04 | 17.96 |
| Senior | Seat 2 | 587 | 78.87 | 7.44 | 432.41 | 741.58 |
| | Seat 3 | 1571 | 108.41 | 14.49 | 1358.52 | 1783.48 |
| | Seat 4 | 1635 | 108.22 | 15.11 | 1422.88 | 1847.11 |
| | Seat 5 | 1690 | 108.72 | 15.54 | 1476.91 | 1903.09 |
| | Wkend | - | - | - | - | |
| Loyalty | Seat 2 | 580 | 92.46 | 6.27 | 398.78 | 761.22 |
| | Seat 3 | 816 | 92.63 | 8.81 | 634.45 | 997.55 |
| | Seat 4 | 1150 | 97.74 | 11.76 | 958.43 | 1341.57 |
| | Seat 5 | 1176 | 96.53 | 12.18 | 986.80 | 1365.20 |
| | Wkend | - | - | - | - | |
| Young | Seat 2 | 210 | 13.88 | 15.13 | 182.80 | 237.20 |
| | Seat 3 | 359 | 19.22 | 18.68 | 321.33 | 396.67 |
| | Seat 4 | 251 | 13.17 | 19.06 | 225.19 | 276.81 |
| | Seat 5 | 227 | 11.94 | 19.01 | 203.60 | 250.40 |
| | Wkend | 35 | 5.88 | 5.95 | 23.48 | 46.52 |
| Foreigner (standard ticket) | Seat 2 | 364 | 25.36 | 14.35 | 314.29 | 413.71 |
| | Seat 3 | 626 | 33.33 | 18.78 | 560.67 | 691.33 |
| | Seat 4 | 961 | 48.13 | 19.97 | 866.67 | 1055.33 |
| | Seat 5 | 1179 | 55.11 | 21.39 | 1070.98 | 1287.02 |
| | Wkend | 281 | 18.19 | 15.45 | 245.35 | 316.65 |

Table 9: WTP based on MNL for switching from Seat1 category and weekday performance

and Willis, 2016).

Secondly, the models are based on an RP data set; hence, we deal with individuals who have already decided to buy a theater ticket. The purchase in itself implies that the WTP is higher than the ticket price (otherwise the individual would not buy the ticket). Conversely, an SP experiment would include the no-purchase option. In any case, it should be pointed out that the choice of which seat category to buy depends on the difference between WTP and the ticket price: for example, looking at the Standard ticket buyers in Table 9, the difference between WTP for the $Seat5$ and $Seat4$ categories is 70DKK. This implies that when the difference in the ticket price between these two seat categories is greater than 70 DKK, the customer will prefer to buy a $Seat4$ category ticket than a $Seat5$ category ticket.

| Category | Attribute | WTP (DKK) | Standard error | *t*-ratio | 95% confidence interval | |
|---|---|---|---|---|---|---|
| Class 1 | Seat 2 | 2759 | 1993.337 | 1.38 | - | |
| | Seat 3 | 3049 | 1993.316 | 1.52 | - | |
| | Seat 4 | 3303 | 1993.374 | 1.66 | - | |
| | Seat 5 | 3613 | 1993.37 | 1.81 | - | |
| | Wkend | 89 | 2.7492 | 32.37 | 83.61 | 94.39 |
| Class 2 | Seat 2 | 614 | 1746.646 | 0.35 | - | |
| | Seat 3 | 18220 | 54062.16 | 0.34 | - | |
| | Seat 4 | 28707 | 85418.41 | 0.34 | - | |
| | Seat 5 | 22991 | 67957.04 | 0.34 | - | |
| | Wkend | - | - | - | - | |
| Class 3 | Seat 2 | 502 | 115.4725 | 4.35 | 275.67 | 728.32 |
| | Seat 3 | 454 | 67.1648 | 6.76 | 322.36 | 585.64 |
| | Seat 4 | - | - | - | - | |
| | Seat 5 | - | - | - | - | |
| | Wkend | 351 | 124.0801 | 2.82 | 107.80 | 594.19 |
| Class 4 | Seat 2 | 68 | 56.87 | 1.19 | - | |
| | Seat 3 | 258 | 6.7481 | 38.23 | 244.77 | 271.22 |
| | Seat 4 | 253 | 6.6976 | 37.77 | 239.87 | 266.13 |
| | Seat 5 | 312 | 6.9706 | 44.76 | 224.27 | 399.73 |
| | Wkend | - | - | - | - | |

Table 10: WTP based on LCM for switching from Seat 1 category and weekday performance

# 7 Conclusions

In a period in which the public funds allocated to cultural organizations are decreasing and performing arts organizations are struggling to attract a broader audience and to achieve a balance between revenue and losses, price discrimination strategy is emerging as a tool to achieve the organizations' revenue and attendance targets. Indeed, offering a schedule of prices according to seat location in the venue, is a practice that allows the theater to discriminate between customers according to their WTP. This paper is a first attempt to develop a discrete choice model that analyzes customers' preference for the attributes connected to the type of tickets, in terms of seat quality and day of performance.

We have employed a data set that includes information on Royal Danish Theatre bookings in the period 2010-2013 with the aim of estimating three discrete choice models that explain ticket purchase behavior. Our analysis reveals some distinguishable patterns that characterize heterogeneous behavior in the choice of theater ticket. This heterogeneity in preferences is strictly connected to customers' characteristics in terms of age, theater attendance frequency and period of purchase. Actually, these customers' characteristics are those that can be inferred from the booking data.

Our findings can provide guidance to policy makers and theater managers in setting prices. Indeed, price is one of the tools with which theater can achieve their aims of increasing both the theater audience and box office revenue. For example, a different pricing policy should be adopted for young and senior customers, as they exhibit opposite patterns in terms of preference and WTP. There is also room for the adoption of dynamic pricing, as the customers who buy a ticket at the end of the sale period are characterized by a higher price sensitivity. Our results also suggest differentiate the pricing policy among production genres, given that the customer's behavior is not homogeneous among genres. Even if some of the above implications can be provided by theater managers based on their experience, our quantitative analysis is needed for different reasons: first, it provides a scientific evidence of these inferences. Second, discrete choice models allow us to measure the extent to which the different ticket attributes affect the choice of a ticket: for example, we have seen that seat category has a greater impact on explaining the choice of ticket than the day of the performance. Third, this approach provides some estimates of the WTP which not only are essential in setting price, but they can also be integrated in more sophisticated optimization models.

Our interpretation of the final results presents a limit: indeed, we do not have access to other socio-economic characteristics, in particular customers' income. This is quite relevant as the utility that each customer maximizes should be subject to the budget (income) constraint. However, we should highlight two facts: first, a theater cannot implement a pricing discrimination policy based solely on the customers' income. Secondly, this paper aims to investigate customer's behavior based on booking data, hence it is based on the information that theatre can normally acquire considering that customers are not required to give information about their income.

Future studies could explore customer behavior in more detail with respect to price differentiation and consider other socio-economic characteristics, such as income, education and family composition. A further study on this topic is important, because theater demand has some peculiar features compared to other industries that adopt RM technique, such as the transportation industries: a multi-objective function that is not limited solely to revenue; a cultural product with personal and subjective value; a product that lacks standardization, and a risk component in demand because of the unknown characteristics of the cultural product before its consumption.

# References


- ABBÉ-DECARROUX, F. (1994). The perception of quality and the demand for services: empirical application to the performing arts. *Journal of Economic Behavior & Organization*, 23(1), 99-107.

- BLIEMER, M. C., ROSE, J. M. (2013). Confidence intervals of willingness-to-pay for random coefficient logit models. *Transportation Research Part B: Methodological*, 58, 199-214.

- BIERLAIRE, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland.

- BILLE, T., BERLE, J.B., SKOT-HANSEN, D. (2015). Theater Management - quality indicators and demand. *Paper presented at the The 13th International Conference on Arts & Cultural Management (AIMAC2015), June 29 Ð July 1*, Aix-en-Provence / Marseilles, France.

- BOXALL, P. C., ADAMOWICZ, W. L. (2002). Understanding heterogeneous preferences in random utility models: a latent class approach. *Environmental and resource economics*, 23, 421-446.

- CARRIER, E. (2008). Modeling the choice of an airline itinerary and fare product using booking and seat availability data. (*Doctoral dissertation), Massachusetts Institute of Technology, Department of Civil and Environmental Engineering.*

- CHAN, T. W., GOLDTHORPE, J. H. (2005). The social stratification of theatre, dance and cinema attendance. *Cultural Trends*, 14, 193-212.

- COLBERT, F., BEAUREGARD, C., VALLÉE, L. (1998). The importance of ticket prices for theatre patrons. *International Journal of Arts Management*, 1, 8-15.

- CORNING, J., LEVY, A. (2002). Demand for live theater with market segmentation and seasonality. *Journal of Cultural Economics*, 26, 217-235.

- DALY, A., HESS, S., DE JONG, G. (2012). Calculating errors for measures derived from choice modelling estimates. *Transportation Research Part B: Methodological*, 46, 333-341.

- DRAKE, M. J., DURAN, S., GRIFFIN, P. M., SWANN, J. L. (2008). Optimal timing of switches between product sales for sports and entertainment tickets. *Naval Research Logistics*, 55, 59-75.

- Felton, M.V. (1994). Evidence of the existence of the cost disease in the performing arts. *Journal of Cultural Economics*, 18, 301-312.

- Fernandez-Blanco, V., Orea, L., Prieto-Rodriguez, J. (2009). Analyzing consumers heterogeneity and self-reported tastes: An approach consistent with the consumerÕs decision making process. *Journal of Economic Psychology*, 30, 622-633.

- Gapinski, J. H. (1984). The economics of performing Shakespeare. *The American Economic Review*, 74, 458-466.

- Greene, W.H., Hensher, D.A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research* B 37, 681-698.

- GrisolŠa, J. M., Willis, K. G. (2011a). An evening at the theatre: using choice experiments to model preferences for theatres and theatrical productions. *Applied Economics*, 43, 3987-3998.

- Grisolia, J. M., Willis, K. G. (2011b). Heterogeneity in willingness to pay for theatre productions: individual specific willingness to pay estimates for theatres, shows and their attributes. *Scottish Journal of Political Economy*, 58, 378-395.

- Grisolia, J. M., Willis, K. G. (2012). A latent class model of theatre demand. *Journal of Cultural Economics*, 36, 113-139.

- GrisolŠa, J. M., Willis, K. G. (2016). Consumer choice of theatrical productions: a combined revealed preference-stated preference approach. *Empirical Economics*, 50, 933-957.

- Hansmann, H. (1981). Nonprofit enterprise in the performing arts. *The Bell Journal of Economics*, 341-361.

- Hetrakul, P., Cirillo, C. (2013). Accommodating taste heterogeneity in railway passenger choice models based on internet booking data. *Journal of Choice Modelling*, 6, 1-16.

- Horowitz, J., 1982. An evaluation of the usefulness of two standard goodness-of fit indicators for comparing non-nested random utility models. *Transportation Research Record*, 874, 19Ð25.

- Huntington, P. A. (1993). Ticket pricing policy and box office revenue. *Journal of Cultural Economics*, 17, 71-87.

- Kamakura, W. A., Russell, G. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, 26, 379-390.

- Laamanen J.P (2013). Estimating demand for opera using sales system data: the case of Finnish National Opera. *Journal of Cultural Economics* 37, 417-432.

- Lancaster K. J. (1966). A new approach to consumer theory. *The Journal of Political Economy*.132-157.

- Leslie, P. (2004). Price discrimination in Broadway theater. *RAND Journal of Economics*, 35, 520-541.

- McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3, 303-328.

- Roeder, K., Lynch, K. G., Nagin, D. S. (1999). Modeling uncertainty in latent class membership: A case study in criminology. *Journal of the American Statistical Association*, 94, 766-776.

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

- Schimmelpfennig, J. (1997). Demand for ballet: a non-parametric analysis of the 1995 Royal Ballet summer season. *Journal of Cultural Economics*, 21, 119-127.

- Seaman, B. A. (2006). Empirical studies of demand for the performing arts. *Handbook of the Economics of Art and Culture* 1, 415-472.

- Sintas, J. L., çlvarez, E. G. (2005). Four characters on the stage playing three games: performing arts consumption in Spain. *Journal of Business Research*, 58, 1446-1455.

- Stigler, G. J., Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67, 76-90.

- Swait, J. (1994). A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. *Journal of Retailing and Consumer Services*, 1, 77-89.

- Swanson, S. R., Davis, C., Zhao, Y. (2008). Art for artÕs sake? An Examination of motives for arts performance attendance. *Nonprofit and Voluntary Sector Quarterly*, 37, 300-323.

- TEREYAGOGLU, N., FADER, P., VEERARAGHAVAN, S. (2012). Filling seats at a theater: Estimating the impact of posted prices and dynamic discounts. Working paper, University of Pennsylvania, Philadelphia.

- THROSBY, D. (1990). Perception of quality in demand for the theatre. *Journal of Cultural Economics*, 14, 65-82.

- URRUTIAGUER, D. (2002). Quality judgements and demand for French public theatre. *Journal of Cultural Economics* 26, 185-202.

- VERMUNT, J. K., MAGIDSON, J. (2005). Technical guide for latent GOLD choice 4.0. Basic and Advanced, *Belmont Massachusetts. Statistical Innovations Inc.*

- WEDEL, M., KAMAKURA, W. A. (2000). *Market segmentation: conceptual and methodological foundations*. Dordrecht: Kluwer.

- WERCK, K., HEYNDELS, B. (2007). Programmatic choices and the demand for theatre: The case of Flemish theatres. *Journal of Cultural Economics*, 31, 25-41.

- WHELAN, G., BATLEY, R., SHIRES, J., WARDMAN, M. (2008). Optimal fares regulation for BritainÕs railways. *Transportation Research Part E: Logistics and Transportation Review*, 44, 807-819.

- WILLIS, K. G., SNOWBALL, J. D. (2009). Investigating how the attributes of live theatre productions influence consumption choices using conjoint analysis: the example of the National Arts Festival, South Africa. *Journal of Cultural Economics*, 33, 167-183.

- ZIEBA, M. (2009). Full-income and price elasticities of demand for German public theatre. *Journal of Cultural Economics*, 33, 85-108.

# Appendix

## Productions considered in the models

| Title | Season | Genre |
|---|---|---|
| Et Folkesagn | 2010/2011 | Ballet |
| Boris Godunov | 2010/2011 | Opera |
| Madame Butterfly | 2010/2011 | Opera |
| Kvinden uden skygge | 2010/2011 | Opera |
| Balletaften | 2010/2011 | Ballet |
| Broadway for en aften | 2011/2012 | Ballet |
| Alceste | 2011/2012 | Opera |
| En skærsommernatsdrøm | 2011/2012 | Play |
| Così fan tutte | 2011/2012 | Opera |
| Kameliadamen | 2011/2012 | Ballet |
| Mågen | 2011/2012 | Play |
| Parsifal | 2011/2012 | Opera |
| Nøddeknækkeren | 2011/2012 | Ballet |
| Den Gerrige | 2011/2012 | Play |
| Albert Harring | 2011/2012 | Opera |
| Tannhäuser | 2012/2013 | Opera |
| Den fiffige lille ræv | 2012/2013 | Opera |
| Romeo & Juliet | 2012/2013 | Ballet |
| Madame Butterfly | 2012/2013 | Opera |
| Vildanden | 2012/2013 | Play |
| La Bayàdere | 2012/2013 | Ballet |
| Kollektivet | 2012/2013 | Play |
| La Ventana / Kermessen i Brügge | 2012/2013 | Ballet |

# A DEA approach for selecting a bundle of tickets for performing arts events

Andrea Baldin[*]

*Department of Management, Ca' Foscari University of Venice, Cannaregio 873, 30121 Venice, Italy*

**Abstract**

Most performing arts organizations offer its customer the choice of either buying event tickets individually or buying a bundle of tickets for two or more events. During the selection of the bundle to be offered, the theatre manager faces several possible combinations of events. In this paper we tackle the issue of identifying the most efficient subset of the events scheduled to offer as a bundle. We formulate this problem following the choice-based network Revenue Management approach. Assuming price as fixed on two type events - lowbrow and highbrow - proposed by the theatre, the purchase decision is modelled on the basis of two random variables: available time and reservation price per perfomance. The super-efficiency DEA model will be implemented in order to find the most efficient combination of events to be bundled, defined as the one that offers the most favourable trade-off between expected revenue, attendance and capacity consumption. A regression of the DEA scores on managerial variables and bundle attributes will allows us to obtain some insights into what determines the efficiency level of a bundle.

**Keywords** Data envelopment analysis (DEA); Revenue Management; Bundling; Performing arts institutions

[*]E-mail: andrea.baldin@unive.it

# 1 Introduction

One of the strategic decisions that performing arts organizations face is the selection of event tickets to offer as a bundle. In the Revenue Management (RM) framework, this issue can be classified into a structural-decision, as it refers to the mechanism to use for selling products - in our case, theatre tickets.

Broadly speaking, bundling is the strategy of marketing two or more products and/or services as a single package at a special price (Guiltinian, 1987). This is a pervasive practice in the marketing of theatres that is realized through subscription offers. The main rationale for bundling is identified in its capacity to implicitly price discriminate. Let consider this example taken from Venkatesh and Mahajan (1993): the reservation prices for two events A and B are respectively \$10 and \$20. If the price of a single ticket is \$19, the consumer will buy only the ticket for the performance B. If a subscription priced at \$28 is offered, the consumer will purchase the bundle of events. In economics terms, bundle allows to transfer the consumer's surplus (i.e the difference between the reservation price and the actual price paid) from the event evaluated more to the less attractive one. The advantage of bundling is, therefore, its ability to segment the market, reducing the variability in demand that derive from consumers' heterogeneous preference. In our case bundling tickets allows to increase both the attendance and revenue of potentially low demand events when they are sold combined with high demand events. There are other motivations behind selling bundle of tickets that are strictly connected with the theater context: to obtain in advance an amount of cash flow with which base future decisions; and to establish a loyal customer base. Subscribers are a very important customer segment for a theatre as they:1) assure a certain level of attendance for the theatre season, 2) are more likely to renew tickets in the future (Duran *et al.*, 2012); and to spread positive word of mouth that induces others customers to buy tickets (Drake *et al.*, 2008).

From the demand side, the incentive for the customer to buy a bundle is given by the savings derived by a lower unit price for ticket and by non-monetary compensation related to the symbolic dimension of a loyalty program (Johnson and Garbarino, 2001).

Given the discussion above, it is fundamental for the theatre to determine the content of the bundle. Most non profit performing arts organizations offer not only events that are very popular and whose tickets are easily sold, but also productions that have a low demand since are difficult to master for the non-expert. The reason behind this cultural policy lies in the need for the non profit performing arts organizations to meet different and conflicting objectives, as illustrated by Hans-

mann(1981): the popular events (denoted in literature as lowbrow events) fulfill the attendance goal, i.e to spread culture to as broad a segment of population as possible; on the contrary, the less popular events (denoted as highbrow events), improve the tastes of a small number of connoisseurs and fulfill the quality goal.
In this paper we study the most efficient bundle composition of a non profit theatre, considering that such organization achieves its aims through a repertoire composed by highbrow and lowbrow titles. Generally speaking, it is known that it is better to bundle high demand events together with low demand events, than bundling events with similar demand, because the high demand events induces customer to attend the low ones. On the other hand, bundling two high demand events may result in a decrease of potential revenue as the unit price for event in a bundle is lower than the single ticket price; similarly, a bundle of low demand events would result not attractive for customers. This is just a generalization, as the best bundle composition depends on factors as the price of single ticket and bundled tickets and the distribution of reservation price of different events. Our aim is to identify a possible methodology in order to identify the most efficient bundle composition.
From the methodological perspective, the novelty of this study is that we treat the bundle composition decision as a network RM problem. Network RM is commonly referred to a quantity-based problem in the airline industries, in which the firms has to manage the capacity of connecting flights in a network. In general, network RM deals with industries that sells product consisting of more than one resource. In our case, we treat each event as a resource with which theatre offers products of one resource (single ticket) or multiple resource (bundle of tickets). This paper relies on Talluri and Vany Ryzin (2004) and Liu and Van Ryzin's (2008) studies that have introduced, in a costumer choice behaviour setting, the notion of efficient offer set as the one that offers the most favourable trade-off between expected revenue and capacity consumption. In order to model the consumer choice behaviour, we adapt the Venkatesh and Mahajan (1993) and Ansari *et al.*'s (1996) framework considering non-homogeneous (in terms of demand) events, in which the consumer decision making process is based on two criteria: available time to attend the events and reservation price per perfomance. In such a way we will obtain the choice probability for each bundle and consequently its expected revenue per consumer, capacity consumption and attendance. These value will constitute the variables of the super-efficiency DEA model that will identify the most efficient offer set. DEA seems to be a suitable tool in our context, given the multi-objective nature of the non profit organizations.
To the best of our knowledge, this is the first paper that tackles the issue of bundle

selection in the performing arts context, designing a new approach to this topic. The analysis so far provides a set of insights in what determines the efficiency level of a bundle, even if our model considers only two type of events and two different pattern of reservation price. Indeed, a more realistic setting would require a significantly higher computationale effort.
The remainder of the paper is organized as follow: Section 2 provides a brief literature review. Section 3 describes the general model including the notion of efficiency, the super-efficiency DEA model and the specification of input/output data based on the multicriteria customer choice behaviour. Section 4 provides a numerical example. Finally Section 5 presents some concluding remarks.

## 2 Literature review

A large amount of RM literature has investigated issues related in particular to the airline industry: such studies include two-fare class problem (Littlewood, 1972), seat allocation for multiple classes (Belobaba, 1987), multi-leg (network RM) problem (Talluri and Van Ryzin, 1998). There is very little empirical research that has been done in the area of RM applied to performing arts organizations. Most of it has focused on the price discrimination practice implemented by theatres. Huntington (1993) considers a variant of the hedonic pricing model to describe price differentiation by seat, and shows that the price discrimination policy leads to a greater profit than the unique price policy. Rosen and Rosenfield (1997) focus on the issue on how the theatre should sort seats in categories, whereas Leslie (2004) analyzes the price discrimination policy for a particular Broadway show in order to estimate a structural econometric model of price discrimination based on the individual consumer behaviour. Tereyagoglu *et al.* (2012) employs a competing hazard framework to model the ticket sales, where the customers race against each other for the same ticket. The aim of their work is to analyze how pricing and discount actions over time affect the timing of customers' purchase as well as the propensity to purchase a ticket of different categories of customers (subscribers and occasional buyers).
A stream of literature related to our work is on bundling strategies, a topic broadly studied from both economic and marketing approach. The first perspective has a normative nature, providing stylized analytical models. In this area we find the first studies on bundling, starting from Stigler (1963) who analyzes the simple case of a mopolist that offers two products with perfectly negatively correlated price, to the seminal paper by Schmalensee (1984) in which the negative correlation of reservation price is relaxed. Subsequent studies have considered the role of competition

(Matutes and Regibeau, 1992), bundles with more than two products (Bakos and Brynjolfsson, 2000) and the role of complement and substitutes product (Venkatesh and Kamakura, 2003).

The marketing approach has contributed to the bundling research in two way: first, developing specific methodologies, such as conjoint analysis (Goldberg *et al.*, 1984), balance model (Bradlow and Rao, 2000), mixed integer linear programming (Hanson and Martin, 1990), probabilistic modelling (Venkatesh and Mahajan, 1993) and combinatorial methods (Chung and Rao, 2003); in order to assist the decision makers in designing and pricing the bundle. Second, highlighting different issues directly related with bundling, such as co-branding (Venkatesh and Mahajan, 1997) and product integration (Stremersch and Tellis, 2002). The existing literature on bundling is very rich: for a detailed review we refer to Venkatesh and Mahajan (2009) and Rafiei *et al.* (2013). One of the main result of the literature is that, in fairly general setting, mixed bundling is likely to be the optimal strategy. However, in order to determine the optimal strategy, factors such as substitutability and complemetarity relations, distribution of the reservation prices and the correlation among them should be considered.

Focusing on the performing arts sector, the work most related with this study is that of Venkatesh and Mahajan (1993): here the authors propose a probabilistic optimization model to derive, for each of the three bundling strategies, the price that maximizes the organization's profit, and evaluate which approach assures the maximum profit. The consumers' decision making is determined by the available time to attend performances and the reservation price, assuming that the density function of these two random variables follows a Weibull distribution whose parameters are estimated empirically through a survey submitted to the customer. The final results indicate that the mixed bundle strategy lead to a greater profit. Ansari *et al.* (1996) extend Venkatesh and Mahajan's (1993) model by considering a non profit performing arts organization whose objective function is the maximization of number of users, subject to a non-deficit constraint. The authors show that also for such organizations mixed bundling is the optimal strategy.

Other papers has analyzed bundling strategy under a RM perspective. Drake *et al* (2008), observing that most sports and entertainment organizations are used to offer bundle of tickets first and put single-event tickets on sale at a later date, find the optimal timing decision for switch the product sale. The authors model the demand as a linear Markovian process and generalize the model considering products with different demand characteristics. Duran *et al* (2012) extend Drake's *et al* (2008) works, considering the dynamic switching time decision under Poisson demand pro-

cess. Yakici *et al* (2014), similarly to our study, consider the problem of the selection of event tickets to include into the bundle. In the case of bundle of three or more events, the authors propose a heuristic approach to create the best bundle. Compared to us, the authors consider not only the demand as a guidelines in bundling, but also the timing dimension of the event. However, the number of events to be included in the bundle is pre-determinated by the decision maker. Finally, it should be considered the study by Ferreira and Wu (2011) since it is the only work that adopt Data Envelopment Analysis as a tool to make the bundle selection decision. In this study the cost of the bundle and its price are used respectively as the input and the output of the DEA model.
This paper aims to contribute to the RM literature under different aspects: first, at best of our knowledge, this is the first paper that considers the bundle selection problems in the performing arts context, in which the organizations pursue different objectives in addition to revenue. Second, we propose a new approach to the bunding problem adopting a network RM perspective and an integrated model that make use of a probabilistic approach and super-efficiency data envelopment analysis.

# 3 Model

## 3.1 The bundling problem

We consider a non profit theatre that offers a portfolio of $m$ events during the performance season of a theatre. Let us assume that it is possible to distinguish two groups of events: high demand and low demand events. The theatre offers both single tickets for each event and a bundle of tickets. The problem considered here is to determine a bundle composition that allows the organization to fulfill its objective: revenue and attendance. We formulate this problem following the choice-based network revenue management approach (Liu and Van Ryzin, 2008). The problem is defined in terms of resources and products. The $m$ events are the resources with which the organization can provide a set $N$ of products which are the single events and all the possible combinations of events that can be offered in bundle. Each product has an associated revenue which corresponds to its selling price: we denote with $p$ the price of the single ticket and $p_B$ the price of the bundle. We assume that the price $p$ is equal across performances. We further assume that there is no form of price discrimination other than bundling. In other words, a unique price is charged for each product without considering, for example, the quality of seat, customer segment or the timing of purchase. Moreover, even if the events differ in terms of demand, we assume that the theatre does not modify the ticket

price according to the popularity of the event. On the contrary, it applies uniform pricing. This assumption is not unrealistic at all. The argument put forward by Orbach & Einav (2007) and Choi *et al.* (2014) regarding the price discrimination in the context of movie theatres can be applied also in that of the performing arts: price discrimination based on the popolarity is seen by customers as unfair. Hence, in the presence of events with different demand characteristics, price bundling is the most accepted way to differentiate price of different events (see Courty, 2000).
We assume that the organization selects a set $S \subseteq N$ of products made up of all the single tickets for the events $i$ and a single bundle of tickets, denoted with $B$, for $k$ different events, so as $S = \{1, \ldots, m, B\}$. $S$ is denoted as the organization's *offer set*. Given the offer set $S$, we denote with $P_i(S)$ and $P_B(S)$ the probability of buying respectively the single ticket of the event $i$ and the bundle $B$, that will be calculated in the section 3.3. Changing the offer set $S$, it will change also the probability of buying either a single ticket or the bundle or neither of them. In fact, given that the model incorporates the customer choice behaviour, it is not assumed that the consumer is a bundle purchaser (i.e subscriber) or a single ticket purchaser independently by the control applied by the seller (as in the traditional RM methods based on the independent demand model); on the contrary the choice of the consumer depends on which is the offer set available (Shen & Su, 2007).
We denote as $R(S)$ the expected revenue when $S$ is offered:

$$R(S) = \sum_{i=1}^{m} P_i(S) \cdot p + P_B(S) \cdot p_B \tag{1}$$

The price of the bundle is always less than the sum of the single prices of the events in the bundle. Moreover, we assume that the unit price of the event in the bundle is decreasing with respect to the cardinality $k$ of the bundle. This assumption derives from the principles of the second degree price discrimination, according to which the unit price varies depending on the quantity demanded. In our case, where all the $k$ events of the bundle are sold at the same price, we hypothesize that the price $p_B$ of the bundle can be formulated as follows:

$$p_B = k[p \cdot (1 - r)^{k-1}], \tag{2}$$

where $r$ is the discount rate [1]. Equation (2) allows us to have a formulation of $p_B$ that is increasing and concave with respect to the size of the bundle $k$.
Further, we denote with $Q_i(S)$ the probability of using a unit of capacity - i.e, a

[1] In other words, the unit price of an event included in a bundle is discounted by $[(1-r)^{k-1} \cdot 100]\%$ respect to a full price single ticket

seat in the theatre - of the event $i$, given by:

$$Q_i(S) = \begin{cases} P_i(S) + P_B(S), & \text{if } i \in B, \\ P_i(S), & \text{if } i \notin B. \end{cases}$$

Hence $Q(S) = (Q_1(S), \ldots, Q_m(S))^T$ is the vector $(m \times 1)$ of capacity consumption probabilities that considers both the single tickets and the bundle sold.
Liu and Van Ryzin (2008) have shown that the efficient offer sets are the ones to use in the network RM problem. Their concept of efficient offer sets is the natural extension of that introduced by Talluri & Van Ryzin (2004) for the choice based single product RM model. The key concept is that the efficient sets are those that provide the most favorable trade-off between expected revenue and capacity consumption. Recalling their definition, a set $\hat{S}$ is *inefficient* with respect to the other offer set, if there exist a convex combination of the alternative offer sets that generate a greater expected revenue consuming the same (or less) capacity.
More formally, denoting with $\alpha(S)$ the weights for convex linear combination, if there exist $\alpha(S)$, $S \subseteq N$, with $\alpha(S) \geq 0$ and $\sum_S \alpha(S)$ =1 such that (Talluri and Van Ryzin, 2004):

$$R(\hat{S}) < \sum_S \alpha(S) R(S) \tag{3}$$

and

$$Q(\hat{S}) \geq \sum_S \alpha(S) Q(S), \tag{4}$$

then the offer set $\hat{S}$ is *inefficient*. It can be shown that the above mentioned conditions of inefficiency are equivalent of those resulting from the dual output-oriented DEA (Data Envelopment Analysis) model with variable return to scale (see Banker *et al.*, 1984). Therefore, a measure of efficiency of the offer set $(S)$ can be obtained implementing the DEA technique (see Section 1.1). DEA is a multicriterial technique that seems to be particularly suitable in the context of non profit organizations, given that the maximization of revenue is not the unique objective. According to Hansmann's (1981) seminal paper, performing arts organizations aim to maximize three objectives: revenue, quality and attendance. The first one is measured as the expected revenue $R(S)$ of the offer set. The second objective is not considered in this study since the concept of quality and its assessment with regard to cultural goods may raise some methodological challenges: indeed, their value is based on abstract, subjective, and experience-related aspects that make difficult to assess their quality (see Caves, 2000). Regarding the third objective, we refer to Section 3.3 for a detailed formulation. Denoting with $A(\hat{S})$ the attendance obtained with the offer set $(\hat{S})$ and with $Y(\hat{S}) = (R(\hat{S}), A(\hat{S}))$ the vector of the outcomes of the offer set $T$, (5) becomes:

$$Y(\hat{S}) < \sum_{S} \alpha(S) Y(S) \tag{5}$$

Apparently there is a linear relationship between $Q(S)$ and $A(S)$. In reality, as it will be shown in Section 3.3, $A(S)$ depends also on the capacity of the theatre, denoting how the efficiency of the bundle depends also on the capacity constraint. Moreover, in calculating the total attendance, we should also consider the situation in which the subscriber does not attend all the events included in the bundle.

## 3.2 Data Envelopment Analysis

DEA is a mathematical optimization model formulated as a linear programming problem, which measures the relative efficiency of decision making units (DMUs) that use different inputs in order to produce different outputs. In the DEA literature, a key assumption is the homogeneity of the DMUs. It means that each DMU uses the same input and output measures (variable in the amount from one DMU to another). In our case, each offer sets can be considered as a DMU that converts a rate of capacity consumption (the input vector $Q(S)$) into an amount of expected revenue and attendance (the output vector $Y(S)$).
Adopting the notation in Section 3.1, the efficiency score of the offer set $\hat{S}$ can be calculated through the following linear programming model (Banker *et al.*, 1984):

$$\max_{\theta, \alpha(S)} \quad \theta_{\hat{S}} \tag{6a}$$

$$\text{subject to} \quad Q(\hat{S}) \geq \sum_{S} \alpha(S) Q(S), \tag{6b}$$

$$\theta_{\hat{S}} Y(\hat{S}) \leq \sum_{S} \alpha(S) Y(S), \tag{6c}$$

$$\sum_{S} \alpha(S) = 1. \tag{6d}$$

Model (6a)-(6d) is formulated following the output orientation, meaning that $\theta_{\hat{S}}$ is a scalar $\geq 1$ which indicates how much the offer set $\hat{S}$ should radially increase their outputs in order to achieve the efficient frontier. Hence, an efficiency measure $\theta_{\hat{S}} = 1$ characterizes an efficient offer set, whereas a value of $\theta_{\hat{S}} \geq 1$ indicates that the offer set is dominated by other ones.
It can be easily verified that, for an inefficient offer set $\hat{S}$ - i.e $\theta_{\hat{S}} \geq 1$ - the conditions

(4) and (5) are satisfied. On the contrary, if the offer set $\hat{S}$ is efficient - i.e $\theta_{\hat{S}} = 1$ - condition (5) is not satisfied as we have an equality relation for at least one element of the output vector $Y(\hat{S})$.
Note that detecting the efficient offer sets via DEA does not violate the basic assumption implied by Liu and Van Ryzin (2008): in particular the main assumptions of convexity of production set that implies concavity of the efficient frontier (see Lemma 1 in Talluri and Van Ryzin, 2004 and p.307 in Liu and Van Ryzin, 2008) is implied also in the DEA methodology (see Charnes *et al.*, 1978). Even if Talluri and Van Ryzin (2004) and Liu and Van Ryzin (2008) propose a different approach to characterize the efficient offer sets, we believe that DEA seems to be suitable in our case for a couple of reasons:

- the multi-objective nature of the nonprofit organizations,
- the possibility to exted the model in order to rank the efficient offer sets.

In particular, the second aspect is crucial: network RM literature has focused on the airline industry, where different efficient offer sets are offered according to the remaining capacity. Performing arts organizations are less flexible in their supply of products: the bundle of tickets is offered before the beginning of the season and its composition does not change during the selling period. Therefore, the organization must choose the "most efficient" offer set among the efficient set.
In order to identify such offer set, we employ an extension of the basic DEA model (6a-6d) denoted as "Super-efficiency DEA", introduced by Andersen and Petersen (1993). This method has the desiderable feature of discriminating the efficient DMUs. Indeed, the standard DEA model assign the equal score to all the efficient units that lie on the efficient frontier. Basically, in order to overcome this problem, the authors propose to exclude from the reference set the DMU under evaluation. Hence, the super-efficiency DEA model that will be used in this paper can be expressed as:

$$\max_{\theta,\alpha(S)} \quad \theta_{\hat{S}} \tag{7a}$$

$$\text{subject to} \quad Q(\hat{S}) \geq \sum_{\substack{S \\ S \neq \hat{S}}} \alpha(S) Q(S), \tag{7b}$$

$$\theta_{\hat{S}} Y(\hat{S}) \leq \sum_{\substack{S \\ S \neq \hat{S}}} \alpha(S) Y(S), \tag{7c}$$

$$\sum_{\substack{S \\ S \neq \hat{S}}} \alpha(S) = 1. \tag{7d}$$

When an efficient DMU is eliminated from the reference set, the efficient frontier created by the remaining DMUs will shrink. The effect of this is that the efficient unit eliminated will have a value of $\phi_0 \leq 1$, indicating the maximum proportional decrease in the outputs that allows the DMU to mantain the "efficient" status with respect to the new frontier; whereas the score for the inefficient DMUs is unaffected. We have three cases: if $\phi_0 < 1$, the DMU is super-efficient, if $\phi_0 = 1$, the DMU is efficient, if $\phi_0 > 1$ the DMU is inefficient. Indeed such DMUs are not able to project onto the efficiency frontier derived by its exclusion, regardless of the rate of output decrease. Some scholars (Lovell and Rouse, 2003; Chen, 2005; Cook *et al.*, 2008; Lee et al., 2011; Cheng and Liang, 2011) have put effort in providing a numerical score for those efficient DMUs for which the super-efficiency DEA model is unfeasible. In this paper we adopt the approach proposed by Chen (2005) that characterizes the super-efficiency of the DMU for which the model (7a)-(7d) provides an infeasible solution, and has the advantage to allow the employment of the conventional DEA software.
Therefore, the super-efficiency DEA model (7a)-(7d) is applied to all the offer sets. Such model provides the efficiency score to all the offer sets for which the solution is feasible. Concerning the other offer sets, the efficiency score is obtained through the following procedure (Chen, 2005): first, the inefficient DMUs are radially projected onto the efficient frontier through the proportional input reduction, using the input-oriented DEA model:

$$\min_{\beta,\alpha(S)} \quad \beta_{\hat{S}} \tag{8a}$$

$$\text{subject to} \quad \beta_{\hat{S}} Q(\hat{S}) \geq \sum_S \alpha(S) Q(S), \tag{8b}$$

$$Y(\hat{S}) \leq \sum_S \alpha(S) Y(S), \tag{8c}$$

$$\sum_S \alpha(S) = 1. \tag{8d}$$

As a second step, the super-efficiency model (7a)-(7d) is implemented replacing $Q_{\hat{S}}$ with the efficient input values $\hat{Q}_{\hat{S}} = \beta_{\hat{S}} \cdot Q_{\hat{S}}$. We obtain a new super efficiency score for the DMU for which the original super efficiency model is infeasible. In case we don't obtain again a feasible solution, it means that this DMU is efficient but not characterized by output surplus, and its score is set equal to 1.

## 3.3 Specification of Input/Output Data

To derive the choice probabilities of the different offer sets, essential for the construction of the input and output data set, we rely on Venkatesh and Mahajan (1993) and Ansari *et al.* (1996). According to them, the consumer's choice is affected by two dimensions: leisure time and price. The central role of these dimensions in the performing arts sector is pointed out by Vogel (1990) and is empirically supported by many econometric demand studies (for a summary see Seaman, 2006). Following Venkatesh and Mahajan (1993), time and money dimensions are captured by two variables, respectively the number of performance a person is likely to attend, and his reservation price. An important assumption is that the two variables are independent each other, whereas it should be natural to think that there is an inverse relationship between them, as a high cost of time (that implies a low level of leisure time) is associated with high income (hence higher reservation price), and viceversa. However, Venkatesh and Mahajan (1993) claim it is possible that members of a high income family may have a large amount of leisure time. On the other hand, people with lower wage rates may have to work more hours, which results in a higher cost of time.
We consider three random variables:

- $X$ is the number of performances $i$ (among $m$) a person can attend,
- $R_h$ and $R_h$ are the performancewise reservation price respectively for the low and high demand events.

Whereas Venkatesh and Mahajan (1993) assume that, at market level, the reservation price distributions are the same across performances, we made a simpler assumption: such distributions are the same across type of performances at individual - and not market - level. It means that for an indivual the reservation price is the same for all the performances of the same type. In such a case, the person's mean reservation price, denoted with $\bar{R}$, is no more modelled as random variable whose parameters are to be estimated, as in Venkatesh and Mahajan (1993), but it is obtained as the average of the performancewise reservation prices of the events included in the bundle. This implies that the reservation price of a bundle is equal to the sum of the reservation price of the individual element: this is the additivity assumption (Venkatesh and Mahajan, 2009) that in our case seems suitable as the products are neither complementary nor substitutive.
Moreover we assume the theatre is a monopolist, so it does not compete with other organizations in attracting the customers.

In our setting the organization has scheduled $m$ events. Among the $m$ events that take place, let $a$ be the number of high demand events (for ease of exposition we denote these as events of type $h$) and $b$ the number of low demand events (denoted as events of type $l$), such that $a + b = m$. Considering the events of the same type indistinguishable from the demand perspective, there exists $n=(a \cdot b)+m-2$ possible bundling combinations[2]. Each possible bundle contains $k$ events, with $2 \leq k \leq m$. As already said, each offer set $S$ consists of single tickets for all the events and one possible bundle. So each offer set differs from each other by the combinations of events that define the bundle $j$.
We denote with $\hat{a}$ and $\hat{b}$ the number of events of type $h$ and $l$ included in the bundle, so that each bundle is denoted by the pair of numbers ($\hat{a}$, $\hat{b}$).
Whereas Venkatesh & Mahajan (1993) consider events homogeneous from the demand side, in this regard we deal with two types of events. This leads to a different formulation of the probability of purchasing either one single ticket or a bundle of tickets. Denting with $\bar{R}$ the mean reservation price, when the events are homogeneous, the sufficient condition of buying a bundle is that the $(p_B/i) < p_s$ and $\bar{R} > (p_B/i)$, that is, the average price paid for each performance attended is less than both the price of the single ticket and the person's mean reservation price. However, in our case this is not sufficient. In fact $\bar{R}$ can varies according to the composition of the bundle and customer's preference about the type of event. Given this, the choice decision on what to purchase is based on what maximizes the difference between the reservation price and the ticket/bundle's face value.

The market can be divided in six market segments. The segmentation is based on:

- The preference on the type of event
  - $Pr(R_h > R_l)$: portion of the market that prefers events of type $h$,
  - $Pr(R_h < R_l)$: portion of the market that prefers events of type $l$.
- The reservation price of the event of type $h$
  - $Pr(R_h > p)$: portion of the market that can buy a single ticket of $h$,
  - $Pr(R_h < p)$: portion of the market that never buys a single ticket of $h$.
- The reservation price of the event of type $l$

[2] $(a \cdot b)$ is the number of possible bundles that contain both events of type $h$ and $l$. In addition to these, there are $a - 1$ bundles that contain only type $h$ events and $b - 1$ bundles that contain only type $l$ events

- $Pr(R_l > p)$: portion of the market that can buy a single ticket of $l$,
- $Pr(R_l < p)$: portion of the market that never buys a single ticket of $l$.

Table 1 illustrates the characteristics of each market segment.
All market segments can potentially buy a bundle, but two of these will buy only

| **Market segment** | **Market portion** | **Purchase choice** |
|---|---|---|
| 1 | $Pr(R_l > p \mid R_h > R_l) \cdot Pr(R_h > R_l)$ | Bundle or single tickets of events $h$ and $l$ |
| 2 | $Pr(R_h > p \mid R_l > R_h) \cdot Pr(R_l > R_h)$ | Bundle or single tickets of events $l$ and $h$ |
| 3 | $Pr(R_h > p \mid R_l < p) \cdot Pr(R_l < p)$ | Bundle or single tickets of events $h$ |
| 4 | $Pr(R_l > p \mid R_h < p) \cdot Pr(R_h < p)$ | Bundle or single tickets of event $l$ |
| 5 | $Pr(R_h < p \mid R_l < R_h) \cdot Pr(R_l < R_h)$ | Bundle or no purchase |
| 6 | $Pr(R_l < p \mid R_h < R_l) \cdot Pr(R_h < R_l)$ | Bundle or no purchase |

Table 1: Characteristics of the market segments

a bundle and not single tickets; whereas there are two market segments that, in addition to the bundle, can buy also single tickets of high and low demand events; one market segment that potentially can buy single tickets of only high demand events and one market segment that can buy single tickets of only low demand events.
Given this, the decision on what to purchase is based on the comparison between, on one side the surplus obtained by the bundle (and eventually $(i-k)$ single tickets (in case $i > k$)) and, on the other side, the surplus derived from the purchasing of only single tickets. Distinguishing the market segments is crucial in this comparison. Indeed, in calculating the surplus of the bundle, for customers for which $i < k$, the number of events to be considered inside the bundle depends on the preference on the type of events. The surplus of several single tickets depends not only on the preference on the type of perfomance, but also on the reservation price of the not preferred event.
Given a bundle $(\hat{a}, \hat{b})$ of size $k = \hat{a} + \hat{b}$, let denote with $\gamma_s(i)$ and $\phi_s(i)$ the surplus derived when the consumer, belonging to the market segment $s = 1, 2, ...6$, purchases respectively only single tickets or the bundle (and eventually single tickets in addition to the bundle).
Let consider the first market segment (i.e $s = 1$): customers for which $i < a$ will attend $i$ events of type $h$ and no events of type $l$; otherwise, if $i > a$, customers will attend $a$ events of type $h$ and $i - a$ events of type $l$. Hence customers of this market segment will attend $min\{a, i\}$ events of type $h$ and $max\{i - a, 0\}$ events of type $l$. We obtain $\gamma_1(i)$ as:

$$\gamma_1(i) = R_h \cdot min\{a, i\} + R_l \cdot max\{i - a, 0\} - p \cdot i$$

If customers of the first segment decide to buy the bundle, we have to distinguish several cases:

- if $i \leq \hat{a}$, customers will attend $\hat{a}$ events of type $h$,
- if $\hat{a} < i \leq k$, customers will attend $\hat{a}$ events of type $h$ and $i - \hat{a}$ events of type $l$ included in the bundle,
- if $i > k$, customers will attend all the events on the bundle and, moreover, they will buy single tickets for $i - k$ events, broken down as follows:
  - if $(i - k) < (a - \hat{a})$, customers will attend $(i - k)$ events of type $h$ that are not included in the bundle, and no events of type $l$;
  - if $(i - k) > (a - \hat{a})$, customers will attend $(a - \hat{a})$ events of type $h$ that are not included in the bundle, and $i - a - \hat{b}$ events of type $l$.

Given this, we can derive $\phi_1(i)$ as:

$$\begin{aligned}\phi_1(i) =& R_h \cdot min\{\hat{a}, i\} + R_l \cdot max\{min\{i, k\} - \hat{a}, 0\} - p_B + R_h \cdot min\{(max\{i - k\}, 0), (a - \hat{a})\} \\ &+ R_l \cdot max\{(i - a - \hat{b}), 0\} - p \cdot \{max(i - k), 0\}.\end{aligned}$$

The third segment differs from the first one only as regards the impossibility to buy single tickets of event of type $l$. Hence $\gamma_3(i)$ and $\phi_3(i)$ are easily obtained from $\gamma_1(i)$ and $\phi_1(i)$ not considering the purchase of single tickets for $l$ events:

$$\gamma_3(i) = R_h \cdot min\{a, i\} - p \cdot min\{a, i\};$$

$$\begin{aligned}&\phi_3(i) = R_h \cdot min\{\hat{a}, i\} + R_l \cdot max\{min\{i, k\} - \hat{a}, 0\} - p_B \\ &+ R_h \cdot min\{(max\{i - k\}, 0), (a - \hat{a})\} - p \cdot min\{(max\{i - k\}, 0), (a - \hat{a})\}.\end{aligned}$$

As regards customers belonging to the fifth market segment, we know that they don't buy single tickets because their surplus would be negative. Hence $\gamma_5(i) = 0$; whereas $\phi_5(i)$ can be easily obtained by $\phi_3(i)$ not considering the purchase of single tickets:

$$\phi_5(i) = R_h \cdot min\{\hat{a}, i\} + R_l \cdot max\{min\{i, k\} - \hat{a}, 0\} - p_B.$$

The second, fourth and sixth market segments are specular respectively to the first, third and fifth market segments, differentiating for a difference preference for the type of events. Hence $\gamma_2(i)$, $\phi_2(i)$, $\gamma_4(i)$, $\phi_4(i)$, $\phi_6(i)$ can be easily obtained respectively from $\gamma_1(i)$, $\phi_1(i)$, $\gamma_3(i)$, $\phi_3(i)$, $\phi_5(i)$ substituting $R_h$, $R_l$, $a$, $\hat{a}$, $b$, $\hat{b}$ respectively with $R_l$, $R_h$, $b$, $\hat{b}$, $a$, $\hat{a}$.

Thus, given a bundle $(\hat{a}, \hat{b})$ of size $k = \hat{a}+\hat{b}$, a consumer belonging to market segment $s = 1, 2, 3, 4$ buys a bundle of tickets if his surplus is greater than the one obtained buying several single tickets, i.e if $\phi_s(i) \geq \gamma_s(i)$, whereas if $s = 5, 6$, $\phi_s(i)$ has to be strictly positive.[3]
We denote with $Pr_s$ the probability that a customer belong to the market segment $s$, with $s = 1, 2...6$; moreover we denote with $T = s$ the event that the customer belongs to the market segment $s$. Now we can define the probabilty $Pr_B$ of buying a bundle as:

$$P_B = \sum_{s=1}^{6} Pr_s \sum_{i=1}^{n} Pr(X = i) \cdot Pr(\phi_s(i) \geq \gamma_s(i) \mid T = s). \tag{9}$$

Concerning the probability of buying a single ticket of an event $h$ and $l$, we have to distinguish among events that belongs to the bundle $B$ and those which don't. Individuals of market segments 1 and 3 buy a single ticket of an event $h \in B$ whether $\phi_s(i) < \gamma_s(i)$. Hence $Pr(\phi_s(i) < \gamma_s(i))$ gives the probability to buy at least one event of type $h \in B$. Assuming that all the events of the same type has the same probability to be bought through a single ticket, the probability of buying a ticket for the specific event $h$ (that can be or not be included in the bundle) for segment 1 and 3 is equal to $min\{1, i/a\}$.
Also an individual of market segment $s = 2$ can buy a single ticket for an event $h_B$ if $i > b$. In this case the probability that a specific event is attendend is equal to $(i - b)/a$.
Thus the probability that a single ticket for a specific event $h \in B$ is sold is given by:

$$\begin{aligned} P(h_B) = &\sum_{s=1,3} Pr_s \sum_{i=1}^{n} \cdot Pr(X = i) \cdot Pr(\phi_s(i) < \gamma_s(i) \mid T = s) \cdot min(1, i/a) \\ &+ \sum_{i=b+1}^{n} Pr(X = i) \cdot Pr_2 \cdot Pr(\phi_2(i) < \gamma_2(i) \mid T = 2) \cdot (i - b)/a. \end{aligned} \tag{10}$$

[3]There is a case for which it is advantageous to buy the bundle even if $\phi_s(i) < \gamma_s(i)$. Let us consider for example market segments 1 and 3 and a bundle that contains both type of events. Whether the following conditions hold: $R_h - R_l > p$, $P_b/\hat{a} < p$, $i > \hat{a}$ and $a \neq \hat{a}$, it results to be convenient to buy the bundle attending the $\hat{a}$ events included in it and replace the events of type $l$ included in the bundle with events of type $h$ outside the bundle. In calculating the probability to buy the bundle and single tickets we should take into account this situation. However, these conditions are verified for a large value of the discount rate $r$. In our numerical example $r$ is not so large such that we can exclude the occurrence of this situation.

Concerning the tickets of events $h \notin B$, in addition to (10), it must be considered the case in which the consumer buys the bundle (i.e when $\phi_s(i) > \gamma_s(i)$). For $s = 1, 3$ the condition is that $i > k$ and the probability to buy a single ticket for a specific event $h \notin B$ is equal to $min\{1, (i - \hat{k})/(a - \hat{a})\}$.
For $s = 2$, a ticket of an event $h \notin B$ is purchased if $i > k+(b-\hat{b})$ and the probability to buy a single ticket for a specific event $h \notin B$ is equal to $(i - k - (b - \hat{b}))/(a - \hat{a})$. Thus, when $a \neq \hat{a}$, the probability that a single ticket for a specific event $h \notin B$ is sold is given by:

$$\begin{aligned} P(h_{\not{B}}) =& P(h_B) + \sum_{s=1,3} Pr_s\{\sum_{i=k}^{n} Pr(X = i) \cdot Pr(\phi_s(i) > \gamma_s(i) \mid T = s) \cdot min\{1, (i - k)/(a - \hat{a})\} \\ &+ \sum_{i=k+(b-\hat{b})}^{n} Pr(X = i) \cdot Pr_2 \cdot Pr(\phi_2(i) > \gamma_2(i) \mid T = 2) \cdot (i - k - (b - \hat{b}))/(a - \hat{a}). \end{aligned} \tag{11}$$

Similarly, $P(l_B)$ and $P(l_{\not{B}})$ can be obtained from (10) and (11) substituting $Pr_1$, $Pr_2$, $Pr_3$, $ha, \hat{a}, b, \hat{b}$ respectively with $Pr_2$, $Pr_1$, $Pr_4$, $b, \hat{b}, a, \hat{a}$.

The output of the DEA model are: expected revenue and attendance. For each offer set, the expected revenue for an arrived customer is given by:

$$R(S) = p_B \cdot P_B + p \cdot (P(h_B) \cdot \hat{a} + P(l_B) \cdot \hat{b} + P(h_{\not{B}}) \cdot (a - \hat{a}) + P(l_{\not{B}}) \cdot (b - \hat{b})), \tag{12}$$

where $p_B$ is given by (2).

In order to calculate the attendance, it should be taken into account not only the capacity of the theatre but also the fact that when $i < k$ a person who buys a bundle do not necessarily attend all the performances included in the bundle: indeed, customers will attend the $i$ events inside the bundle according to his/her preference. Taking for example the attendance of the events of type $h$ that belongs to the bundle: customers who prefer events of type $h$ and at the same time buy a bundle, will not attend $(\hat{a} - i)$ events of type $h \in B$ in case $\hat{a} > i$. Similarly, for those who prefer events of type $l$ and buy a bundle, there are $(k - i)$ events of type $h \in B$ not attended when $\hat{b} < i < k$.
We denote with $Z$ the amount of seats not occupied for events $h$ belonging to the bundle, by customers who have bought the bundle. For the first, third and fifth market segment, we have to consider the case in which $\hat{a} > i$: in this situation,

when $s = 1, 3$, $\phi_s(i) > \gamma_s(i)$ when $p_B/i < p$, whereas $\phi_5(i) > \gamma_5(1)$ when $R_h \cdot i > p_B$. For the second, fourth and sixth market segments, we have to consider the case in which $\hat{b} < i < k$. Hence $Z$ is given by:

$$\begin{aligned} Z =& M \cdot j\{[\sum_{i=1}^{\hat{a}} P(X=i) \cdot (\hat{a} - i)] \cdot [(Pr_1 + Pr_3) \cdot v + Pr_5 \cdot Prob(R_h \cdot i > p_b \mid T = 5)] \\ &+ \sum_{s=2,4,6} Pr_s \sum_{i=\hat{b}}^{k} P(X=i) \cdot (k - i) \cdot Pr(\phi_s(i) > \gamma_s(i) \mid T)\}, \end{aligned} \tag{13}$$

with:

$$j = \begin{cases} 1 & \text{if } \hat{a} \neq 0, \\ 0 & \text{otherwise;} \end{cases}$$

$$v = \begin{cases} 1 & \text{if } (p_B/i) < p, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, we define $W$ as the amount of seats not occupied by customers who have bought the bundle for events $l$ belonging to the bundle:

$$\begin{aligned} W =& M \cdot t \cdot \{[\sum_{i=1}^{\hat{b}} P(X=i) \cdot (\hat{b} - i)] \cdot [(Pr_2 + Pr_4) \cdot v + Pr_6 \cdot Prob(R_l \cdot i > p_b \mid s = 6)] \\ &+ \sum_{s=1,3,5} Pr_S[\sum_{i=\hat{a}}^{k} P(X=i) \cdot (k - i) \cdot Pr(\phi_s(i) > \gamma_s(i) \mid s)\}), \end{aligned} \tag{14}$$

with:

$$t = \begin{cases} 1 & \text{if } \hat{b} \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

and $v$ as defined in (13).

Given this, the total attendance is given by:

$$\begin{aligned} A(S) =& min\{c, [M \cdot P(h_{\not B}) + w \cdot max\{0, (M \cdot (P(h_B) + P_B) - c - Z)\} \cdot \hat{a}/(a - \hat{a})]\} \cdot (a - \hat{a}) + \\ & min\{c, [M \cdot P(l_{\not B}) + u \cdot max\{0, (M \cdot (P(l_B) + P_B) - c - Y)\} \cdot \hat{b}/(b - \hat{b})]\} \cdot (b - \hat{b}) + \\ & min\{c \cdot \hat{a}, M \cdot (P(h_B) + P_B) \cdot \hat{a} - Z\} + min\{c \cdot \hat{b}, M \cdot (P(l_B) + P_B) \cdot \hat{b} - Y\}, \end{aligned} \tag{15}$$

with: with:

$$w = \begin{cases} 1 & \text{if } \hat{a} \neq a, \\ 0 & \text{otherwise;} \end{cases}$$

$$u = \begin{cases} 1 & \text{if } \hat{b} \neq b, \\ 0 & \text{otherwise.} \end{cases}$$

where $M$ is the market share and $c$ the capacity of the theatre.
Notice that for the attendance of events not included in the bundle we have considered the possibility that people who can attend an event included in the bundle through single ticket, will switch to an event non included in the bundle cause of the sold out of events admitted to subscribers. For example, events of type $h$ included in the bundle are sold out whether $M \cdot (P(h_B) + P_B) - Z > c$. So $max\{0, (M \cdot (P(h_B) + P_B) - Z - c)\} \cdot \hat{a}$ represents the "excess" of customers considering all the events $h \in B$. In formulating (15) we assume that this "excess" is equally allocated among the $(a - \hat{a})$ events of type $h$. A similar argument holds considering the events of type $l$.

# 4 Numerical example

In this section, we present a simulated numerical study of the model presented in Section 3.
The following setted is used. We consider a theatre which offers 10 events, 5 lowbrow events and 5 highbrow events. Hence we have: $m = 10$, $a = b = 5$, $n = 33$. The aim is to identify which among the possible 33 combinations of bundle is the most efficient.
First of all, it is necessary to estimate the probability density function (pdf) of the random variables $X, R_l, R_h$. Concerning the first one, at best of our knowledge Venkatesh and Mahajan (1993) is the only study that has modelled the distribution of the number of performances a person is likely to attend. Specifically, in their study this variable follows a Weibull distribution with parameters (2.66, 5.42), where the first term is the shape parameter and the second one is the scale parameter. We adopt this distribution for the $X$ random variable.
Denoting with *f(x)* the pdf of the variable X, the probability of attending exactly $i$ performances is:

$$\int_i^{i+1} f(x)\, dx$$

whose results is shown in Table 1.

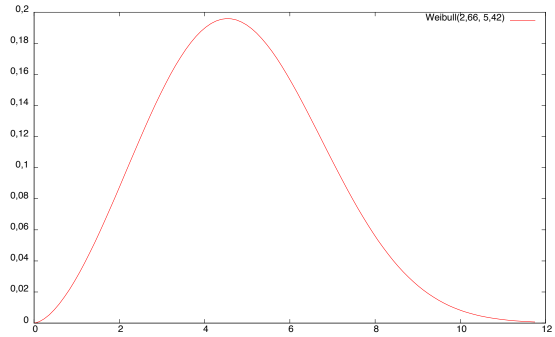


Figure 1: pdf of the likely number of performances a person attend

| **Performances** | **Probability** |
|---|---|
| 0 | 0.011 |
| 1 | 0.057 |
| 2 | 0.119 |
| 3 | 0.172 |
| 4 | 0.194 |
| 5 | 0.176 |
| 6 | 0.131 |
| 7 | 0.079 |
| 8 | 0.039 |
| 9 | 0.015 |
| $\geq 10$ | 0.007 |

Table 2: Probability that a person attends $i$ performance

In the literature, the most employed distribution that characterizes the reservation price are: uniform distribution (Matutesh and Regibeau, 1992; Venkatesh and Kamakura, 2003), normal distribution (Bakos and Brynjolfsson, 1999; Schmalensee, 1984) and, for skewed data, lognormal distribution (Cantono and Silverberg, 2008). Focusing on the reservation price distribution of performing arts ticket, the literature is scarce. Grisolia and Willis (2011) derive the individual willingness to pay of single attributes of an event in a stated preference setting, but their distribution in the whole sample is not defined. Again, Venkatesh and Mahajan (1993) help us. They estimate the distribution of the performancewise reservation price of a serie of music/dance perfomances, that we classified as highbrow (low demand) events. The auhtors adopt a Weibull distribution because is very flexible to accomodate a wide range of distributions. However, as our formulation involves linear combination of probability distributions, the Weibull distributions would make problematic the estimation of the choice probabilities. Using their parameter values of the distribution, we generate around 300 random numbers for the Weibull distribution and test if they can be approximated with a normal or log-normal distribution. Starting for the perfomancewise reservation price distribution (that will be the $R_l$ random variable), we generate a Weibull distribution with parameters (2.38, 15.08). The data fits with a Normal distribution with mean $\lambda_{Rl} = 12.72$ and variance $\sigma^2_{Rl} = 30.38$, as the Jarque-Bera[4] test confirmed[5]. We run the same data transforming the data in their logarithm, in order to test if the data follow a log-normal distribution, but the Jarque-Bera test reject the null Hypothesis.

Concerning the high demand events, we have hypothesized the parameters of the distribution such that its mean is higher than that of the low demand events, and its variance is lower. Indeed a greater variance of the low demand events can found a justification from the fact that consumers that are attracted by highbrow events belong to the upper-class with a greater availability of money, whereas the lower and middle social class is more attracted by lowerbrow events[6]. In such a way we assure the the values of the right tale of the distribution for the low demand events is greater than those of the lowbrow events. Therefore we assume that $R_h \sim N(18, 5.7)$.

Figure 2 plots the pdf of the performancewise reservation price for both high and low demand events. We assume that there is not correlation between the reservation

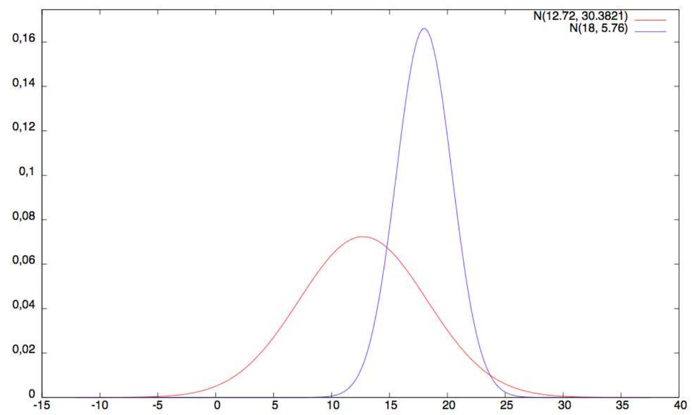


Figure 2: pdf of the performancewise reservation price for high and low demand events

price of high and low demand events: indeed, for the cluster of consumers that shows an omnivorous pattern of consumption (see Sintas and Alvarez, 2005) it can be argued that there is a positive correlation between reservation price of high and low demand events. On the contrary, consumers who present an univore consumption patterns (as the snob or popular consumers) are likely to show a negative correlation between the reservation price of highbrow and lowbrow events. Given the simultaneous existence of this two contradictory patterns of consumption, we assume that, at

[4]Our sample size is big enough to confirm the validity of the test, as Frain(2007) finds that the test lacks in power when the sample is smaller than 50

[5]The null hypothesis is that data are from a normal distribution. The test statistic is 1.5330 and it is smaller than the critical value that is 5.7734.

[6]For a cluster analysis of the performing arts consuption see Sintas and Alvarez (2005)

the market level, there is no correlation between the reservation price of the high and low demand events. The calculation of the probability of buying a bundle and/or single tickets deals with different bivariate normal distributions. As an example, consider (10) with $s = 3$ in which we have : $Pr_3 \cdot Pr(\phi_3(i) < \gamma_3(i) \mid s = 3)$. It can be written as:

$$Pr(R_h > p, R_l < p, x \cdot R_h + y \cdot R_l < z), \tag{16}$$

where:
$x = [min(\hat{a}, i) + min[(max(i-k), 0), (a-\hat{a})] - min(a, i)]$; $y = [max(min(i, k) - \hat{a}, 0)]$;
$z = p_b + p \cdot min[(max(i-k), 0), (a-\hat{a})] - p \cdot min(a, i)$.
We denote with Y the third variable such that

$$V \sim N(18 \cdot x + 12.72 \cdot y, 5.7 \cdot x^2 + 30.38 \cdot y^2).$$

So $F = (R_h, R_l, Y)$ is the multivariate normal distribution we are interested in. Since its support lies in a 2-dimensional space, and $R_h < \frac{z - y \cdot R_l}{x}$, (16) can be obtained integrating over probability density function of the bivariate normal [7]

$$\int_{-\infty}^{p} \int_{p}^{\frac{z-y\cdot r_l}{x}} \mathbb{1}_{\{\frac{z-y\cdot r_l}{x} \geq p\}} f_{R_h}(r_h) f_{R_l}(r_l)\, dr_h dr_l, \tag{17}$$

where $\mathbb{1}_{\{\}}$ is the indicator function, $f_{R_h}$ and $f_{R_l}$ are respectively the density of $R_h$ and $R_l$.
Once calculated the probability of purchasing a single ticket and the bundle [8], the input and the output values for the DEA model can be obtained.
It is well known that the capability of DEA to discriminate between efficient and inefficient DMUs depends on the relationship between the number of input and output and the number of DMUs. In literature there are different rule of thumbs: Golany and Roll [1989] suggest that the number of DMUs should be greater than $2 \cdot (e + f)$ where $e$ and $f$ are respectively the number of input and output used. According to Dyson *et al.* [2001],instead, the number of DMUs should be greater than $2 \cdot e \cdot f$. In order to further discriminate the DMUs, we have exploit the units invariant property of DEA and the fact that our inputs are expressed with the same units of measure to reduce the number of input from 10 to 2. As inputs we use the average value of $Q(S)$ for high demand and low demand events. This choice implies, for events of the same type, the same value judgments about the occupancy of the

[7] This is for $x > 0$, otherwise the integration region changes
[8] The probability values are obtained through the package "mvtnorm" in the R software

theatre, that seems a reasonable assumption
We calculate the efficiency score considering 3 different values of the discount rate $r$ (0.03; 0.04; 0.05) and 3 different values of the capacity of the theatre $c$ (400; 600; 800), obtainin in total 9 situations.
Table 3 shows the efficiency score of each $(\hat{a}, \hat{b})$ possible bundle [9] . The efficiency score reported is the reciprocal of the efficiency score obtained in the output-oriented approach, so that the greater is the score, the more efficient is the DMU.
The results shown in Table 3 shows that there is quite small variation in the DEA coefficients. In relation to this, it may be concluded that a chenge in the bundle policy has a small impact in the expected revenue and attendance of the theatre, but actually the interpretation of the table reveals how the choice of the bundle can have a significant impact on the theatre's objective. For instance, let us consider the situation in which $r = 0.04$ and $c = 600$. In this situation, the bundle (0,2) produces, *ceteris paribus* $[(1 - (1/1.039)) \cdot 100] = 3.75\%$ more output (expected revenue and attendance) than the reference set in the efficient frontier. Assuming $p = 15$€ and $M = 1000$, it results that, given the same capacity consumption, the bundle (0,2) increases the attendance of 139 people and the total revenue of 289€ respect to the efficient frontier. Conversely, the bundle (5,0) produces $[(1/0.945) - 1) \cdot 100] = 5.82\%$ less output than the reference set in the efficient frontier, meaning a loss in attendance of 210 people and a loss in total revenue of 158€.

As Table 3 shows, the super-efficiency scores for a DMU takes different value according to which value of $c$ and $r$ are considered. In order to derive the possible relationship between score, capacity, discount rate and composition of the bundle, a simple OLS regression is estimated. To this purpose, the use of the super-efficiency model presents advantages over the traditional DEA scores. The latters, in fact, are censored at 1, forcing the adoption of a Tobit regression which requires the restricted assumption of normality of the underlying (uncensored) scores (Nahra *et al.*, 2009). On the contrary, using as dependent variable the supper-efficiency scores, we can use the OLS regression that relaxes the assumption on the distribution of the scores. To characterize the composition of the bundle, we use two variables: the first, called *size*, is an index that denotes the magnitude of the bundle and is defined as the ratio between the events included in the bundle and the total number of events. The second, called *popularity* (*pop*), denotes the portion in the bundle of high demand events respect to low demand events. It is calculated in such a way [10]:

[9] The efficiency scores are calculated using the "Benchmarking" package in R

[10] This index is inspired by the accounting literature on sustainability report: an index used to assess how sustainable is a firm is given by $\frac{realvalue-minimum}{maximum-minimum}$ where *real value* is the difference between positive information and negative information argued by the financial report; *minimum* is

| **DMU** | $r = 0.03$ $c = 400$ | $r = 0.03$ $c = 600$ | $r = 0.03$ $c = 800$ | $r = 0.04$ $c = 400$ | $r = 0.04$ $c = 600$ | $r = 0.04$ $c = 800$ | $r = 0.05$ $c = 400$ | $r = 0.05$ $c = 600$ | $r = 0.05$ $c = 800$ |
|---|---|---|---|---|---|---|---|---|---|
| (1,1) | 0.997 | 0.997 | 0.997 | **1.005** | **1.005** | **1.005** | 1.006 | **1.006** | **1.006** |
| (2,0) | **1.037** | **1.004** | **1.004** | 0.990 | 0.995 | 1.000 | 0.988 | 0.990 | 1.000 |
| (3,0) | 0.980 | 0.980 | 0.981 | 0.975 | 0.975 | 1.000 | 0.968 | 0.968 | 1.000 |
| (4,0) | 0.968 | 0.968 | 0.983 | 0.957 | 0.957 | 1.000 | 0.945 | 0.945 | 1.000 |
| (5,0) | 0.960 | 0.960 | 0.984 | 0.945 | 0.945 | 1.000 | 0.929 | 0.932 | 1.000 |
| (2,1) | 0.999 | 0.999 | 0.999 | 1.002 | 1.003 | 1.003 | 1.002 | 1.004 | 1.004 |
| (3,1) | 0.998 | 0.998 | 0.998 | 0.996 | 0.994 | 1.000 | 0.991 | 0.992 | 1.000 |
| (4,1) | 0.993 | 0.992 | 0.998 | 0.985 | 0.983 | 1.003 | 0.973 | 0.978 | 1.004 |
| (5,1) | 0.993 | 0.993 | **1.007** | 0.980 | 0.980 | 0.995 | 0.960 | 0.964 | 0.992 |
| (0,2) | 0.991 | 0.990 | 0.990 | **1.060** | **1.039** | 1.000 | **1.019** | **1.039** | 1.000 |
| (1,2) | 0.995 | 0.995 | 0.995 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| (2,2) | 0.998 | 0.997 | 0.997 | 1.000 | 1.001 | 1.000 | 1.000 | 0.998 | 0.997 |
| (3,2) | 1.000 | 0.997 | 0.997 | 0.999 | 0.997 | 1.001 | 0.997 | 1.000 | 1.002 |
| (4,2) | **1.003** | 1.000 | **1.002** | 0.999 | 0.995 | **1.004** | 0.992 | 0.990 | 1.001 |
| (5,2) | 1.001 | **1.004** | 1.001 | 1.000 | 1.002 | **1.007** | 0.987 | 0.995 | **1.006** |
| (0,3) | 0.992 | 0.986 | 0.986 | 1.002 | 1.000 | 1.000 | 1.001 | 1.000 | 1.000 |
| (1,3) | 0.995 | 0.990 | 0.990 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| (2,3) | 0.998 | 0.994 | 0.992 | 1.000 | 1.002 | 1.000 | 1.000 | **1.007** | 1.000 |
| (3,3) | 1.001 | 0.996 | 0.996 | 1.002 | 0.997 | 0.998 | 1.000 | 0.996 | 0.998 |
| (4,3) | 0,996 | 0.995 | 0.995 | 1.001 | 0.995 | 0.998 | 0.998 | 0.996 | 1.000 |
| (5,3) | 1.001 | **1.001** | 1.001 | 0.998 | **1.006** | 0.999 | 0.999 | 1.003 | **1.006** |
| (0,4) | 0.997 | 0.982 | 0.982 | 1.003 | 1.000 | 1.000 | 1.004 | 1.000 | 1.000 |
| (1,4) | 0.998 | 0.986 | 0.985 | 1.001 | 1.000 | 1.000 | 1.001 | 1.000 | 1.000 |
| (2,4) | 1.001 | 0.993 | 0.990 | 1.001 | 0.996 | 0.996 | 1.001 | 0.993 | 0.991 |
| (3,4) | 0.995 | 0.990 | 0.991 | 0.997 | 0.991 | 0.993 | 1.000 | 0.993 | 0.995 |
| (4,4) | 0.993 | 0.992 | 0.992 | 0.993 | 0.994 | 0.991 | 0.998 | 0.993 | 0.994 |
| (5,4) | 0.997 | 0.997 | 0.997 | 1.003 | 1.003 | 1.003 | 0.997 | 1.001 | 0.997 |
| (0,5) | **1.009** | 0.979 | 0.979 | **1.012** | 1.000 | 1.000 | **1.014** | 1.000 | 1.000 |
| (1,5) | 1.002 | 0.986 | 0.985 | 1.004 | 0.991 | 0.991 | **1.012** | 0.989 | 0.989 |
| (2,5) | 0.995 | 0.987 | 0.987 | 0.998 | 0.988 | 0.990 | 1.000 | 0.990 | 0.989 |
| (3,5) | 0.991 | 0.988 | 0.988 | 0.994 | 0.991 | 1.000 | 0.998 | 0.990 | 0.989 |
| (4,5) | 0.986 | 0.986 | 0.987 | 0.993 | 0.990 | 1.000 | 0.993 | 0.991 | 0.989 |
| (5,5) | 0.998 | 0.998 | 0.998 | 0.997 | 0.997 | 0.997 | 1.004 | 1.004 | 1.004 |
| Efficient DMUs | 9 | 4 | 5 | 16 | 14 | 23 | 17 | 14 | 22 |

Table 3: Efficiency score; in bold the three most efficient bundle for each situation.

$$Popularity = \frac{(\hat{a} - \hat{b}) + 5}{10}$$

We considers together all the 9 situations in order to verify the relationship between the efficiency scores, characteristics of the bundle, capacity and discount rate.
The RESET test suggests us to transform all the independent variable in logarithms [11], hence we estimated the following equation:

$$\begin{aligned} score =& \alpha + \beta_1 ln(popularity) + \beta_2 ln(size) + \beta_3 ln(c) \cdot ln(popularity) + \beta_4 ln(r) \cdot ln(popularity) \\ &+ \beta_5 ln(c) \cdot ln(size) + \beta_6 ln(r) \cdot ln(size) + \beta_7 ln(popularity) \cdot ln(size) \end{aligned} \tag{18}$$

The bivariate correlation among the variables $ln(c)$, $ln(r)$, $ln(popularity)$ and $ln(score)$ are examined in order to detect co-linearity; however, the results don't show significant correlation among variables. As there is evidence for heteroskedasticity ($p - value < 0.10$ for the White's test), we use the robust standard error estimates. We drop the main effect of $c$ and $r$ in the model because we know that the capacity and the discount rate don't affect the efficiency score in themselves but in their interaction with the characterizes of the bundle.
Table 4 shows the result of the OLS regression: Being (18) a level-log equation, the marginal effect are interpreted in terms of semi-elasticity.
For example, rearranging the derivative of the score with respect to the capacity, we have:

$$\frac{\partial score}{\partial c} \cdot \frac{c}{1} = \beta_2 \cdot ln(pop) + \beta_4 ln(size), \tag{19}$$

where the right hand side indicates the change in units of the efficiency score when the capacity increases by 1%.
The coefficient estimations suggest that an increase in the capacity should be combined with an increase in the popularity of the bundle ($\beta_3 > 0$). This is not surprising: when the capacity is small, it is more likely that the theatre is sold out when high demand events are proposed, even if they are not discounted through

the total number of indicator with a the negative sign; *maximum* is the total number of indicator with positive sign. As this index is used to indicate quality of sustainability, our index is used to indicate popularity of the bundle. In our case, 5 is used as it is the maximum number of possible low demand events that a bundle can include, and 10 is the maximum number of events a bundle can include. We prefer to use this index, instead of a simple difference between high and low demand events, for a couple of reason: first, this index is always positive between 0 and 1; second, being positive it can be easily integrated with the *size* to compose an unique index that consider both the quantity of events in the bundle and its popularity

[11] As a consequence, the bundle (0,5) is dropped from the regression because its popularity score is equal to 0 and its logarithms tend to minus infinity

| Independent variables | coefficient | Robust st.error |
|---|---|---|
| $ln(popularity)$ | $-0.1241^{****}$ | 0.0244 |
| $ln(size)$ | $0.1116^{***}$ | 0.0416 |
| $ln(c) \cdot ln(popularity)$ | $0.0131^{****}$ | 0.0030 |
| $ln(r) \cdot ln(popularity)$ | $-0.0143^{***}$ | 0.0046 |
| $ln(c) \cdot ln(size)$ | $-0.0127^{**}$ | 0.0057 |
| $ln(r) \cdot ln(size)$ | 0.0070 | 0.0077 |
| $ln(popularity) \cdot ln(size)$ | $0.0144^{***}$ | 0.0046 |
| $constant$ | $0.9966^{****}$ | 0.0027 |
| $R^2$ | 0.1707 | |
| $F$-model | $6.78^{***}$ | |
| No. of observations | 288 | |

$^{*}p < .10$ $^{**}p < .05$ $^{***}p = .01$ $^{****}p < .001$

Table 4: Results of OLS

bundles. On the contrary when the capacity is large, it is more difficult to sold out high demand events and, in this sense, including high demand events can help for this purpose.

On the other hand, when the discount rate increase, it is suggested to decrease the popularity of the bundle ($\beta_4 < 0$). This result can be explained considering that a high discount rate should incentivate the attendance of low demand events: in fact a high discount rate associated to high demand events may result in a decrease of potential revenue.

Finally it should be noted that the coefficient of the interaction term between popularity and size is positive ($\beta_7 > 0$), denoting that the increase in popularity of the bundle yields an increase in the score only for bundles that include many events. As the coefficient associated with the popularity of bundle is negative ($\beta_1 < 0$), we can deduce that an increase of popular events in the bundle should be combined with the inclusion of low demand events in the bundle.

# 5 Conclusions

In this paper we analyze a structural-based RM problem in the performing arts organization, related to the composition of a bundle of tickets to be offered. The decision problem consists in the determination of the most efficient bundle in term of quantity and ratio between high demand and low demand events. For this purpose, we adopt a network RM approach [Liu and Van Ryzin, 2008] in defining the efficient bundle and subsequently a proposed a super-efficient DEA model in order to detect the most efficient bundle among all the possible combination of events. In order to derive the choice probability for each possible bundle, and so the input and output of the DEA model, we model the purchase decision on the basis of two random variables: availability of time and reservation price per each type of performance. In the numerical example presented we consider nine situations that differ each other by the value of capacity of the theatre and discount rate of the bundle. Once obtained the efficiency score, we regress them on variables that characterize the bundles and subsequently on managerial variables (capacity and discount rate) in order to obtain some insights into what determines the efficiency level of a bundle. The results obtained can provide some recommendations for decision-makers. In particular, given a subscription, the theater manager can consider the possibility to change the composition of the bundle to be offered to the customers in the following season, on the basis of these advices: if the discount rate of the bundle increases, the theater manager should promote a greater inclusion of low demand events in the bundle, and viceversa. If the theater owns more than one stage, the theater manager should propose a bundle with a large share of popular events for the theater with the largest capacity, whereas a subscription that includes a large share of low demand events should be associated with the smallest theater.

This paper presents as main limitation the fact that we consider only two type of events with the assumption that events of the same type have the same distribution of the reservation price. Future research should overcome this limitation, considering events that present a unique pattern of reservation price.

# References


- Andersen, P., Petersen, N.C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science* 39, 1261-1264.

- Ansari, A., Siddarth, S., Weinberg, C.B. (1996). Pricing a bundle of products or services: The case of nonprofits. *Journal of Marketing Research*, 86-93.

- Bakos, Y., Brynjolfsson, E. (2000). Bundling and Competition on the Internet. *Marketing science*, 19(1), 63-82.

- Banker, R.D., Charnes, A., Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30, 1078-1092.

- Belobaba, P. P. (1987). Survey paper-airline yield management an overview of seat inventory control. *Transportation Science*, 21(2), 63-73.

- Bradlow, E. T., Rao, V. R. (2000). A hierarchical Bayes model for assortment choice. *Journal of Marketing Research*, 37(2), 259-268.

- Cantono, S., Silverberg, G. (2009). A percolation model of eco-innovation diffusion: the relationship between diffusion, learning economies and subsidies. *Technological forecasting and social change*, 76(4), 487-496.

- Caves, R. (2000). *Economics of the creative industries*. Cambridge, Mass: Harvard.

- Charnes, A., Cooper, W.W., Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research* 2, 429-444.

- Chen, Y. (2005). Measuring super-efficiency in DEA in the presence of infeasibility. *European Journal of Operational Research* 161, 545-551.

- Chen, Y., Liang, L. (2011). Super-efficiency DEA in the presence of infeasibility: One model approach. *European Journal of Operational Research*, 213(1), 359-360.

- Choi, C., Jeong, M., Mattila, A.S. (2014). Revenue management in the context of movie theaters: Is it fair? *Journal of Revenue & Pricing management.*

- Chung, J., Rao, V. R. (2003). A general choice model for bundles with multiple-category products: Application to market segmentation and optimal pricing for bundles. *Journal of Marketing Research*, 40(2), 115-130.

- Cook, W.D., Liang, L., Zha, Y., Zhu, J. (2008). A modified super-efficiency DEA model for infeasibility. *Journal of the Operational Research Society*, 69, 276-281.

- Courty, P. (2000). An economic guide to ticket pricing in the entertainment industry. *Recherches Čconomiques de Louvain/Louvain Economic Review*, 167-192.

- Drake, M.J., Duran, S., Griffin, P.M., Swann, J.L. (2008). Optimal timing of switches between product sales for sports and entertainment tickets. *Naval Research Logistics*, 55(1), 59-75.

- Duran, S., Swann, J.L., Yakõcõ, E. (2012). Dynamic switching times from season to single tickets in sports and entertainment. *Optimization Letters*, 6(6), 1185-1206.

- Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S., Shale, E.A. (2001). Pitfalls and protocols in DEA. *European Journal of operational research*, 132(2), 245-259.

- Ferreira, K. D., Wu, D. D. (2011). An integrated product planning model for pricing and bundle selection using Markov decision processes and data envelope analysis. *International Journal of Production Economics*, 134(1), 95-107.

- Frain, J.C. (2007). Small sample power of tests of normality when the alternative is an $\alpha$-stable distribution. *Trinity Economics Papers*, (207)

- Golany, B., Roll, Y. (1989). An application procedure for DEA. *Omega*, 17(3), 237-250.

- Goldberg, S. M., Green, P. E., Wind, Y. (1984). Conjoint analysis of price premiums for hotel amenities. *Journal of Business*, 111-132.

- Grisolia, J. M., Willis, K. G. (2011b). Heterogeneity In Willingness To Pay For Theatre Productions: Individual Specific Willingness To Pay Estimates For Theatres, Shows And Their Attributes. *Scottish Journal of Political Economy*, 58(3), 378-395.

- Guiltinian, J.P. (1987). The price bundling of services: A normative framework. *The Journal of Marketing*, 74-85.

- Hansmann, H. (1981). Nonprofit enterprise in the performing arts. *The Bell Journal of Economics*, 341-361.

- HANSON, W., MARTIN, R. K. (1990). Optimal bundle pricing. *Management Science*, 36(2), 155-174.

- HUNTINGTON, P. A. (1993). Ticket pricing policy and box office revenue. *Journal of Cultural Economics*, 17(1), 71-87.

- JOHNSON, S.M., GARBARINO, E. (2001). Customers of perform- ing arts organizations: Are subscribers different from non- subscribers? *International Journal of Nonprofit and Voluntary Sector Management*, 6, 61-77.

- LEE, H. S., CHU, C. W., ZHU, J. (2011). Super-efficiency DEA in the presence of infeasibility. *European Journal of Operational Research*, 212(1), 141-147.

- LESLIE, P. (2004). Price discrimination in Broadway theater. *RAND Journal of Economics*, 520-541.

- LITTLEWOOD, K. (1972). Forecasting and control of passenger bookings. *AGIFORS Symposium Proceedings,* 95Ð117.

- LIU, Q., VAN RYZIN, G. J (2008). On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management*, 10(2), 288-310.

- LOVELL C.A.K, ROUSE A.P.B. (2003). Equivalent standard DEA models to provide super-efficiency scores. *Journal of the Operational Research Society* , 54, 101-108.

- MATUTES, C., REGIBEAU, P. (1992). Compatibility and bundling of complementary goods in a duopoly. *The Journal of Industrial Economics*, 37-54.

- NAHRA, T. A., MENDEZ, D., *Alexander*, J. A. (2009). Employing super-efficiency analysis as an alternative to DEA: An application in outpatient substance abuse treatment. *European Journal of Operational Research*, 196(3), 1097-1106.

- ORBACH, B.Y., EINAV, L. (2007). Uniform prices for differentiated goods: The case of the movie-theater industry. *International Review of Law and Economics*, 27(2), 129-153.

- ROSEN, S., ROSENFIELD, A. M. (1997). Ticket Pricing 1. *The Journal of Law and Economics*, 40(2), 351-376.

- RAFIEI, H., RABBANI, M., RAZMI, J., JOLAI, F. (2013). Product Bundle Pricing in the New Millennium: A Literature Review. *International Journal of Advances in Management Science*

- Schmalensee, R. (1984). Gaussian demand and commodity bundling. *Journal of Business*, S211-S230.

- Seaman, B.A. (2006). Empirical studies of demand for the performing arts. *Handbook of the economics of art and culture*, 1, 415-472.

- Shen, Z. J. M., Su, X. (2007). Customer behavior modeling in revenue management and auctions: A review and new research opportunities. *Production and operations management*, 16(6), 713-728.

- Sintas, J. L.,çlvarez, E. G. (2005). Four characters on the stage playing three games: performing arts consumption in Spain. *Journal of Business Research*, 58(10), 1446-1455.

- Stigler, G. J. (1963). United States v. Loew's Inc.: A note on block-booking. *Sup. Ct. Rev.*, 152.

- Stremersch, S., Tellis, G. J. (2002). Strategic bundling of products and prices: A new synthesis for marketing. *Journal of Marketing*, 66(1), 55-72.

- Talluri, K. T., & Van Ryzin, G. (1998). An analysis of bid-price controls for network revenue management. *Management Science*, 44(11-part-1), 1577-1593.

- Talluri, K. T., Van Ryzin, G. J (2004). Revenue management under a general discrete choice model of consumer behaviour. *Management Science*, 50(1), 15-33.

- Talluri, K. T., Van Ryzin, G. J (2006) *The theory and practice of revenue management*, (Vol. 68). Springer.

- Tereyagoglu, N., Fader, P., Veeraraghavan, S. (2012). Filling seats at a theater: Estimating the impact of posted prices and dynamic discounts. *The Wharton School, University of Pennsylvania*, Philadelphia, PA, 19104.

- Venkatesh, R., Kamakura, W. (2003). Optimal bundling and pricing under a monopoly: Contrasting complements and substitutes from independently valued products. *The Journal of business*, 76(2), 211-231.

- Venkatesh, R., Mahajan, V. (1993). A probabilistic approach to pricing a bundle of products or services. *Journal of Marketing Research*, 3,494-508.

- Venkatesh, R., Mahajan, V. (1997). Products with branded components: An approach for premium pricing and partner selection. *Marketing Science*, 16(2), 146-165.

- Venkatesh, R., Mahajan, V. (2009). The design and pricing of bundles: a review of normative guidelines and practical approaches *Handbook of pricing research in marketing*, 232.

- Vogel, H.L. (1990). *Entertainment industry economics: A guide for financial analysis.* Cambridge University Press.

- Yakici, E., Ězener, O. Ě., Duran, S. (2014). Selection of event tickets for bundling in sports and entertainment industry. *Computers & Industrial Engineering*, 74, 257-269.

# Multiobjective optimization model for pricing and seat allocation problem in non profit performing arts organization


Andrea Baldin*[1], Trine Bille†[2], Andrea Ellero‡[3], and Daniela Favaretto§[4]

[1,3,4] *Department of Management, Ca' Foscari University of Venice, Cannaregio 873, 30121 Venice, Italy*
[2] *Department of Management, Politics and Philosophy, Copenhagen Business School, Porcelænshaven 18B, 3.floor, 2000 Frederiksberg, Denmark*



**Abstract**

The implementation of Revenue Management (RM) techniques in non profit performing arts organizations presents new challenges compared to other sectors, such as transportion or hospitality industries, in which these techniques are more consolidated. Indeed, performing arts organizations are characterized by a multi-objective function that is not solely limited to revenue. On the one hand, theatres aim to increase revenue from box office as a consequence of the systematic reduction of public funds; on the other hand they pursue the objective to increase its attendance. A common practice by theatres is to incentive the customers to discriminate among themselves according to their reservation price, offering a schedule of different prices corresponding to different seats in the venue. In this context, price and allocation of the theatre seating area are decision variables that allow theatre managers to manage these two conflicting goals pursued. In this paper we introduce a multi-objective optimization model that jointly considers pricing and seat allocation. The framework proposed integrates a choice model estimated by



*E-mail: andrea.baldin@unive.it
†E-mail: tbi.mpp@cbs.dk
‡E-mail: ellero@unive.it
§E-mail: favaret@unive.it


Authors' contributions: Andrea Baldin contributes to the paper by conceiving and defining the design of the study, doing part of the literature review, estimating the model, interpreting the results, and most of the writing. Trine Bille contributes to the paper by doing part of the literature review and writing the section related to the Royal Danish Theatre. Andrea Ellero and Daniela Favaretto contribute to the paper by presenting the algorithm and revising the paper critically.

multinomial logit model and the demand forecast, taking into account the impact of heterogeneity among customer categories in both choice and demand. The proposed model is validated with booking data referring to the Royal Danish Theatre during the period 2010-2015.

**Keywords** Multi-objective optimization; Pricing; Seat allocation; Multinomial logit model; Theatre demand

# 1 Introduction

In the seminal article by Baumol and Bowen (1966) the authors claim how theatres will be more and more dependent on subsidies, due to their productivity lag. However, the last decades' tendency shows that public funds allocated to non profit performing arts organizations in Western countries (Marco-Serrano, 2006) are decreasing[1]. This fact has forced theatres to increase other sources of revenue, including box office revenue. In addition, such organizations pursue the aim to increase the attendance, for a couple of reason: first, they feel the mission to spread culture to as broad a segment of the populace as possible (Hansmann, 1981) legitimizing their social value; second, they prefer to avoid empty seats in the venue that can have negative effect on the reputation of the theatre.
In this context, managers of the performing arts organizations can implement Revenue Management (RM) techniques in order to balance between the rate of occupancy and the profitability of theatre. The most common among these techniques is realized through market segmentation based on the price leverage, that leads to different pricing scheme. For istance, price reductions are offered to customers segments, such as students and senior citizens, who are supposed to be less able to pay. Discounts are offered also to those customers - subscribers - who buy in advance a bundle of tickets, assuring a long-term commitment toward the theatre. Due to heterogeneity in price sensibility within the same customer segment, one usual practice by theatre is to use a non-linear tariff system offering a schedule of different prices according to the quality of the product. In this case, different prices are charged according to the seat location in the venue in order to better capture consumers' willingness to pay. Indeed, this mechanism incentives customers to discriminate themselves by choosing the seating area they prefer. So, beside the pricing strategy,

[1]This framework holds also for our case study: the Royal Danish Theatre. According to the National Danish Statistics (http://www.statbank.dk), the public subsidy to the Royal Danish Theatre decreases from 608 675 Danish crowns in the 2011/2012 season, to 573 900 Danish crowns in the 2014/2015 season.

also the seat allocation across these fare classes (i.e seating area) represents a decision that may foster an orientation by the theater towards either the maximization of the total attendance or the maximization of revenue. In the first case we expect that theatre would increase the accessibility of the most expensive seating areas for all the customers: to do this, it is convenient to propose a scheme in which the prices of the different seating are closer downward. This mechanism will lead to an increase of the size of the expensive seating area and, in addition, can favorite a customer buy-up behaviour (i.e buying a ticket for a more expensive fare class when the ticket for the required seating area is not available). In the second case, we expect that theatre would strengthen the self-discrimination operated by customers. Thus, the allocation policy will strongly depend on the type of customer that attends the performance: if the performance attracts an audience group (as young customers) that is supposed to be highly price sensitive, the theatre would enlarge the cheapest seating area in order to prevent a loss in revenue. In the opposite case, the theater would take advantage of the inelastic demand by enlarging the expensive seating area.

With respect to this pricing and allocation strategy, it becomes essential not only the demand forecasting, but also the understanding of the customers' behavior with respect to price discrimination by seating area. Since Talluri and Van Ryzin's (2004) paper, discrete choice models have emerged as a standard approach in the RM literature to incorporate the buy-up and buy-down behavior.

This paper proposes an optimization model that jointly considers the pricing and allocation problem in the performing arts context. To this end, the demand forecasting is integrated with a customer choice model. In order to accomodate for heterogeneity in preference over seating areas, we adopt a multinomial logit model (MNL) using customer's characteristics and performance-production attributes as variables to be interacted with the characteristics of the choice alternatives.

We aim to contribute to the RM and theatre demand literature by demonstrating that discrete choice and optimization models can effectively be employed to assist theatre managers in both setting price and seating allocation. Indeed, theatres, as providers of performing arts services, meet the requirements for the implementation of RM techniques, such as the capacity constraint and the perishability issue (i.e the ticket value deteriorates as time goes on and it is null when the performance starts). However, as already said, subsidized theatres, compared to the dominant sector in RM literature like the transportation and hospitality industries, are organizations which pursue other objectives in addition to the revenue maximization. Therefore, we employ a bi-objective optimization model to study the optimal pricing and seat

allocation with conflicting objectives related to revenue and total attendance.
Our model has been implemented to a data set provided by the Royal Danish Theatre which refers to the period 2010-2015. A simulation is conducted considering three performances that differ from each other by characteristics that affect the demand.
To the best of our knowledge, there are no studies in the RM literature that has considered jointly the pricing and allocation problem in a context in which the organization present an objective function that is not limited solely limited to revenue. In this sense, we believe that this is the added value of this article, providing guidances to theatre managers in adapting the price and size of seating area according to the weight given to each objective to pursue.
The remainder of the paper is organized as follows: Section 2 presents the relevant literature on RM in the performing arts context; Section 3 describes the research framework, whereas Sections 4 and 5 present respectively the demand estimation and the choice model. Section 6 describes the optimization model, whereas Section 7 presents the results of our simulation. Finally Section 8 provides some conclusions.

# 2 Literature review

Opposed to a huge amount of RM research devoted to transportation industry, the issues related to pricing and allocation problems in performing arts organizations have not received a great deal of attention. This disparity is due to different factors: first, RM is not a pervasive practice in most of these organizations. Second, this kind of organizations is characterized by peculiar features, in primis the multidimensional nature of their objectives.
The literature of cultural economics has been dealing with the objectives of performing arts institutions. Since most performing arts institutions are nonprofit firms this taps into a more general literature on the objectives of nonprofit firms (e.g. Hansmann (1980) and Steinberg (1986)). Steinberg (1986) suggests that nonprofit firms are either service maximizers or budget maximizers or something in between. However, in the performing arts, the concept of service is not straight forward. Several authors (e.g. Throsby and Wither, 1979, Throsby, 1994, and Hansmann, 1981) have suggested three different measures of output: 1) Quality, 2) Audience size and 3) Budget. Several empirical studies have shown that the performing arts are primarily output maximizers (either quality or quantity), and less budget maximizers (see e.g. Luksetich and Lange, 1995; and Gapinski, 1985). For an overview of the literature see Brooks (2006). To our knowledge no studies have been made, dealing with the

optimization decisions in the performing arts, when the repertoire is planned (based on quality decisions), and the theatre wants to make the optimal decision on how to maximize attendance as well as revenue, based on decision on prices and seat categories.
Most of the research related to the demand-management decision in the theatre sector has focused on the price discrimination practice. Hansmann (1981) claims that in the nonprofit performing arts sector, price discrimination is not effective due to the difficulty to identify customers with inelastic demand. Therefore, according to the author, the only form of discrimination that nonprofit enterprise can apply is by asking for a voluntary donation, in order to extract a part of consumer's surplus. Seaman (1985) raises some doubts about Hansmann's (1981) hypotheses: the author measures the degree of price discrimination (such as: the number of different prices charged and the standard deviation of the prices charged) to a set of non profit performing art organization. He concludes that price discrimination varies significantly across art forms (opera, ballet, theater, symphony concert) and that the organizations that discriminate more are characterized by a high ratio between fixed cost and attendance. Huntington (1993) justifies the adoption of price discrimination by seating area, by referring to the Rosen's (1974) utility model (i.e the hedonic price model), as there are observable differences between different seats. Moreover, the author compares the box office revenue between theaters operating a single price policy and those operating a discrimination pricing policy: he finds that the price range policy is statistically significant and positively correlated with the revenue of the theatre, controlling for seat capacity and the number of performances per year. Rosen and Rosenfield (1997) describe a model in which theater venue has two types of seats: (high and low quality), and the theatre manager knows the distribution of reservation price for both seat category. First, the authors solve the pricing problem, given the quantity of seats for each category. Second, the authors solves the allocation problem, given the optimal pricing policy. Leslie (2004) considers the Broadway show "Seven Guitars" and estimates a structural econometric model of price discrimination based on a individual consumer behavior model, that incorporate all the types of price discrimination (by seating area and social category). The model allows him to perform different experiment using alternative pricing policies. Tereyagoglu *et al.* (2012) use the data from the ticket purchase transaction of the shows of a symphony orchestra in the northeast region of the US, in order to employ a proportional hazard framework to analyze how pricing and discount actions over time affect the timing of customers purchase.

# 3 Research framework

## 3.1 The Royal Danish Theater

The Royal Danish Theatre was founded in 1748 and is the Danish national theatre. It has three main Stages in Copenhagen. The Old Stage from 1874, a new Royal Opera House from 2005 and a new Royal Playhouse from 2008. The Opera House and the Playhouse has a main stage and smaller stages for experimental productions. It is one of the few theaters in the world offering both opera, ballet and theatre performances as well as classical concerts. Today The Old Stage is the house, where ballet is performed.

The law of the Royal Danish Theatre states that it is the national theatre for the whole country and the entire population. Besides, it has an obligation to produce a broad repertoire of high artistic quality within ballet, opera and plays. It is obligated to continue the classical traditions as well as developing the performing arts in new and contemporary ways. A special concern is on productions of Danish origin. The Royal Danish Theatre is on the state budget under the Ministry of Culture, and has a number of more specific obligations in agreement with the current Minister of Culture. Included in these obligations are general cultural policy goals, like having special productions for children and youth, and to keep prices to a level that make the theatre accessible for all socio-economic groups.

In 2015 the theatre had a total budget of 705,4 million DKK (94 million Euros), of which 76 percent were public support from the Government. The theatre had 165,8 million DKK (22 million Euros) in own earnings, of which 69 percent (15 million Euros) was from ticket sales, the rest was income from sponsors etc.

Due to its obligations as a national theatre, it has to decide its repertoire based on a number of parameters, namely quality and variety, understood as a fairly large number of different productions from the classical repertoire as well as new productions, developing the performing arts, and Danish as well as international production from the world repertoire. Besides, it has to decide the number of performances of each production during the season and how they are scheduled on weekday and weekends. It will create a loss in earnings if a given production is played less than demanded by the audience as well as if a performance is played more times than demanded by the audiences (empty seats). There are high fixed costs take a new producing on stage (due to rehearsal time, designing the staging etc.), but the costs to prolong a production with extra performances are small, and the marginal costs are lower than the marginal revenue (Bille Hansen, 1991).

Finally, it has to decide its price policy, including price differentiation based on different audience groups (like young, subscribers and senior) as well as seat categories, time of the performance, the type of the performance, the production costs etc.

## 3.2 Problem description

In this paper we assume that the repertoire decisions are already determined by the theater, both with regard to the variety of productions and the number of performances of each production during a season. With this restriction the theatre has to decide on the price and the allocation of seat categories for the individual performances. It is assumed, that the theatre wants to optimize both attendance and revenue, where the former finds an upper limit in the theatre capacity. Our bi-objective optimization model incorporates the demand forecast and the customers' seat choice model. The latter is estimated with a multinomial logit (MNL) model that predicts the probability to choose a particular seating area as a function of price and performance characteristics. From Baldin and Bille (2016) we know that some audience groups (especially young people) are quite price sensitive, while other groups are very insensitive to price (e.g., subscribers). Therefore, we estimate one demand forecast for each customer category; whereas the choice model accounts for heterogeneity by including choice-invariant variables that accounts also for the customer category.
The methodological procedure in this paper follows the study by Hetrakul and Cirillo (2014) that proposes, in a railway setting, an optimization model in which discrete choice models and demand function are integrated, in order to calculate the price and fraction of the demand to be accepted for each origin-destination pair.

# 4 Demand forecast

## 4.1 Sample selection

The demand estimation is based on booking data from the sale system of the Royal Danish Theater for the period 2010/2011 to 2014/2015. The sample is constituted by 401 opera performances which took place during that period. We estimate a demand function for each customer category identified, with which we refer to the price type applied by the theater in the price discrimination process across buyers. Hence, we assume that the market segmentation is solely based on the price leverage. Among the numerous price type existing (including customers with a loyalty card, employees, group sales, disabled...), we consider the three main customer

categories that together account for nearly 80% of the total tickets sold: standard ticket buyers (45.9%) who pay the full price for the ticket; young (under 25 years)-student customers (6.1%) for which tickets are discounted by 50%; and subscribers (26.1%). Regarding the subscribers categories, Royal Danish Theatre applies two type of subscription: a fixed subscription, in which the bundle of events included is predetermined by the theatre, and a "choose your own" subscription that allows the customers to choose the productions they want to see. In the first case, a discount of 15% is applied, whereas in the second case the discount drops to 10%. In order to simplify the optimization model, we merge the two types of subscription, considering the average discount of 12.5%. In this category we include also the additional tickets that a subscriber can purchase, besides his subscription. For example, when a subscriber buys a performance ticket of a production that it is not included in the subscription, also this ticket is discounted by 10%.
For the purpose of model simplicity, there are some remarkable categories that, given their low number of attendees per performance, are not considered. For example, tickets for senior customers, which are entitled to a discount of 50%, represents only 2.5% of the tickets sold just because this discount is made available only for some performances decided by theatre management. Indeed, as many senior customers are subscribers, it does not result convenient to offer this discount for all the performances. We exclude also the young/student subscribers, which accounts for 0.74% of the total theater market: their discount is 65% for a fixed subscription and 60% for a "choose your own" subscription.

## 4.2 Demand estimation

Following the literature, we adopt a double-log specification, which is the most popular functional form adopted in estimating theatre-attendance demand (Seaman, 2006). For each category $j$, the following demand function is estimated:

$$ln(D_j) = \alpha_j + \beta_j ln(p_j) + \gamma_j' z + \epsilon_j \tag{1}$$

so as:

$$D_j = exp(\alpha_j + \beta_j ln(p_j) + \gamma_j' z + \epsilon_j) \tag{2}$$

where, for given a performance, $D_j$ is the number of tickets sold to category $j$, $p_j$ is the average price of ticket deflated by CPI[2] charged to category $j$: in particular, we take the average price of the different seat categories offered by the theatres. $z$ is a vector of performance and production characteristics, while $\epsilon_j$ is an error term.

[2] CPI data are collected by *Statistics Denmark*: http://www.dst.dk/en

Concerning the performances scheduling, we include three dummy variables to take into account the weekly seasonal effect: WKDAY denotes performances run during weekdays (from Monday morning to Friday morning); WKEND indicates performances run during Friday and Saturday evening or during the evening before a public holiday. Finally SUNDAY denotes performances that take place on Sunday or in a public holiday day. This latter group of performances are "matinee" as no evening performances take place on Sunday. Except on Sundays, in the other days of week performances can take place either on monday-afternoon or during the evening. We denote with EVE performances that take place during the evening.
In order to capture the seasonality effect, we construct month dummies variables for each month of the year, except for July and August when the theatre is closed. In addition, following Corning and Levi (2002) we include also REMAIN and TOTPERF denoting respectively the number of remaining and total perfomances of a given production. We find also a significant interaction between these two variables: indeed, this interaction term allows to weight the amount of remaining performance: for instance, the effect of the second to last performance changes when the total number of performances are twenty or, for example, five.
We also control for the production characteristics: to capture the popularity of an opera show, we introduce the variable POP measured as the number of times the production is performed worldwide during the same year it has been performed in the Royal Danish Theatre [3]. However, it should be considered that some Danish production (e.g *Maskarade*, *Livlægens besøg*) are popular in Denmark but not worldwide. To control for this aspect, we include the dummy DANISH, denoting Danish productions. Moreover, the dummy variable NEWDKT controls for productions that take place for the first time at Royal Danish Theater.
We also control for the year the production has been created by introducing three dummies: 1920-2015, 1850-1919, BEFORE 1850.
As our analysis is based on performances running in 5 years, we include a time trend $t$ variable. Finally, considering that the total capacity of the theatre can change due to production requirements, we add the variable CAPACITY indicating the number of the available seats for the show.
In estimating the demand function for subscribers, we add a new variable SUBYEAR as the log number of subscribers in the current season. Indeed subscriptions are sold in advance and the number of subscribers is known to the theatre before the season starts. However, for customers who buy a fixed subscription, it is unknown

[3] We collect these data through "Operabase", a website designed to collect statistics about operatic activity worldwide: http://operabase.com

their distribution among performances, given a production.
Table 1 provides a descriptive statistics of the data.

| **Variable** | **Mean** | **SD** | **Min** | **Max** |
|---|---|---|---|---|
| Price (standard ticket) | 456.06 | 74.93 | 208.96 | 661.13 |
| Standard tickets sold | 562.34 | 252.11 | 62 | 1117 |
| Young tickets sold | 73.46 | 63.65 | 0 | 576 |
| Subscribers ticket sold | 310.33 | 151.37 | 0 | 716 |
| REMAIN | 7.49 | 5.38 | 1 | 30 |
| TOTPERF | 14 | 6.07 | 6 | 30 |
| CAPACITY | 1482.89 | 45.51 | 1297 | 1529 |
| POP | 213.17 | 186.00 | 1 | 507 |
| SUNDAY | 0.174 | 0.380 | 0 | 1 |
| WKEND | 0.257 | 0.437 | 0 | 1 |
| WKDAY | 0.568 | 0.496 | 0 | 1 |
| EVE | 0.733 | 0.443 | 0 | 1 |
| JANUARY | 0.157 | 0.364 | 0 | 1 |
| FEBRUARY | 0.117 | 0.322 | 0 | 1 |
| MARCH | 0.149 | 0.357 | 0 | 1 |
| APRIL | 0.115 | 0.319 | 0 | 1 |
| MAY | 0.147 | 0.355 | 0 | 1 |
| JUNE | 0.047 | 0.2127 | 0 | 1 |
| SEPTEMBER | 0.027 | 0.163 | 0 | 1 |
| OCTOBER | 0.085 | 0.279 | 0 | 1 |
| NOVEMBER | 0.125 | 0.331 | 0 | 1 |
| DECEMBER | 0.030 | 0.171 | 0 | 1 |
| 1920-2015 | 0.160 | 0.366 | 0 | 1 |
| 1850-1919 | 0.486 | 0.500 | 0 | 1 |
| BEFORE 1850 | 0.354 | 0.479 | 0 | 1 |
| DANISH | 0.027 | 0.163 | 0 | 1 |
| NEW DKT | 0.651 | 0.477 | 0 | 1 |
| $t$ | 3.06 | 1.295 | 1 | 5 |

401 observations

Table 1: Descriptive statistics of OLS variables

We estimate (1) by OLS with robust standard-error. Although more sophisticated models are available for a forecast analysis (Ainslie *et al.*, 2005), not necessarily such techniques provide a significant improvement (Andrews *et al.*, 2008; Eliashberg *et al.*, 2009).
We have also checked for multicollinearity issues that does not seem to arise.
Table 2 shows the estimation results of the demand function for all the categories considered.
Results of the demand estimation reveal that price elasticity differs across customer category. Young customers is the audience group most price sensitive: a 1 % increase in ticket price results in approximately 1.84 % decline in demand. Standard ticket buyers are less price sensitive as the price elasticity is less than unity: a 1 % increase in ticket price results in approximately 0.49 % decline in demand. Consistent with

| Variable | Single tickets | Young | Subscribers |
|---|---|---|---|
| Intercept | 2.3538** | 6.7917**** | −0.1729 |
| | *(1.017)* | *(1.986)* | *(3.1428)* |
| Log price | −0.4904*** | −1.8440**** | −0.1315 |
| | *(0.1811)* | *(0.3994)* | *(0.7346)* |
| SUNDAY | 0.22741**** | −0.0459 | 0.2652 |
| | *(0.0654)* | *(0.1321)* | *(0.1980)* |
| WKEND | 0.4620**** | 0.0083 | −0.0338 |
| | *(0.0357)* | *(0.0644)* | *(0.0926)* |
| EVE | −0.1205** | −0.0220 | 0.1626 |
| | *(0.0596)* | *(0.1106)* | *(0.1838)* |
| REMAIN | −0.0575**** | −0.0483*** | −0.0122 |
| | *(0.0089)* | *(0.0173)* | *(0.0249)* |
| TOTPERF | 0.0365**** | −0.0081 | −0.0472*** |
| | *(0.0049)* | *(0.0097)* | *(0.0154)* |
| REMAIN x TOTPERF | 0.0020**** | 0.0018** | 0.0013 |
| | *(0.0004)* | *(0.0007* | *(0.0009)* |
| JANUARY | 0.2743* | −0.0458 | 1.0828** |
| | *(0.1603)* | *(0.2493)* | *(0.5359)* |
| FEBRUARY | 0.3370** | 0.1374 | 1.1611** |
| | *(0.1590)* | *(0.2415)* | *(0.5034)* |
| MARCH | 0.3383** | −0.0316 | 1.2604** |
| | *(0.1568)* | *(0.2350)* | *(0.4980)* |
| APRIL | 0.4957*** | −0.0492 | 1.2728** |
| | *(0.1580)* | *(0.2384)* | *(0.5195)* |
| MAY | 0.5726**** | −0.1028 | 1.3395*** |
| | *(0.1542)* | *(0.2318)* | *(0.5044)* |
| JUNE | 0.5192*** | −0.2148 | 1.1711** |
| | *(0.1631)* | *(0.2764)* | *(0.4956)* |
| SEPTEMBER | −0.2083 | −0.9979*** | 1.5058*** |
| | *(0.1949)* | *(0.3619)* | *(0.5473)* |
| OCTOBER | 0.0237 | −0.3690 | 1.3119** |
| | *(0.1597)* | *(0.2449)* | *(0.5255)* |
| NOVEMBER | 0.0575 | −0.2394 | 1.0168** |
| | *(0.1554)* | *(0.2315)* | *(0.5030)* |
| POP | 0.0007**** | 0.0022**** | −0.0013**** |
| | *(0.0001)* | *(0.0002)* | *(0.0002)* |
| 1850-1919 | 0.6935**** | 0.1236 | 0.6683*** |
| | *(0.0758)* | *(0.1903)* | *(0.2374)* |
| BEFORE 1850 | 0.6385**** | 0.1108 | 0.8165**** |
| | *(0.0743)* | *(0.1851)* | *(0.2292)* |
| DANISH | −0.1132 | 0.8734**** | −0.5989*** |
| | *(0.0858)* | *(0.1247)* | *(0.1997)* |
| NEWDKT | −0.0648* | 0.1970*** | −0.0524 |
| | *(0.0376)* | *(0.0771)* | *(0.0774)* |
| CAPACITY | 0.0037**** | 0.0045**** | 0.0008 |
| | *(0.0004)* | *(0.0009)* | *(0.0010)* |
| $t$ | −0.0101 | 0.0606** | 0.1007** |
| | *(0.0164)* | *(0.0303)* | *(0.0508)* |
| SUBYEAR | | | 0.5192**** |
| | | | *(0.1348)* |
| R-square | 0.7512 | 0.4213 | 0.3830 |
| Model F-value | 51.64**** | 13.22**** | 5.02**** |
| No. of observations | 401 | 401 | 401 |

Robust st.error listed under coefficients
$^{****}p < 0.001$ $^{***}p < 0.01$ $^{**}p < 0.05$ $^{*}p < 0.10$

Table 2: Estimation results of demand functions

previous results in literature (Felton, 1994; Baldin and Bille, 2016), subscribers are the least price sensitive: for our sample the price coefficient is even not statistically significative. This result is not surprisingly as literature has shown in some cases even a positive price elasticity in the demand for performing arts, configuring the theatrical experience as a Veblen good (Laamanen, 2013).

The results for the single ticket buyers show a strong explanatory power ($R^2 = 0.75$) and almost all variables are statistically significant. In particular, Table 2 shows that, for this type of customers, the demand is higher for Friday/Saturday evening performances. The number of times a title is rerun (TOTPERF), which is supposed to be an indicator of the total expected demand for that production, has a positive impact on the demand for a single performance. Moreover, given the same production, each performance has a 5.75 % higher demand than the previous, holding fixed the number of times a performance is rerun. This is probably due to a word-of-mouth effect (Laamanen, 2013). Furthermore, we can deduce that single ticket buyers prefer traditional and less risky productions than the more experimental ones: indeed the productions that take place for the first time at Royal Danish Theatre have a negative impact on demand; whereas popularity score has a positive impact, as well as those production composed before 1919.

Results for young customers and subscribers have a lower explanatory power ($R^2$ is respectively 0.42 and 0.38). For the former, there is a positive word-of-mouth and time trend effect. Furthermore, the Danish productions have a strong positive effect on demand, as well as the popularity of the production worldwide; but also the productions that take place for the first time at Royal Danish Theater seem to be appealing to young customers.

Concerning subscribers, we note a significant month-seasonality and time trend effect. Contrary to single ticket buyers, subscribers seem to appreciate less conventional productions, as the coefficient associated with the popularity score is negative. On the other side, productions composed before 1850 seem to be preferred by this audience group.

Table 3 compares the actual attendance with the values predicted by the demand functions. The prediction capability of the model is measured with different indicator, such as root mean squared error, mean absolute error, average error and Pearson correlation between predicted and actual. In addition we perform the out of sample validation. We consider 74 performances run during season 2015/2016 that are not included in our sample. The demand functions for such performances are estimated using the coefficients obtained for our initial sample, and their final estimations are compared with the actual attendance.

Whereas the average errors is decidedly higher for the out of sample performances than the sample performances; the other measures are similar among the two groups of performances.

| 2010/2011-2014/2015 | **Root mean squared errors** | **Mean absolute errors** | **Pearson correlation** | **Average errors** |
|---|---|---|---|---|
| Single tickets | 148.33 | 114.55 | 0.78 | 12.76 |
| Young | 52.09 | 28.36 | 0.68 | 9.92 |
| Subscribers | 139.22 | 106.44 | 0.56 | 26.20 |
| 2015/2016 | | | | |
| Single tickets | 155.31 | 133.06 | 0.83 | -86.11 |
| Young | 53.52 | 30.69 | 0.56 | 12.11 |
| Subscribers | 111.02 | 100.84 | 0.57 | -33.58 |

Table 3: Predictive performance of the demand functions

# 5 Customer choice model

## 5.1 Sample selection

The choice model concerns the price discrimination across seating area. The theatre policy has been refined in the last years. In 2010 the *OperaHouse* offered 5 different price zones, 6 price zones in 2011 and 8 seating area from 2012 onwards (Figure 1).

The subdivision is not physically evident: for example, zone "price A" includes both stall seats and first balcony seats; whereas zones "price B" comprehend stall seats as well as first and second balcony seats, and so forth. This allows the theatre manager to be quite flexible in the subdivision of the venue.
Since the number of price zones changed through the period under examination, we aggregated productions with more than five price zones into five seat categories. The procedure adopted follows Baldin and Bille (2016), to which we refer for details.
For logistic reason, it has not been possible to collect data for the choice model estimation for the whole sample considered in the demand function. Our sample consists in 103322 bookings which involve 11 opera productions and 122 performances.

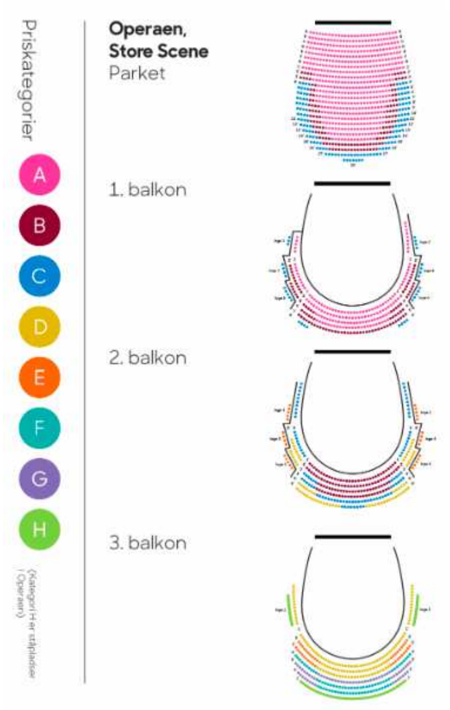


Figure 1: Price zones at the Opera House

## 5.2 MNL estimation

After estimating the demand for each performance, in this section we propose a choice model for the seating area decision. To this aim, we adopt a multinomial logit (MNL) approach. Hence, we assume that each customer chooses the seat that maximizes his utility. The independent variables that enter in the model as choice's attributes are: *price* and a dummy variable for each seat category. These variables aim to capture the tradeoff behavior between cheap seats with low visibility and/or acoustics and more expensive high quality seats. Moreover, in order to address heterogeneity, we allow the price sensitivity and the marginal utility of the seating areas to vary across customer categories. The price coefficient interacts also with variables related to the performance characteristics.
The utility of a customer that buys a ticket which refers to the seating area $s$, for the performance $i$, can be formulated as:

$$U_{sj} = V_{sj} + \epsilon_{sj} \tag{3}$$

with

$$V_{sj} = p_{sj} \cdot (\beta_1 + \beta_2 \cdot young + \beta_3 \cdot sub + \gamma' z) + seat_s \cdot (\delta_1 + \delta_2 \cdot young + \delta_3 \cdot sub) \tag{4}$$

where $young$ and $sub$ are dummy variables denoting whether the customer is respectively a young customer or a subscriber. This implies that single ticket buyers are treated as the base category. $z$ is a vector of performance and production characteristics. In our estimation, such characteristics are represented by the dummy variables SUNDAY and WEEKEND, already defined in the demand function. Moreover, we used the number of times the production is performed worldwide during the same year, in order to define three dummy variables that denote the degree of popularity of the production: $LowPopularity$ (for productions run less than 50 times worldwide) treated as base variable; $MediumPopularity$ (for productions run between 50 and 150 times worldwide) and $HighPopularity$ (for productions run more than 150 times worldwide). Finally $seat_s$ is a dummy variable denoting whether the seat belongs to area $s$ or not.

Assuming that the error components in (2) are independent and identically distributed according to a Gumbel distribution, the probability of a customer belonging to category $j$ purchasing a ticket of seating area $s$ (among the 5 seating areas) is given by:

$$Pr(s \mid j) = \frac{exp[V_{sj}]}{\sum_{t=1}^{5} exp[V_{tj}]} \tag{5}$$

Estimation results for the MNL model are displayed in Table 4.

As expected, young customers are highly price sensitive, followed by standard-ticket buyers and subscribers. In addition, the price coefficient increases significantly when we consider popular productions as well as, surprisingly, performances that take place on Sunday.

With regard to the seat quality, the coefficients reflect an expected pattern: keeping the price fixed, an increase of the quality of the seat leads to a greater utility. This pattern holds for all the customer categories considered. Contrary to Baldin and Bille (2016), we can not compare the marginal utility of the seat categories across customers categories because each category has its own price coefficient. However, in terms of willingness to pay (WTP), i.e the ratio between the coefficient of the attribute and the price coefficient, it results that this value is greater for subscribers, followed by standard-ticket buyers and young customers.

# 6 Optimization model

The optimization model we propose considers the two objectives of the theatre, i.e., to maximize revenue and attendance, in a constrained bi-objective maximization framework. It incorporates both the demand function and the customers' seat

| | Coefficient | *t-stat* |
|---|---|---|
| **Price** | **-0.00129**$^{****}$ | $-11.88$ |
| Price-Young | $-0.0109^{****}$ | -26.32 |
| Price-Subscribers | $0.00074^{****}$ | 3.87 |
| Price-Popularity Medium | 0.00005 | $-1.03$ |
| Price-Popularity high | $0.00036^{****}$ | $-1.93$ |
| Price-Wkend | 0.00004 | 0.91 |
| Price-Sunday | $0.00023^{****}$ | 5.55 |
| **Seat 2** | **0.782**$^{****}$ | 28.93 |
| Seat 2 - Young | $0.427^{****}$ | 7.09 |
| Seat 2 - Subscriber | 1.51 | 22.84 |
| **Seat 3** | **1.37**$^{****}$ | 34.83 |
| Seat 3 - Young | $0.474^{****}$ | 5.44 |
| Seat 3 - Subscriber | $1.56^{****}$ | 18.78 |
| **Seat 4** | **1.87**$^{****}$ | 36.04 |
| Seat 4 - Young | $1.12^{****}$ | 9.81 |
| Seat 4 - Subscriber | $1.73^{****}$ | 16.92 |
| **Seat 5** | **1.95**$^{****}$ | 29.37 |
| Seat 5 - Young | $1.56^{****}$ | 11.36 |
| Seat 5 - Subscriber | $1.86^{****}$ | 14.58 |
| No. of observations | | 103322 |
| $\rho^2$ | | 0.102 |
| Adjusted $\rho^2$ | | 0.102 |
| Null log-likelihood | | - 166290.344 |
| Final log-likelihood | | - 149291.813 |

$^{****}p < .001$

Table 4: Estimation of multinomial logit model

choices described in the previous section. The decision variables are the prices $p_{sj}$, for each seating area $s$ and each customer category $j$. As these prices affect the demand and the customers' seat choice (as described by (formula...-.)), the optimal prices determine the optimal splitting into fare classes of the seats in the theatre. The expected revenue and attendance can be written as, respectively,

$$Revenue = \sum_{j=1}^{3} D_j(p_j) \cdot \left[ \sum_{s=1}^{5} Pr(s \mid j) \cdot p_{sj} \right] \tag{6}$$

and

$$Attendance = \sum_{j=1}^{3} D_j(p_j) \cdot \left[ \sum_{s=1}^{5} Pr(s \mid j) \right], \tag{7}$$

where $D_j$ is the number of tickets sold to category $j$, defined by the estimated demand function (2); $p_j$ is the average price for a customer belonging to category $j$; $Pr(s \mid j)$ is the probability of buying a ticket of seating area $s$, given the customer category $j$, for the considered performance, as defined by (5). The maximum number of seats that can be sold is bounded by the capacity of the theatre $C$:

$$\sum_{j=1}^{3} D_j(p_j) \cdot \left[\sum_{s=1}^{5} Pr(s \mid j)\right] \leq C \ . \tag{8}$$

Moreover, we have to consider a set of constraints that are required by the theatre policy:

$$p_{(s-1)j} < p_{sj} < p_{(s+1)j}, \quad \text{for each } j \text{ and } s \tag{9}$$

and

$$p_{s(j-1)} < p_{sj} < p_{s(j+1)}, \quad \text{for each } j \text{ and } s. \tag{10}$$

As seen in Section 4.1, both the ticket price for a young customer and for a subscribers are obtained discounting the standard ticket price, given a seating area $s$.
However, we allow for a more flexible relationship:

$$0.4 \cdot p_{standardticket} < p_{young} < 0.6 \cdot p_{standardticket}, \tag{11}$$

$$0.7 \cdot p_{standardticket} < p_{subscriber} < 0.9 \cdot p_{standardticket}. \tag{12}$$

Finally, we have the constraint that defines the relation between $p_{sj}$ and $p_j$

$$p_j = \frac{1}{5} \sum_{s=1}^{5} p_{sj}. \tag{13}$$

# 7 Optimization results

The bi-objective optimization model we solved consists in maximizing the two objectives, *Revenue* and *Attendance*, under the above defined constraints: the solution of such a problem is the set of Pareto optimal points, the so-called Pareto frontier of the problem. We observe that we are facing a nonlinear bi-objective problem, due to the exponential term both in the demand function and in the formulation of the probability in the multinomial logit model. As usual in multi-objective optimization, in particular in the non-linear case, it is convenient to look for some points of the Pareto frontier; those points should be interesting from the point of view of the decision maker, in our case the direction of the Theatre.
We solved the problem by means of the Synchronous Approach adopted by Miettinen and Mäkelä (2006). Their model, called NIMBUS (Nondifferentiable Interactive Multiobjective BUndle-based optimization System), allows to deal with nondifferen-

tiable and nonconvex multiobjective optimisation problems. The approach is based on the interaction between the decision maker and the solution algorithm and is realized via the Internet based system WWW-NIMBUS (https://wwwnimbus.it.jyu.fi). The single steps of the solution approach consist in the solution of single objective (sub)problems via classical subgradients methods (see, e.g., Clarke, 1983). Successive single optimisation subproblems are then solved under the guidance of the decision maker: each successive solution is a Pareto optimal solution of the multiobjective problem. At each iteration the decision maker can indicate the preferred way to navigate the set of Pareto optimal solutions, choosing the objectives which value should be improved and, at the same time, which objectives should pay the cost of such improvement. This way the most appropriate solutions from the decision maker point of view are selected from the Pareto optimal solutions set. The software is free for the academic community and is operated directly on an Internet site, requiring neither the download of any software nor huge computing capabilities of the client computer.

As case studies we consider three performances that differ each other by characteristics that affect the demand, in order to verify how different levels of theater occupancy require different pricing and allocation policies, in particular considering the peak-load pricing issue (i.e differentiating prices charged depending on peak and off-peak periods). For purpose of better comparison between the actual pricing and allocation policies and those resulting from the optimization model, we have choosen three performances that show a fitted value of the demand very close to the real demand. The first performance is a high demand performance, namely a Saturday evening performance of *La Tosca* that fills up to 91.05% capacity . The second performance analyzed is a low demand performance, *Djævlene fra Loudun*, run in a weekday: this performance fills less than half of the total capacity (41.98%). The third performance is a medium-popular production, namely *Rusalka*, with 67.86% of the total capacity filled.

We are therefore able to compare the price[4] and seat allocation results of the bi-objective optimization model with the actual results and those resulting from two single objective optimization models: one that maximizes only the Revenue and one that maximizes only the total Attendance (see Tables 5-7).

Some remarks about the implementation of the models: first, for a realistic comparison with the actual data we have considered three performances whose fitted value of the demand is very close to the actual value. Second, for the purpose of realism we have imposed a lower and an upper bound to the 15 decision variables,

[4]Price are expressed in Danish crown (DKK):$1DKK \approx 0.13$€

equal respectively to the half and twice value of the actual price. Third, we subtract from the value of the capacity $C$ the number of tickets sold to other categories not considered, which is considered as already known by the theatre manager.
Table 5 considers the results obtained for a Saturday evening performance of *La Tosca*. It is a high-demand event almost (but not completely) sold-out.

| Seat | Actual | | Bi-objective | | Revenue max. | | Attendance max. | |
|---|---|---|---|---|---|---|---|---|
| | Price | no. of seats | Price | no. of seats | Price | no. of seats | Price | no. of seats |
| Seat1-standard | 160 | 46 | 300 | 59 | 320 | 67 | 176 | 60 |
| Seat2-standard | 345 | 79 | 341 | 123 | 690 | 105 | 399 | 108 |
| Seat3-standard | 525 | 144 | 625 | 173 | 1050 | 137 | 408 | 193 |
| Seat4-standard | 720 | 366 | 963 | 211 | 1440 | 160 | 536 | 285 |
| Seat5-standard | 895 | 258 | 1360 | 160 | 1710 | 136 | 565 | 301 |
| Seat1-young | 80 | 13 | 134 | 11 | 160 | 16 | 74 | 18 |
| Seat2-young | 173 | 23 | 170 | 24 | 299 | 10 | 172 | 19 |
| Seat3-young | 263 | 14 | 252 | 17 | 420 | 5 | 226 | 19 |
| Seat4-young | 360 | 14 | 386 | 11 | 576 | 2 | 294 | 27 |
| Seat5-young | 448 | 8 | 544 | 3 | 684 | 1 | 311 | 37 |
| Seat1-subscribers | 140 | 7 | 226 | 1 | 224 | 1 | 153 | 1 |
| Seat2-subscribers | 302 | 13 | 293 | 10 | 604 | 9 | 319 | 10 |
| Seat3-subscribers | 459 | 24 | 540 | 17 | 918 | 17 | 349 | 18 |
| Seat4-subscribers | 630 | 18 | 842 | 33 | 1260 | 31 | 428 | 35 |
| Seat5-subscribers | 783 | 42 | 1215 | 38 | 1539 | 37 | 497 | 43 |
| Total | 682 118 | 1069 | 691 167 | 891 | 826 714 | 733 | 529 468 | 1174 |
| % improve (Revenue and attendance) | | | +1.33 | -16.65 | +21.20 | -31.43 | -22.38 | +9.83 |

Table 5: Revenue and attendance comparison. Case study: *La Tosca*

The bi-objective optimization model solutions shown in Table 5 (as well as all the other solutions shown in Table 6 and 7) represent one of the points of the Pareto frontier. Hence there are other possible solutions. Figure 2 shows some alternative solutions of the bi-objective model. The solution proposed in Table 5 leads to an increment in revenue of 1.33% and a decrease of the total attendance of 16.65%.

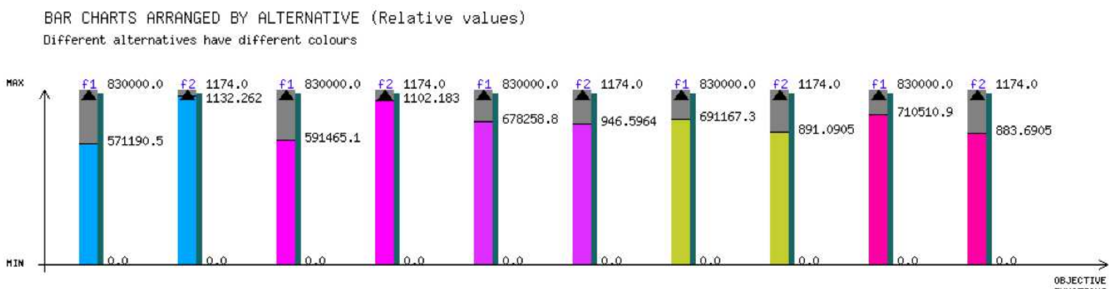


Figure 2: Some alternative optimal values of the bi-objective model

From Figure 2 it is evident the existence of a trade-off among the two objectives: an increase in revenue is associated with a lower value of the attendance and viceversa. It is interesting to observe how price and seat allocation change according to the orientation of the theatre manager towards the two objectives. From Table 5 we can deduce that when the objective is the maximization of the revenue, the theatre exploits the inelasticity that characterizes subscribers and standard tickets by increasing the price to the upper bound. As young customers are price sensitive, their price increases until the loss of young customers is not balanced by a higher revenue per seat. In the attendance maximization perspective, since the performance almost reaches the capacity constraint, the objective is achieved by lowering only the prices of the most expensive seat category.

In relation to the allocation policy, we notice that when the theater is "attendance maximizer" customers are more likely to shift to a higher seat quality (buy-up behaviour) as a consequence of a generalized price reduction. Viceversa, if the theater is "revenue maximizer" customers are more likely to buy a ticket for a cheap seat because they are not willing to pay more. This behaviour is evident when we refer to price sensitive customers. On the contrary, price insensitive customers are not influenced by the theatre policy in their choice of the seating area, which is confirmed in Figures 3 and 4 - respectively for young customers and subscriptions - showing how the probability of buying a ticket of certain seating area changes according to the theatre policy. Thus, the optimal pricing and allocation policy depends on the type of customers the theatre is expected to accomodate.

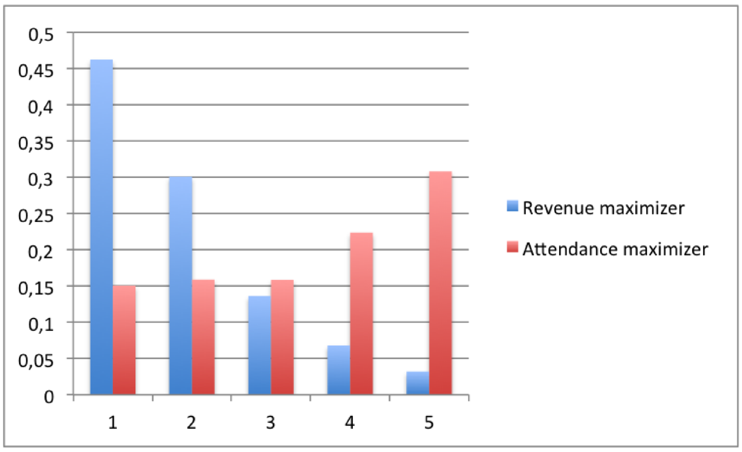


Figure 3: Young customers' choice probabilities

Table 6 considers the results obtained for a weekday performance of *Djævlene fra Loudun*. It is a low-demand event in which the theater is usually occupied approximately only a little bit more than a third of its capacity. One solution obtained solving the bi-objective model allows an increase in attendance of 13.26% and a decrease in revenue of 12.25%. Compared to the previous case, here the theater is forced to reduce prices to the lower bound when it aims to maximize attendance.

Concerning the allocation policy, the pattern previously described is even more ev-

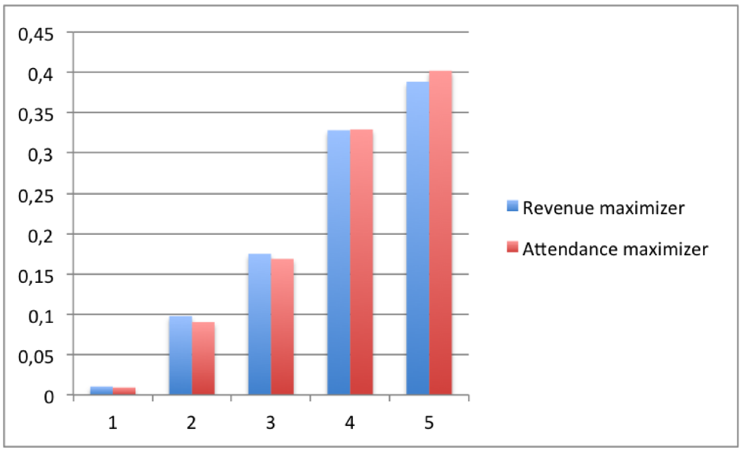


Figure 4: Subscribers' choice probabilities

| Seat | Actual | | Bi-objective | | Revenue max. | | Attendance max. | |
|---|---|---|---|---|---|---|---|---|
| | Price | no. of seats | Price | no. of seats | Price | no. of seats | Price | no. of seats |
| Seat1-standard | 160 | 41 | 110 | 12 | 320 | 8 | 80 | 13 |
| Seat2-standard | 295 | 23 | 216 | 24 | 590 | 20 | 148 | 26 |
| Seat3-standard | 425 | 29 | 366 | 36 | 850 | 26 | 213 | 44 |
| Seat4-standard | 545 | 50 | 505 | 50 | 1090 | 31 | 307 | 66 |
| Seat5-standard | 695 | 42 | 560 | 50 | 1352 | 24 | 348 | 69 |
| Seat1-young | 80 | 16 | 47 | 9 | 160 | 6 | 40 | 9 |
| Seat2-young | 148 | 11 | 91 | 17 | 277 | 4 | 74 | 21 |
| Seat3-young | 213 | 1 | 148 | 16 | 340 | 4 | 107 | 26 |
| Seat4-young | 273 | 9 | 202 | 27 | 436 | 4 | 137 | 58 |
| Seat5-young | 348 | 3 | 226 | 33 | 541 | 2 | 174 | 62 |
| Seat1-subscribers | 140 | 3 | 79 | 2 | 224 | 3 | 70 | 2 |
| Seat2-subscribers | 258 | 11 | 162 | 23 | 516 | 23 | 129 | 22 |
| Seat3-subscribers | 372 | 13 | 287 | 41 | 744 | 38 | 186 | 41 |
| Seat4-subscribers | 551 | 86 | 443 | 74 | 981 | 67 | 276 | 79 |
| Seat5-subscribers | 608 | 107 | 503 | 88 | 1216 | 72 | 304 | 97 |
| Total | 209 268 | 445 | 183 420 | 504 | 302 384 | 331 | 144 313 | 637 |
| % improve (Revenue and attendance) | | | -12.35 | +13.26 | +44.50 | -25.61 | -31.04 | +43.15 |

Table 6: Revenue and attendance comparison. Case study: *Djævlene fra Loudun*

ident as the price coefficient of the MNL model decreases when a no popular event is considered. Hence, if the theatre is attendance (revenue) maximizer, it is suggested to increase(decrease) the size of the most expensive (cheapest) seating area, especially when it is expected to attract a price-insensitive audience group (as yung customers).

Table 7 considers the results obtained for a Sunday performance of *Rusalka*. This is an intermediate case compared to the previous two. This case is interesting as the bi-optimization model provides a solution that dominates the current value of the objectives. Indeed, the solution proposed allows an increase in revenue of 1.90% and, at the same time, an increase in attendance of 1.83%.

| Seat | Actual | | Bi-objective | | Revenue max. | | Attendance max. | |
|---|---|---|---|---|---|---|---|---|
| | Price | no. of seats | Price | no. of seats | Price | no. of seats | Price | no. of seats |
| Seat1-standard | 160 | 43 | 172 | 35 | 320 | 40 | 84 | 39 |
| Seat2-standard | 345 | 70 | 292 | 68 | 692 | 59 | 204 | 75 |
| Seat3-standard | 525 | 72 | 619 | 88 | 1052 | 74 | 315 | 120 |
| Seat4-standard | 720 | 150 | 740 | 129 | 1440 | 83 | 408 | 181 |
| Seat5-standard | 895 | 109 | 782 | 134 | 1743 | 66 | 483 | 182 |
| Seat1-young | 80 | 6 | 79 | 8 | 160 | 7 | 45 | 9 |
| Seat2-young | 173 | 18 | 127 | 15 | 301 | 4 | 92 | 17 |
| Seat3-young | 263 | 7 | 249 | 6 | 423 | 2 | 183 | 11 |
| Seat4-young | 360 | 8 | 296 | 12 | 580 | 1 | 184 | 35 |
| Seat5-young | 448 | 2 | 420 | 4 | 702 | 0 | 224 | 36 |
| Seat1-subscribers | 140 | 19 | 149 | 5 | 224 | 5 | 71 | 5 |
| Seat2-subscribers | 302 | 27 | 259 | 44 | 622 | 43 | 151 | 45 |
| Seat3-subscribers | 459 | 35 | 542 | 76 | 920 | 74 | 266 | 82 |
| Seat4-subscribers | 630 | 163 | 639 | 146 | 1260 | 134 | 322 | 160 |
| Seat5-subscribers | 783 | 175 | 703 | 177 | 1568 | 152 | 397 | 194 |
| Total | 550 190 | 929 | 560 664 | 946 | 874 000 | 745 | 394 026 | 1191 |
| % improve (Revenue and attendance) | | | +1.90 | +1.83 | +58.85 | -19.91 | -28.38 | +28.20 |

Table 7: Revenue and attendance comparison. Case study: *Rusalka*

# 8 Conclusions

This paper has proposed a model that simultaneously optimizes the pricing and seating-allocation policy of a theater. In particular, we present a bi-objective optimization model that integrates the demand forecast and a choice model, where the customer chooses one among different seating areas which differ each other in price and quality. The multi-objective nature of our model reflects the multi-dimensional nature of nonprofit performing arts organizations. In our case, the objectives we assume to be maximized are revenue and attendance. The approach adopted allows also to take into account heterogeneity in price sensitivity and choice behaviour across different customer categories. The proposed model is applied to booking data provided by the Royal Danish Theater referring to the period 2010-2015. More precisely we consider three different performances in order to explore the potentialities of the model.

From a management perspective, the model can provide to theatre managers insightful policy implications in terms of demand-management decision. The results obtained confirm the existence of a trade-off between the two theater objectives. When the theater is revenue maximizer, prices charged to price insensitive customers are raised, and the cheapest seating area is enlarged to prevent a loss of revenue. Viceversa, when the theater is audience maximizer, prices are set at lower level, in particular the ones associated to the most expensive seating area. As a consequence, it is recommended to increase the number of seats allocated to the most expensive area, in order to encourage a shift of consumer choices to higher quality seats. The allocation policy just described is particularly effective when a performance is expected to attract customers with an elastic demand, since these customers are more sensitive to price changes in their seat choice; and also when the performance will probably not attract a large audience. Moreover, in one case the bi-objective model provides a solution that gives an improvement in both revenue and attendance from the current situation.

Overall, our examples clarify that both price and capacity allocation are leverages with which a theatre can calibrate its objectives, even when revenue is not considered as the main goal to pursue.

## References


- Ainslie, A., Dreze, X., Zufryden, F. (2005). Modeling movie lifecycles and market share. *Marketing Science*, 24, 508-517.

- Andrews, R. L., Currim, I. S., Leeflang, P., Lim, J. (2008). Estimating the Scan Pro model of store sales: HB, FM or just OLS? *International Journal of Research in Marketing*, 25, 22-33.

- Baldin, A., Bille, T. (2016). Modelling preference heterogeneity for theater tickets: a discrete choice modelling approach on Royal Danish Theater booking data., *Paper presented at the 19th International Conference on CUltural Economics (ACEI 2016), Valladolid (Spain).*

- Baumol, W. J., Bowen, W.G. (1966). *Performing Arts: The Economics Dilemma*, Twentieth Century Fund, New York.

- Bille Hansen, T. (1991). *Det Kgl. Teater - et kulturøkonomisk studie*, AKF Forlaget, Copenhagen.

- Brooks, A.C. (2006). Nonprofit firms in the performing arts, *Handbook of the Economics of Art and Culture*, 1, 473-506.

- Clarke, F.H. (1983). *Optimization and nonsmooth analysis*, Wiley, New York.

- Corning, J., Levy, A. (2002). Demand for live theater with market segmentation and seasonality. *Journal of Cultural Economics*, 26, 217-235.

- Eliashberg, J., Hegie, Q.,Ho, J., Huisman, D., Miller, S. J., Swami, S., Weinberg, C.B., Wierenga, B. (2009). Demand-driven scheduling of movies in a multiplex. *International Journal of Research in Marketing*, 26, 75-88.

- Felton, M.V. (1994). Evidence of the existence of the cost disease in the performing arts. *Journal of Cultural Economics*, 18, 301-312.

- Gapinski, J.H. (1985). Do the nonprofit performing arts optimize? The moral form Shakespeare, *Quarterly Review of Economics and Business*, 25, 27-37.

- Hansmann, H. (1980): The role of nonprofit enterprise, *Yale Law Journal*,89, 835-898

- Hansmann, H. (1981). Nonprofit enterprise in the performing arts.*The Bell Journal of Economics*, 341-361.

- Hetrakul, P., Cirillo, C. (2014). A latent class choice based model system for railway optimal pricing and seat allocation. *Transportation Research Part E: Logistics and Transportation Review*, 61, 68-83.

- Huntington, P. A. (1993). Ticket pricing policy and box office revenue *Journal of Cultural Economics*, 17, 71-87.

- Laamanen, J.P. (2013). Estimating demand for opera using sales system data: the case of Finnish National Opera. *Journal of Cultural Economics* 37, 417-432.

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

- Luksetich, W.A., Lange, W.A. (1995). A simultaneous model of nonprofit symphony, *Journal of Cultural Economcs*, 19, 49-68

- Marco-Serrano, F. (2006). Monitoring managerial efficiency in the performing arts: A regional theatres network perspective. *Annals of Operations Research*, 145, 167-181.

- Miettinen, K., Mäkelä, M.M. (2006). Synchronous Approach in Interactive Multiobjective Optimization. *European Journal of Operational Research*,170, 909-922

- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *The journal of political economy* 82, 34-55.

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

- Seaman, B. A. (1985). Price discrimination in the arts. *Managerial Economics for the Arts. Association for Cultural Economics, University of Akron*, 47-60.

- Seaman, B. A. (2006). Empirical studies of demand for the performing arts. *Handbook of the economics of art and culture*, 1, 415-472.

- Steinberg, R. (1986). The revealed objective functions of nonprofit firms, *Rand Journal of Economics*,17, 508-526.

- Talluri, K., Van Ryzin, G. (2004). Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1), 15-33.

- Tereyagoglu, N., Fader, P., Veeraraghavan, S. (2012). Filling seats at a theater: Estimating the impact of posted prices and dynamic discounts *Working paper, University of Pennsylvania, Philadelphia.*

- Throsby, D. (1994). The production and consumption of the arts: A view of cultural economics, *Journal of Economics Literature*, 32, 1-29

- Throsby, D., Withers, G.A. (1979). *The Economics of the Performing Arts*, St. Martin's Press, New York.