

Descriptive and predictive models of consumer demand for Perm opera

Alina Ozhegova Evgeniy M. Ozhegov

National Research University Higher School of Economics, Perm
Research Group for Applied Markets and Enterprises Studies (GAMES)

15.11.2016
GAMES Research Seminar

Introduction

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Perm Opera and Ballet Theatre

- ▲ One of Russia's oldest theatres (Foundation in 1870).
- ▲ The theater has won the variety of awards (Golden Masks).
- ▲ The theatre is organizer of the International Diaghilev Festival and the Ekaterina Maximova Arabesque Ballet Competition.
- ▲ Since 2011 Teodor Currentzis is the theatre's Artistic Director.

Introduction

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Assumptions for Research Question

- ▲ The theater is nonprofit.
- ▲ The goal is to attract the residents to cultural life.
- ▲ The revenue from performances should cover at least 10% of expenditures.
- ▲ Objective function: maximization of revenue and maximization of attendance.
- ▲ The seats in a house are heterogeneous (price, quality of view, sound, prestige).
- ▲ Consumer demands for a particular seat (seating area) and for performance.
- ▲ Price varies over seats, performance and play characteristics.

Introduction

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Goal to

- ▲ Estimate the demand function for performing arts
- ▲ Increase the efficiency of pricing mechanism over seats and performances

Research Question

- ▲ Do the effects of a price and characteristics vary across the performances and the seats in a house?

Literature Review

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Studies on stated preferences

- ▲ Consumer preferences (Bille-Hansen, 1997);
- ▲ Utility function (Grisolia & Willis, 2012);
- ▲ Consumer surplus (Train, 2003);
- ▲ Willingness to pay (Levy-Garboua & Montmarquette, 1996);
- ▲ Patron (Baumol & Bowen, 1966).

Literature Review

Opera
Demand

Introduction

**Literature
Review**

Data

Methodology

Results for
explanatory
model

Methodoly
cont.

Results for
predictive
model

Comparison

Conclusion

Studies on revealed preferences

- ▲ Price elasticity of demand (Moore, 1966; Throsby & Withers, 1979);
- ▲ Income elasticity of demand (Greckel & Felton, 1987; Felton, 1994);
- ▲ Cross price elasticity of demand (Withers, 1980; Gapinski, 1984);
- ▲ The effect of quality (Throsby, 1983; Corning & Levy, 2002).

Data Description

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

- ▲ Data Source: Sales System of Perm Opera and Ballet Theatre
- ▲ 4 seasons (2011-2012/2014-2015)
- ▲ 985 performances
- ▲ 170 unique productions
- ▲ 680000 sales operations

Data Description

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

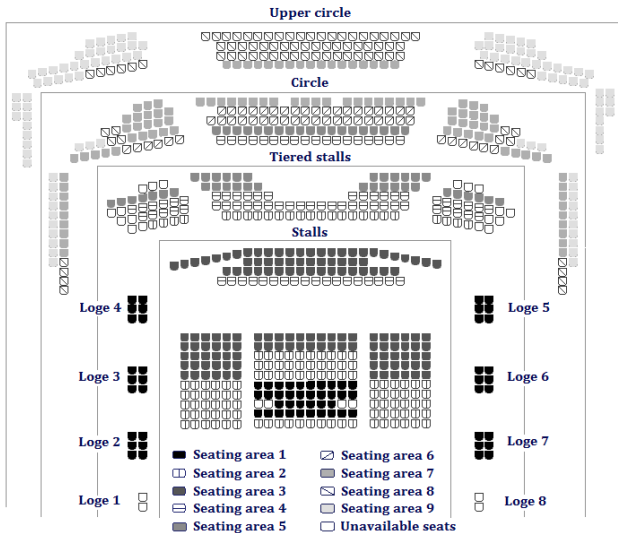
Results for
predictive
model

Comparison

Conclusion

- ▲ The name of performance;
- ▲ time: season, year, month, day of week, hour;
- ▲ the basic price of a ticket;
- ▲ the sector (loge, the stalls, tiered stalls, circle, upper circle);
- ▲ row and seat;
- ▲ seating area.

Scheme of a house



Data description. Characteristics of performances

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

- ▲ Type of performance (opera/ballet);
- ▲ the date of writing (classical/modern);
- ▲ the author (the rating, the nationality);
- ▲ the world rating of production;
- ▲ the date of premiere in a theatre;
- ▲ the duration of performance;
- ▲ the director and band director of production;
- ▲ chorus-master/ballet-master;
- ▲ the number of awards;
- ▲ age recommended for attendance.

Descriptive statistics: performance characteristics

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Obs	Share
Type of performance	2682	
Ballet	954	36.6
Opera	1728	64.4
The date of writing	2682	
Before 1990	2304	85.9
1990 and later on	378	14.1
Recommended age group	2682	
From 0	1107	41.3
From 12	1170	43.6
From 16	405	15.1
The presence of nomination in Golden Mask	2682	
Presence	386	14.4
Absence	2196	85.6
The presence of winning in Golden Mask	2682	
Presence	144	5.4
Absence	2538	96.6
The author's nationality	2682	
Russian	1521	56.7
Other	1161	43.3

Descriptive statistics: performance characteristics

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Description	Obs	Mean	St.dev.	Min	Max
Length	Length in minutes	1971	152.9	46.60	60	230
Antracts	The number of antracts	1971	1.60	0.6	1	3
Acts	The number of acts	2268	2.7	0.66	1	4
Rating of opera	1/(rating of opera)	2682	0.08	0.22	0.01	1
Rating of composer	1/(rating of composer)	2682	0.09	0.21	0.01	1
Rating of ballet	1/(rating of ballet)	2682	0.09	0.22	0.01	1

Descriptive statistics: price

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Obs	Mean	St.dev.	Min	Max
Basic price	2682	412	381	100	2000
Basic price (area1)	298	903.4	503.2	300	2000
Basic price (area2)	298	619.8	655.1	250	1400
Basic price (area3)	298	525.2	364.8	210	1300
Basic price (area4)	298	462.1	336.9	180	1200
Basic price (area5)	298	378.1	278.3	160	1000
Basic price (area6)	298	299.4	222.2	140	800
Basic price (area7)	298	239.9	158.6	120	600
Basic price (area8)	298	180.4	94.5	110	400
Basic price (area9)	298	100	0	100	100

Descriptive statistics: attendance rate

Opera
Demand

Introduction

Literature
Review

Data

Methodology

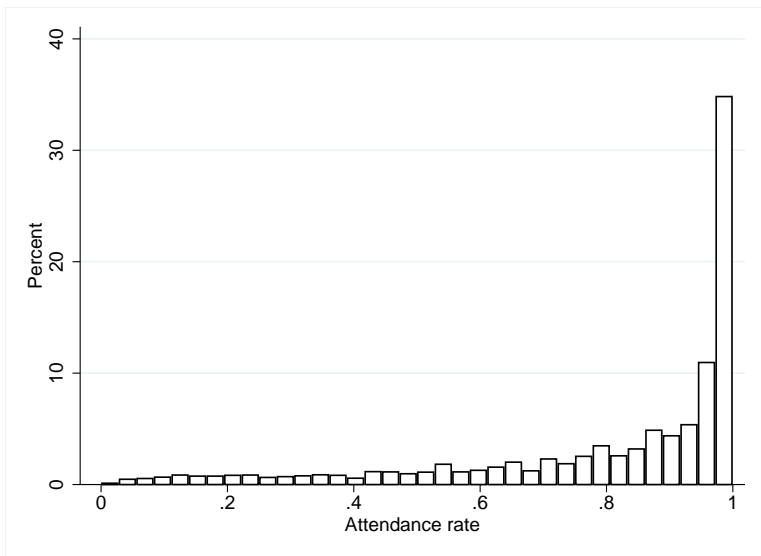
Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion



Descriptive statistics: attendance rate

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Obs	Mean	St.dev.	Min	Max
Attendance rate	2682	0.80	0.25	0	1
Attendance (area1)	298	0.85	0.16	0.11	1
Attendance (area2)	298	0.89	0.14	0.35	1
Attendance (area3)	298	0.89	0.15	0.35	1
Attendance (area4)	298	0.90	0.15	0.11	1
Attendance (area5)	298	0.84	0.21	0.11	1
Attendance (area6)	298	0.80	0.25	0.06	1
Attendance (area7)	298	0.70	0.32	0.02	1
Attendance (area8)	298	0.65	0.34	0	1
Attendance (area9)	298	0.72	0.31	0	1

Descriptive statistics: attendance rate

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Operas	Ballets	Total
Attendance rate	0.72	0.96	0.80
Attendance (area1)	0.83	0.92	0.85
Attendance (area2)	0.87	0.96	0.89
Attendance (area3)	0.85	0.96	0.89
Attendance (area4)	0.86	0.97	0.90
Attendance (area5)	0.76	0.98	0.84
Attendance (area6)	0.68	0.98	0.80
Attendance (area7)	0.52	0.97	0.70
Attendance (area8)	0.46	0.95	0.65
Attendance (area9)	0.64	0.87	0.72

Descriptive statistics: attendance rate

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Operas	Ballets	Total
Attendance rate	0.72	0.96	0.80
Price for area1 = 300	0.79		0.79
Price for area1 = 500	0.70	0.98	0.72
Price for area1 = 600		1.00	1.00
Price for area1 = 700	0.67	0.93	0.68
Price for area1 = 800	0.59	0.95	0.92
Price for area1 = 1000	0.79	0.96	0.90
Price for area1 = 1500		0.97	0.97
Price for area1 = 2000	0.88	0.94	0.92

Descriptive statistics: additional variables

Variable	Mean	St.D.	Min	Max	Play with max
<i>Operas</i>					
Plays in prev. year	2.2	2.4	0	13	Малахитовая шкатулка
Plays in ± 15 days	1.1	1.2	0	8	Малахитовая шкатулка
Plays in ± 7 days	1.0	1.0	0	4	Малахитовая шкатулка
Mean attendance	608	92	372	841	Свадьба Фигаро Севильский цирюльник
Attended	2078	2134	0	10758	Севильский цирюльник
<i>Ballets</i>					
Plays in prev. year	3.1	3.2	0	14	Щелкунчик
Plays in ± 15 days	1.4	1.7	0	7	Щелкунчик
Plays in ± 7 days	1.3	1.7	0	7	Щелкунчик
Mean attendance	809	35	695	839	Ромео и Джульетта
Attended	3295	3137	0	14068	Щелкунчик

Data Analysis

Opera
Demand

Introduction

Literature
Review

Data

Methodology

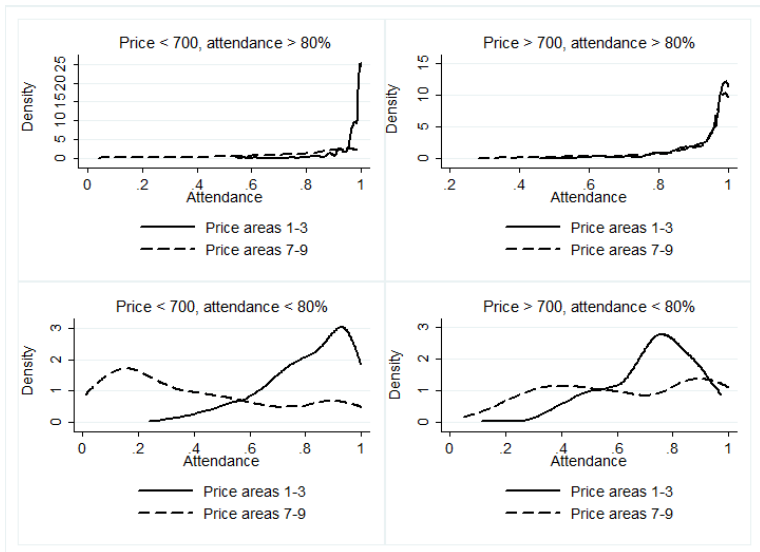
Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion



Methodology

Model of censored quantile regression (Chernozhukov & Hong, 2011):

$$Q_{y_{ij}^*|x_{ij},p_{ij}}(\alpha) = x_{ij}\beta(\alpha) + p_{ij}\gamma(\alpha),$$
$$Q_{y_{ij}|x_{ij},p_{ij}}(\alpha) = \begin{cases} Q_{y_{ij}^*|x_{ij},p_{ij}}(\alpha), & y_{ij}^* \leq 1 \\ 1, & y_{ij}^* > 1 \end{cases}, \quad (1)$$

where

y_{ij} is the observed demand on performance i in seating area j ;

y_{ij}^* is a potential demand on performance i in seating area j ;

$Q(\alpha)$ is a conditional quantile function of level α ;

α is a level of quantile, $\alpha \in [0..1]$;

p_{ij} is the price of a ticket on performance i in a j -th seating area;

x_{ij} are the characteristics of performance i in a j -th seating area.

Estimation procedure

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Step 1. Estimation of probability to be censored for each observation

$$\hat{P}_{ij} = E[Prob(y_{ij} > 1) | x_{ij}, p_{ij}], \forall (ij)$$

Step 2. Estimation of model on the sample J_0

$$\check{y}_{ij} = Q_{y_{ij}|x_{ij},p_{ij}}(\alpha), \text{ where } (ij) \in J_0$$

$$J_0 = \{(ij) | \hat{P}_{ij} \leq F_{0.9}^{-1}(\hat{P}_{ij} < (1 - \alpha))\}$$

Step 3. Final estimation of model on the sample J_1

$$\hat{y}_{ij} = Q_{y_{ij}|x_{ij},p_{ij}}(\alpha), \text{ where } (ij) \in J_1$$

$$J_1 = \{(ij) | \check{y}_{ij} \leq F_{0.97}^{-1}(\check{y}_{ij} \leq 1)\}$$

Results. Regressions

Variable	OLS	Median regression	Censored median regression
Basic price/100	-0.022*** (0.002)	-0.015*** (0.003)	-0.028*** (0.004)
The russian author	0.060*** (0.010)	0.052*** (0.017)	0.090*** (0.018)
Premiere	0.110*** (0.014)	0.112*** (0.023)	0.154*** (0.026)
Rating of opera	0.034 (0.027)	0.068 (0.044)	0.020 (0.044)
Rating of ballet	0.107*** (0.023)	0.072* (0.038)	0.252*** (0.054)
Type: Ballet	0.333*** (0.013)	0.257*** (0.021)	0.414*** (0.024)
Number of awards in GM	0.045*** (0.011)	0.053*** (0.019)	0.059*** (0.020)
Band director T.Currentzis	0.039** (0.019)	0.020 (0.031)	0.054* (0.032)
Recommended age: from 12 y.o.	0.039*** (0.011)	0.007 (0.018)	0.043** (0.019)
Recommended age: from 16 y.o.	-0.078*** (0.018)	-0.128*** (0.030)	-0.098*** (0.030)
The time of day: evening	-0.028** (0.014)	-0.010 (0.023)	-0.026 (0.026)
Constant	0.793*** (0.025)	0.813*** (0.041)	0.847*** (0.047)
Number of observations	2682	2682	2682
Number of parameters	35	35	35
R^2	0.467	0.452	0.502

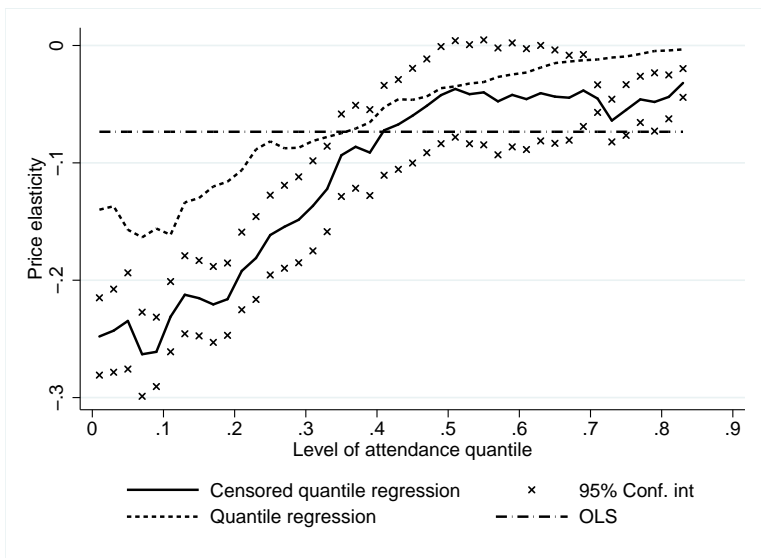
Results. Censored quantile regression on different quantiles

Variable	$\alpha=0.2$	$\alpha=0.4$	$\alpha=0.6$	$\alpha=0.8$
Basic price/100	-0.036*** (0.003)	-0.031*** (0.003)	-0.026*** (0.004)	-0.012*** (0.002)
The russian author	0.104*** (0.016)	0.107*** (0.016)	0.086*** (0.020)	0.046*** (0.011)
Premiere	0.182*** (0.022)	0.174*** (0.023)	0.135*** (0.029)	0.056*** (0.017)
Rating of opera	0.105*** (0.040)	0.071* (0.040)	0.018 (0.049)	0.013 (0.029)
Rating of ballet	0.267*** (0.045)	0.195*** (0.046)	0.233*** (0.060)	0.157*** (0.049)
Type: Ballet	0.533*** (0.021)	0.466*** (0.022)	0.363*** (0.028)	0.135*** (0.015)
The number of awards in GM	0.063*** (0.017)	0.074*** (0.017)	0.030 (0.021)	0.015 (0.016)
Band director T. Currentzis	0.009 (0.029)	0.036 (0.029)	0.008 (0.035)	-0.004 (0.021)
Recommended age: from 12 y.o.	0.049*** (0.017)	0.038** (0.017)	0.048** (0.022)	0.017 (0.012)
Recommended age: from 16 y.o.	-0.095*** (0.027)	-0.117*** (0.027)	-0.046 (0.033)	-0.030 (0.019)
The time of day: evening	-0.068*** (0.022)	-0.018 (0.023)	-0.018 (0.029)	-0.003 (0.016)
Constant	0.674*** (0.040)	0.795*** (0.042)	0.891*** (0.053)	0.973*** (0.029)
Number of observations	2682	2682	2682	2682
Number of parameters	35	35	35	35

Results. Censored quantile regression on different quantiles

Variable	$\alpha=0.2$	$\alpha=0.4$	$\alpha=0.6$	$\alpha=0.8$
Log. of price	-0.222*** (0.018)	-0.131*** (0.019)	-0.070*** (0.025)	-0.032*** (0.012)
The russian author	0.081*** (0.015)	0.101*** (0.016)	0.067*** (0.019)	0.039*** (0.010)
Premiere	0.146*** (0.021)	0.182*** (0.023)	0.108*** (0.030)	0.040** (0.017)
Rating of opera	0.122*** (0.037)	0.093** (0.039)	0.045 (0.048)	0.038 (0.024)
Rating of ballet	0.269*** (0.040)	0.170*** (0.042)	0.159*** (0.056)	0.171*** (0.045)
Type: Ballet	0.535*** (0.021)	0.421*** (0.023)	0.266*** (0.028)	0.092*** (0.014)
The number of awards in GM	0.007 (0.016)	0.057*** (0.016)	0.033 (0.021)	0.018* (0.011)
Band director T. Currentzis	0.020 (0.026)	0.034 (0.029)	0.058 (0.036)	0.007 (0.018)
Age recommended: from 12 y.o.	0.024 (0.016)	0.057*** (0.017)	0.066*** (0.021)	0.029*** (0.010)
Age recommended: from 16 y.o.	-0.099*** (0.025)	-0.132*** (0.025)	-0.105*** (0.031)	-0.063*** (0.015)
Time of day: evening	-0.052** (0.021)	-0.000 (0.021)	-0.010 (0.026)	-0.015 (0.014)
Constant	1.785*** (0.109)	1.391*** (0.118)	1.204*** (0.152)	1.121*** (0.074)
Number of observations	2682	2682	2682	2682
Number of parameters	35	35	35	35

Results. Comparison of marginal effects



Endogeneity of Price

Opera
Demand

Introduction

Literature
Review

Data

Methodology

**Results for
explanatory
model**

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

- ▲ Possible correlation between unobservable performance quality and price
- ▲ The variation of price within the performance as an instrument for the price (Hausman & Taylor, 1981)
- ▲ Two-step estimation procedure (Chernozhukov, Kowalski & Fernandez-Val, 2015)
 - ▶ Regression of price on within price and performance characteristics
 - ▶ Regression of attendance rate on price, performance characteristics and residuals

Endogeneity of Price. Estimation

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Step 1. Regression of price on within price and performance characteristics

$\hat{p}_{ijk} = Q_{p_{ijk}|\tilde{p}_{ijk}, x_{ijk}}(\alpha)$, where $\tilde{p}_{ijk} = p_{ijk} - \bar{p}_{.jk}$

$\hat{e}_{ijk} = p_{ijk} - \hat{p}_{ijk}$ - residuals

p_{ijk} - the price on play i in seating area j for performance k

\tilde{p}_{ijk} - price within, an instrument for p_{ijk} ,

$\bar{p}_{.jk}$ - the average price over plays on performance k in seating area j .

Step 2. Regression of attendance rate on price, performance characteristics and residuals from the first step

$\hat{y}_{ijk} = Q_{y_{ijk}|p_{ijk}, x_{ijk}, \hat{e}_{ijk}}(\alpha)$

Endogeneity of Price. Results

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

	$\alpha=0.3$	$\alpha=0.3$	$\alpha=0.5$	$\alpha=0.5$	$\alpha=0.7$	$\alpha=0.7$
	CQIV	CQR	CQIV	CQR	CQIV	CQR
Basic price	-0.041*** (0.003)	-0.040*** (0.003)	-0.038*** (0.004)	-0.037*** (0.004)	-0.031*** (0.004)	-0.030*** (0.004)
$\hat{\epsilon}$	-0.000 (0.000)	-	0.000 (0.000)	-	0.000 (0.000)	-
N	2384	2682	2384	2682	2384	2682
K	35	35	35	35	35	35

Note: standard errors are in parenthesis. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. In the model we also include as control variables: the nationality of author, the rating and type of performance, the premier year, the band director, the recommended age, the time of day, the number of Golden Mask, the year and month of play and seating area dummies.

Brief summary

Opera
Demand

Introduction

Literature
Review

Data

Methodology

**Results for
explanatory
model**

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

- ▲ The necessity to employ the quantile regression
- ▲ The necessity to account for demand censoring
- ▲ Consumers are elastic by price on average
- ▲ Price elasticity is decreasing with the quantile in absolute values
- ▲ Robustness check has revealed an absence of endogeneity problem.

Methodology for prediction

Data
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

**Methodology
cont.**

Results for
predictive
model

Comparison

Conclusion

- ▲ Censored quantile regression:
 - + Accounts for heterogeneity of effects
 - + Accounts for censoring by calibrating on (probably) uncensored data
 - + More robust to outliers and nonnormal distributions of y and ϵ
 - Heterogeneity is based on dependent variable: useless when predict
- ▲ Regression (gradient stochastic) boosting:
 - + Implementable for any regression model
 - + Accounts for influential observations
 - (Was) no implemented algorithm for CQR
- ▲ Regression trees:
 - + Accounts for heterogeneity based on RHS variables
 - Hardly (but) implementable algorithm in Stata
 - Can not use many leaves when data is not so big
 - + May choose when to stop splitting using economic sense and analysis of predictive accuracy in leaves

Methodology. Idea of regression boosting

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

**Methodology
cont.**

Results for
predictive
model

Comparison

Conclusion

In a model

$$y = f + e$$

$\nabla L(e)$ is a gradient for $\sum_i L(y_i, f)$

Ex. $L(y_i, f) = (y_i - f)^2 = e_i^2$

Update f by a loss gradient helps to optimize loss function

Methodology. Idea of regression boosting

Set the initial guess for f :

$$f_0 = q_y(0.5)$$

For steps $m = 1..M$:

- ▶ calculate the residuals

$$e_i^{m-1} = y_i - f^{m-1}$$

- ▶ regress the loss gradient on explanatory variables

$$\nabla L(e_i^{m-1}) = g^m(x_i) + \eta_i$$

- ▶ predict the residuals

$$\hat{e}_i^{m-1} = \nabla^{-1}(\hat{g}^m(x_i))$$

- ▶ update f_{m-1} :

$$f_m = f_{m-1} + \hat{e}_i^{m-1}$$

Methodology. Idea of regression boosting

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

**Methodology
cont.**

Results for
predictive
model

Comparison

Conclusion

Cool ideas:

- ▲ go in a direction of gradient with a step length $0 < \lambda < 1$

$$f_m = f_{m-1} + \lambda \hat{e}_i^{m-1}$$

- ▲ use the stochastic sample of observations ($I_m \in \{0, 1\}$)

$$f_m = f_{m-1} + I_m \lambda \hat{e}_i^{m-1}$$

- ▲ train a model for residuals on a subsample of observations and do the cross-validation

Methodology. Idea of CQR boosting with heterogeneity

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

More cool ideas:

- ▲ use three-step CQR feature: train a model for residuals on the uncensored observations only
- ▲ split the sample on subsamples by
 - ▶ type of performance: operas and ballets
 - ▶ quality of seat: areas 1-5 and 6-9
 - ▶ quality of performance by price signal: price for area1 < 1000 and 1000+
- ▲ if convergence is not achieved in some subsample then unsplit the subsamples

Results. Prediction accuracy

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Operas	Ballets	Total
Accuracy	0.89	0.90	0.89
Accuracy (area1)	0.94	0.88	0.92
Accuracy (area2)	0.96	0.90	0.94
Accuracy (area3)	0.94	0.94	0.94
Accuracy (area4)	0.95	0.95	0.95
Accuracy (area5)	0.84	0.90	0.89
Accuracy (area6)	0.89	0.92	0.90
Accuracy (area7)	0.84	0.94	0.88
Accuracy (area8)	0.86	0.87	0.86
Accuracy (area9)	0.65	0.69	0.66

Results.Prediction accuracy

Opera
Demand

Introduction

Literature
Review

Data

Methodology

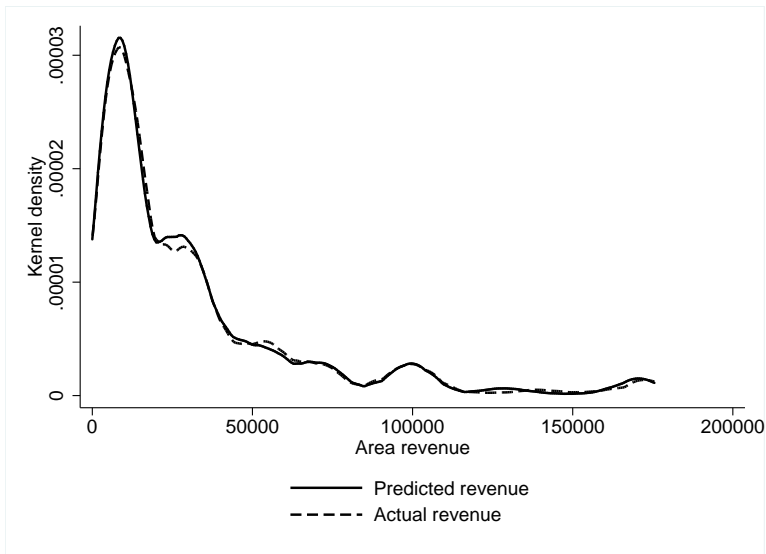
Results for
explanatory
model

Methodoly
cont.

**Results for
predictive
model**

Comparison

Conclusion



Results. Prediction accuracy

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

Variable	Training sample (80%)		Testing sample (20%)	
	Operas	Ballets	Operas	Ballets
Accuracy	0.89	0.90	0.86	0.87
Accuracy(area1)	0.95	0.89	0.92	0.77
Accuracy(area2)	0.97	0.89	0.94	1.00
Accuracy(area3)	0.95	0.92	0.95	0.89
Accuracy(area4)	0.96	0.94	0.88	1.00
Accuracy(area5)	0.86	0.97	0.90	1.00
Accuracy(area6)	0.89	0.94	0.88	0.90
Accuracy(area7)	0.90	0.96	0.84	0.88
Accuracy(area8)	0.87	0.91	0.89	0.80
Accuracy(area9)	0.67	0.70	0.50	0.61
Accuracy		0.89		0.86
R^2		0.758		0.809
RMSE		0.06		0.07
RMSE, %		8.5		9.4

Results. Price elasticity

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

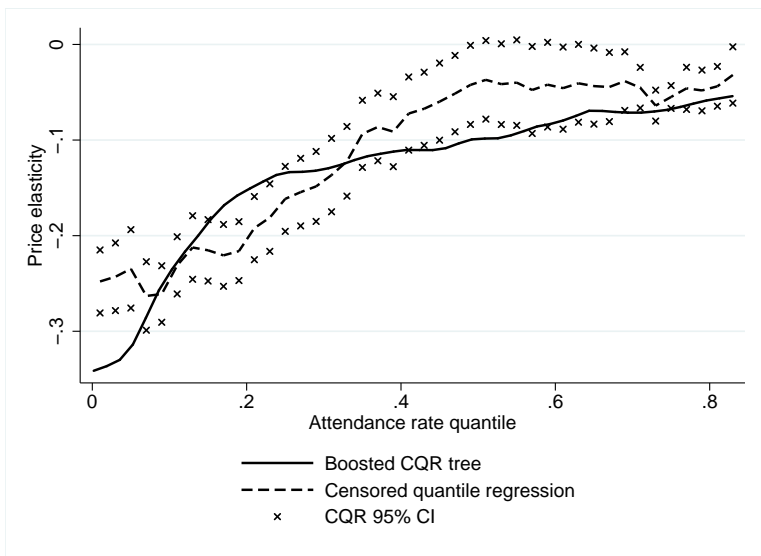
Results for
predictive
model

Comparison

Conclusion

Variable	Operas	Ballets	Total
Price elasticity	-0.172	-0.024	-0.124
Price elasticity (area1)	-0.076	-0.106	-0.087
Price elasticity (area2)	-0.119	-0.061	-0.104
Price elasticity (area3)	-0.136	-0.056	-0.116
Price elasticity (area4)	-0.083	-0.050	-0.075
Price elasticity (area5)	-0.211	-0.050	-0.171
Price elasticity (area6)	-0.309	-0.012	-0.147
Price elasticity (area7)	-0.162	-0.001	-0.057
Price elasticity (area8)	-0.327	-0.001	-0.114
Price elasticity (area9)	-0.451	-0.024	-0.179

Results. Comparison of models



Key results

Opera
Demand

Introduction

Literature
Review

Data

Methodology

Results for
explanatory
model

Methodology
cont.

Results for
predictive
model

Comparison

Conclusion

- ▲ Heterogeneous effect of price on attendance
- ▲ Model is helpful for:
 - ▶ "What if" analysis
 - ▶ Prediction of attendance
 - ▶ Prediction of latent attendance and price optimization