The impact of social media activities in theater demand

Andrea Baldin*[1], Trine Bille†[2], Raghava Rao Mukkamala‡[3], and Ravi Vatrapu§[4]

[1,2] *Department of Management, Politics and Philosophy, Copenhagen Business School, Porcelænshaven 18B, 2000 Frederiksberg, Denmark*

[3,4] *Department of Digitalization, Howitzvej 60, 2000 Frederiksberg, Denmark*

Paper to be presented at the 20th International Conference on Cultural Economics, Melbourne, June 26-29, 2018.

Abstract

It is well known how theatre performances, as experiential goods, are subject to the "nobody knows" property (Caves, 2000) which leads to uncertainty in theatre demand. As such, many studies on theatre demand have included subjective quality indicators, like professional reviews or friends and relatives' evaluations (word-of-mouth mechanisms), in order to correct a prior misspecification of the theatre demand equation. Following this stream of literature, this paper considers an emerging source of information, which is the electronic word of mouth provided by social media. Compared to conventional word of mouth mechanism, social media are endowed with desirable features that may reduce the uncertainty brought by the "nobody knows" property, as they propagate an enormous, enduring and in real time amount of information and opinions. This paper aims to test the potentiality of social media in understanding theatre demand. We do this by combining booking data for the period 2010-2016 from the sale system of the Royal Danish Theatre with volumetric data extracted by the official Facebook page of the theatre. In particular, we take into account the double role of the feedback provide by social media (in terms of likes and comments), looking at two perspectives. The

*E-mail: ab.mpp@cbs.dk
†E-mail: tbi.mpp@cbs.dk
‡E-mail: rrm.digi@cbs.dk
§E-mail: rv.digi@cbs.dk

first one considers the distinction between the predictor and influencer role of the eWOM, whereas the second perspective pertains to the interdependent relationship between eWOM and total audience. These different perspectives are analyzed respectively through two modelling approaches: a panel data model and a simultaneous equation system. Final results suggest how the n° of likes, rather than the Facebook comments, have both a predictor and influencer role. In addition, the interdependence between Facebook "likes" and the amount of ticket sold is demonstrated.

Keywords Theatre demand, Social media, word of mouth, panel data

1 Introduction

The purpose of this paper is to analyze the effect of social media in the demand for performing arts. This issue can be particularly relevant in analysing the consumption of performing arts. Indeed, performing arts provide an illustrative example of how the consumer behaviour, in the cultural sector, does not tie up with the standard assumptions of the neoclassical economic theory. One of the main assumption of the neoclassical theory considers customers as fully and costlessly informed about the market. However, it is well known how theatre performances are experience goods whose quality is not known before it is experienced. This leads to the uncertain nature of the demand for creative goods, which is the "nobody knows" property as referred by Richard Caves (2000). From the supply perspective it becomes difficult to understand why people choose to consume what they do, and hence to predict the attendance of a given show. Similarly, for customers it is difficult to assess the quality and the value of a cultural product before committing to consume it, thus all the information acquired about the quality of a theatrical production play a crucial role in the purchase decision context, determining its box-office success or failure. Consumers may therefore use different quality indicators in order to try and evaluate the quality beforehand. These quality indicators can be either objective facts (such as the title, the author, the director or the actors etc.) or subjective quality indicators such as professional reviews (e.g. in newspapers) or friends and relatives evaluations (word-of-mouth mechanisms).

Because of the rapid development of Internet and social media, a new communication model known as electronic WOM (henceforth, eWOM) based on social media and online community has gained an increasing attention by researchers who analyze the WOM mechanism. Social media can be construed as a form of collective

wisdom (Asur and Huberman, 2010) since they propagate an enormous, highly variable and in real time amount of information through large user communities; they allow users to share information about the feelings, emotional moods and opinions regarding whatever experience, including attending a theatrical performance.
Given that social media are sources for big data, several researchers have explored the possibility to exploit the information inferred by social media (Twitter, Facebook, Youtube..) and web searches (Google Trends, Yahoo Search Query Logs..), to predict real-world outcomes. Such studies include the prediction of iPhone sales (Lassen et al. ,2014); car sales (Voortman, 2015); and political election (Franch, 2013). In the cultural field, applications can be found in the context of movies, in which it is investigated the effect of online review (Duan *et al.*, 2008; Chu *et. al*, 2016) collected through (for instance, Youtube, Yahoo! Movies, etc.) on the movie revenue and audience; and the predictive power of social media such as Twitter (Asur and Huberman, 2010). We can presume that the application of these tools might be more effective in application area characterized by product with a global (or almost) market, such as Iphone, cars and movies. This is not the case of performing arts: each production refers to the institution that supplies it. A *Traviata* that takes place in a theatre in a given country might be different than the same show performed in a theatre in another country, as well as the potential audience is different. Information regarding the quality perceived of a theatrical performance can be taken through critics review (reflecting the experts' opinion) or surveys addressed to the audience; but whereas the effect of the former is found to be not significant on the audience (see Urrutiaguer, 2002); the latter is a time consuming activity for the theater, and not easily available for researchers. Despite its limitation in the performing arts context, compared to other application areas, social media may still represent an additional instrument to collect easily a higher amount of information concerning the audience impression of the performance: this information is exchanged in the community user, leading us to suspect that the diffusion of eWOM is faster and more effective than the traditional WOM mechanism. Moreover, given the ample audience of online communities, social media may have a role in strengthening the so-called "bandwagon effect", a typical characteristics of cultural consumption that essentially indicates how consumers, in order to reduce the search and information costs, follow the crowd and make choices on the base of the choices of the majority.
This paper aim to test whether social media data related to theater productions have an impact on the amount of tickets sold. We do this by combining booking data from the Royal Danish Theater and information come from the theater official Facebook page, from the season 2010/2011 to the season 2015/2016. In particular, we

take into account the double role of the feedback provide by social media (in terms of likes and comments), looking at two perspectives. The first one considers the distinction between the predictor and influencer role of the eWOM: data are predictor whether a correlation between eWOM and attendance reflects production-specifc characteristics (e.g director, title etc...), so that such characteristics simultaneously determine the audience quality perception and the demand for that given production. Alternatively, eWOM has an influencer role whether the relation between eWOM and audience suggests a casuality relation: eWOM provides information to the customers that affect their decision to attend or not attend the performance. The second perspective pertains to the interdependent relationship between eWOM and total audience. According to this framework, eWOM influences current and future audience, which in turn affects the current and future eWOM, and so on. Hence, eWOM both influences and are influenced by the theatre audience.
These different perspective are analyzed respectively through two modelling approaches: a panel data approach and a simultaneous equation system. The structure of the paper can be outlined as follows: Section 2 reviews the literature concerning the adoption of quality indicator in estimating the theatre demand; Section 3 provides the context of our study whereas Section 4 presents the methodology and Section 5 discusses the results obtained. Finally some conclusions are provided.

2 Literature review

The demand for the performing arts is a widely analyzed topic in cultural economics. One of the main difficulty in the empirical studies concerns the inclusion of the qualitative characteristics of the production that, in the case of "higher arts", seem to be decisive for the audience. In literature, different quality indicators have been used, which are summarized in Table 1.
Some authors (Jenkins and Austen-Smith, 1987; Krebs and Pommerehne, 1995) have focused on the concept of "highbrow" and "lowbrow" productions as a quality indicator, providing evidence of a positive impact of "lowbrowness" programmes (measured as the ratio of works with more than 75 performances to all works played in a season) on theatre demand.
Concerning the inclusion of subjective quality indicators, Throsby's (1980, 1990) pioneering works include the reviewers' opinions, developing four variables that assess the press review opinions related to different "technical" dimensions (source material, production, acting and design) and showing their impact on the demand.

Variable	**Description**	**Authors**
Quality review	Press review using a scale from 1(very poor) to 5(excellent)	Throsby (1990); Abbé-Decarroux (1994); Colbert *et al.* (1998); Corning and Levi (2002)
In-house production	Whether it is an in-house production	Abbé-Decarroux (1994); Colbert *et al.* (1998)
Reputation of the author, play, producer and cast	Known/unknown	Abbé-Decarroux (1994), Willis and Snowball (2009)
Reputation of the theatre	Perception of the quality of the theatre (through loyalty of attendance)	Urrutiaguer(2002)
Reputation of the director	Whether the director manages a theatrical institution or not	Urrutiaguer (2002); Willis and Snowball (2009)
Quality of the design and costumes	Expenses for decor and costumes per production	Zieba (2009)
Size of the company		Colbert et al. (1998), Akdede and King (2006)
Repertoire classification	Classic, Modern, Contemporary or Atypical play	Abbé-Decarroux (1994); Corning and Levi(2002); Urrutiaguer(2002)

Table 1: Quality indicators used in literature. Source:Grisolia and Willis (2012)

Tobias (2004) proposes a new measure to aggregate ordinal expert judgments into a single measure, using it into regression in order to find some relations between the quality measure and economic variables. The author finds that production costs is positively related to the quality measure for what concerns opera and ballet, giving evidence of positive but decreasing marginal returns; whereas for drama productions it results that the production-costs are incapable of explaining experts' perception of quality.

Urrutiaguer (2002) proposes that opposing opinions on quality are the reason why regression models often reveal a low statistical significance for the quality variables that are used to explain theatrical demand. A regression equation is constructed in order to explain demand, with continuous variables for price and volume, and with dummy variables for drama critics, directors, growth in funding by public authorities and repertoire classification. Detailed data on demand for French "theatrical institutions" in 1995 and 1996 are used to test the model. To some extent, the results support the hypothesis that the media reputation of shows, as expressed in the form of drama reviews, and the artistic reputation of directors, which are listed

in the program, has opposite effects on attendance. However, the most reliable sign of quality remains the reputation of the theatrical institution.
Finally, Baldin and Bille (2017) assess the impact of both professional reviewers' evaluation (on a scale from 1-5 through reviews in the newspaper) and audiences' evaluation (through surveys to the participants), concluding that while the audiences' positive evaluations have a significant positive impact on attendance, reviewers' evaluations don't.
Following this line of research, this paper aims to verify which is the role of social media data with respect to their capacity in predicting and/or influencing theatre demand. In particular, we consider the Royal Danish Theatre as case study. To the best of our knowledge, this is the first study that considers the role of social media in the performing arts context. Remaining in the broad cultural sector, the effect of the eWOM and social media has been analyzed with respect to the movie industry: Asur and Huberman (2010) predict with high accuracy the box-office revenues of movies in advance of their release using the rate of chatter from Twitter and sentiment analysis of the tweets. Duan *et al.* (2008) develop a simultaneous equation system to capture the interrelationship between eWOM, collected through online reviws, and movie sales, demonstrating that eWOM is both precursor and outcome of movie sales. Chu *et al.* (2016) discuss the role of audience's movie reviewer, collected through online forum, on movie attendance, distinguishig between the influencer and predictor role. The authors conclude that positive online reviews have both a positive prediction and influence effects; the negative reviews have a negative prediction effect and a positive influence effect and the neutral reviews have both negative prediction and influence effects. Their modelling strategy used to distinguish the two effects will be used as well in this paper in Section 4.2.

3 The Royal Danish Theatre

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 theatres 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 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 2017 the theatre had a total budget of 689,2 million DKK (93 million Euros), of which 75 percent were public support from the Government. The theatre had 171,4 million DKK (23 million Euros) in own earnings, of which 74 percent (17 million Euros) was from ticket sales, the rest was income from sponsors etc.
In total 488.668 people attended performances at the theatre in 2017. The main driver of attendees was plays, which accounted for 41 percent (200.163 people) of the total number of attendees. Opera accounted for 25 percent (124.052 people), while ballet accounted for 23 percent (112.913 people). At last 11 percent (51.540 people) of all attendees went to concerts.
Due to its obligations as a national theatre, it has to decide its repertoire based on a number of parameter, 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 and senior) as well as time of the performance, seat categories, the type of the performance, the production costs etc.

4 Methodology

4.1 Dataset

Attendance data are based on booking data from the sale system of the Royal Danish Theater for the period 2010/2011 to 2015/2016. The sample consists of 128 productions (all of which has been mentioned in at least one Facebook post) which took place during that period. Volumetric eWOM data are extracted from the official Facebook page of the Royal Danish Theatre (Det Kongelige Teater). Currently, the page has approximately 64,000 followers. The Social Data are collected using our custom-built and in-house tool: Social Data Analytics Tool (SODATO) (Hussain and Vatrapu, 2014a; Hussain and Vatrapu, 2014b; Vatrapu *et al.*, 2015). The raw dataset consisted of around 311220 data points. Each row is equivalent to an action on Royal theatre's Facebook page. The data are ordered into dimensions, which include both ordered and categorical data and contains information about actor ID, timestamp, event name (whether it is a post, comment, like), actor name, type of post and if relevant links and text value of post and comments. The dataset contains around 4000 posts, around 35000 comments, around 10000 comment replies, likes on comment and likes on comment replies. The rest of the data items are likes on the posts which are around 260 000 and likes are most dominant social media action that was performed by the Facebook users of Royal theatre. The total downloaded dataset contains information from 2008 until 2017, however, we have used the relevant social data with respective to the performances and productions that were considered for this paper.

We choose Facebook (henceforth FB) instead of other social media (such as Twitter and Instagram) simply because it is the one with the higher number of followers.

As shown in Figure 1 and in Table 2, we can notice that the number of productions with at least one FB post increases over the years, especially in the last two season considered. This is more evident looking at the percentage of productions, among all the productions that took place on a given season, with at least one FB post, passing from 26% in the season 2010/2011, to 67.2% in the season 2015/2016. This indicates the growing importance and awareness of social media as a marketing tool.

We carry on two different analysis: in the first one, similarly to Chu *et al.* (2016), we distinguish the predictor and influencer effect through a panel data approach. In the second one, similarly to Duan *et al.* (2008) we estimates a simultaneous equation system that takes into account the fact that eWOM, inferred through FB data, both influence and are influenced by the amount of tickets sold. The variables of the two models will be presented separately.

Season	No. of productions	Production with FB post	Percentage
2010-2011	30	8	26%
2011-2012	48	18	37.5%
2012-2013	41	14	34.1%
2013-2014	55	20	36.4%
2014-2015	54	27	50%
2015-2016	61	41	67.2%

Table 2: Productions with at least on FB post

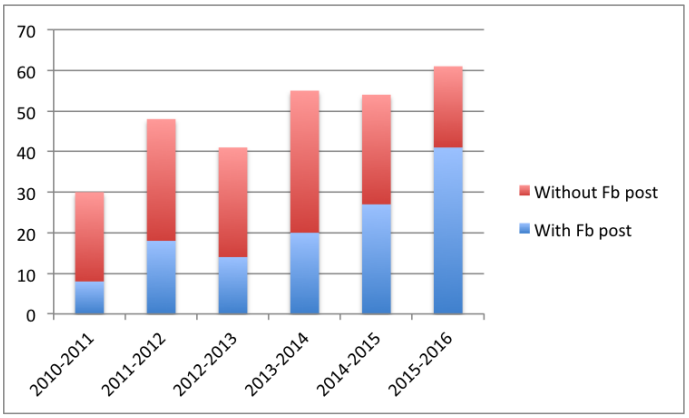

Figure 1: Productions with at least on FB post

4.2 Influencer or predictor?

In this section we investigate the role of FB data. Do they have an influence or a predictive role?Or both? A possible relationship between FB data and attendance can suggest casuality or simply reflects the impact of unobservable characteristics of the product (i.e theater performance). In the former case, FB data has an influencer role, meaning that it provides information to consumers that affect his purchasing decision. For example, many positive "likes" on Facebook concerning a production may encourage customers to see the performance. In the latter case, FB data has a predictor role: a correlation between positive opinions and high demand reflect positive performance-specific characteristics (e.g director, newness..). According to this perspective, such performances-characteristics simultaneously cause FB activities and high demand.

In order to distinguish between these two effects, we rely on the idea by Sher *et al.* (2016), considering a panel dataset estimated by both random effect and fixed effect. If FB data have a predictor effect and reflect the characteristics of a production, that are invariant over time, then the change of FB data and the change of demand over time do not have a systematic relationship. On the contrary, if FB data have a influencer effect, the change over time of demand is related to a change over time of FB data, regardless of the time-invariant production characteristics.

As dependent variable, we consider the weekly number of ticket sold of a given production: For some productions the sale opens for a pre-sale in May; whereas for most production the tickets start to be sold in August. We consider the whole sale period as the FB post related to a production can be published in anytime during this period. Figure 2 shows the trend of tickets sold for the ballet *Sylfiden* during

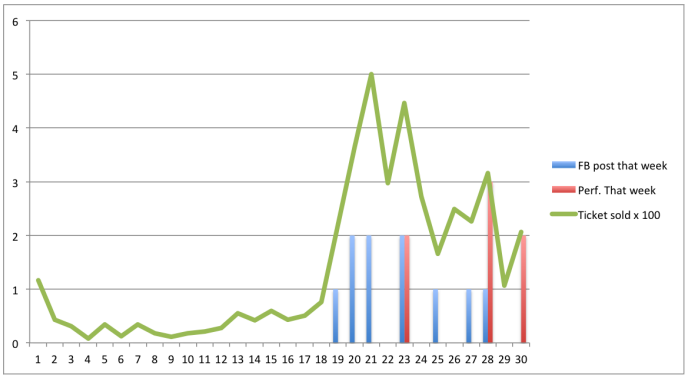

Figure 2: Ticket sold, performance and FB post for the production *Sylfiden*

the 30 weeks of sale. During this period, different FB posts have been published, both before the first performance and between one performance and another. In this example, the first post has been published in the 19th week of the sale period; whereas the first performance took place in the 23rd week. Clearly, the sale period ends with the last performance of that production scheduled.

The fixed effect model implies that any changes in the dependent variable must be due to influences other than these fixed characteristics. So, the parameters estimated of the FB data (that varies over time) are interpreted as influence effect as the performance-specific characteristics cancel (being them invariant over time).
In this case we aim to estimate the following equation:

$$(y_{it} - \bar{y_i}) = \beta'(A'_{SMit} - \bar{A_{SMi}}) + \gamma'(X'_{it} - \bar{X_i}) + (\epsilon_{it} - \bar{\epsilon_i}) \tag{1}$$

where y_{it} is the number of attenders for the performance i run the t^{th} time, A'_{SM} is a vector denoting the social media data, X'_{it} is the vector of control variables that are time-variant, ϵ_{it} is the error term and the overline represent the average values. In the random effect model, on the contrary, it is possible to include time invariant variables, as it is assumed that differences across entities have some influence on your dependent variable. In this case, the coefficients of the FB data are interpreted as both prediction and influence effect.

In this case, the equation to estimate is the following:

$$y_{it} = \alpha + \beta' A_{SMit} + \gamma' X_{it} + \delta' Z_i + \epsilon_{it} \tag{2}$$

where Z_i denotes the time-invariant variables.
As social media activities we consider:

- Total n^o *of likes* on facebook post related to the performance i in week t
- Total n^o *of comments* on facebook post related to the performance i in week t
- Total n^o *of post* related to the performance i in week t

Clearly, the total n° of likes and comments depend on the n° of post. Hence, we consider the ratios n° of likes / n° of post and n° of comments / n° of post, denoting the n° of likes and comments per post. However, the FB post can be considered as a "marketing" action undertaken by the theatre that can affect the n° of ticket sold. Hence in a second model we consider, besides the main effect of the three variables, also the interaction terms between n° of likes- comments and n° of post. In this way, we verify how the effect of FB likes and comments differs according to the n° of post, and viceversa.
We do not perform a sentiment analysis on the text of the comments: an analysis of the text not reported in this paper has shown that Facebook followers tend to express mainly positive opinion ; in some cases neutral opinions but never a negative opinion: we suspect that possible negative opinions are shared through word of mouth rather than published in the Theatre's Facebook page. So, including sentiment analysis in this paper would be misleading. We can thus consider each like and comment as a positive word of mouth.
As time-variant variables we consider: a dummy $PRESALE$ equal to 1 if the tickets are sold during the presale period; $WEEKSALE$, a variable that consider the number of weeks passed since the beginning of the sale period; $PERF_t$, that denotes the number of performances that take place in the week considered, and $PERF_{t+1}$ that considers the number of performances scheduled in the following week: indeed we have verified that, for a single performance, there is an increase of the number of tickets sold in the last two weeks. Finally, the variable $REMAINING$ denotes the number of remaining performances of a given production.
As time-invariant variables we consider the production characteristics: two dummies that identify the genre ($OPERA$ and $BALLET$, with $PLAY$ as base variable); we include the dummy variable $DANISH$, denoting Danish productions and $INTERNATIONAL$ denoting International productions ($MIXED$ productions as

base variable); the dummy variable $NEWDKT$ controls for productions that take place for the first time at Royal Danish Theater. We also control for the year in which the production was created by introducing three dummies: 1920-1965, 1850-1919, $BEFORE$ 1850 (where $AFTER$ 1965 is the base variable). As our analysis is based on performances running throughout 6 years, we include a time trend variable t. Finally $PRICE$ is the average price (in Danish crown) of a standard ticket charged for the production.
Table 3 provides a descriptive statistics of the data.

Variable	**Mean**	**SD**	**Min**	**Max**
Ticket sold	264.124	378.402	0	4552
n° of likes	16.635	90.272	0	2004
n° of comments	2.868	34.165	0	1059
n° of post	0.201	0.614	0	6
$PRESALE$	0.053	0.224	0	1
$WEEKSALE$	16.371	11.711	0	55
$PERF_t$	0.457	1.153	0	10
$PERF_{t+1}$	0.457	1.153	0	10
$REMAINING$	12.984	9.942	1	47
$OPERA$	0.398	0.489	0	1
$BALLET$	0.382	0.486	0	1
$DANISH$	0.209	0.407	0	1
$INTERNATIONAL$	0.668	0.471	0	1
$NEWDKT$	0.546	0.498	0	1
1920-1965	0.134	0.341	0	1
1850-1919	0.438	0.496	0	1
$BEFORE$ 1850	0.137	0.343	0	1
t	3.979	1.678	1	6
$PRICE$	435.22	141.22	81.29	723.33

3045 observations

Table 3: Descriptive statistics

4.3 Simultaneous equation system

In this paragraph we model the eWOM effect considering the interdependence between FB activities and ticket sold. This approach is adapted by the conceptual framework proposed by Duan *et al.* (2008) in the movie context. We assume that eWOM (in the forms of FB "likes" and comments) influence the current amount of tickets sold. However, also people who have bought a ticket for a given performance may contribute to positive feedback through social media activities, which in turn leads to an increase in the amount of ticket sold.
First, in a single-fixed effect equation that relates the amount of ticket sold with

social media effect, we verify whether there is a simultaneity issue, i.e whether the variables n^o *of likes* and n^o *of comments* are endogenous. From the Durbin-Watson test it results that only n^o *of likes* is endogenous (*p-value* < 0.001); whereas n^o *of comments* is not.
Hence, we model this framework through a two equation system:

$$\text{Ticket sold}_{it} = \alpha_1 + \beta_1 \text{n}^\text{o}\text{of likes}_{it} + \beta_2 \text{n}^\text{o}\text{of comments}_{it} + \gamma' X_{it} + \mu_i + \epsilon_{it} \quad (3)$$

$$\text{N}^\text{o}\text{of likes}_{it} = \alpha_2 + \eta' \text{POST TYPE}'_{it} + \theta \text{ticket sold}_{it} + \delta_1 \text{FB Week} + \delta_2 \text{CUMPERF}_{it} + \lambda_i + \upsilon_{it} \quad (4)$$

In equation (3) X is the same vector of time-variant characteristics as in (1). The theatre productions factors, which are time-invariant, are controlled by the fixed effect, so production-specific dummy variables are included in both equations.
In equation (4), we relate the n^o *of likes* in time t to the amount of ticket sold in the same week, the amount of performances of the given productions that already had taken place ($CUMPERF$) and the n^o of performance in the current week. In addition, we control also for the amount of different types of post. In particular, we distinguish 5 types of FB post: link to a web page, status, photo, video and events. Finally, we assume that the total amount of likes are related to the n^o of weeks passed since the FB page has been opened by the theatre (FB Week): indeed, the n^o of FB followers (so the potential n^o of people who potentially can give a "like") normally increases during time.
The three-stage least-square (3SLS) procedure is used to simultaneously estimate the system of two equations.

5 Results and discussion

Table 4 shows the estimation of both panel data models, estimated with fixed and random effect.
Results show some expected findings: the amount of tickets sold is lower during the presale period, and increases over time during the sale period (coefficient of the variable $WEEKSALE$ is positive and statistically significant). As expected, the n^o of performances scheduled in the current and in the following week have a positive impact on the number of ticket sold, as well as the n^o of performances remaining.
What interest us more, for the purpose of the paper, are the coefficient of the social media activities. We can notice that the n^o of likes have a positive and highly significant effect on the amount of ticket sold, both in the fixed and random effect model, suggesting that this variable has both an influencer and a predictor role. In

Variable	Fixed effect		Random effect	
	Model 1	**Model 2**	**Model 1**	**Model 2**
N° of likes	-	0.444****	-	0.498****
		(0.139)		*(0.141)*
N° of post	-	66.678****	-	75.811****
		(10.98)		*(11.05)*
N° of comments	-	−2.754****	-	−2.785****
		(0.768)		*(0.780)*
N° of likes x N° of post	-	−0.155**	-	−0.174***
		(0.062)		*(0.063)*
N° of comments x N° of post	-	2.507****	-	2.523****
		(0.739)		*(0.751)*
Likes per post	0.809****	-	0.893****	
	(0.110)		*(0.111)*	
Comments per post	−0.243	-	−0.254*	
	(0.148)		*(0.150)*	
$PRESALE$	−281.85****	−279.76****	−250.73****	−248.74****
	(27.99)	*(27.66)*	*(27.27)*	*(26.91)*
$WEEKSALE$	5.675****	5.161****	3.337****	2.898****
	(0.600)	*(0.598)*	*(0.573)*	*(0.568)*
$PERF_t$	86.255****	74.733****	87.602****	76.023****
	(5.821)	*(5.927)*	*(5.751)*	*(5.829)*
$PERF_{t+1}$	67.472****	61.865****	73.678****	67.394****
	(5.297)	*(5.275)*	*(5.339)*	*(5.358)*
$REMAINING$	28.739****	26.367****	24.696****	22.789****
	(1.438)	*(1.455)*	*(1.236)*	*(1.239)*
$OPERA$	-	-	284.37****	238.09****
			(50.85)	*(50.34)*
$BALLET$	-	-	353.67****	321.65****
			(41.10)	*(40.61)*
$DANISH$	-	-	109.36**	104.02**
			(47.14)	*(46.34)*
$INTERNATIONAL$	-	-	126.10***	126.29***
			(44.78)	*(44.01)*
$NEWDKT$	-	-	−63.288**	−66.755***
			(24.82)	*(24.39)*
1920-1965	-	-	−98.689**	−99.057**
			(42.67)	*(41.93)*
1850-1919	-	-	$-14,664$	−14.989
			(41.56)	*(40.85)*
$BEFORE$ 1850	-	-	−95.914**	−92.975**
			(46.87)	*(46.06)*
t	-	-	1.397	0.877
			(7.387)	*(7.257)*
$PRICE$	-	-	0.146	0.199
			(0.147)	*(0.144)*
Intercept	−266.87****	−229.98****	−513.07	−473.91****
	(24.91)	*(24.91)*	*(67.85)*	*(66.87)*
R^2 within	0.321	0.337	0.316	0.333
R^2 overall	0.225	0.245	0.372	0.392
N° of observations				3045
N° of groups				128

$^{****}p < 0.001$ $^{***}p < 0.01$

$^{**}p < 0.05$ $^{*}p < 0.10$

Robust standard errors in italics under
the estimated coefficientsare the

Table 4: Estimation results of panel data

particular, the effect of the n° of like per post (Model 1) is positive, as well as the main effect of the n° of likes (Model 2). However, the effect of the total n° of likes decreases as the n° of post increases: this is not surprising, as it is reasonable to suppose that the effect of many "likes" achievable with few post is greater than the same amount of "like" achievable with many post. The magnitude of the coefficient of the main term is almost three times than the coefficient of the interaction term. Surprisingly we have the opposite effect when we consider the effect of FB comments: in the first model, the n° of comments per post is not statisically significant (in the fixed effect model) or negative but poorly significant (in the random-effect model). In the second model, the effect of the total n° of comment is negative and statistically significant, however its effect increases as the n° of post increase. As the coefficient of the main term is almost equal, the effect of the n° of comments is slightly negative when there is only one post published during the week, and positive when there are two or more post published.

Table 5 shows the estimation of the 3SLS model.

	Coefficient	**SE**
Ticket sold equation		
N° of likes	0.855****	0.089
N° of comments	-0.049	0.146
$PRESALE$	-271.15****	27.072
$WEEKSALE$	5.253****	0.578
$PERF_t$	81.987****	5.778
$PERF_{t+1}$	59.897****	5.131
$REMAINING$	26.442****	1.440
$Intercept$	-331.40****	65.328
R^2	51.65	
Like equation		
$Events$	-59.527	38.808
$Link$	26.205***	8.217
$Status$	28.398****	4.309
$Photo$	97.681****	3.960
$Video$	158.73****	5.007
$Ticket\ sold$	0.016***	0.006
$FB\ Week$	0.043***	0.014
$CUM\ PERF$	0.361	0.322
$Intercept$	-15.684****	4.412
R^2	45.40	

****$p < 0.001$ ***$p < 0.01$
**$p < 0.05$ *$p < 0.10$

Table 5: Estimation results of the 3SLS model

For the ticket sold equation we note that, similarly to the panel data model, the n°

of likes is a significant predictor. The coefficient related to the n° of comments is not statistically significant: considering this model together with the panel data models, it seems that the FB likes are much more important in understanding theater demand rather than the FB comments. This is due probably to the fact that, as we can notice from the descriptive statistics, FB users prefer to give a "like" rather than write a comment. The sign of the other coefficients of the first equation are the same than in the panel data model: the amount of tickets sold is lower during the presale period and increases over time during the sale period; in addition there is a positive relationship between the amount of tickets sold and the n° of performances in the current and in the following week.
Concerning the second equation, we can notice that some types of post attract a higher amount of likes than other ones: in particular, "videos" and "photos" bring more "likes" than "status" and "link"; whereas the "event" post is not statistically significant. The amount of tickets sold is a significant predictor of the n° of likes, confirming their interdependence. As expected, there is a positive relationship between the n° of weeks passed since the FB page has been opened and the n° of likes.

6 Conclusion

In this paper we show that Facebook data can constitute a new effective tool for understanding theatre demand. This has crucial implications as it allows a further comprehension of audience preference and can constitute a mean by which assist theatre management as well as theatre policy in taking decision concerning, for example, pricing and organization of the theatre season programme.
In particular, the Facebook "likes" given to the Facebook post have both a prediction and an influence role, whereas the Facebook "comments" have a small impact on theater demand. In addition, there is an interdependence between the n° of "likes" and the amount of tickets sold.
The adoption of social media data in understanding theatre demand is not as effective as it can be in other applications area, characterized by the supply of products with a global (or almost) market and audience, such as Iphone, cars and movies. However, it represents an additional instrument that allows to collect easily a higher amount of information concerning the audience impression of a theater performance. Given the growing usage of social media by consumers, we believe that this stream of research will benefit from further development.

References

- Abbé-Decarrroux, F. (1994): The Perception of Quality and the Demand for Services - Empirical Application to the Performing Arts, *Journal of Economic Behavior and Organization*, vol. 23, pp 99-197.

- Akdede, S. H., King, J. T. (2006). Demand for and productivity analysis of Turkish public theater. *Journal of Cultural Economics*, 30(3), 219-231.

- Asur, S., Huberman, B. A. (2010). Predicting the future with social media. *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference* Vol. 1, pp. 492-499. IEEE.

- Baldin, A., Bille, T. (2017) *Theater Management - quality indicators and demand*, Paper presented at the 8th European Workshop on Applied Cultural Economics, Cracow, Poland.

- Bille Hansen, T. (1991): *Det Kgl. Teater Ð et kultur£konomisk studie*, AKF Forlaget, Copenhagen.

- Caves, R. (2000) *Economics of the Creative Industries*, Cambridge MA, Harvard University Press.

- Chu, P.C., Liu, Y.H., Sher, C.Y (2016) *The impact of amateur film reviews on movie theater attendance: predictors versus influencers*, Paper presented at the 19th International Conference on Cultural Economics, Valladolid, Spain.

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

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

- Duan, W., Gu, B., Whinston, A. B. (2008). Do online reviews matter?ÑAn empirical investigation of panel data. *Decision support systems*, 45(4), 1007-1016.

- Grisolia, J. M., Willis, K. G. (2012). A latent class model of theatre demand. *Journal of Cultural Economics*, 36(2), 113-139.

- Hussain, A., Vatrapu, R. (2014a). Social data analytics tool: Design, development, and demonstrative case studies. In *Enterprise Distributed Object Computing Conference Workshops and Demonstrations (EDOCW), 2014 IEEE 18th International* (pp. 414-417). IEEE.

- Hussain, A., Vatrapu, R. (2014b). Social data analytics tool (sodato). In *International Conference on Design Science Research in Information Systems* (pp. 368-372). Springer, Cham.

- Jenkins, S., Austen-Smith, D. (1987). Interdependent decision-making in non-profit industries: A simultaneous equation analysis of English provincial theatre. *International Journal of Industrial Organization*, 5(2), 149-174.

- Krebs, S., Pommerehne, W. W. (1995). Politico-economic interactions of German public performing arts institutions. *Journal of Cultural Economics*, 19(1), 17-32.

- Lassen, N. B., Madsen, R., Vatrapu, R. (2014).Predicting iphone sales from iphone tweets. *Enterprise Distributed Object Computing Conference (EDOC)*, 81-90.

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

- Throsby, C.D., Nielsen, E. (1980): *Product Quality Decisions in Nonprofit Performing Arts Firms*, Macquarie University, School of Economic and Financial Studies, Research paper no. 215

- Throsby, D.C. (1990): Perception of Quality in Demand for Theatre, *Journal of Cultural Economics*, vol. 14, pp 65-82.

- Tobias, S. (2004). Quality in the performing arts: Aggregating and rationalizing expert opinion. *Journal of Cultural Economics*, 28(2), 109-124.

- Urrutiaguer, D. (2002): Quality Judgements and Demand for French Public Theatre, *Journal of Cultural Economics*, 26, 185-202.

- Vatrapu, R., Hussain, A., Lassen, N. B., Mukkamala, R. R., Flesch, B., Madsen, R. (2015). Social set analysis: four demonstrative case studies. In *Proceedings of the 2015 International Conference on Social Media & Society* (p. 3). ACM.

- 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(3), 167-183.

• Zieba, M. (2009). Full-income and price elasticities of demand for German public theatre. *Journal of Cultural Economics,* 33(2), 85-108.